Mikhail Yu. Khachay, Natalia Konstantinova, Alexander Panchenko, Radhakrishnan Delhibabu, Nikita Spirin, Valeri G. Labunets (Eds.)

AIST'2015 — Analysis of Images, Social Networks and Texts

Supplementary Proceedings of the 4th International Conference on Analysis of Images, Social Networks and Texts (AIST'2015)
April 2015, Yekaterinburg, Russia

The proceedings are published online on the CEUR-Workshop web site in a series with ISSN 1613-0073, Vol-1452.

Preface

This volume contains proceedings of the fourth conference on Analysis of Images, Social Networks and Texts (AIST'2015)[1]. The first three conferences in 2012–2014 attracted a significant number of students, researchers, academics and engineers working on interdisciplinary data analysis of images, texts, and social networks.

The broad scope of AIST makes it an event where researchers from different domains, such as image and text processing, exploiting various data analysis techniques, can meet and exchange ideas. We strongly believe that this may lead to crossfertilisation of ideas between researchers relying on modern data analysis machinery. Therefore, AIST brings together all kinds of applications of data mining and machine learning techniques. The conference allows specialists from different fields to meet each other, present their work, and discuss both theoretical and practical aspects of their data analysis problems. Another important aim of the conference is to stimulate scientists and people from the industry to benefit from the knowledge exchange and identify possible grounds for fruitful collaboration.

The conference was held during April 9–11, 2015. Following an already established tradition, the conference was organised in Yekaterinburg, a cross-roads between European and Asian parts of Russia, the capital of Urals region.The key topics of AIST are analysis of images and videos; natural language processing and computational linguistics; social network analysis; pattern recognition, machine learning and data mining; recommender systems and collaborative technologies; semantic web, ontologies and their applications.

The Program Committee and the reviewers of the conference included well-known experts in data mining and machine learning, natural language processing, image processing, social network analysis, and related areas from leading institutions of 22 countries including Australia, Bangladesh, Belgium, Brazil, Cyprus, Egypt, Finland, France, Germany, Greece, India, Ireland, Italy, Luxembourg, Poland, Qatar, Russia, Spain, The Netherlands, UK, USA and Ukraine.

This year the number of submission has doubled and we have received 140 submissions mostly from Russia but also from Algeria, Bangladesh, Belgium, India, Kazakhstan, Mexico, Norway, Tunisia, Ukraine, and USA. Out of 140 only 32 papers were accepted as regular oral papers (24 long and 8 short). Thus, the acceptance rate of this volume was around 23%. In order to encourage young practitioners and researchers we included 5 industry papers to the main volume and 25 papers to the supplementary proceedings. Each submission was reviewed by at least three reviewers, experts in their fields, in order to supply detailed and helpful comments.

The conference also featured several invited talks and tutorials, as well as an industry session dedicated to current trends and challenges.

[1] http://aistconf.org/

Invited talks:

- Pavel Braslavski (Ural Federal University, Yekaterinburg, Russia), Questions Online: What, Where, and Why Should we Care?
- Michael Khachay (Krasovsky Institute of Mathematics and Mechanics UB RAS & Ural Federal University, Yekaterinburg, Russia), Machine Learning in Combinatorial Optimization: Boosting of Polynomial Time Approximation Algorithms.
- Valeri Labunets (Ural Federal University, Yekaterinburg, Russia), Is the Human Brain a Quantum Computer?
- Sergey Nikolenko (National Research University Higher School of Economics & Steklov Mathematical Institute, St. Petersburg, Russia), Probabilistic Rating Systems
- Andrey Savchenko (National Research University Higher School of Economics, Nizhny Novgorod, Russia), Sequential Hierarchical Image Recognition based on the Pyramid Histograms of Oriented Gradients with Small Samples
- Alexander Semenov (International laboratory for Applied Network Research at HSE, Moscow, Russia), Attributive and Network Features of the Users of Suicide and Depression Groups of Vk.com

Tutorials:

- Alexander Panchenko (Technische Universität Darmstadt, Germany), Computational Lexical Semantics: Methods and Applications
- Artem Lukanin (South Ural State University, Chelyabinsk, Russia), Text Processing with Finite State Transducers in Unitex

The industry speakers also covered a wide variety of topics:

- Dmitry Bugaichenko (OK.ru), Does Size Matter? Smart Data at OK.ru
- Mikhail Dubov (National Research University Higher School of Economics, Moscow, Russia), Text Analysis with Enhanced Annotated Suffix Trees: Algorithmic Base and Industrial Usage
- Nikita Kazeev (Yandex Data Factory), Role of Machine Learning in High Energy Physics Research at LHC
- Artem Kuznetsov (SKB Kontur), Family Businesses: Relation Extraction between Companies by Means of Wikipedia
- Alexey Natekin (Data Mining Labs), ATM Maintenance Cost Optimization with Machine Learning Techniques
- Konstantin Obukhov (Clever Data), Customer Experience Technologies: Problems of Feedback Modeling and Client Churn Control
- Alexandra Shilova (Centre IT), Centre of Information Technologies: Data Analysis and Processing for Large-Scale Information Systems

We would also like to mention the best conference paper selected by the Program Committee. It was written by Oleg Ivanov and Sergey Bartunov and is entitled “Learning Representations in Directed Networks”.

We would like to thank the authors for submitting their papers and the members of the Program Committee for their efforts in providing exhaustive reviews. We would also like to express special gratitude to all the invited speakers and industry representatives.

We deeply thank all the partners and sponsors, and owe our gratitude to the Ural Federal University for substantial financial support of the whole conference, namely, the Center of Excellence in Quantum and Video Information Technologies: from Computer Vision to Video Analystics (QVIT: CV $\rightarrow$ VA). We would like to acknowledge the Scientific Fund of Higher School of Economics for providing AIST participants with travel grants. Our special thanks goes to Springer editors who helped us, starting from the first conference call to the final version of the proceedings. Last but not least, we are grateful to all organisers, especially to Eugeniya Vlasova and Dmitry Ustalov, and the volunteers, whose endless energy saved us at the most critical stages of the conference preparation.

Traditionally, we would like to mention the Russian word "aist" is more than just a simple abbreviation (in Cyrillic), it means a "stork". Since it is a wonderful free bird, a symbol of happiness and peace, this stork brought us the inspiration to organise the AIST conference. So we believe that this young and rapidly growing conference will be bringing inspiration to data scientists around the World!

April, 2015

Mikhail Yu. Khachay
Natalia Konstantinova
Alexander Panchenko
Radhakrishnan Delhibabu
Nikita Spirin
Valeri G. Labunets

Organisation

The conference was organized by a joint team from Ural Federal University (Yekaterinburg, Russia), Krasovsky Institute of Mathematics and Mechanics, Ural Branch of Russian Academy of Sciences (Yekaterinburg, Russia), and the National Research University Higher School of Economics (Moscow, Russia). It was supported by a special grant from Ural Federal University for the Center of Excellence in Quantum and Video Information Technologies: from Computer Vision to Video Analystics (QVIT: CV $\rightarrow$ VA).

Program Committee Chairs

Mikhail Khachay	Krasovsky Institute of Mathematics and Mechanics of UB RAS, Russia
Natalia Konstantinova	University of Wolverhampton, UK
Alexander Panchenko	Technische Universität Darmstadt, Germany & Université catholique de Louvain, Belgium

General Chair

Valeri G. Labunets	Ural Federal University, Russia

Organising Chair

Eugeniya Vlasova	National Research University Higher School of Economics, Moscow

Proceedings Chair

Dmitry I. Ignatov	National Research University Higher School of Economics, Russia

Poster Chairs

Nikita Spirin	University of Illinois at Urbana-Champaign, USA
Dmitry Ustalov	Krasovsky Institute of Mathematics and Mechanics & Ural Federal University, Russia

International Liaison Chair

Radhakrishnan Delhibabu	Kazan Federal University, Russia

Organising Committee and Volunteers

Alexandra Barysheva	National Research University Higher School of Economics, Russia
Liliya Galimzyanova	Ural Federal University, Russia
Anna Golubtsova	National Research University Higher School of Economics, Russia
Vyacheslav Novikov	National Research University Higher School of Economics, Russia
Natalia Papulovskaya	Ural Federal University, Yekaterinburg
Yuri Pekov	Moscow State University, Russia
Evgeniy Tsymbalov	National Research University Higher School of Economics, Russia
Andrey Savchenko	National Research University Higher School of Economics, Russia
Dmitry Ustalov	Krasovsky Institute of Mathematics and Mechanics & Ural Federal University, Russia
Rostislav Yavorsky	National Research University Higher School of Economics, Russia

Industry Session Organisers

Ekaterina Chernyak	National Research University Higher School of Economics, Russia
Alexander Semenov	National Research University Higher School of Economics, Russia

Program Committee

Mikhail Ageev	Lomonosov Moscow State University, Russia
Atiqur Rahman Ahad	University of Dhaka, Bangladesh
Igor Andreev	Mail.Ru, Russia
Nikolay Arefiev	Moscow State University & Digital Society Lab, Russia
Jaume Baixeries	Politechnic University of Catalonia, Spain
Pedro Paulo Balage	Universidade de São Paulo, Brazil
Sergey Bartunov	Lomonosov Moscow State University & National Research University Higher School of Economics, Russia
Malay Bhattacharyya	Indian Institute of Engineering Science and Technology, India
Vladimir Bobrikov	Imhonet.ru, Russia
Victor Bocharov	OpenCorpora & Yandex, Russia
Daria Bogdanova	Dublin City University, Ireland
Elena Bolshakova	Lomonosov Moscow State University, Russia
Aurélien Bossard	Orange Labs, France
Pavel Botov	Moscow Institute of Physics and Technology, Russia
Jean-Leon Bouraoui	Université Catholique de Louvain, Belgium
Leonid Boytsov	Carnegie Mellon University, USA
Pavel Braslavski	Ural Federal University & Kontur Labs, Russia
Andrey Bronevich	National Research University Higher School of Economics, Russia
Aleksey Buzmakov	LORIA (CNRS-Inria-Université de Lorraine), France
Artem Chernodub	Institute of Mathematical Machines and Systems of NASU, Ukraine
Vladimir Chernov	Image Processing Systems Institute of RAS, Russia
Ekaterina Chernyak	National Research University Higher School of Economics, Russia
Marina Chicheva	Image Processing Systems Institute of RAS, Russia
Miranda Chong	University of Wolverhampton, UK
Hernani Costa	University of Malaga, Spain
Florent Domenach	University of Nicosia, Cyprus
Alexey Drutsa	Lomonosov Moscow State University & Yandex, Russia
Maxim Dubinin	NextGIS, Russia
Julia Efremova	Eindhoven University of Technology, The Netherlands
Shervin Emami	NVIDIA, Australia
Maria Eskevich	Dublin City University, Ireland
Victor Fedoseev	Samara State Aerospace University, Russia
Mark Fishel	University of Zurich, Germany

Thomas Francois	Université catholique de Louvain, Belgium
Oleksandr Frei	Schlumberger, Norway
Binyam Gebrekidan Gebre	Max Planck Computing & Data Facility, Germany
Dmitry Granovsky	Yandex, Russia
Mena Habib	University of Twente, The Netherlands
Dmitry Ignatov	National Research University Higher School of Economics, Russia
Dmitry Ilvovsky	National Research University Higher School of Economics, Russia
Vladimir Ivanov	Kazan Federal University, Russia
Sujay Jauhar	Carnegie Mellon University, USA
Dmitry Kan	AlphaSense Inc., USA
Nikolay Karpov	National Research University Higher School of Economics, Russia
Yury Katkov	Blue Brain Project, Switzerland
Mehdi Kaytoue	INSA de Lyon, France
Laurent Kevers	DbiT, Luxembourg
Michael Khachay	Krasovsky Institute of Mathematics and Mechanics UB RAS, Russia
Evgeny Kharitonov	Moscow Institute of Physics and Technology, Russia
Vitaly Khudobakhshov	Saint Petersburg University & National Research University of Information Technologies, Mechanics and Optics, Russia
Ilya Kitaev	iBinom, Russia & Voronezh State University, Russia
Ekaterina Kochmar	University of Cambridge, UK
Sergei Koltcov	National Research University Higher School of Economics, Russia
Olessia Koltsova	National Research University Higher School of Economics, Russia
Natalia Konstantinova	University of Wolverhampton, UK
Anton Konushin	Lomonosov Moscow State University, Russia
Andrey Kopylov	Tula State University, Russia
Kirill Kornyakov	Itseez, Russia & University of Nizhny Novgorod, Russia
Maxim Korolev	Ural State University, Russia
Anton Korshunov	Institute of System Programming of RAS, Russia
Yuri Kudryavcev	PM Square, Australia
Valentina Kuskova	National Research University Higher School of Economics, Russia
Sergei O. Kuznetsov	National Research University Higher School of Economics, Russia
Valeri G. Labunets	Ural Federal University, Russia
Alexander Lepskiy	National Research University Higher School of Economics, Russia
Benjamin Lind	National Research University Higher School of Economics, Russia

Natalia Loukachevitch	Research Computing Center of Moscow State University, Russia
Ilya Markov	University of Amsterdam, The Netherlands
Luis Marujo	Carnegie Mellon University, USA & Universidade de Lisboa, Portugal
Sérgio Matos	University of Aveiro, Portugal
Julian Mcauley	The University of California, San Diego, USA
Yelena Mejova	Qatar Computing Research Institute, Qatar
Vlado Menkovski	Eindhoven University of Technology, The Netherlands
Christian M. Meyer	Technische Universität Darmstadt, Germany
Olga Mitrofanova	St. Petersburg State University, Russia
Nenad Mladenovic	Brunel University, UK
Vladimir Mokeyev	South Ural State university, Russia
Gyorgy Mora	Prezi Inc., USA
Andrea Moro	Universita di Roma, Italy
Sergey Nikolenko	Steklov Mathematical Institute & National Research University Higher School of Economics, Russia
Vasilina Nikoulina	Xerox Research Center Europe, France
Damien Nouvel	National Institute for Oriental Languages and Civilizations, France
Dmitry Novitski	Institute of Cybernetics of NASU
Georgios Paltoglou	University of Wolverhampton, UK
Alexander Panchenko	Université catholique de Louvain, Belgium
Denis Perevalov	Krasovsky Institute of Mathematics and Mechanics, Russia
Georgios Petasis	National Centre of Scientific Research "Demokritos", Greece
Andrey Philippovich	Bauman Moscow State Technical University, Russia
Leonidas Pitsoulis	Aristotle University of Thessaloniki, Greece
Lidia Pivovarova	University of Helsinki, Finland
Vladimir Pleshko	RCO, Russia
Jonas Poelmans	Alumni of Katholieke Universiteit Leuven, Belgium
Alexander Porshnev	National Research University Higher School of Economics, Russia
Surya Prasath	University of Missouri-Columbia, USA
Delhibabu Radhakrishnan	Kazan Federal University and Innopolis, Russia
Carlos Ramisch	Aix Marseille University, France
Alexandra Roshchina	Institute of Technology Tallaght Dublin, Ireland
Eugen Ruppert	TU Darmstadt, Germany
Mohammed Abdel-Mgeed M. Salem	Ain Shams University, Egypt
Grigory Sapunov	Stepic, Russia
Sheikh Muhammad Sarwar	University of Dhaka, Bangladesh
Andrey Savchenko	National Research University Higher School of Economics, Russia

Marijn Schraagen	Utrecht University, The Netherlands
Vladimir Selegey	ABBYY, Russia
Alexander Semenov	National Research University Higher School of Economics, Russia
Oleg Seredin	Tula State University, Russia
Vladislav Sergeev	Image Processing Systems Institute of the RAS, Russia
Andrey Shcherbakov	Intel, Russia
Dominik Ślęzak	University of Warsaw, Poland & Infobright Inc.
Gleb Solobub	Agent.ru, Russia
Andrei Sosnovskii	Ural Federal University, Russia
Nikita Spirin	University of Illinois at Urbana-Champaign, USA
Sanja Stajner	University of Lisbon, Portugal
Rustam Tagiew	Alumni of TU Freiberg, Germany
Irina Temnikova	Qatar Computing Research Institute, Qatar
Christos Tryfonopoulos	University of Peloponnisos, Greece
Alexander Ulanov	HP Labs, Russia
Dmitry Ustalov	Krasovsky Institute of Mathematics and Mechanics & Ural Federal University, Russia
Natalia Vassilieva	HP Labs, Russia
Yannick Versley	Heidelberg University, Germany
Evgeniya Vlasova	Higher School of Economics
Svitlana Volkova	Johns Hopkins University, USA
Konstantin Vorontsov	Forecsys & Dorodnicyn Computing Center of RAS, Russia
Ekaterina Vylomova	Bauman Moscow State Technical University, Moscow
Patrick Watrin	Université catholique de Louvain, Belgium
Rostislav Yavorsky	National Research University Higher School of Economics, Russia
Roman Zakharov	Université catholique de Louvain, Belgium
Marcos Zampieri	Saarland University, Germany
Sergei M. Zraenko	Ural Federal University, Russia
Olga Zvereva	Ural Federal University, Russia

Invited Reviewers

Sujoy Chatterjee	University of Kalyani, India
Alexander Goncharov	CVisionLab, Russia
Vasiliy Kopenkov	Image Processing Systems Institute of RAS, Russia
Alexis Moinet	University of Mons, Belgium
Ekaterina Ostheimer	Capricat LLC, USA
Sergey V. Porshnev	Ural Federal University, Russia
Paraskevi Raftopoulou	Technical University of Crete, Greece
Ali Tayari	Technical University of Crete, Greece

Sponsors and Partners

Ural Federal University
Krasovsky Institute of Mathematics and Mechanics
GraphiCon
Exactpro
IT Centre
SKB Kontur
JetBrains
Yandex
Ural IT Cluster
NLPub
Digital Society Laboratory
CLAIM

Table of Contents

Mathematical Model of the Impulses Transformation Processes in Natural Neurons for Biologically Inspired Control Systems Development

Bakhshiev A.V., Gundelakh F.V.

Russian State Scientific Center for Robotics and Technical Cybernetics (RTC), Saint-Petersburg, Russian Federation
{alexab, f.gundelakh}@rtc.ru

Abstract. One of the trends in the development of control systems for autonomous mobile robots is the approach of using neural networks with biologically plausible architecture. Formal neurons do not take into account some important properties of a biological neuron, which are necessary for this task. Namely - a consideration of the dynamics of data changing in neural networks; difficulties in describing the structure of the network, which cannot be reduced to the known regular architectures; as well as difficulties in the implementation of biologically plausible learning algorithms for such networks. Existing neurophysiological models of neurons describe chemical processes occurring in a cell, which is too low level of abstraction.
The paper proposes a neuron's model, which is devoid of disadvantages described above. The feature of this model is description cell possibility with tree-structured architecture dendrites. All functional changes are formed by modifying structural organization of membrane and synapses instead of parametric tuning. The paper also contains some examples of neural structures for motion control based on this model of a neuron and similar to biological structures of the peripheral nervous system.

Keywords: neural network, natural neuron model, control system, biologically inspired neural network, motion control

1 Introduction

Nowadays, a lot of attention is paid to the study of the nervous system's functioning principles in the problems of motion control and data processing and the creation of biologically inspired technical analogues for robotics [1,2,3].

At the same time borrowing just part of the data processing cycle inherent to natural neural structures, seems to be ineffective. In this case, we can't avoid the step of converting the "inner world's picture" of our model, expressed in the structure and set of the neural network's parameters, set up in the narrow context in the terms of current problem. Such conversion can nullify the effectiveness of the approach. It is necessary to start with a construction of simple self-contained systems that function in an

environment model, and then gradually complicate them. For example, it is possible to synthesize the control system functionally similar to the reflex arc of human nervous system (Fig. 1).

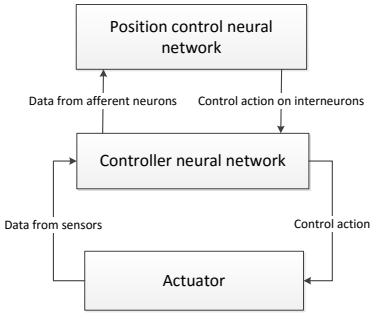


Fig. 1. Control system similar to reflex arc of human nervous system

In this case, position control neural network has input and output layers of neurons, as well as several hidden layers. Input and output layers have connections with neurons of other neural networks, while neurons of the hidden layers are connected only to the neurons of current neural network [4].

However, the most promising is the development of full-scale systems that implement all phases of the data transformation from sensors to effectors inherent to natural prototypes.

There are many models of neuronal and neural networks. These models may be quite clearly divided into two groups: for applied engineering problems (derived from the formal neuron model) [5], and models, designed for the most complete quantitative description of the processes occurring in biological neurons and neural networks [6,7] .

Considering modeling of natural neuron, we investigate the transition from formal neuron models to more complex models of neurons as a dynamic system for data transformation [8] suitable for control tasks (Fig. 2).

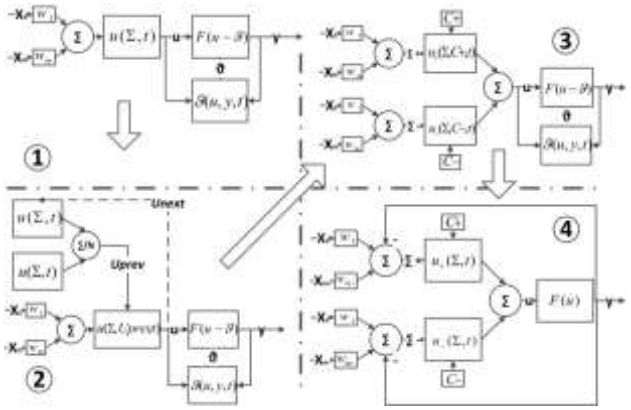


Fig. 2. Evolution of neuron model

Where $x_1 - x_m$ - neuron input signals;

$w_1 - w_m$ - weights;

y - neuron output signal;

u - membrane potential value;

ϑ - threshold function;

F - activation function;

N - number of membrane segments at the dendrite branching node;

$C+, C-$ - constants for expected level of membrane potential contribution;

$u+, u-$ - contributions to the membrane potential from depolarizing and hyperpolarizing ionic mechanisms;

Figure 2-1 represents a universal model of the formal neuron in general. Classic formal neurons can be derived from this model, if we abandon the temporal summation of signals to establish a fixed threshold and choose, for example, a sigmoid activation function.

Further development of this model may be adding a description of the structural organization of the neuron membrane (Fig. 2-2), with a separate calculation of the contribution to the total potential (Fig. 2-3) to provide at each site the ability to integrate information about the processes occurring with different speeds, as well as rejection of the an explicit threshold setting and move to the representation of the signal in the neural network as a stream of pulses (Fig. 2-4). As a result, the potential value of the neuron membrane segment is derived not only from the values of the inputs and the weights of synapses, but also from the average value of the membrane potential of other connected membrane segments. This will simulate the structure of the dendritic and synaptic apparatus of neurons and carry out more complex calculations of the spatial and temporal summation of signals on the membrane of the neuron. Thus, membrane segment should be considered as the minimal functional element of the neural network.

Given the existence of temporal summation of signals, the structural organization allows to implement separate processing of signals with different functionality on a

single neuron. To do this, may be selected a single dendrite, which will provide, for example, only the summation of signals on the current position of the control object formed by afferent neurons, as well as to the signal of corrections to position, that formed by the highest level of control. The individual dendrite will implement similar behavior, for example, the speed of the object and the body of the neuron will provide the integral combination of these control loops, which otherwise would require adding an additional neuron.

2 Neuron model

It is assumed that the inputs of the model get pulsed streams, which are converted by synapses into the analog values that describe the processes of releasing and metabolizing of the neurotransmitter in the synaptic cleft. The model assumes that the input and output signals of the neuron is zero for the absence of a pulse, and constant for the duration of the pulse. The pulse duration is determined by the time parameters of the neuron's membrane. Membrane of soma and dendrites is represented by a set of pairs of ionic mechanisms' models that describe the function of depolarization and hyperpolarization mechanisms, respectively. The outputs of the ionic mechanisms' models represent the total contribution to the intracellular potential of depolarization and hyperpolarization processes occurring in the cell. The signals from the synapses modifies the ionic mechanisms' activity in the direction of weakening their functions, which simulates the change in the concentration of ions inside the cell under the influence of external influences. It is proposed to distinguish the type of ionic mechanism in the sign of the output signal. A positive value of the output characterizes depolarizing influence, while negative characterizes hyperpolarization. Thus, the total value of the output values will characterize the magnitude of the membrane segment contribution to the total intracellular neuron potential [9].

The role of synaptic apparatus in the model is the primary processing of the input signals. It should be noted that the pattern of excitatory and inhibitory synapses are also identical to each other, and the difference in their effects on cell's membranes is determined by which of the ionic mechanisms each particular synapse is connected to. Each synapse in this model describes a group of natural neuron synapses.

More detailed model of the membrane is shown in Fig. 3.

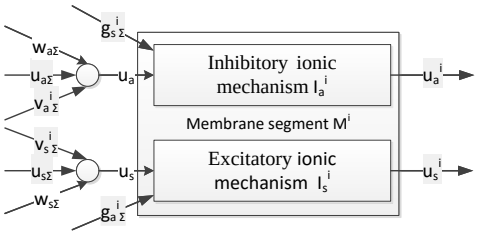


Fig. 3. Functional diagram of the *i*-th membrane segment model *Mi*

Each membrane segment $M^i, i = \overline{1, L}$ consists of a pair of mechanisms - hyperpolarization mechanism ($I_a{}^i$), and depolarization mechanism ($I_s{}^i$). Output of the membrane's segment is a pair of the contribution values of hyperpolarization (u_a) and depolarization (u_s), which determines the contribution to the total intracellular potential.

Each membrane's segment M^i can be connected to previous membrane's segment M^j taking its values $\{u_a{}^j, u_s{}^j\}$ as inputs. When specified membrane's segment is the last in the chain (the end of the dendrite or the segment of soma), as signals $\{u_a{}^j, u_s{}^j\}$ stands pair of fixed values {*-Em, Em*} simulating some of the normal concentration of ions in the cell in a fully unexcited state.

Excitatory $\{x_{sk}^i\}, k = \overline{1, M_i}$ and inhibitory $\{x_{ak}^i\}, k = \overline{1, N_i}$ neuron's inputs are inputs of many models of excitatory $\{S_{sk}^i\}, k = \overline{1, M_i}$ and inhibitory $\{S_{ak}^i\}, k = \overline{1, N_i}$ synapses, for each of the membrane's segments M^i.

The resulting values of the effective influence on the mechanisms of synaptic hyperpolarization ($g_s{}^{i}{}_{\Sigma}$) and depolarization ($g_a{}^{i}{}_{\Sigma}$) are obtained by summation:

$$g_s{}^{i}{}_{\Sigma} = \sum_{k=1}^{M_i} g_{sk}{}^{i}, g_a{}^{i}{}_{\Sigma} = \sum_{k=1}^{N_i} g_{ak}{}^{i}. \quad (1)$$

Outputs of all membrane segment models are summed by following formula:

$$u_{\Sigma} = \frac{1}{L} \sum_{i=1}^{L} u^i$$

The resulting signal is assumed as total intracellular potential of the neuron. Each pair (depolarization and hyperpolarization mechanisms), depending on their internal properties, can be regarded as model of dendrite segment or soma segment. Increasing the number of pairs of such mechanisms automatically increases the size of the neuron, and allows simulating a neuron with a complex organization of synaptic and dendritic apparatus.

Similarly, the summation of signals at branching nodes of dendrites - the total contribution of the hyperpolarization and depolarization mechanisms $\{u_a{}^j, u_s{}^j\}$ are divided by their number.

Fig. 4 contains a general view of the neuron's membrane structure [10].

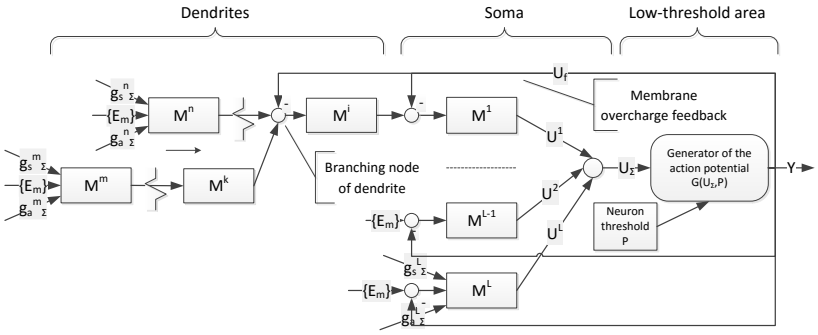


Fig. 4. Structural diagram of the neuron membrane

The body of the neuron (soma), we assume those parts of the membrane that are covered by feedback from the generator of the action potential. It should also be noted that the closer a membrane's segment located to the generator, the more effective its contribution to the overall picture of synapses in neuronal excitation.

Thus, in terms of the model:

1. carried out on the dendrites spatial and temporal summation of signals over long periods of time (a small contribution to the excitation of the neuron from each synapse), and accumulation of potential does not depend on the neuron discharges;
2. in the soma of a neuron produced summation of signals at short intervals of time (a big contribution to the excitation of the neuron from each synapse) and accumulated potential is lost when the neuron discharges;
3. in low-threshold area is carried impulse formation on reaching the threshold of generation and signal of membrane recharge.

The following discloses the mathematical description of the neuron model elements.

Synapse model. It is known that the processes of releasing and metabolizing of the neurotransmitter are exponential, and besides the process of releasing neurotransmitter, usually is much faster than the metabolizing process.

Another important factor is the effect presynaptic inhibition consists in that, when the concentration of the neurotransmitter exceeds certain limit values, synaptic influence on ion channel starts to decrease rapidly - despite the fact that the ion channel is fully open. Reaching the limit concentration is possible when synapse is stimulated by the pulsed streams with high pulse frequency.

Model that implements all three main features of the synapse's functioning can be described by the following equations:

$$\left.\begin{aligned} T_S \cdot \frac{d\rho}{dt} + \rho(t) = E_y^{-1} \cdot x, \\ T_S = \begin{cases} \tau_s, \text{ при } x > 0, \\ \tau_d, \text{ при } x \le 0. \end{cases} \end{aligned}\right| \left.\begin{aligned} g_* = 4 \cdot \xi \cdot (\rho - \xi \cdot \rho^2) \\ g = \begin{cases} R_s^{-1} \cdot g_*, \text{ при } g_* > 0, \\ 0, \text{ при } g_* \le 0. \end{cases} \end{aligned}\right\} \quad (2)$$

Where τ_s - time constant of releasing neurotransmitter,

τ_d - time constant of metabolizing neurotransmitter,

$\xi \in [0.5, \infty)$ - limit value of neurotransmitter's concentration needed to presynaptic inhibition effect,

$R_S > 0$ - synapse's resistance ("weight"), that characterizes the efficiency of synapse's influence on the ionic mechanism,

E_y - the amplitude of the input signal.

Initial conditions: $\rho(0) = 0$.

Model's input is a discrete signal *x(t)*, which is a sequence of pulses with a duration of 1 ms and an amplitude *E*. The releasing and metabolizing processes of the neurotransmitter are proposed to simulate the first order inertial element with logic control by time constant. Variable ρ characterizes the concentration of neurotransmitter released in response to a pulse. Usage of variable g_* allows us to simulate presynaptic inhibition effect.

Model's output *g(t)* is an efficiency of influence on ionic mechanism and it is proportional to the synapse's conduction. Thus, in the absence of input actions synapse conductance tends to zero, which corresponds to the open switch in the equivalent circuit of the membrane.

Model of membrane's ionic mechanism. It is known that the ion channel can be represented by an equivalent electrical circuit [11], which has three major characteristics - the resistance R_m, capacitance C_m and ion concentration $E_m = v$ maintained within the cell membrane pump function. Product $T_m = R_m C_m$ characterizes inertia of the channel that defines the rate of recovery of the normal concentration of ions E_m in the cell. Synapse's influence on the ionic mechanism consists in the loss of efficiency of the channel's pumping function and reducing the ions' concentration in the cell, with the time constant of the process:

$$T = R^I C_m . \quad (3)$$

Resistance R^I is determined from the relation:

$$\frac{1}{R^I} = g_1 + g_2 + ... + g_n + \frac{1}{R_m} = g_\Sigma + \frac{1}{R_m} . \quad (4)$$

Where $g_1, g_2, ..., g_n$ - conductions of active synapses' models that have an influence on the current ionic channel. Reduction in ions' concentration at the same time is proportional to the product $g_\Sigma \cdot R_m$ and the less, the lower the ions' concentration in the cell is.

Fig. 5a shows the dependence of the synapse's contribution in changing the membrane potential on the ratio of the synapse's channel and postsynaptic membrane's resistance. It can be seen that the effective control range of the synapse's resistance is in the range [0.1: 10] of membrane's resistance. Fig. 5b shows the change in the potential contribution to the number of active synapses in the ratio Rs/Rm = 10 (dashed line) and 1 (solid line).

The ordinate axis in both graphs - normalized postsynaptic membrane potential change in proportion to its nominal value. Fig. 5a: the dependence of the efficiency on the ratio of the synapse's channel and the membrane's resistance. Fig. 5b: the dependence of the efficiency on the number of synapses.

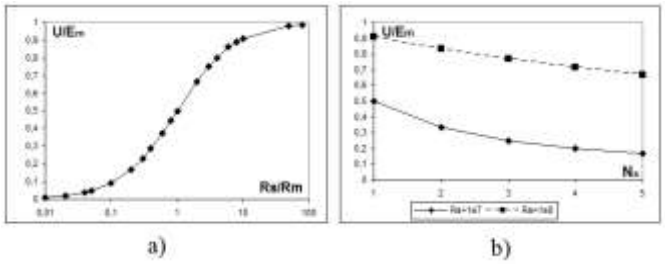


Fig. 5. Current efficiency of the synapse's model

Inertial properties of the ionic mechanism's model are proposed to describe as an aperiodic element with logic control by time constant. For the ionic mechanism of depolarization equations have the following form:

$$\begin{cases} \left.\begin{aligned} &T_I \cdot \frac{du}{dt} + (1 + g_{s\Sigma} \cdot R_m) \cdot u = v \\ &T_I = \frac{C_m}{g_{s\Sigma} + R_m^{-1}} \end{aligned}\right\}, g_{a\Sigma} = 0 \\ u = E_m^-, g_{a\Sigma} \neq 0 \end{cases} \quad (5)$$

Where $g_{a\Sigma}$ - the total efficiency of synapses influence on the hyperpolarization mechanism,

$g_{s\Sigma}$ - the total efficiency of synapses influence on the depolarization mechanism,

$R_m > 0$ - membrane's resistance,

$C_m > 0$ - membrane's capacitance,

v - the expected contribution of the model in the value of the intracellular potential in the absence of external excitation. This value is determined by the activity of neighboring membrane segments,

u - a real model's contribution to the value of the intracellular potential.

Initial conditions: *u(0)=0.*

For ionic mechanism of hyperpolarization equations are analogous up to relocation of the effects of excitatory and inhibitory synapses and Em^- on Em^+.

Action's potential generator's model. Generator's model performs the formation of rectangular pulses of given amplitude E_y as a result of exceeding fixed threshold P by the potential u_Σ. The model can be described by the following equations:

$$\left.\begin{aligned} T_G \cdot \frac{du_*}{dt} + u_* = u_\Sigma, \\ y = F_G(u_*). \end{aligned}\right\} \quad (6)$$

Where $P > 0$ – neuron's threshold,

T_G - time constant, which determines the duration of the feedback overcharging membrane and characterizing pulse durations,

$F_G(u_*)$ - Function describing the hysteresis. The output of the function is E_y, if $u_* \geq P$ and zero if $u_* \leq 0$.

Initial conditions: $u_*(0) = 0$.

Output signal *y(t)* goes to overcharge feedbacks of cell's soma.

3 Research

Setting the model's parameters was based on experimental data on the time parameters of the processes occurring in the natural neuron [10].

Fig. 6 shows a typical response of a neuron model to the exciting pulse. In the graph of intracellular potential (2) can be seen a typical region of the neuron membrane depolarization is preceded by the formation of an action potential, the zone of hyperpolarization after pulse generation and residual membrane depolarization at the end of the generation's pattern.

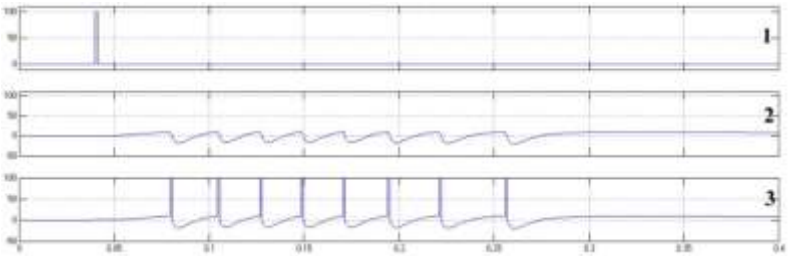


Fig. 6. Neuron with synapse on its dendrite (1 - stimulating effect 2 - intracellular membrane potential on the generator of the action potential, 3 - neuron responses combined with the graph of the intracellular potential)

One of the main characteristics of the natural neuron qualitatively affects the transformation of the pulsed streams is the size of the membrane. Unlike small neuron large neuron is less sensitive to the effects of input and generates a pulse sequence typically in a lower frequency range and generally corresponds to input effects with single pulses.

The developed model allows to build neurons with different membrane structure and location of synapses on it. Changing the number of the membrane segments neurons of different sizes can be modeled, without changing the values of the parameters.

With the increasing size of the soma at the same stimulation of the neuron number of pulses in the pattern of neuron response decreases and the interval between them increases. Fig. 7a demonstrates dependence of the response's average frequency from the number of pulses *Np* in it. Fig. 7b demonstrates dependence of response's average frequency from the number of neuron's soma segments *L*.

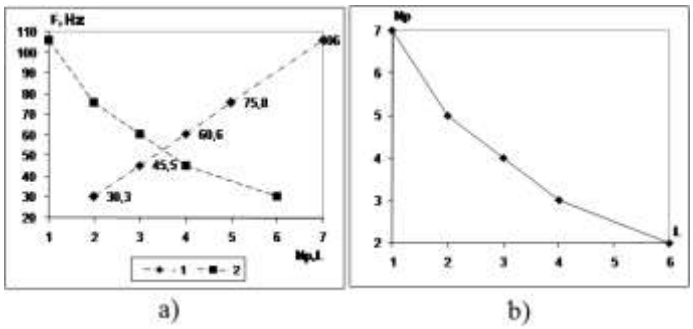


Fig. 7. Discharge frequency, depending on the neuron's soma size

As a simple neural structures with feedback considered element, which is a widely held in the nervous system connection excitatory inhibitory neurons, first studied in neurophysiological experiments, the interaction of motoneuron and Renshaw's cells (Fig. 8).

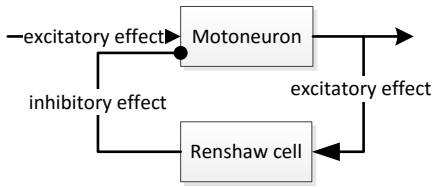


Fig. 8. The scheme of recurrent inhibition by the example of the regulation of motoneuron discharges

There are two mechanisms for increasing the strength of muscle contraction. The first is to increase the pulse repetition frequency at the output of motoneuron. Second - increasing the number of active motoneurons, the axons of which are connected to the muscle fibers of the muscle. Specialized inhibition neuron in the chain of recurrent inhibition - Renshaw cell - limits and stabilizes the frequency of motoneuron discharges. Example of such a structure shows an analog model (Fig. 9), the behavior of which corresponds to neurophysiological data [11].

Fig. 9. Recording pulsed streams in studying the interaction of motoneuron and Renshaw cells motoneuron at the excitation frequency of 20Hz (a) and 50 Hz (b): 1 - excitatory motoneuron input; 2 - Renshaw cell's discharges; 3 - motoneuron output pulses. Above - the time stamp 10 ms

The graphs show that the frequency of motoneuron stimulation enhances the inhibitory effect on Renshaw cells with motoneuron, causing, in turn, decrease the frequency of motoneuron discharges. Thus, when the frequency of motoneuron stimulation increases, the frequency of the pulses at the output of the first moments increases and then stabilizes at a low level with a duration of interpulse intervals determined by the duration of the Renshaw cell's discharge. It is essential that this limit is dependent on whether the motoneuron by recurrent inhibition "own" Renshaw cells or not. Computer simulation has allowed a more detailed study of the interaction of neurons.

The results of the experiment are shown in Fig. 10, where the top-down plotted input pulsed stream at the input of motoneurons and pulsed streams of motoneuron Renshaw cell with recurrent inhibition and, accordingly, these neurons without feedback when motoneuron excites Renshaw cell, but it does not slow motoneuron.

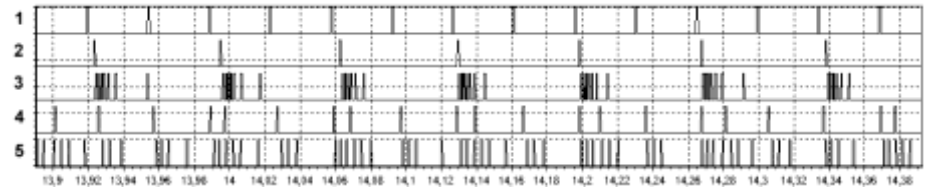


Fig. 10. Reactions of structure "motoneuron-Renshaw cell" upon excitation of motoneurons pulsed stream at 50 Hz: 1 - input pulsed stream; 2 – motoneuron's reaction with enabled FB; 3 - Renshaw cell responses with enabled FB; 4 – motoneuron's reaction without FB; 5 - Renshaw cell responses without FB

4 Conclusion

The paper presents a model of a neuron, which can serve as the basis for constructing models of neural networks of living organisms and study their applicability in solving the problems of motion control of robotic systems. The model allows to describe the structure of the neuron's membrane (dendritic and synaptic apparatus).

Plasticity model is also based primarily on changes in the structure of the membrane, rather than adjusting the parameters of the model (synapse weights, neuron's threshold, etc.), which simplifies the construction of models of specific known biological neural structures.

5 Sources

1. McKinstry, J. L., Edelman, G. M., Krichmar, J. L.: A cerebellar model for predictive motor control tested in a brain-based device. PNAS, February 28, 2006, vol. 103, No.9, pp. 3387–3392 (2006)
2. Hugo de Garis, Chen Shuo, Ben Goertzel, Lian Ruiting.: A world survey of artificial brain projects, Part I: Large-scale brain simulations. Neurocomputing 74, pp. 3–29 (2010)
3. Bakhshiev, A.V., Klochkov, I.V., Kosareva, V.L., Stankevich, L.A.: Neuromorphic robot control systems. Robotic and Technical Cybernetics No. 2(3)/2014, pp.40–44. Russia, Saint-Petersburg, RTC (2014)
4. Bakhshiev, A.V., Gundelakh. F.V.: Investigation of biosimilar neural network model for motion control of robotic systems. Robotics and Artificial Intelligence: Proceedings of the VI Russian Scientific Conference with international participation, 13 december 2014, Zheleznogorsk, Russia (2014)
5. McCulloch, W. S., Pitts W.: A logical calculus of the ideas immanent in nervous activity // Bulletin of Mathematical Biophysics, vol. 5, pp. 115-133 (1943)
6. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. J. Physiology, 117, pp. 500–544 (1952)
7. Izhikevich, E.M.: Simple model of spiking neurons. IEEE transactions on neural networks. A publication of the IEEE Neural Networks Council, vol. 14, No. 6, pp. 1569–1572 (2003)
8. Romanov, S.P.: Neuron model. Some problems of the Biological Cybernetics. Russia, pp. 276-282 (1972)
9. Bakhshiev, A.V., Romanov, S.P.: Neuron with arbitrary structure of dendrite, mathematical models of biological prototypes. Neurocomputers: development, application, Russia, No.3, pp. 71-80 (2009)
10. Bakhshiev, A.V., Romanov, S.P.: Reproduction of the reactions of biological neurons as a result of modeling structural and functional properties membrane and synaptic structural organization. Neurocomputers: development, application, Russia, No.7, pp. 25–35 (2012)
11. John Carew Eccles. The Physiology of Synapses. Springer-Verlag (1964)

The Construction of Images Using Minimal Spline

Burova I.G. and Bezrukavaya O.V.

St. Petersburg State University, St. Petersburg, Russia,
i.g.burova@spbu.ru,

Abstract. Tasks of data compression, transmission, subsequent recovery with a given accuracy are of great practical importance. In this paper we consider a problem of constructing graphical information on a plane with the help of a parametric defined splines with different properties. Here we compare the polynomial and the trigonometric splines of the first and the second order, the polynomial integro-differential splines, the trigonometric integro-differential splines. We consider a compression of the image to a relatively small number of points, and a restoration of graphic information with the given accuracy.

Keywords Image Construction, Polynomial Splines, Trigonometrical Splines, Integro-differential Splines, Interpolation

1 Introduction

Plotting functions by means of splines is widely used in practice [3–8]. Here we compare the polynomial and the trigonometric splines of the first and the second order, the polynomial integro-differential splines, the trigonometric integro-differential splines. These splines are characterized by the fact that the approximation of a function is constructed at each grid interval separately as a linear combination of values of the functions in neighboring grid nodes and some basic functions (see [1, 2]). The image can be compressed to a small number of points, which we call the control points (points of interpolation). The result of the image compression has the form of the control points and information of the applied basic splines. If it is necessary, the user can restore the image through an appropriate algorithm.

2 Right polynomial, right trigonometric splines

Let n be natural number, a, b be real numbers, $\{t_j\}$ be ordered equidistant set of nodes on $[a, b]$, $h = t_{j+1} - t_j$.

Let function u be such that $u \in C^3[a, b]$. We use the approximation for $u(t)$ in the form

$$\widetilde{u}(t) = u(t_j)\omega_j(t) + u(t_{j+1})\omega_{j+1}(t) + u(t_{j+2})\omega_{j+2}(t), \quad t \in [t_j, t_{j+1}], \qquad (1)$$

where $\omega_j(t)$, $\omega_{j+1}(t)$, $\omega_{j+2}(t)$ we determine from the system

$$\widetilde{u}(x) = u(x), \quad u(x) = \varphi_i(x), \ i = 1, 2, 3. \tag{2}$$

Here $\varphi_i(x)$, $i = 1, 2, 3$, is Chebyshev system on $[a, b]$, $\varphi_i \in C^3[a, b]$.

2.1 Right polynomial splines

In polynomial case we take $\varphi_i(x) = x^{i-1}$, So we have from (2)

$$\omega_j(t) = \frac{(t - t_{j+1})}{(t_j - t_{j+1})} \cdot \frac{(t - t_{j+2})}{(t_j - t_{j+2})}, \quad \omega_{j+1}(t) = \frac{(t - t_j)}{(t_{j+1} - t_j)} \cdot \frac{(t - t_{j+2})}{(t_{j+1} - t_{j+2})}, \tag{3}$$

$$\omega_{j+2}(t) = \frac{(t - t_j)}{(t_{j+2} - t_j)} \cdot \frac{(t - t_{j+1})}{(t_{j+2} - t_{j+1})}. \tag{4}$$

We obtain for $t \in [t_j, t_{j+1}]$ the estimation of the error of the approximation by the polynomial splines (1), (3) – (4): $|\widetilde{u}(t) - u(t)| \le K_1 h^3 \|u'''\|_{[t_j, t_{j+2}]}$, $K_1 = 0.0642$.

2.2 Right trigonometric splines

In trigonometric case we take $\varphi_1 = 1$, $\varphi_2 = \sin(x)$, $\varphi_3 = \cos(x)$. So we have from the system (2):

$$\omega_j(t) = \frac{\sin(t/2 - t_{j+1}/2)}{\sin(t_j/2 - t_{j+1}/2)} \cdot \frac{\sin(t/2 - t_{j+2}/2)}{\sin(t_j/2 - t_{j+2}/2)}, \tag{5}$$

$$\omega_{j+1}(t) = \frac{\sin(t/2 - t_j/2)}{\sin(t_{j+1}/2 - t_j/2)} \cdot \frac{\sin(t/2 - t_{j+2}/2)}{\sin(t_{j+1}/2 - t_{j+2}/2)}, \tag{6}$$

$$\omega_{j+2}(t) = \frac{\sin(t/2 - t_j/2)}{\sin(t_{j+2}/2 - t_j/2)} \cdot \frac{\sin(t/2 - t_{j+1}/2)}{\sin(t_{j+2}/2 - t_{j+1}/2)}. \tag{7}$$

The error of the approximation $u(t)$ by (1), (5)–(7) is the next:

$$|\widetilde{u}(t) - u(t)| \le K_2 h^3 \|u' + u'''\|_{[t_j, t_{j+2}]}, \quad K_2 > 0, \quad t \in [t_j, t_{j+1}].$$

3 Integro-differential splines

Integro-differential polynomial splines were invented by Kireev V.I [3]. The way of constructing the nonpolynomial integro-differential splines is in [2]. The integro-differential right spline of the third order has the form:

$$\widetilde{u}(t) = u(t_j)w_j(t) + u(t_{j+1})w_{j+1}(t) + \int_{t_j}^{t_{j+2}} u(t)dt \ w_j^{<1>}(t), \ t \in [t_j, t_{j+1}], \tag{8}$$

where $\omega_j(t)$, $\omega_{j+1}(t)$, $w_j^{<1>}(t)$ we determine from the system (2).

3.1 Integro-differential right polynomial splines

In polynomial case we have

$$w_j(t) = \frac{A_1}{(t_j - t_{j+2})(t_j - t_{j+1})(t_j - 3t_{j+1} + 2t_{j+2})}, \tag{9}$$

$A_1 = (-t_{j+1}+t)(3t_{j+2}t+3tt_j-6t_{j+1}t-2t_j^2-2t_jt_{j+2}-2t_{j+2}^2+3t_{j+1}t_{j+2}+3t_{j+1}t_j)$,

$$w_{j+1}(t) = \frac{(-t_j + t)(3t - t_j - 2t_{j+2})}{(t_j - t_{j+1})(t_j - 3t_{j+1} + 2t_{j+2})}, \tag{10}$$

$$w_j^{<1>}(t) = \frac{6(-t_{j+1} + t)(-t_j + t)}{(t_j - t_{j+2})^2(t_j - 3t_{j+1} + 2t_{j+2})}. \tag{11}$$

We can use in (8) the next formula:

$$\int_{t_j}^{t_{j+2}} u(t)dt \approx (t_{j+2} - t_j)(u(t_j) + 4u(t_{j+1}) + u(t_{j+2}))/6.$$

We obtain $|\widetilde{u}(t) - u(t)| \le K_3 h^3 \|u'''\|_{[t_j, t_{j+2}]}, \quad K_3 > 0,\ t \in [t_j, t_{j+1}]$.

3.2 Integro-differential right trigonometrical splines

In trigonometric case we have

$$w_j(t) = \frac{A_3}{B_3}, \quad w_{j+1}(t) = \frac{A_4}{B_4}, \tag{12}$$

where

$A_3 = (\cos(-t_{j+1}+t_j) - \cos(t_{j+1}-t_{j+2}) - t_{j+2}\sin(t-t_{j+1}) + t_j\sin(t-t_{j+1}) - \cos(t-t_j) + \cos(t-t_{j+2}))$,

$B_3 = (\cos(-t_{j+1}+t_j) - \cos(t_{j+1}-t_{j+2}) - t_{j+2}\sin(-t_{j+1}+t_j) + t_j\sin(-t_{j+1}+t_j) - 1 + \cos(t_j - t_{j+2}))$,

$A_4 = (\cos(t-t_j) - \cos(t-t_{j+2}) + t_{j+2}\sin(t-t_j) - t_j\sin(t-t_j) - 1 + \cos(t_j - t_{j+2}))$,

$B_4 = (\cos(-t_{j+1}+t_j) - \cos(t_{j+1}-t_{j+2}) - t_{j+2}\sin(-t_{j+1}+t_j) + t_j\sin(-t_{j+1}+t_j) - 1 + \cos(t_j - t_{j+2}))$,

$$w_j^{<1>}(t) = (\sin(t - t_{j+1}) - \sin(-t_{j+1} + t_j) - \sin(t - t_j))/B_5, \tag{13}$$

$B_5 = (\cos(-t_{j+1}+t_j) - \cos(t_{j+1}-t_{j+2}) - t_{j+2}\sin(-t_{j+1}+t_j) + t_j\sin(-t_{j+1}+t_j) - 1 + \cos(t_j - t_{j+2}))$.

If we know only the values $u(t_j)$, $u(t_{j-1}) = u(t_j - h)$, $u(t_{j+2}) = u(t_j + 2h)$, then we can use in (8) the formula:

$$I_t = \int_{t_j}^{t_{j+2}} u(t)dt = u(t_{j-1})\frac{2h\cos(h) - 2\sin(h)}{\cos(h) - \cos(2h)} - u(t_j)\frac{-2h\cos(h) + \sin(h) + h}{\cos(h) - 1} +$$

$$+u(t_{j+2})\frac{2\sin(h)\cos(h) - h - \sin(h)}{-\cos(h) - 1 + 2\cos^2(h)} + R_1.$$

It can be shown that $R_1 = 0$, if $u(t) = 1, \sin(t), \cos(t)$, and

$I_t = (-(4/9)u(t_{j-1}) + (5/3)u(t_j) + (7/9)u(t_{j+2}))\,h + O(h^3)$.

4 Constructing approximation on the plane

4.1 Piecewise linear set of parametric spline

Let function u be such that $u \in C^2[a,b]$.

We build the approximation of $u(t)$ in the form

$$\tilde{u}(t) = -A\,(u(t_j) - u(t_{j+1})) + u(t_j), \quad A = \frac{t - t_j}{t_{j+1} - t_j}, \tag{14}$$

Here $t \in [t_j, t_{j+1}]$, $j = 0, \ldots, n-1$. We can obtain for $t \in [t_j, t_{j+1}]$

$$|\widetilde{u}(t) - u(t)| \le K_0 h^2 \|u''\|_{[t_j, t_{j+1}]}, \quad K_0 = 0.125.$$

Consider the approximation of the curve on the plane using the linear splines. Let n points z_1, z_2, ..., z_n are given on the plane. Suppose point z_i has coordinates (x_i, y_i). Then we can construct the next approximations:

$\tilde{x}(t) = -A\,(x(t_j) - x(t_{j+1})) + x(t_j)$, $\tilde{y}(t) = -A\,(y(t_j) - y(t_{j+1})) + y(t_j)$, where $A = (t - t_j)/(t_{j+1} - t_j)$, if $t \in [t_j, t_{j+1}]$. The error of the approximation on the plain is the next: $R(t) = \sqrt{|\tilde{x}(t) - x(t)|^2 + |\tilde{y}(t) - y(t)|^2}$.

4.2 Minimal quadratic set of the right polynomial parametric spline

Consider the approximation of the curve on the plane with the help of quadratic splines (1), (3)–(4). Let functions $x = x(t)$ and $y = y(t)$ be such that $x, y \in C^3[a,b]$, $x(t_j)$ is the value of x in the node t_j, $y(t_j)$ is the value of y in the node t_j. Then we can use the following formulas:

$\tilde{x}(t) = ACx(t_j) - ABx(t_{j+1}) + BCx(t_{j+2})$, $\tilde{y}(t) = ACy(t_j) - ABy(t_{j+1}) + BCy(t_{j+2})$ on $[t_j, t_{j+1}]$, $j = 1, \ldots, n-1$, where $A = (t - t_{j+2})/(t_j - t_{j+1})$, $B = (t - t_j)/(t_{j+1} - t_{j+2})$, $C = (t - t_{j+1})/(t_j - t_{j+2})$.

5 Numerical experiments

Let function $z(t) = (x(t), y(t)$ be such that $x(t) = \sin(t), y(t) = \cos(t)$. Suppose we have $z_j = (x(j), y(j))$, $j = 1, 2, \ldots, 9$. We construct $\tilde{z} = (\widetilde{x}(t), \widetilde{y}(t))$, with the help of splines (1), (3)–(4), (1), (5)–(7), (8), (9)–(11), (8), (12)–(13). The results of application the splines (1), (3)–(4), and the splines (1), (5)–(7) are presented on graphs 1a, 1b. The results of application the splines (8), (9)–(11), and the splines (8), (12)–(13) are presented on graphs 2a, 2b.

Now we take $x(t) = t - 2\sin(t)$, $y(t) = 1 - 2\cos(t)$. The result of application the trigonometric splines (1), (5)–(7) is presented on graph 3a. The result of application the polynomial splines (8), (9)–(11) is presented on graph 3b.

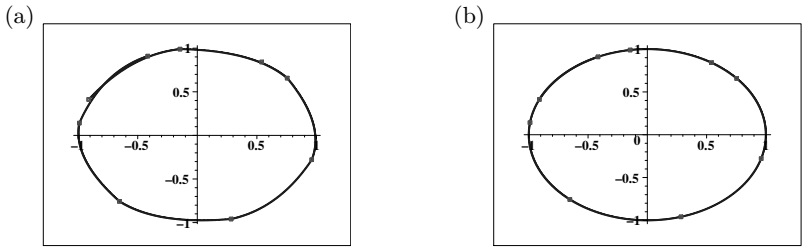


Fig. 1. Graphs of approximation by the minimal polynomial splines (1), (3)–(4): (a); by the trigonometric splines (1), (5)–(7): (b)

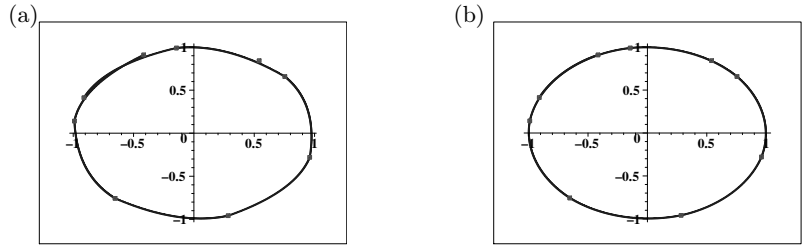


Fig. 2. Graphs of approximation by the right polynomial integro-differential splines (8), (9)–(11): (a), by the right trigonometric integro-differential splines (8), (12)–(13): (b)

6 Imaging of letters using a piecewise linear splines

Here we construct, compress and restore the image of the letters using splines.

For example, consider the construction and compression of the letter A. Coordinates of points for the letter "A" $z_i, i = 1, 2, 3, 4, 5$, we take in the form:

```
x[1]:=2:x[2]:=3:x[3]:=4:x[4]:=3.5:x[5]:=2.5:
y[1]:=2:y[2]:=4:y[3]:=2:y[4]:=3:y[5]:=3:
t[1]:=1:t[2]:=2:t[3]:=3:t[4]:=4:t[5]:=5:
```

Coordinates of points for the letter "E" $z_i, i = 1, 2, 3, 4, 5$ are given as:

Figure 4a shows the letter "A" which is constructed with the help of the control points: (2;2),(3;4),(4;2),(3.5;3), (2.5;3), and Fig. 4b shows the letter "E" which is constructed with the help of the points: (4;4), (2;4), (2;3), (3;3), (2;3), (2;2), (4;2) and the splines (14).

Each letter is given by a minimum number of the control points. For different letters the number of control points is different. Now we can compress the image and have only the control points and the information about the basis splines. We can hold or send the information someone. The recipient can restor the letters using the control points and information about the splines.

(a) (b)

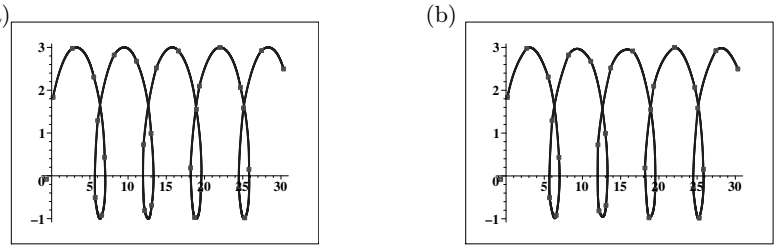


Fig. 3. Graphs of the approximation $x(t) = t - 2\sin(t)$, $y(t) = 1 - 2\cos(t)$ by the right trigonometric splines (1), (5)–(7): (a); by the right polynomial integro-differential splines (8), (9)–(11): (b)

(a) (b)

Fig. 4. Plot of the letter "A" (the control points: (2;2),(3;4),(4;2),(3.5;3), (2.5;3)) (a). Plot of the letter "E" (the control points: (4;4), (2;4), (2;3), (3;3), (2;3), (2;2), (4;2)): (b).

References

1. Burova I.G., Demyanovich Yu.K. Minimal Splines and theirs Applications. Spb. (2010) (Russian).
2. Burova Irina. On Integro Differential Splines Construction. Advances in Applied and Pure Mathematics. Proceedinngs of the 7-th International Conference on Finite Differences, Finite Elements, Finite Volumes, Boundary Elements (F-and-B'14). Gdansk. Poland. pp. 57–61 (May 15-17, 2014)
3. Kireev V.I., Panteleev A.V. Numerical methods in examples and tasks. M. 480 p. (2008)(in Russian)
4. Ruzanski, E. P., Chandrasekar, V. Weather radar data interpolation using a kernel-based lagrangian nowcasting technique. IEEE Transactions on Geoscience and Remote Sensing. Vol. 53, Issue 6(1), pp. 3073–3083 (June 2015)
5. Mariani, M. C., Basu, K. Spline interpolation techniques applied to the study of geophysical data. Physica A: Statistical Mechanics and its Applications. Vol. 428, 15. pp. 68–79 (June 2015)
6. Parker, W.D., Umrigar, C.J., Alfe, D., Petruzielo, F.R., Hennig, R.G., Wilkins, J.W. Comparison of polynomial approximations to speed up planewave-based quantum Monte Carlo calculations. Journal of Computational Physics. Vol. 287, pp. 77–87 (April 05, 2015)
7. Tzivelekis, C.A., Yiotis, L.S., Fountas, N.A., Krimpenis, A.A. Parametrically automated 3D design and manufacturing for spiral-type free-form models in an interactive CAD/CAM environment. International Journal on Interactive Design and Manufacturing. 10 p. (10 February 2015) (Articles not published yet)
8. Zavjalov Yu.S., Kvasov B.I., Miroshnichenko V.L. Metody spline-functions. M.(1980) (Russian)

Fast Infinitesimal Fourier Transform for Signal and Image Processing via Multiparametric and Fractional Fourier Transforms

Ekaterina Ostheimer[1], Valeriy Labunets[2], and Stepan Martyugin[3]

[1] Capricat LLC, 1340 S., Ocean Blvd., Suite 209, Pompano Beach, 33062 Florida, USA
katya@capricat.com,
[2] Ural Federal University, pr. Mira, 19,Yekaterinburg, 620002, Russian Federation
vlabunets05@yahoo.com,
[3] SPA Automatics, named after Academician N.A. Semikhatov, Mamina Sibiryaka, 145, Yekaterinburg, 620002, Russian Federation
stmart2608@gmail.com

Abstract. The fractional Fourier transforms (FrFTs) is one-parametric family of unitary transformations $\{\mathcal{F}^{\alpha}\}_{\alpha=0}^{2\pi}$. FrFTs found a lot of applications in signal and image processing. The identical and classical Fourier transformations are both the special cases of the FrFTs. They correspond to $\alpha = 0$ ($\mathcal{F}^0 = I$) and $\alpha = \pi/2$ ($\mathcal{F}^{\pi/2} = \mathcal{F}$), respectively. Up to now, the fractional Fourier spectra $F^{\alpha_i} = \mathcal{F}^{\alpha_i}\{f\}$, $i = 1, 2, ..., M$, has been digitally computed using classical approach based on the fast discrete Fourier transform. This method maps the N samples of the original function f to the N samples of the set of spectra $\{F^{\alpha_i}\}_{i=1}^{M}$, which requires $MN(2 + \log_2 N)$ multiplications and $MN \log_2 N$ additions. This paper develops a new numerical algorithm, which requires $2MN$ multiplications and $3MN$ additions and which is based on the infinitesimal Fourier transform.

Keywords: Fast fractional Fourier transform, infinitesimal Fourier transform, Schrödinger operator, signal and image analysis

1 Introduction

The idea of fractional powers of the Fourier operator $\{\mathcal{F}^a\}_{a=0}^{4}$ appeared in the mathematical literature [1,2,3,4]. The idea is to consider the eigen-value decomposition of the Fourier transform $\mathcal{F}$ in terms of the eigen-values $\lambda_n = e^{jn\pi/2}$ and eigen-functions in the form of the Hermite functions. The family of FrFT $\{\mathcal{F}^a\}_{a=0}^{4}$ is constructed by replacing the n-th eigen-value $\lambda_n = e^{jn\pi/2}$ by its a-th power $\lambda_n^a = e^{jn\pi a/2}$ for a between 0 and 4. This value is called the transform order. There is the angle parameterization $\{\mathcal{F}^{\alpha}\}_{\alpha=0}^{2\pi}$, where $\alpha = \pi a/2$ is a new angle parameter. Since this family depends on a single parameter,

the fractional operators $\{\mathcal{F}^a\}_{a=0}^4$ (or $\{\mathcal{F}^\alpha\}_{\alpha=0}^{2\pi}$) form the Fourier-Hermite one-parameter strongly continuous unitary multiplicative group $\mathcal{F}^a\mathcal{F}^b = \mathcal{F}^{\overset{a\oplus b}{4}}$ (or $\mathcal{F}^\alpha\mathcal{F}^\beta = \mathcal{F}^{\overset{\alpha\oplus\beta}{2\pi}}$), where $a\overset{a\oplus b}{4}b = (a+b)\,\mathrm{mod}4$ (or $\alpha\overset{\alpha\oplus\beta}{2\pi}\beta = (\alpha+\beta)\,\mathrm{mod}2\pi$) and $\mathcal{F}^0 = I$. The identical and classical Fourier transformations are both the special cases of the FrFTs. They correspond to $\alpha = 0$ ($\mathcal{F}^0 = I$) and $\alpha = \pi/2$ ($\mathcal{F}^{\pi/2} = \mathcal{F}$), respectively.

In 1980, Namias reinvented the fractional Fourier transform (FrFT) again in his paper [6]. He used the FrFT in the context of quantum mechanics as a way to solve certain problems involving quantum harmonic oscillators. He not only stated the standard definition for the FrFT, but, additionally, developed an operational calculus for this new transform. This approach was extended by McBride and Kerr [7]. Then Mendlovic and Ozaktas introduced the FrFT into the field of optics [8] in 1993. Afterwards, Lohmann [9] reinvented the FrFT based on the Wigner-distribution function and opened the FrFT to bulk-optics applications. It has been rediscovered in signal and image processing [10]. In these cases, the FrFT allows us to extract time-frequency information from the signal. A recent state of the art can be found in [11]. In the series of papers [12,13,14,15,16], we developed a wide class of classical and quantum fractional transforms.

In this paper, the infinitesimal Fourier transforms are introduced, and the relationship of the fractional Fourier transform with the Schrödinger operator of the quantum harmonic oscillator is discussed. Up to now, the fractional Fourier spectra $F^{\alpha_i} = \mathcal{F}^{\alpha_i}\{f\}$, $i = 1, 2, ..., M$, have been digitally computed using classical approach based on the fast discrete Fourier transform. This method maps the N samples of the original function f to the NM samples of the set of spectra $\{F^{\alpha_i}\}_{i=1}^M$, which requires $MN(2+\log_2 N)$ multiplications and $MN\log_2 N$ additions. This paper develops a new numerical algorithm, which requires $2MN$ multiplications and $3MN$ additions and which is based on the infinitesimal Fourier transform.

2 Eigen-decomposition and Fractional Discrete Transforms

Let $\mathcal{F} = [F_k(i)]_{k,i=0}^{N-1}$ be an arbitrary discrete unitary $(N \times N)$-transform, λ_n and $\Psi_n(t)$ $n = 0, 1, \ldots, N-1$ be its eigen-values and eigen-vectors, respectively. Let $\mathbf{U} = \left[\Psi_0(i)|\Psi_1(i)|\vdots|\Psi_{N-1}(i)\right]$ be the matrix of the $\mathcal{F}$-transform eigen-vectors. Then $\mathbf{U}^{-1}\cdot\mathcal{F}\cdot\mathbf{U} = \mathbf{Diag}\{\lambda_n\}$. Hence, we have the following eigen-decomposition: $\mathcal{F} = [F_k(i)] = \mathbf{U}\cdot\mathbf{\Lambda}\cdot\mathbf{U}^{-1} = \mathbf{U}\cdot\mathbf{Diag}\{\lambda_n\}\cdot\mathbf{U}^{-1}$.

Definition 1. [12]. *For an arbitrary real numbers $a_0, \ldots, a_{N-1}$, we introduce the multi-parametric $\mathcal{F}$-transform*

$$\mathcal{F}^{(a_0,\ldots,a_{N-1})} := \mathbf{U}\left\{\mathbf{diag}\left(\lambda_0^{a_0}, \ldots, \lambda_{N-1}^{a_{N-1}}\right)\right\}\mathbf{U}^{-1}. \tag{1}$$

If $a_0 = \ldots = a_{N-1} \equiv a$ then this transform is called fractional $\mathcal{F}$-transform [12,13,14,15,16]. For this transform we have

$$\mathcal{F}^a := \mathbf{U}\left\{\mathbf{diag}\left(\lambda_0^a, \ldots, \lambda_{N-1}^a\right)\right\}\mathbf{U}^{-1} = \mathbf{U}\mathbf{\Lambda}^a\mathbf{U}^{-1}. \tag{2}$$

The zero-th-order fractional $\mathcal{F}$-transform is equal to the identity transform: $\mathcal{F}^0 = \mathbf{U}\mathbf{\Lambda}^0\mathbf{U}^{-1} = \mathbf{U}\mathbf{U}^{-1} = \mathbf{I}$, and the first-order fractional Fourier transform operator $\mathcal{F}^1 = \mathcal{F}$ is equal to the initial $\mathcal{F}$-transform $\mathcal{F}^1 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$.

The families $\left\{\mathcal{F}^{(\alpha_0,\ldots,\alpha_{N-1})}\right\}_{(\alpha_0,\ldots,\alpha_{N-1})\in\mathbf{R}^N}$ and $\{\mathcal{F}^a\}_{a\in\mathbf{R}}$ form multi- and one-parameter continuous unitary groups, respectively, with multiplication rules

$$\mathcal{F}^{(a_0,\ldots,a_{N-1})}\mathcal{F}^{(b_0,\ldots,b_{N-1})} = \mathcal{F}^{(a_0+b_0,\ldots,a_{N-1}+b_{N-1})} \quad \text{and} \quad \mathcal{F}^a\mathcal{F}^b = \mathcal{F}^{a+b}.$$

Indeed, $\mathcal{F}^a\mathcal{F}^b = \mathbf{U}\mathbf{\Lambda}^a\mathbf{U}^{-1}\cdot\mathbf{U}\mathbf{\Lambda}^b\mathbf{U}^{-1} = \mathbf{U}\mathbf{\Lambda}^{a+b}\mathbf{U}^{-1} = \mathcal{F}^{a+b}$ and

$$\mathcal{F}^{(a_0,\ldots,a_{N-1})}\mathcal{F}^{(b_0,\ldots,b_{N-1})} =$$

$$= \mathbf{U}\left\{\mathbf{diag}\left(\lambda_0^{a_0}, \ldots, \lambda_{N-1}^{a_{N-1}}\right)\right\}\mathbf{U}^{-1}\cdot\mathbf{U}\left\{\mathbf{diag}\left(\lambda_0^{b_0}, \ldots, \lambda_{N-1}^{b_{N-1}}\right)\right\}\mathbf{U}^{-1} =$$

$$= \mathbf{U}\left\{\mathbf{diag}\left(\lambda_0^{a_0+b_0}, \ldots, \lambda_{N-1}^{a_{N-1}+b_{N-1}}\right)\right\}\mathbf{U}^{-1} = \mathcal{F}^{(a_0+b_0,\ldots,a_{N-1}+b_{N-1})}.$$

Let $\mathcal{F} = [F_k(i)]_{k,i=0}^{N-1}$ be a discrete Fourier $(N\times N)$-transform (DFT), then $\lambda_n = e^{j\pi n/2} \in \{\pm 1, \pm j\}$, where $j = \sqrt{-1}$ and $\{\Psi_n(t)\}_{n=0}^{N-1}$ are the Kravchuk polynomials.

Definition 2. *The multi-parametric and fractional DFT are*

$$\mathcal{F}^{(a_0,\ldots,a_{N-1})} := \mathbf{U}\left\{\mathbf{diag}\left(e^{j\pi 0a_0/2}, e^{j\pi 1a_1/2}, \ldots, e^{j\pi(N-1)a_{N-1}/2}\right)\right\}\mathbf{U}^{-1},$$

$$\mathcal{F}^a := \mathbf{U}\left\{\mathbf{diag}\left(e^{j\pi na/2}\right)\right\}\mathbf{U}^{-1}$$

and

$$\mathcal{F}^{(\alpha_0,\ldots,\alpha_{N-1})} := \mathbf{U}\left\{\mathbf{diag}\left(e^{j0\alpha_0}, e^{j1\alpha_1}, \ldots, e^{j(N-1)\alpha_{N-1}}\right)\right\}\mathbf{U}^{-1},$$

$$\mathcal{F}^{\alpha} := \mathbf{U}\left\{\mathbf{diag}\left(e^{jn\alpha}\right)\right\}\mathbf{U}^{-1}$$

in a- and α-parameterizations, respectively, where $\alpha = \pi a/2$.

The parameters $(a_0, \ldots, a_{N-1})$ and a can be any real values. However, the operators $\mathcal{F}^{(a_0,\ldots,a_{N-1})}$ and $\mathcal{F}^a$ are periodic in each parameter with period 4 since $\mathcal{F}^4 = I$. Hence, $\mathcal{F}^{(a_0,\ldots,a_{N-1})}\mathcal{F}^{(b_0,\ldots,b_{N-1}} = \mathcal{F}^{\substack{(a_0\oplus b_0,\ldots,a_{N-1}\oplus b_{N-1})\\ 4}}$ and $\mathcal{F}^a\mathcal{F}^b = \mathcal{F}^{\substack{a\oplus b\\ 4}}$, where $a_i\underset{4}{\oplus}b_i = (a_i+b_i)\bmod 4$, $\forall i = 0, 1, ..., N-1$. Therefore, the ranges of $(a_0, \ldots, a_{N-1})$ and a are $(\mathbf{Z}/4\mathbf{Z})^N = [0,4]^N = [-2,2]^N$ and $\mathbf{Z}/4\mathbf{Z} = [0,4] = [-2,2]$, respectively.

In the case of α-parameterization, we have $\alpha_i\underset{2\pi}{\oplus}\beta_i = (\alpha_i+\beta_i)\bmod 2\pi$, $\forall i = 0, 1, ..., N-1$. So, the ranges of $(\alpha_0, \ldots, \alpha_{N-1})$ and α are $(\mathbf{Z}/2\pi\mathbf{Z})^N = [0, 2\pi]^N = [-\pi, \pi]^N$ and $\mathbf{Z}/2\pi\mathbf{Z} = [0, 2\pi] = [-\pi, \pi]$, respectively.

3 Canonical FrFT

The continuous Fourier transform is a unitary operator $\mathcal{F}$ that maps square-integrable functions on square-integrable ones and is represented on these functions $f(x)$ by the well-known integral

$$F(y) = (\mathcal{F}f)(y) = \frac{1}{\sqrt{2\pi}} \int_{x \in \mathbf{R}} f(x) e^{-jyx} dx. \tag{3}$$

Relevant properties are that the square $\left(\mathcal{F}^2 f\right)(x) = f(-x)$ is the inversion operator, and that its fourth power $\left(\mathcal{F}^4 f\right)(x) = f(x)$ is the identity. Hence, $\mathcal{F}^3 = \mathcal{F}^{-1}$. Thus, the operator $\mathcal{F}$ generates a cyclic group of the order 4. In 1961, Bargmann extended the Fourier transform in his paper [5] where he gave definition of the FrFT that was based on the Hermite polynomials as an integral transformation. If $H_n(x)$ is a Hermite polynomial of order n, where $H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{x^2}$, then for $n \in \mathbf{N}_0$, functions $\Psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x) e^{-x^2/2}$ are the eigen-functions of the Fourier transform

$$\mathcal{F}\left[\Psi_n(x)\right] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Psi_n(x) e^{2\pi \mathrm{j} y x} dx = \lambda_n \Psi_n(y) = e^{-j\frac{\pi}{2}n} \Psi_n(y)$$

with $\lambda_n = j^n = e^{-j\frac{\pi}{2}n}$ being the eigen-value corresponding to the n-th eigen-function. According to Bargmann, the fractional Fourier transform $\mathcal{F}^\alpha = [K^\alpha(x,y)]$ is defined through its eigen-functions as

$$K^\alpha(x,y) := \mathbf{U}\left\{\mathbf{diag}\left(\mathbf{e}^{-\mathbf{j}\alpha\mathbf{n}}\right)\right\}\mathbf{U}^{-1} = \sum_{n=0}^{\infty} e^{-j\alpha n} \Psi_n(x) \Psi_n(y). \tag{4}$$

Hence,

$$K^\alpha(x,y) := \sum_{n=0}^{\infty} e^{-j\alpha n} \Psi_n(x) \Psi_n(y) = e^{-\left(x^2+y^2\right)} \sum_{n=0}^{\infty} \frac{e^{-j\alpha n} H_n(x) H_n(y)}{2^n n! \sqrt{\pi}} =$$

$$= \frac{1}{\sqrt{\pi}\sqrt{1-e^{-2j\alpha}}} \cdot \exp\left\{\frac{2xye^{-j\alpha} - e^{-2j\alpha}\left(x^2+y^2\right)}{1-e^{-2j\alpha}}\right\} \exp\left\{-\frac{\left(x^2+y^2\right)}{2}\right\}, \tag{5}$$

where $K^\alpha(x,y)$ is the kernel of the FrFT. In the last step we used the Mehler formula [19]

$$\sum_{n=0}^{\infty} \frac{e^{-j\alpha n} H_n(x) H_n(y)}{2^n n! \sqrt{\pi}} = \frac{1}{\sqrt{\pi}\sqrt{1-e^{-2j\alpha}}} \exp\left\{\frac{2xye^{-j\alpha} - e^{-2j\alpha}\left(x^2+y^2\right)}{1-e^{-2j\alpha}}\right\}.$$

Expression (5) can be rewritten as

$$K^\alpha(x,y) = \sqrt{\frac{1-j\cot\alpha}{2\pi}} \exp\left\{\frac{j}{2\sin\alpha}\left[\left(x^2+y^2\right)\cos\alpha - 2xy\right]\right\},$$

where $\alpha \neq \pi\mathbf{Z}$ (or $a \neq 2\mathbf{Z}$). Obviously, functions $\Psi_n(x)$ are eigen-functions of the fractional Fourier transform $\mathcal{F}^\alpha[\Psi_n(x)] = e^{jn\alpha}\Psi_n(x)$ corresponding to the n-th eigen-values $e^{jn\alpha}$, $n = 0, 1, 2, \ldots$. The FrFT $\mathcal{F}^\alpha$ is a unitary operator that maps square-integrable functions $f(x)$ on square-integrable ones

$$F^\alpha(y) = (\mathcal{F}^\alpha f)(y) = \int_{x\in\mathbf{R}} f(x)K^\alpha(x,y)dx =$$

$$= \frac{e^{-\frac{j}{2}\left(\frac{\pi}{2}\hat{\alpha}-\alpha\right)}}{\sqrt{2\pi\,|\sin\alpha|}} \int_{\mathbf{R}} f(x)\exp\left\{\frac{j}{2\sin\alpha}\left[\left(x^2+y^2\right)\cos\alpha - 2xy\right]\right\}dx.$$

There exist several algorithms for fast calculation of spectrum of the fractional Fourier transform $F^\alpha(y)$. But all of them are based on the following transform of the FrFT:

$$F^\alpha(y) = (\mathcal{F}^\alpha f)(y) = \frac{e^{-\frac{j}{2}\left(\frac{\pi}{2}\hat{\alpha}-\alpha\right)}e^{jy^2\frac{\cos\alpha}{2\sin\alpha}}}{\sqrt{2\pi\,|\sin\alpha|}} \int_{\mathbf{R}} \left[f(x)e^{j\frac{x^2}{2}\cot\alpha}\right]e^{-jxy}dx =$$

$$= A_\alpha(y)\cdot\mathcal{F}\left\{f(x)\cdot B_\alpha(x)\right\}(y),$$

where $A_\alpha(y) = \frac{e^{-\frac{j}{2}\left(\frac{\pi}{2}\hat{\alpha}-\alpha\right)}e^{jy^2\frac{\cos\alpha}{2\sin\alpha}}}{\sqrt{2\pi|\sin\alpha|}}$, $B_\alpha(x) = e^{j\frac{x^2}{2}\cot\alpha}$.

Let us introduce the uniform discretization of the angle parameter α on M discrete values $\{\alpha_0, \alpha_1, ..., \alpha_i, \alpha_{i+1}, ..., \alpha_{M-1}\}$, where $\alpha_{i+1} = \alpha_i + \Delta\alpha$, $\alpha_i = i\Delta\alpha$ and $\Delta\alpha = 2\pi/M$.

The set of M spectra $\{F^{\alpha_0}(y), F^{\alpha_1}(y), ..., F^{\alpha_{M-1}}(y)\}$ can be computed by applying the following sequence of steps for all $\{\alpha_0, \alpha_1, ..., \alpha_{M-1}\}$:

1. Compute products $f(x)B_{\alpha_k}(x)$, which require N multiplications.

2. Compute the Fast Fourier Transform ($N\log_2 N$ multiplications and additions).

3. Multiply the result by $A_\alpha(y)$ (N multiplications).

This numerical algorithm requires $MN\log_2 N$ additions and $MN(2+\log_2 N)$ multiplications.

4 Infinitesimal Fourier Transform

In order to construct fast multi-parametric $\mathcal{F}$-transform and fractional Fourier transform algorithms we turn our attention to notion of a semigroup and its generator (infinitesimal operator). Let $\mathrm{L}_2(\mathbf{R}, \mathbf{C})$ be a space of complex-valued functions (signals), and let $\mathbf{Op}(\mathrm{L}_2)$ be the Banach algebra of all bounded linear operators on $\mathrm{L}_2(\mathbf{R}, \mathbf{C})$ endowed with the operator norm. A family $\{\mathbf{U}(\alpha)\}_{\alpha\in\mathbf{R}} \subset \mathbf{Op}(\mathrm{L}_2)$ is called the Hermite group on $\mathrm{L}_2(\mathbf{R}, \mathbf{C})$ if it satisfies the Abel functional equations $\mathbf{U}(\alpha+\beta) = \mathbf{U}(\alpha)\mathbf{U}(\beta)$, $\alpha, \beta \in \mathbf{R}$ and $\mathbf{U}(0) = \mathbf{I}$, and the orbit maps $\alpha \to F^\alpha = \mathbf{U}(\alpha)\{f\}$ are continuous from $\mathbf{R}$ into $\mathrm{L}_2(\mathbf{R}, \mathbf{C})$ for every $f \in \mathrm{L}_2(\mathbf{R}, \mathbf{C})$.

Definition 3. *The infinitesimal generator* $\mathbf{A}(0)$ *of the group* $\{\mathbf{U}(\alpha)\}_{\alpha\in\mathbf{R}}$ *and the infinitesimal transform* $\mathbf{U}(d\alpha)$ *are defined as follows* [18,19]*:*

$$\mathbf{A}(0)=\left.\frac{\partial\mathbf{U}(\alpha)}{\partial\alpha}\right|_{\alpha=0},\quad \mathbf{U}(d\alpha)=\mathbf{I}+d\mathbf{U}(0)=\mathbf{I}+\mathbf{A}(0)d\alpha.$$

Obviously,

$$\mathbf{U}(\alpha_0+d\alpha)=\mathbf{U}(\alpha_0)+d\mathbf{U}(\alpha_0)=\mathbf{U}(\alpha_0)+\left.\frac{\partial\mathbf{U}(\alpha)}{\partial\alpha}\right|_{\alpha_0}d\alpha=$$

$$=\mathbf{U}(\alpha_0)+\mathbf{A}(\alpha_0)d\alpha.$$

But

$$\mathbf{U}(\alpha_0+d\alpha)=\mathbf{U}(d\alpha_0)\mathbf{U}(\alpha_0)=\left[\mathbf{I}+d\mathbf{U}(0)\right]\mathbf{U}(\alpha_0)=$$

$$=\mathbf{U}(\alpha_0)+\left.\frac{\partial\mathbf{U}(\alpha)}{\partial\alpha}\right|_{\alpha=0}\mathbf{U}(\alpha_0)d\alpha=$$

$$=\mathbf{U}(\alpha_0)+\mathbf{A}(0)\mathbf{U}(\alpha_0)d\alpha=\left[\mathbf{I}+\mathbf{A}(0)\right]\mathbf{U}(\alpha_0)d\alpha.$$

Hence, $\mathbf{A}(\alpha_0)=\mathbf{A}(0)\mathbf{U}(\alpha_0)$ and $F^{\alpha_0+d\alpha}(y)=\left[\mathbf{I}+\mathbf{A}(0)\right]F^{\alpha_0}(y)d\alpha$.

Define now the linear operator $\mathcal{H}=\frac{1}{2}\left(\frac{d^2}{dx^2}-x^2+1\right)$. It is known that

$$\mathcal{H}\Psi_n(x)=\frac{1}{2}\left(\frac{d^2}{dx^2}-x^2+1\right)\Psi_n(x)=n\Psi_n(x). \tag{6}$$

From (4) and (6) we have

$$j\frac{\partial}{\partial\alpha}F^{\alpha}(y)\Big|_{\alpha=0}=j\frac{\partial}{\partial\alpha}\left\{\mathcal{F}^{\alpha}F\right\}(y)\Big|_{\alpha=0}=\sum_{n=0}^{\infty}n\Psi_n(y)\int_{\mathbf{R}}\Psi_n(x)f(x)dx,$$

$$\mathcal{H}F^{\alpha}(x)=\sum_{n=0}^{\infty}n\Psi_n(y)\int_{\mathbf{R}}\Psi_n(x)f(x)dx.$$

Therefore, $j\frac{\partial F^{\alpha}(x)}{\partial\alpha}=\mathcal{H}F^{\alpha}(y)$, $\frac{\partial F^{\alpha}(x)}{F^{\alpha}(x)}=-j\mathcal{H}\partial\alpha$. The solution of this equation is given by $F^{\alpha}(x)=\left\{e^{-j\alpha\mathcal{H}}F\right\}$ and $\mathcal{F}^{\alpha}=e^{-j\alpha\mathcal{H}}=e^{-j\alpha\left[\frac{1}{2}\left(\frac{d^2}{dx^2}-x^2+1\right)\right]}$. Obviously,

$$\mathcal{F}^{\alpha+d\alpha}=\mathcal{F}^{d\alpha}\mathcal{F}^{\alpha}\simeq\left(\mathbf{I}+d\mathcal{F}^{\alpha}\right)\exp\left[-j\alpha\mathcal{H}\right]=$$

$$=\left(\mathbf{I}+\frac{\partial\mathcal{F}^{\alpha}}{\partial\alpha}d\alpha\right)\exp\left(-j\alpha\mathcal{H}\right)=\left(\mathbf{I}-j\mathcal{H}d\alpha\right)\exp\left(-j\alpha\mathcal{H}\right),$$

where the operator

$$\mathcal{F}^{d\alpha}=\left(\mathbf{I}-j\mathcal{H}d\alpha\right)=\mathbf{I}-j\frac{1}{2}\left(\frac{d^2}{dx^2}-x^2+1\right)d\alpha \tag{7}$$

is called the *infinitesimal Fourier transform* or the *generator* of the fractional Fourier transforms [17,18].

Let us introduce operators $(M_x f)(x) := xf(x)$ and $(M_y F)(y) := yF(y)$. Using the Fourier transform (3), the first of ones may be written as $M_x = \mathcal{F}^{-1}\left(j\frac{d}{dy}\right)\mathcal{F}$. Obviously, $x^2 = M_x^2 = -\mathcal{F}^1\left(\frac{d^2}{dy^2}\right)\mathcal{F}$. Then

$$\mathcal{F}^{d\alpha} = \mathbf{I} - j\frac{1}{2}\left(\frac{d^2}{dx^2} + \mathcal{F}^{-1}\left(\frac{d^2}{dy^2}\right)\mathcal{F} + 1\right)d\alpha.$$

Discretization of x-domain with the interval discretization Δx is equal to the periodization of y-domain

$$\frac{d^2}{dx^2} + \mathcal{F}^{-1}\left(\frac{d^2}{dy^2}\right)\mathcal{F} + 1 \longrightarrow D_{\Delta x}\left[\frac{d^2}{dx^2}\right] + \mathcal{F}^{-1}\left(P_{2\pi/\Delta x}\left[\frac{d^2}{dy^2}\right]\right)\mathcal{F} + 1.$$

Discretization of y-domain with the interval discretization Δy is equal to the periodization of x-domain

$$D_{\Delta x}\left[\frac{d^2}{dx^2}\right] + \mathcal{F}^{-1}\left(P_{2\pi/\Delta x}\left[\frac{d^2}{dy^2}\right]\right)\mathcal{F} + 1 \longrightarrow$$

$$\longrightarrow P_{2\pi/\Delta y}D_{\Delta x}\left[\frac{d^2}{dx^2}\right] + \mathcal{F}^{-1}\left(P_{2\pi/\Delta x}D_{\Delta y}\left[\frac{d^2}{dy^2}\right]\right)\mathcal{F} + 1.$$

An approximation for the second derivative can be given by the second order central difference operator

$$\frac{d^2}{dx^2}f(x) \approx f(n\underset{N}{\ominus}1) - 2f(n) + F(n\underset{N}{\oplus}1), \quad \frac{d^2}{dy^2}F(y) \approx F(k\underset{N}{\ominus}1) - 2F(k) + F(k\underset{N}{\oplus}1),$$

where $N = 2\pi/\Delta x\Delta y$. On the other hand,

$$\mathcal{F}^{-1}\left(\frac{d^2}{dy^2}F(y)\right)\mathcal{F} \approx \mathcal{F}^{-1}\left[F(k\underset{N}{\ominus}1) - 2F(k) + F(k\underset{N}{\oplus}1)\right]\mathcal{F} =$$

$$= \left(f(n)e^{-j\frac{2\pi}{N}n} - 2f(n) + f(n)e^{j\frac{2\pi}{N}n}\right) = 2f(n)\left(\cos\frac{2\pi}{N}n - 1\right).$$

These allow one to give the approximation for $\mathcal{H} = \frac{1}{2}\left(\frac{d^2}{dx^2} - x^2 + 1\right)$ as follows:

$$\mathcal{H}f(x) = \left[\frac{1}{2}\left(\frac{d^2}{dx^2} - x^2 + 1\right)\right]f(x) \approx$$

$$\approx \frac{1}{2}\left\{\left[f(n\underset{N}{\ominus}1) - 2f(n) + f(n\underset{N}{\oplus}1)\right] + 2f(n)\left(\cos\frac{2\pi}{N}n - 1\right) + f(n)\right\} =$$

$$= -\left[\cos\frac{2\pi}{N}n - 3/2\right]f(n) + \frac{1}{2}\left[f(n\underset{N}{\ominus}1) + f(n\underset{N}{\oplus}1)\right].$$

In the N-diagonal basis we have

$$\mathcal{F}^{d\alpha} f(x) \approx \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \\ \vdots \\ f(N-1) \end{bmatrix} + j\Delta\alpha \times$$

$$\times \left[\begin{array}{c|c|c|c|c|c} -1/2 & 1/2 & . & . & . & 1/2 \\ \hline 1/2 & \cos(1\Omega) - 3/2 & 1/2 & . & . & . \\ \hline . & 1/2 & \cos(2\Omega) - 3/2 & 1/2 & . & . \\ \hline . & . & 1/2 & \cos(3\Omega) - 3/2 & 1/2 & . \\ \hline . & . & . & 1/2 & \ddots & 1/2 \\ \hline 1/2 & . & . & . & 1/2 & -1/2 \end{array}\right] \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \\ \vdots \\ f(N-1) \end{bmatrix}, \tag{8}$$

where $\Omega = 2\pi/N$.

Let us introduce the uniform discretization of the angle parameter α on M discrete values $\{\alpha_0, \alpha_1, ..., \alpha_i, \alpha_{i+1}, ..., \alpha_{M-1}\}$, where $\alpha_{i+1} = \alpha_i + \Delta\alpha$, $\alpha_i = i\Delta\alpha$ and $\Delta\alpha = 2\pi/M$. Then

$$F^{\alpha_{i+1}}(y) = F^{\alpha_i + \Delta\alpha}(y) \approx F^{\alpha_i}(k) + j\Delta\alpha \times$$

$$\times \left\{ \left[\cos\frac{2\pi}{N} k - 3/2 \right] F^{\alpha_i}(k) + \frac{1}{2} \left[F^{\alpha_i}(k \underset{N}{\ominus} 1) + F^{\alpha_i}(k \underset{N}{\oplus} 1) \right] \right\}. \tag{9}$$

It is easy to see that this algorithm requires only $2MN$ multiplications and $3MN$ additions vs. $MN(2 + \log_2 N)$ multiplications and $MN \log_2 N$ additions in the classical algorithm. In (8), we used $\mathcal{O}(h^2)$ approximation $\left(\frac{d^2}{dx^2} f\right)(k) \approx (f(k-1) - 2f(k) + f(k+1))$. More fine approximations $\mathcal{O}(h^{2k})$ also can be used [19].

5 Conclusions

In this work, we introduce a new algorithm of computing for Fractional Fourier transforms based on the infinitesimal Fourier transform. It requires $2MN$ multiplications and $3MN$ additions vs. $MN(2 + \log_2 N)$ multiplications and $MN \log_2 N$ additions in the classical algorithm. Presented algorithm can be utilized for fast computation in most applications of signal and image processing. We have presented a definition of the infinitesimal Fourier transform that exactly satisfies the properties of the Schrodinger Equation for quantum harmonic oscillator.

Acknowledgment. This work was supported by grants the RFBR Nos. 13-07-12168, and 13-07-00785.

References

1. Mitra, S.K.: Nonlinear Image Processing. Academic Press Series in Communications, Networking, and Multimedia, San Diego, New York, 248 (2001)
2. Wiener, N.: Hermitian polynomials and Fourier analysis. J. Math. Phys. 8, 70–73 (1929).
3. Condon, E.U.: Immersion of the Fourier transform in a continuous group of functional transforms. Proc. Nat. Acad. Sci. 23, 158–164 (1937)
4. Kober, H.: Wurzeln aus der Hankel- und Fourier und anderen stetigen Transformationen. Quart. J. Math. Oxford Ser. 10, 45–49 (1939)
5. Bargmann, V.: On a Hilbert space of analytic functions and an associated integral transform. Part 1. Commun. Pure Appl. Math. 14, 187–214 (1961)
6. Namias, V.: The fractional order Fourier transform and its application to quantum mechanics. J. Inst. Math. Appl. 25, 241–265 (1980)
7. McBride, A.C., Kerr, F.H.: On Namias' fractional Fourier transforms. IMA J. Appl. Math. 39, 159–265 (1987).
8. Ozaktas, H.M., Mendlovic, D.: Fourier transform of fractional order and their optical interpretation. Opt. Commun. 110, 163–169 (1993)
9. Lohmann, A.W.: Image rotation, Wigner rotation, and the fractional order Fourier transform. J. Opt. Soc. Am. A. 10, 2181–2186 (1993)
10. Almeida, L.B.: The fractional Fourier transform and time-frequency representation. IEEE Trans. Sig. Proc. 42, 3084–3091 (1994)
11. Ozaktas, H.M., Zalevsky, Z., Kutay M.A.: The fractional Fourier transform. Wiley, Chichester, (2001)
12. Rundblad–Labunets, E., Labunets, V., Astola, J., Egiazarian, K.: Fast fractional Fourier-Clifford transforms. Second International Workshop on Transforms and Filter Banks. Tampere, Finland, TICSP Series, 5, 376–405 (1999)
13. Rundblad, E., Labunets, V., Astola, J., Egiazarian, K., Polovnev, S.: Fast fractional unitary transforms. Proc. of Conf. Computer Science and Information Technologies, Yerevan, Armenia, 223–226 (1999)
14. Creutzburg, R., Rundblad, E., Labunets, V.: Fast algorithms for fractional Fourier transforms. Proc. of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing. Antalya, Turkey, June 20–23, 383–387 (1999)
15. Rundblad, E., Labunets, V., Astola, J., Egiazarian, K., Smaga, A.: Fast fractional Fourier and Hartley transforms. Proc. of the 1999 Finnish Signal Processing Symposium. Oulu, Finland, 291–297 (1999)
16. Labunets, E., Labunets, V.: Fast fractional Fourier transforms. Proc. of Eusipco-98, Rhodes, Greece, 8–11 Sept. 1757–1760 (1998)
17. Goldstein, J.: Semigroups of Linear Operators and Applications, Oxford U. Press, New York, (1985)
18. Griffiths, R.B.: Consistent Quantum Theory. `http://www.cambridge.org`
19. Candan C.: The Discrete Fractional Fourier Transform, M.S. thesis, Bilkent Univ., Ankara, Turkey, (1998)

Approximating Social Ties Based on Call Logs: Whom Should We Prioritize?

Mohammad Erfan, Alim Ul Gias, Sheikh Muhammad Sarwar, and Kazi Sakib

Institute of Information Technology
University of Dhaka, Dhaka–1000, Bangladesh
{bit0326,alim,smsarwar,sakib}@iit.du.ac.bd

Abstract. Telecommunication service providers could analyze the call logs of individual users to provide personalized services like recommending Friends and Family (FnF) numbers, that could be registered for having lower tariff during communication. However, suggesting such numbers is not simple as the decision making process can often be misled by factors like overall talk-time or call frequency. This paper proposes a technique to assist mobile phone users for effectively identifying FnF numbers using the underlying social network of a call graph. It resembles to a star network, having the user at the center, where each of the edges are weighted using three call log attributes: call duration, call location and call period. The weighted call graph is then used to calculate user's social closeness based on which FnF numbers are suggested. The experiment was conducted on a set of real life data and it is seen that the proposed technique can effectively suggest FnF numbers with an F-measure value of 0.67.

Keywords: Social network, mobile phone call graph, call log attributes, social closeness

1 Introduction

Personalized services can be defined as context-specific services to each individual client [1]. Since mobile phone can always be carried by the client, it has the potential to be an ideal medium of such personalization. This potential allows telecommunication service providers to offer multiple personalized services like customized ring tone suggestions or different location based information. Providing such services can increase the popularity of the service providers [2] and help them to sustain in the competitive market. Thus, it is high time telecommunication companies focused more on providing personalized services to each user for eventually increasing their overall revenue.

Recommending Friends and Family (FnF) numbers [3], that can be registered by the user for having lower call rate, could be an example of customized personal services. To provide such service, it is essential to retrieve each user's calling pattern by analyzing their respective call logs. However, analysis and interpretation of these call logs is particularly difficult and introduces major research issues [4].

A possible solution could be, utilizing the underlying social network [5] of those call logs. Nevertheless, constructing such network involves multiple challenges like relative prioritization of in and out degrees, and assigning weight to edges by using different call log attributes.

Although multiple research has been conducted based on mobile phone call logs, none has focused directly to provide personalized services to each individual user. There have been some work that focus on investigating social ties based on different call attributes [6], [7]. Some of the researchers have focused on predicting user behavior [8] and community dynamics [9]. Moreover, there are works which forecast incoming calls [10] and correlate international calls with the export-import business of the country [11]. However, it is arguable that whether these state of the arts can be used to provide effective personalized services that can satisfy each individual user.

This paper proposes a methodology to create a weighted call graph based on individual call logs and use the graph's underlying social network for predicting FnF numbers. To weight the edges, three call log attributes are used which are - call duration, call period and call location. The technique empirically measures the effect of these attributes towards social closeness and calculates their relative significance. The summation of these relative significances, that prevents any of the attributes being predominated, is considered as the weight. A closeness score is then calculated by aggregating the product of the weights with values respect to their call being either incoming or outgoing. This score is used by the method to rank and suggest numbers for registering as FnF.

The method was implemented and its performance was evaluated based on a real life dataset. The dataset includes all required call log attributes along with a user provided score which indicates the strength of social tie with the person of that particular call. These scores were considered as the ground truth when weights for edges of the graph are being learned. The social closeness was then measured, based on which numbers were ranked, to predict FnF numbers. The performance of the proposed technique was compared along with traditional approaches based on overall talk-time and call frequency. It seen that the method outperforms the call frequency and overall talk-time based approaches with 38% and 10% higher F-measure value.

Rest of the paper is organized as follows: Section 2 highlights state of the art techniques for analyzing call logs to derive different social contexts. The proposed technique to predict FnF numbers is presented in Section 3. Section 4 includes the experimental details and performance of the proposed technique. Finally, this paper is concluded in Section 5 with a discussion about our work and future research directions.

2 Literature Review

Although log analysis has been a major research concern for a considerable amount of time [12], its application in mobile phone call pattern analysis has started comparatively later. To the best of authors' knowledge, no work directly

focus on utilizing the underlying social network obtained from a single user's call logs to provide customized services. The work that have been done mostly focus on constructing social networks from call logs, predicting the user behavior and identifying multiple social contexts like social affinity or community dynamics. This section highlights these contributions and discusses how they contrast to our problem domain.

There have been some significant work which are concerned with predicting the user behavior. For an example, Dong et. al. proposed a technique to create an undirected binary graph from the call log to find the human behavior of different age and gender [8]. The researchers in [13] proposed a method to derive the user relation using Bayesian network. Phithakkitnukoon et. al. analyzed multiple call log attributes such as call location, talk-time, calling time and inter-arrival time to infer behavioral dependencies [10]. The researcher in [14] inferred friendship network by utilizing the behavioral and survey data. Zhang et. al. have focused on information retrieval techniques to find human behavior patterns and community dynamics [9]. There have also been work to assess the strength of social tie based on call duration [6].

Identifying different social contexts has been another field of major research interest. Phithakkitnukoon et. al. proposed a method to identify social network based on the user phone call log [15]. They have used correlation coefficient to select call log features and then extracted user's social behavioral pattern from that network. Another work of Phithakkitnukoon et. al. has been presented in [16] where they created a network graph to derive socially closest and distant members. The researchers in [11] analyzed international phone calls from Rwanda, between January 2005 to December 2008, to derive the interpersonal connections between the people of different countries. They have used those findings to measure and evaluate the social tie between nations.

Researchers have also focused on the construction of call graph by utilizing different call log attributes. The researchers in [17] created a social network graph using call duration, source number, destination number and physical location of the user. They have extracted these information from Call Detail Record (CDR) [18] files and and generated the adjacency matrix. The edges were weighted using the call duration value. Motahari et. al. have also used the CDR files and created multiple affinity networks such as family members, utility network, friends and co-workers [7]. They have used CDR files containing 4.3 million phone call data which was collected from 360,000 subscribers.

Review of the state of the art shows that these works are generally focused on predicting users' behavioral pattern rather than utilizing those patterns for providing better user experience. It is arguable that whether these methods can be used for determining user's social tie along with utilizing the findings in a separate strategy to suggest FnF numbers. From different research, it is also seen that the edges of the constructed call graph is weighted using call duration. Nevertheless, it may not be the correct representation of social affinity. The reason behind this is though frequent or long communication likely indicates a strong tie, little or no communication does not necessarily indicate a weak tie [6].

Thus it is required to incorporate other call log attributes, to assign weights on graph edges, for actual representation of individual's social tie. This weighted graph can be easily used to infer the socially closest members and their numbers can be suggested as FnF.

3 Proposed Method

This research proposes a methodology to create a social network graph using the phone call log. The network will resemble to a star network where the phone user will be the center node and individuals with whom communication has occurred are the peripheral nodes. There will be weighted edges between the central and peripheral nodes which will indicate the strength of their social ties.The graph will have two type of edges, one for the incoming calls and other for the outgoing ones. The frequency of calls will not effect the number of edges since it will be considered during the weight calculation. A sample social network is presented in Figure 1. From the figure it can be seen that communications are represented by weighted edges where the outgoing call to individual p_n has a weight w_{n1} and the incoming call has a weight w_{n2}.

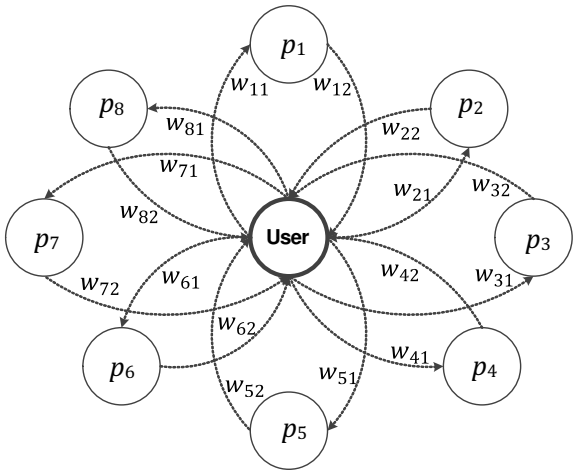


Fig. 1. Call graph generated from a single user's call logs

Initially call log records will be taken from the phone. The call logs should have the following information: call duration, call period (a period of the day like 3:00 AM to 6:00 AM) and call location (the place from which the call was made). It is essential to utilize all these information for weighting graph edges since using only call duration can often misled regarding the approximation of social closeness. For an example, a user may talk with a person at his office for a long time. However, that person could be a client instead of being a friend

or relative. So it will be wrong to consider that person as a socially close one based on the call duration. Moreover, talking with a someone during the office hours (like 9:00 AM to 5:00 PM) for a long time might misled in estimating the nature of social affinity. Thus it would be more reasonable to assign a weight by combining all these three information. The proposed method utilizes these call log attributes to assign a weight which is defined in Equation (1).

$$\omega_{ij} = \sum_{k=1}^{N} (\alpha \cdot \delta_{ij_k} + \beta_{q(ij_k)} + \gamma_{r(ij_k)}) \tag{1}$$

In Equation (1), ω_{ij} refers to the weight of the edge between the person p_i and phone's user for all calls of type j (like incoming, outgoing or missed). The set of all j type calls, with respect to person p_i is represented by $\mathcal{C}_{ij}$. If $|\mathcal{C}_{ij}| = N$, the weight ω_{ij} will be the summation of N values. The values will depend on call duration (δ), call period and call location. The value of call duration is used by multiplying it with a weight α that reflects the significance of call duration towards social ties. To represent the significance of call period and call location, two sets of values, $\mathcal{S}$ and $\mathcal{L}$ has been introduced. The elements of those two sets are represented by β and γ respectively. To define their relation with the set of calls $\mathcal{C}_{ij}$, two mapping functions $q : \mathcal{C}_{ij} \rightarrow \mathcal{S}$ and $r : \mathcal{C}_{ij} \rightarrow \mathcal{L}$ are defined.

In a nutshell, the weight will be calculated by considering each of the calls of type j from person p_i. For each particular call k, its duration will be multiplied by a significance factor α. This multiplied value will be added with the corresponding β and γ value of that particular call. However, the system must at first calculate those significance values. The social closeness should then have to be calculated for suggesting the FnF numbers. The procedure for these tasks are discussed in details in the following subsections.

3.1 Determining the significance values

Since longer communications likely indicate a strong tie [6], the proposed method uses the correlation coefficient between the call duration and social closeness as the value of α. However, information regarding the social closeness of people from each call is a pre-requisite to calculate the coefficient. To determine the values of β and γ, one has to first fix the number of call periods and locations that should be considered. The set of these periods and locations are defined as T and L respectively. For each of the elements in set T and L, the method generates a separate value of β and γ respectively. These values, combined together, are then represented by two sets $\mathcal{S}$ and $\mathcal{L}$ respectively.

The first step in calculating the values of β and γ requires to define a threshold value σ. This value is used to identify the people who are socially closest to the phone user. If the social closeness value of a person is above the threshold, that individual is considered as having a strong social tie. The second steps involves calculating the percentage of people over σ for each of the call periods and locations. These percentages provide an insight regarding the significance of each period and location. However, to consider those as weights, their relative

significance is needed to be considered which is calculated at the final step. This whole procedure is illustrated in Algorithm 1.

Algorithm 1 Determining the significance values

Input: The set of all selected call periods T and locations L
Output: α, $\mathcal{S}$ and $\mathcal{L}$

Begin
Set the value of α to the correlation coefficient between call duration and social closeness
Define a threshold σ for social closeness to determine socially closest persons
Calculate the percentage of calls from people whose social closeness value is over σ on each call period $t_i \in T$ and location $l_i \in L$
Initialize $\mathcal{S} \leftarrow \emptyset$ and $\mathcal{L} \leftarrow \emptyset$
for each $t_i \in T$ **do**
    $\beta_i = \frac{\text{percentage of people over } \sigma \text{ for period } t_i}{\text{summation of percentage values } \forall_{t_i \in T}}$
    $\mathcal{S} \cup \{\beta_i\}$
end for
for each $l_i \in L$ **do**
    $\gamma_i = \frac{\text{percentage of people over } \sigma \text{ for location } l_i}{\text{summation of percentage values } \forall_{l_i \in L}}$
    $\mathcal{L} \cup \{\gamma_i\}$
end for
End

3.2 Calculating Social Closeness and Suggesting FnF numbers

After determining the values of α, β and γ, the weight ω_{ij} of call type j from person p_i can be calculated using Equation (1) . The weights ω_{ij} for each of the individuals are included in the set $\mathcal{W}_{ij}$ which is used by the proposed method to represent a social network as shown in Figure 1. To calculate the social closeness of the person p_i, the method uses a mapping function $\phi : \mathcal{J} \rightarrow \mathcal{I}$, where $\mathcal{J}$ and $\mathcal{I}$ represent the set of call types and their significances respectively. This mapping function ϕ is used to determine the weight of different call types. Assigning different weights to different call types is important since the significance of different call types (incoming or outgoing) could be different. During the calculation of social closeness for person p_i, as shown in Equation (2), the weights of that person ω_{ij} for call type j is multiplied by its respective weight provided by the mapping function ϕ.

$$\theta_i = \sum_{j=1}^{M} \phi(j) \cdot \omega_{ij} \tag{2}$$

The process of determining the social closeness and ranking the individuals based on that score is presented in Algorithm 2. Initially, as used in Algorithm 1,

a threshold σ is defined to identify socially closest persons. This threshold is used to calculate the percentage of socially closest persons for each call of type j. The percentage values are then used to determine the relative significance of different call types. Using these relative significance values, the method calculates the social closeness as defined in Equation (2). Each of the individuals, with whom communication has occurred, are sorted in descending order of social closeness. The method then suggests the numbers of certain individuals, at the top of the list, as FnF numbers.

Algorithm 2 Calculating Social Closeness

Input: The set $\mathcal{W}_{ij}$, $\mathcal{C}_{ij}$, $\mathcal{J}$ and the mapping function ϕ
Output: List of individuals sorted in descending order of social closeness

Begin
Define a threshold σ for social closeness to determine socially closest persons
Calculate the percentage of calls from people whose social closeness value is over σ for each j type of calls
Initialize $\mathcal{I} \leftarrow \emptyset$
for each $j \in \mathcal{J}$ **do**
    $\lambda_j = \frac{\text{percentage of people over } \sigma \text{ for call type } j}{\text{summation of percentage values } \forall_{j \in \mathcal{J}}}$
    $\mathcal{I} \cup \{\lambda_j\}$
end for
Use Equation (2) to calculate social closeness for each of individual
Sort the list of individuals in descending order of social closeness
End

4 Experimental Setup and Results

This section discusses about different experimental parameters that were used during the evaluation of our proposed approach. It also includes the performance of the proposed technique with a comparison along with two baseline approaches to suggest FnF numbers.

4.1 Experimental Parameters

For experimental purpose the call logs were collected from 18 individuals. The call logs included the person's mobile phone number, call duration and call time. The call location was manually collected from them. The participants had also provided a social closeness score, to each of the persons from the call log, in a scale of 1-10. Moreover, each of them were also asked to select 3 numbers from these call logs which they could register as FnF numbers. 5 of the participants provided those numbers and thus their data were used for the testing purpose. Rest of the data were used to train the proposed method that involves learning the values of α, β, γ and λ. During the calculation of these weights, the value of

σ was set to 8 which is considerably higher enough in the social closeness scale for recognizing socially closest persons.

The value of α was calculated by determining the correlation coefficient between call duration and social closeness. It has been found that the call duration and social closeness has a positive relationship, though not that much strong. This has been illustrated in Figure 2 by a scatter diagram where social closeness is plotted against call duration. The value of correlation coefficient was found to be 0.14 and it was set as the value of α.

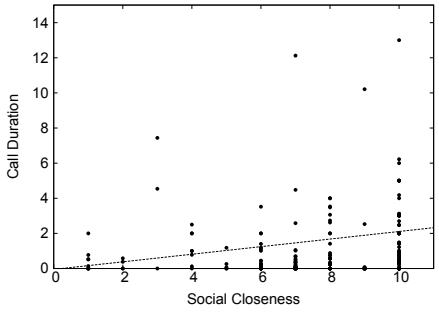


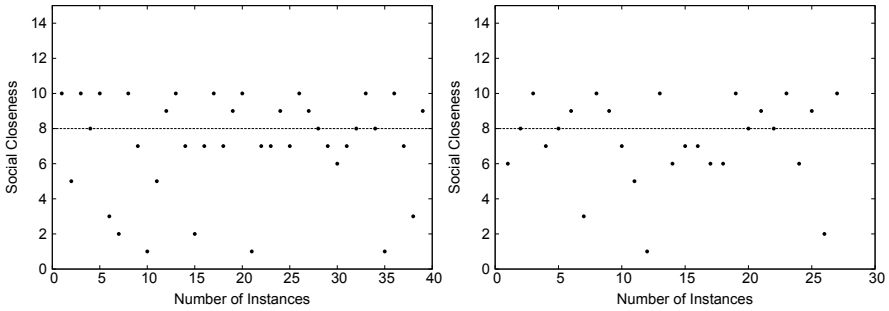
Fig. 2. Scatter diagram representing the correlation coefficient between social closeness and call duration

For the calculation of β, we have divided a day into 8 time periods. The percentage of socially closest calls and their relative weight is presented in Table 1. A scatter diagram is presented in Figure 3 to illustrate the percentage of socially closest people in two different time periods.

Table 1. Relative significance of different time periods used in the experiment

Period	Percentage of calls above σ	Relative weight
12:00 AM to 3:00 AM	60%	0.163
3:00 AM to 6:00 AM	67%	0.18
6:00 AM to 9:00 AM	25%	0.07
9:00 AM to 12:00 PM	55%	0.15
12:00 PM to 3:00 PM	26%	0.071
3:00 PM to 6:00 PM	41%	0.112
6:00 PM to 9:00 PM	56%	0.153
9:00 PM to 12:00 AM	37%	0.101

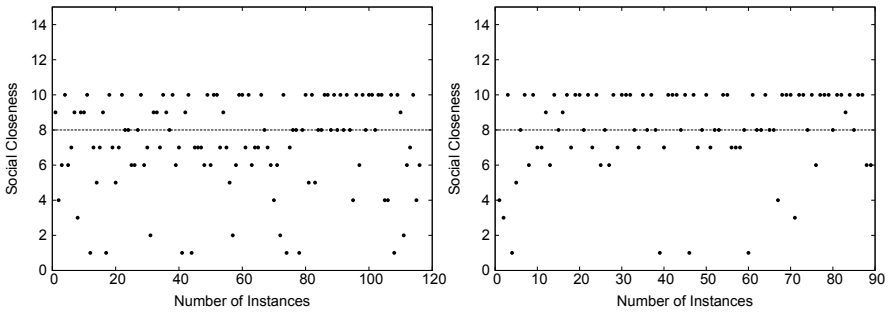
For calculating the value of γ, the locations were categorized before collecting the data. For each of the call logs, the location was selected as either home or work place. Although there were instances where users were neither in home nor



(a) Time period 3:00 PM to 6:00 PM (b) Time period 9:00 PM to 12:00 AM

Fig. 3. Visualization of calls from socially closest persons in two different time periods

in work place, still they had used various measures two provide the location information. Examples include selecting the place nearest to their current position or based on the nature of call and its period. These data, as presented in Figure 4, was used to calculate the value of γ. For the home and workplace, the value turned out to be 0.44 and 0.56 respectively.

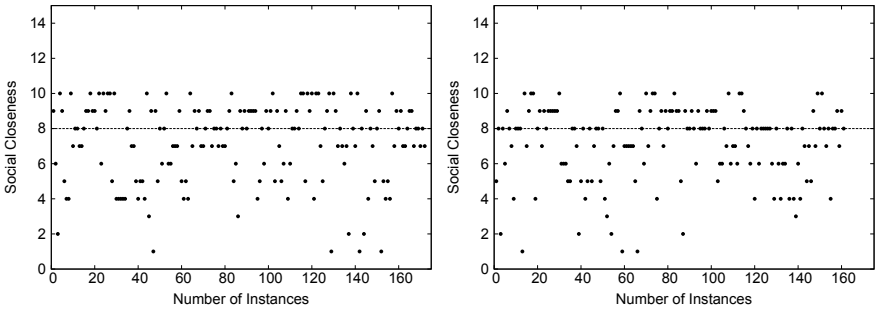


(a) Location - Home (b) Location - Workplace

Fig. 4. Visualization of calls from socially closest persons in two separate locations

In this work, ignoring the missed calls, incoming and outgoing calls has been considered to measure the value of λ. The percentage of calls from socially closest persons for outgoing and incoming calls can be visualized from Figure 5. Since there were only two call types, after calculating the weight of work place, the other weight was calculated by Equation 3. The weights for the incoming and outgoing calls has been calculated as 0.44 and 0.56 respectively.

$$\phi(j) = \begin{cases} \lambda & \text{if the value of j is 1} \\ 1-\lambda & \text{otherwise} \end{cases} \tag{3}$$



(a) Outgoing calls (b) Incoming calls

Fig. 5. Visualization of incoming and outgoing calls from socially closest persons

4.2 Performance Evaluation

The proposed technique tried to match 2 out of 3 FnF numbers provided by each individuals. The system suggested 5 numbers, one after another, and stopped at the point where two numbers had being matched. Thus for 5 individuals, the system tried to match 10 FnF numbers with at most 25 attempts. We have compared the performance of our system with two baseline approaches which are based on overall call duration and call frequency. The evaluation technique for these two approaches remained the same. It is seen that the proposed method has identified 7 FnF numbers with 22 attempts. On the other hand, the duration-based and frequency based approaches has identified 6 and 3 numbers with 22 and 25 attempts respectively.

Table 2. Performance comparison of the proposed technique with respect to duration-based and frequency based approaches

Approaches	Precision	Recall	F-measure
Duration-based	0.27	0.6	0.57
Frequency-based	0.8	0.2	0.19
Proposed technique	0.32	0.7	0.67

The performance of the technique, as shown in Table 2, is interpreted in terms of precision, recall and F-measure. It is seen that proposed method outperforms the other two approaches though the value of the precision is low. However, in this case, recall is more important than precision. This is because for each of the individuals, we are suggesting at most 5 numbers to match with 2 numbers. Due to this, the precision will always be lower due to having more false positives. Nevertheless, the user will always be satisfied if the system chooses more from the set of relevant numbers. So during the calculation of F-measure, as done in [19], the method needs to vary the weights of precision and recall. Thus in our

case, recall has been weighted five times higher than the precision during the calculation of F-measure using Equation (4).

$$F_B = (1 + B^2) \cdot \frac{precision \cdot recall}{B^2 \cdot precision + recall} \tag{4}$$

For the individuals with whom the performance was evaluated, the pattern of score provided by the system and the social closeness score provided by the phone users is illustrated in Figure 6. Although the scale of those two scores are different, it is seen that their curves are very much similar. As the system suggested FnF numbers based on calculated closeness scores, the performance is better than the other two approaches.

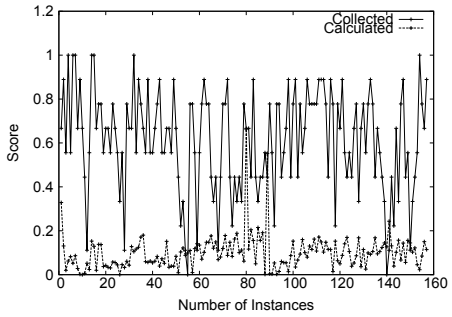


Fig. 6. The pattern of the scores collected from the individuals and calculated by the proposed method

5 Conclusion and Future Work

This paper introduces a technique to prioritize individuals based on call logs and suggest FnF numbers using that prioritize list. The method utilizes the underlying social network of the call graph for a single phone user. The edges of the call graph has been weighted during the prioritization process to incorporate different call log attributes like duration, period and location. The proposed methodology was implemented and tested against a set of real life data. The system has illustrated its effectiveness in identifying FnF numbers with an F-measure value of 0.67. This research also introduces a set of new research ideas for the research communities. This includes assessing the existence of a nonlinear relationship, between duration and social closeness, to improve the performance of the system. Moreover, one could utilize other call log attributes and optimize their total number to get better performance.

Acknowledgement

This work is supported by the University Grant Commission, Bangladesh under the Dhaka University Teachers Research Grant No-Regi/Admn-3/2015/48747.

References

1. Ho, S.Y., Kwok, S.H.: The attraction of personalized service for users in mobile commerce: an empirical study. ACM SIGecom Exchanges 3(4), 10–18 (2002)
2. Witteman, M.: Efficient proximity detection among mobile clients using the GSM network. Master's thesis, University of Twente (2007)
3. Dey, B., Newman, D., Prendergast, R.: Analysing appropriation and usability in social and occupational lives: An investigation of bangladeshi farmers' use of mobile telephony. Information Technology & People 24(1), 46–63 (2011)
4. De Melo, P.O.V., Akoglu, L., Faloutsos, C., Loureiro, A.A.: Surprising patterns for the call duration distribution of mobile phone users. In: Machine learning and knowledge discovery in databases, pp. 354–369. Springer (2010)
5. Chin, A., Zhang, D.: Mobile Social Networking. Springer (2014)
6. Wiese, J., Min, J.K., Hong, J.I., Zimmerman, J.: Assessing Call and SMS Logs as an Indication of Tie Strength. Tech. rep., Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University (2014)
7. Motahari, S., Mengshoel, O.J., Reuther, P., Appala, S., Zoia, L., Shah, J.: The impact of social affinity on phone calling patterns: Categorizing social ties from call data records. In: Proceedings of the Sixth Workshop on Social Network Mining and Analysis. pp. 1–9. ACM (2012)
8. Dong, Z.B., Song, G.J., Xie, K.Q., Wang, J.Y.: An experimental study of large-scale mobile social network. In: Proceedings of the 18th International Conference on World Wide Web. pp. 1175–1176. ACM (2009)
9. Zhang, D., Guo, B., Yu, Z.: The emergence of social and community intelligence. Computer 44(7), 21–28 (2011)
10. Phithakkitnukoon, S., Dantu, R., Claxton, R., Eagle, N.: Behavior-based adaptive call predictor. ACM Transactions on Autonomous and Adaptive Systems (TAAS) 6(3), 21 (2011)
11. Blumenstock, J.E.: Using mobile phone data to measure the ties between nations. In: Proceedings of the 2011 iConference. pp. 195–202. ACM (2011)
12. Oliner, A., Ganapathi, A., Xu, W.: Advances and challenges in log analysis. Communications of the ACM 55(2), 55–61 (2012)
13. Park, H.S., Cho, S.B.: Building mobile social network with semantic relation using Bayesian network-based life-log mining. In: Second International Conference on Social Computing (SocialCom). pp. 401–406. IEEE (2010)
14. Eagle, N., Pentland, A.S., Lazer, D.: Inferring friendship network structure by using mobile phone data. Proceedings of the National Academy of Sciences 106(36), 15274–15278 (2009)
15. Phithakkitnukoon, S., Dantu, R.: Inferring social groups using call logs. In: On the Move to Meaningful Internet Systems: OTM 2008 Workshops. pp. 200–210. Springer (2008)
16. Phithakkitnukoon, S., Dantu, R.: Mobile social group sizes and scaling ratio. AI & society 26(1), 71–85 (2011)
17. Zargari, H.: Social network modelling: Retrieval correlated graphs by mobile phone's chronological billing files. In: The Second International conference "Problems of Cybernetics and Informatics". pp. 67–70 (2008)
18. Horak, R.: Telecommunications and Data Communications Handbook. John Wiley & Sons (2007)
19. Li, X., Wang, Y.Y., Acero, A.: Learning query intent from regularized click graphs. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 339–346. ACM (2008)

Identification of Three-Dimensional Crystal Lattices by Estimation of Their Unit Cell Parameters*

Dmitriy Kirsh[1,2] and Alexander Kupriyanov[1,2]

[1] Samara State Aerospace University (National Research University), Samara, Russian Federation
[2] Image Processing Systems Institute of Russian Academy of Sciences, Samara, Russian Federation

Abstract. The problem of the identification of three-dimensional crystal lattices is considered in the article. Two matching methods based on estimation of unit cell parameters were developed to solve this problem. The first method estimates and compares main parameters of Bravais unit cells. The second method estimates and compares volumes of Wigner-Seitz unit cells. Both methods include normalised similarity measures: an edge similarity measure and an angle similarity measure for Bravais cells and a volume similarity measure for Wigner-Seitz cells. The results of computational experiments on the large set of simulated lattices showed that the developed methods allowed to achieve the identification accuracy above 90% for four lattice systems.

Keywords: crystal lattice, unit cell parameters, Monte Carlo method, similarity measure, structural identification

1 Introduction

One of the basic problems related to X-ray diffraction analysis is the identification of crystal lattices [5]. It is usually solved by comparing estimated parameters of analysed lattice with those of selected sample [2]. The lattice parameters either previously investigated or derived by modeling can be used as samples. Therefore the accurate identification of a crystal lattice requires a large data base of the preselected sample parameters.

Among the main methods for identification of three-dimensional crystal lattices the following ones can be singled out: NIST lattice comparator [1], identification on the basis of atomic packing factor [4] and centered cubic lattice

* The work was financially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the worlds leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 14-01-00369-a, # 14-07-97040-p_ povolzh'e_a); by the ONIT RAS program # 6 Bioinformatics, modern information technologies and mathematical methods in medicine 2015.

comparison method [3]. These methods have a number of drawbacks limiting their application field: complexity of crystal preparation, high error in comparison of lattices, which are similar in volume, etc.

The new approach based on estimation of unit cell parameters attempts to avoid these drawbacks. The algorithms proposed in the work allow calculating a similarity measure for any two crystal lattices.

2 Models of Crystal Lattices

There are several methods to describe crystal lattices. The most widespread method was offered by Auguste Bravais [5]. Bravais unit cell is characterised by a set of six parameters: lengths of the three edges l_1, l_2, l_3 and values of the three angles between the edges $\alpha_1, \alpha_2, \alpha_3$ (Fig. 1).

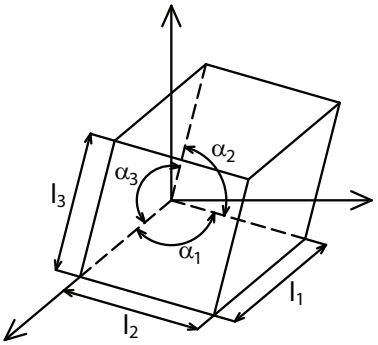


Fig. 1. The main parameters of Bravais unit cell

All Bravais lattices are subdivided into seven lattice systems. Table 1 shows characteristics of their unit cells.

Table 1. Characteristics of lattice systems

Lattice system	Symbol	Edge lengths	Angles
Triclinic	aP	$l_1 \neq l_2 \neq l_3$	$\alpha_1 \neq \alpha_2 \neq \alpha_3$
Monoclinic	mP	$l_1 \neq l_2 \neq l_3$	$\alpha_1 = \alpha_2 = 90^\circ \neq \alpha_3$
Orthorhombic	oP	$l_1 \neq l_2 \neq l_3$	$\alpha_1 = \alpha_2 = \alpha_3 = 90^\circ$
Tetragonal	tP	$l_1 = l_2 \neq l_3$	$\alpha_1 = \alpha_2 = \alpha_3 = 90^\circ$
Cubic	cP	$l_1 = l_2 = l_3$	$\alpha_1 = \alpha_2 = \alpha_3 = 90^\circ$
Rhombohedral	hR	$l_1 = l_2 = l_3$	$\alpha_1 = \alpha_2 = \alpha_3 \neq 90^\circ$
Hexagonal	hP	$l_1 = l_2 \neq l_3$	$\alpha_1 = 120^\circ; \alpha_2 = \alpha_3 = 90^\circ$

Another model of crystal lattices was offered by Jeno Wigner and Frederick Seitz [6]. Wigner-Seitz unit cell is characterised by a set of normal vectors which are drawn to limiting planes. In a three-dimensional space it is a polyhedron which contains inside itself one lattice node (Fig. 2).

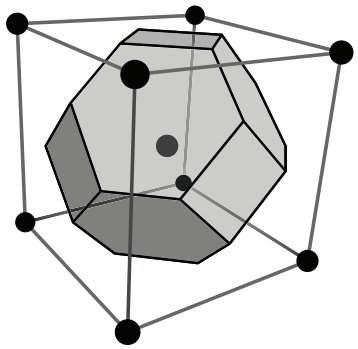


Fig. 2. Body-centered cubic Wigner-Seitz cell

3 Method of Crystal Lattice Identification on the Basis of Bravais Unit Cell Parameter Estimation

The initial data of the identification method are the set of radius-vectors determining the spatial position of crystal lattice nodes.

The following algorithm was designed to calculate the six main parameters of the Bravais unit cell:

1. Center the lattice.
2. Superpose the first radius-vector of minimal length with the axis OX.
3. Transfer the second radius-vector of minimal length into the plane XOY.
4. Select the third radius-vector of minimal length.
5. Calculate the main six parameters: $l_1, l_2, l_3, \alpha_1, \alpha_2, \alpha_3$.

Two normalised measures were introduced to determine separately the degree of the edge similarity and the degree of the angle similarity of two Bravais unit cells.

Similarity measure of edges:

$$\|\boldsymbol{l}_1 - \boldsymbol{l}_2\| = 1 - \frac{\sqrt{(l_{11} - l_{21})^2 + (l_{12} - l_{22})^2 + (l_{13} - l_{23})^2}}{\max\left\{\sqrt{(l_{11})^2 + (l_{12})^2 + (l_{13})^2}, \sqrt{(l_{21})^2 + (l_{22})^2 + (l_{23})^2}\right\}} \quad (1)$$

Similarity measure of angles:

$$\|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2\| = 1 - \max\left\{\sin\left(|\alpha_{11} - \alpha_{21}|\right), \sin\left(|\alpha_{12} - \alpha_{22}|\right), \sin\left(|\alpha_{13} - \alpha_{23}|\right)\right\} \quad (2)$$

4 Method of Crystal Lattice Identification on the Basis of Wigner-Seitz Unit Cell Volume Estimation

The initial data of the identification method are the number of scattering points L and the set of radius-vectors determining the spatial position of crystal lattice nodes.

The following algorithm was designed to calculate the volume of the Wigner-Seitz cell:

1. Center the lattice.
2. Determine the normal vectors from central lattice node to the planes limiting Wigner-Seitz cell.
3. Calculate the volume of cell limited by planes with the use of the Monte Carlo method.
   (a) Generate L-values of three-dimensional random vectors which are uniformly distributed in the whole lattice volume.
   (b) Count the number of vectors that hit in the region limited by planes and calculate the volume of cell based on the fact that the probability of hit in the Wigner-Seitz cell region is proportional to its measure (volume).

A normalised measure was introduced to determine the degree of the volume similarity for two Wigner-Seitz unit cells:

$$\|V_1 - V_2\| = 1 - \frac{\sqrt{(V_1 - V_2)^2}}{\max\{V_1, V_2\}} \quad (3)$$

The following computational experiments of crystal lattice identification on the large set of simulated three-dimensional lattices were conducted to analyse the efficiency of the introduced similarity measures.

5 Results of Experimental Computations

The initial data for experiments were 7,000 lattices (1000 lattices of each lattice system) obtained by simulation. The lengths of edges and values of angles were determined by values of a uniformly distributed random variable.

Each lattice was matched with all the rest lattices in pairs: two lattices were considered to be similar in edges or in angles, if the value of the corresponding similarity measure was no less than 0.95. Selection of this limiting value relates to the fact that currently the error of lattice parameter determination is no less than 5% [1].

Two lattices in the first experiment were brought into comparison only by the value of the edge similarity measure. Matching of two lattices in the second experiment was carried out only by the value of the angle similarity measure. However, the derived values of the percentage of lattice exact identification were near 14% for all lattice system in both experiments. These results show that the partitioning of all Bravais crystal lattices into seven lattice systems by using only the edge similarity measure or the angle similarity measure is not uniform and separable.

Consequently more exact lattice identification requires the simultaneous application of both similarity measures. For this reason in the third experiment two lattices were considered to be similar if values of the edge similarity measure and the angle similarity measure were no less than 0.95. Table 2 shows the results.

Table 2. Percentage of lattice exact identification for the main lattice systems by simultaneous edges and angles comparison

Estimated cell	Sample cell						
	aP	*mP*	*oP*	*tP*	*cP*	*hR*	*hP*
aP	98	1	1	0	0	0	0
mP	7	39	36	14	1	0	3
oP	6	36	39	14	1	1	3
tP	5	26	26	34	4	4	1
cP	4	12	12	23	26	23	0
hR	11	2	2	3	4	78	0
hP	12	3	3	2	0	0	80

Table data show average percent of coincidence of the estimated lattices with sample lattices (both similarity measures are no less than 0.95). For example the set of sample lattices which have coincided with one of the rhombohedral lattices in the third experiment consists on the average of 11% triclinic, of 2% monoclinic, of 2% orthorhombic, of 3% tetragonal, of 4% cubic, of 78% rhombohedral and of 0% hexagonal.

The use of the volume similarity measure of Wigner-Seitz unit cells was the last step to increase the identification accuracy. Matching of two lattices in the final fourth experiment was conducted with the application of all three similarity measures simultaneously: edges and angles of Bravais unit cells and volumes of Wigner-Seitz unit cells. Table 3 shows the results of the experiment.

According to Table 2 and Table 3 data it can be concluded that the accuracy of lattice identification increased by average 15%. The maximum increase of accuracy was achieved for monoclinic lattices by 51%. However, there is a separate group (orthorhombic, tetragonal, cubic) with the low percentage of lattice exact identification and, therefore, the problem of delimiting these three lattice systems cannot be considered as solved.

Table 3. Percentage of lattice exact identification for the main lattice systems by simultaneous edges, angles and volumes comparison

Estimated cell	Sample cell						
	aP	*mP*	*oP*	*tP*	*cP*	*hR*	*hP*
aP	99	1	0	0	0	0	0
mP	2	90	4	3	0	0	1
oP	2	22	54	20	1	1	0
tP	1	10	34	44	6	5	0
cP	0	3	15	26	30	26	0
hR	1	0	2	3	4	90	0
hP	1	7	0	0	0	0	92

6 Conclusion

The developed methods of crystal lattice identification allowed to achieve the identification accuracy above 90% for four lattice systems (triclinic, monoclinic, rhombohedral and hexagonal).

Basing on the performed calculation it can be concluded that the best way to identify the lattice system for the generated set of 7,000 lattices is simultaneous application of all three introduced similarity measures.

Identification accuracy of the remaining three lattice systems (orthorhombic, tetragonal and cubic) is not still sufficiently high. They require further research for the purpose of finding additional similarity measures (for example, comparison of isosurfaces, tensor representation of unit cells, etc.).

References

1. Kessler, E., Henins, A., Deslattes, R., Nielsen, L., Arif, M.: Precision Comparison of the Lattice Parameters of Silicon Monocrystals. Journal of Research of the National Institute of Standards and Technology 99, 1–18 (1994)
2. Kupriyanov, A.: Texture analysis and identification of the crystal lattice type upon the nanoscale images. Computer Optics 35, 128–135 (2011) (in Russian)
3. Patera, J., Skala, V.: Centered cubic lattice method comparison. In: 17th ALGORITMY Conference on Scientific Computing, pp. 309–319. Slovakia (2005)
4. Smith, W.: Foundations of Materials Science and Engineering. McGraw-Hill, New York (2004)
5. Soifer, V., Kupriyanov, A.: Analysis and recognition of the nanoscale images: conventional approach and novel problem statement. Computer Optics 35, 136–144 (2011) (in Russian)
6. Tilley, R.: Crystals and crystal structures. John Wiley & Sons, Chichester (2006)

Construction of Adaptive Educational Forums Based on Intellectual Analysis of Structural and Semantics Features of Messages

Alexander Kozko

Chelyabinsk State University,
Chelyabinsk, Russia
alkozko@gmail.com

Abstract. The paper deals with organization of interaction and communication between subjects of the distance learning process through forums and comment system. Considered existing software tools, their structure and disadvantages. Propose a model of adaptive educational forums, as well as the structural and semantic similarity metrics for extracting dialogues and thematic discussions from arrays of individual comments, as a basis for the construction of adaptive educational forums.

Keywords: distance learning · online discussion forums · semantic similarity

1 Introduction

Development of information technologies makes distance education more available for everyone, and technology of management and support of education process becomes more and more important. Many scientific works dedicated to LMS (Learning Management System), for example works [1, 2, 3], give a description of LMS, and works [4, 5] devoted to review of existing LMS products. However, these works consider of learning management tools mainly from the following perspectives: development, management, courses and distribution of education materials, but not paying enough attention to the issues of communication and information exchange between the different subjects of education process.

At present, modern learning theories such as the theory of connectivism proposed by George Siemens and Stephen Downes, indicate as the basic conditions for successful learning activities not only communication with the teacher, but also the interaction between students and the exchange between them so-called “sensemaking artifacts” [6]. Sensemaking artifacts means blog posts, notes, podcasts, and other educational materials that created by a student for discussion with other students. In the paper [7] analyzed this learning approach and claimed it more effective in comparison with traditional.

Thus, the theme of the organization of communication and interaction between subjects of distance education process is extremely relevant today, but it isn't considered in the existing works.

2 Instruments of communication and informational exchange for education process

Consider the instruments of communication that are used in distance education.

On the one hand LMS, like Moodle, ILIAS, and Sakai and the other hand, of the platform for organization MOOCs (Massive open online course), such as Stepic, Udemy, Coursera, EdX, Udacity propose to use for interaction within the course online forums, blogs, chat rooms, questions and answers websites, wiki-pages and in addition allowed commenting learning materials. For some courses, authors suggest to use third party sites such as Reddit, StackExchange or social networks instead of internal platform tools.

Irrespective of the specifics and type of a course the main tools that are used for communication within the course are standard forums, blogs, questions and answers system, as well as various materials commenting system. However, such instruments do not take into account the specificity of education process and do not provide any applicable analytical functions for organization of education.

3 Overview of informational exchange tools

3.1 Interaction tools

Forum, in generally, - is an online tool for website's visitors communication. The essence of any forum is to create topics with its subsequent discussion. Users can comment on created topics, ask questions and receive answers, and answer other forum user’s questions of the forum and give them advice. Thus, each topic is an initial entry with a set of comments.

Blogs are a set of copyright entries, sorted by creation time, usually from the newest to the oldest. Blogs characterized by the ability to publish reviews (comments) by other users, and that makes blogs an instrument for information exchange. Technically, each blog entry as well as forum topic is the initial post with a set of comments to it.

Comment systems additionally used in distance education platforms to allow discuss of any course materials, such as videos or articles.

3.2 Comments structure

Structuring of post's comments may also vary depending on the implementation. There are three main types of structuring:

• Tree comments - comments list is presented as a hierarchical tree view. New message is placed right after the previous one (quoting it isn't necessary). New comment can also start its own discussion branch.

• Linear (flat) comments - comments within the same topic are published under each other, as they become available; new message is placed last (usually at the bottom); interactive relationship between comments is based on specially decorated citations of references to the author and other means.

• Hybrid comments - represent a cross between a tree and a linear structure, now, it is the most popular form of comments submission.

Comments are usually ordered by date, popularity and the number of votes. Each comment has text but it also has the following attributes: Author Name and Timestamp.

4 Adaptive learning forums

As noted earlier, the above instruments of online communication do not take into account the specifics of education process and do not provide any specific functions for it. Furthermore, they are not designed for a large number of participants, which is usual for MOOC. Problems such as a large number of overlapping topics, unanswered questions, incorrectly exposed statuses and tags, are an evidence that the forum is not an effective communication instrument, we lack of effectiveness of the learning process.

To increase the effectiveness of forum usage, as an instrument of communication and information exchange within education process we propose develop the technology of adaptive educational forums, based on data mining.

Adaptive educational forum is an online forum, whose structure is rebuilt depending on student educational trajectory, his information needs, the features of the course and users activity. For teachers, the analysis of information from educational forum can be a source of data for implicit course quality feedback, student's problems with the understanding of educational materials, student's activity evaluation, etc. Thus, the technology of adaptive educational forums could be a way to increase the quality of distance learning.

5 Analysis of comments on online discussion forum

Based on the overview, it can be argued that in such systems, the main content is often not the whole topic and not individual comments, but thematic discussions and dialogues, consisting of comments and united by one common theme. Therefore, the primary tasks are extracting individual discussions of the whole comments list, and the problem of determining the semantic similarity of discussions among themselves. For dialogues extracting of the whole comments list we offer use at both the semantics of messages and their location in the structure of comments.

5.1 Extracting tree commenting relations from linear comments

Because linear structure's main distinction is implicit comments relations, we propose an approach to the reduction of the linear structure to explicit comments tree structure.

For comments linear structure conversion to the tree and find comments relations, proposed focus on the comment text and message attributes, for example, author nickname, post timestamp, user's nickname which are responsible, text of citation, position in comments list.

Using these attributes, and the semantics of messages for analyze, for linear comments conversion into a tree we propose determine the pairwise comments relations. In general, the numerical metric defines an association between two messages comments represented by equation 1.

$$rel(c_1, c_2) = k_{sem} \cdot sim_{sem}(t_1, t_2) + k_{attr} \cdot sim_{attr}(a_1, a_2) \quad (1)$$

$rel(c_1, c_2)$ – relations between the two commentaries, $sim_{sem}(t_1, t_2)$ – semantic similarity of the two comments texts, $sim_{attr}(a_1, a_2)$ – attribute similarity of the two comments, defined by their attributes, k_{sem} и k_{attr} – coefficients.

To extract tree structure, it is necessary to calculate each comment's c1 degree of relation with all previous time comments c2, then c2 comment that has a maximum relation score becomes the parent of the current comment c1 in a tree structure.

5.2 Extraction of thematic discussions

In a discussion with a large number of participants some of the participants might drift away from the main theme and start to discuss unrelated topics. Such thematic discussions (subtopics) within the main topic can also provide useful information for the user, but their detection difficult. Therefore we propose to allocate such subtopics in separate entities and use them in the construction of adaptive educational forums.

In tree-like comment systems, a message has the following attributes: author nickname, post timestamp, comment depth in tree. Based on this, we proposed similarity metric for two comments, represented by equation 2.

$$sim(c_1, c_2) = k_{sem} \cdot sim_{sem}(t_1, t_2) * \left(\frac{1}{dist(c_1, c_2)} + k_{attr} \cdot sim_{attr}(a_1, a_2) \right) \quad (2)$$

$sim(c_1, c_2)$ – similarity between the two commentaries, $sim_{sem}(t_1, t_2)$ – semantic similarity of the two comments texts, $sim_{attr}(a_1, a_2)$ – attribute similarity of the two comments, $dist(c_1, c_2)$ – the distance between the two comments in the tree, k_{sem}, k_{attr} – coefficients.

To unite a group of comments in a thematic discussion we proposed to use an approach based on clustering algorithms, the result of which (clusters of comments) should be used as boundaries for thematic subtopics.

5.3 Semantic similarity metrics for comments

We consider separately the question of calculating the semantic similarity between the two comments texts. First it is worth noting the specifics of messages - basically all posts extremely short - from a few words to 2-3 sentences, so the use of methods for calculating the semantic similarity of documents can be difficult. It is proposed to determine the similarity of the documents, as similarity of containing concepts.

Existing semantic similarity metrics can be divided into several classes. In work [8] it is proposed to the following classification:

1) Measures based on a corpus of texts: for example, LSA, Web-based - NGD and PMIIR.

2) Measures based on ontologies: measures by Wu and Palmer, Leackock and Chodorow, Resnik, Lin and others.

3) Measures based on definitions: ExtendedLesk, GlossVectors.

Another group of metrics that can be allocated -

4) The metrics based on Wikipedia: Wikipedia Link-based Measure (WLM), Explicit Semantic Analysis (ESA), WikiWalk, WikiRelate! and others.

Choosing a semantic similarity metric for concepts it's necessary to pay attention to the specific learning courses. Themes of learning courses may be beyond ontology but in domain that is disclosed in course may contain specific concepts which well-known ontologies as WordNet don't include and specific ontologies for course does not exist. Therefore, a similarity measure metrics based on ontologies can be used to analyze educational forums, only if a teacher will build the ontology of course by himself.

For this reason assumed to use the metrics based on the online encyclopedia Wikipedia. The advantages of using Wikipedia as a source of data are a volume and wide range of different themes, relevance and partial structure.

6 Related Work

Forums analysis is not widely discussed in scientific papers, however it is possible to highlight the following investigation.

Work [9] devoted to the study groups Usenet, it proposes a method for measuring the similarity of the different groups on the activity of participants in them, as well as introduce a measure for evaluation post belonging specific group, which allows exclude cross-posts from the analysis.

In study [10] proposes a model to estimate the probability of involvement or non-involvement of a user in a specific online discussion based on the activity of his friends and his interests, and the list of friends and user's interests based on the previous activity in other topics.

In the work [11] is proposed an approach for extracting context information, as well as questions and answers from the topic using the method of SVM, in [12] considered other methods used for this purpose.

Finally, in [13], the authors propose an approach that combines both structural and semantic analysis for search discussions and find of key messages in the threads.

7 Conclusion

Thanks to the Internet, distance education has been made available for millions of people around the world, and now it's getting increasingly important role of supporting communication within education process. Number of students enrolled to courses is increased, and growth of the distance education percentage in the education process realm sets new requirements that existing communication tools, do not respond, it reduces the efficiency of education process as a whole.

We propose the concept and the model of adaptive educational forums, the use of which, in our opinion, will increase the efficiency of interaction between students, as well as enhance the role of communication environment in distance education process and give teachers the ability for automatically collect information about students as well as the quality of the course.

For creation of adaptive educational forums, we propose to use an approach that includes allocation of individual thematic subtopics that in turn bases on messages structural and semantic features.

References

1. Dalsgaard, C.: Social software: E-learning beyond learning management systems. European Journal of Open, Distance and E-Learning, 2 (2006)
2. Sclater, N.: Web 2.0, personal learning environments, and the future of learning management systems. Research Bulletin, 13 (2008)
3. Coates, H., James, R., & Baldwin, G.: A critical examination of the effects of learning management systems on university teaching and learning. Tertiary education and management, 11, 19-36 (2005)
4. Ketcham G., Landa K., Brown K., Charuk K., DeFranco T., Heise M., McCabe R., Youngs-Maher P.: Learning Management Systems Review. (2011) http://openscholar.purchase.edu/sites/default/files/keith_landa/files/doodle_lmsreport_final.pdf
5. Siemens G.: Learning or Management Systems? A Review of Learning Management System Reviews. Learning Technologies Centre, University of Manitoba (2006)
6. Siemens G.: Connectivism: Design and Delivery of Social Networked Learning. The International Review of Research in Open and Distance Learning 2011; 12:3-11. (2011)
7. Mott J., Wiley D.: Open for Learning: The CMS and the Open Learning Network. In education, 15(2). (2009)
8. Panchenko A.: Similarity Measures for Semantic Relation Extraction (Ph.D. Thesis). Université catholique de Louvain & Bauman Moscow State Technical University, (2012-2013)
9. McGlohon M., Hurst M.: Community Structure and Information Flow in Usenet: Improving Analysis with a Thread Ownership Model. ICWSM. (2009).
10. Wu, H., Bu, J., Chen, C., Wang, C., Qiu, G., Zhang, L., Shen, J.: Modeling Dynamic Multi-Topic Discussions in Online Forums. In AAAI. (2009)
11. Cao, Y., Yang, W. Y., Lin, C. Y., Yu, Y.: A structural support vector method for extracting contexts and answers of questions from online forums. Information Processing & Management, 47(6), 886-898. (2011)
12. Cong, G., Wang, L., Lin, C. Y., Song, Y. I., Sun, Y.: Finding question-answer pairs from online forums. In Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval (pp. 467-474). ACM. (2008)
13. Lin, C., Yang, J. M., Cai, R., Wang, X. J., Wang, W.: Simultaneously modeling semantics and structure of threaded discussions: a sparse coding approach and its applications. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (pp. 131-138). ACM. (2009)

The Investigation of Deep Data Representations Based on Decision Tree Ensembles for Classification Problems

Pavel Druzhkov and Valentina Kustikova

Lobachevsky State University of Nizhni Novgorod,
Nizhni Novgorod, Russian Federation
{itlab.ml@cs.vmk.unn.ru}

Abstract. A classification method based on deep representation of input data and ensembles of decision trees is introduced and evaluated solving the problem of vehicle classification and image classification with large number of categories.

Keywords: classification; decision trees; random forest, deep learning.

1 Introduction

Deep learning studies machine learning algorithms that enable automatic construction of feature description hierarchies thus making data representations more efficient in terms of the solved problem. Currently, the deep learning methods for solving image classification and object detection problems are making good headway (see, for example, a review of [1] and [2]) as a part of neural network (NN) models.

Although the deep learning ideas are for now best reflected in NN approaches, the integration of hierarchical data representation into other methods such as decision trees and bag-of-words is of special interest, because these combinations may help to reduce complexity of learning and/or the use of deep models and give a better understanding of their success. There were attempts to use decision trees for induction of new feature descriptions [3, 4]. Nevertheless, no examples of successful applications of such models stacks are known to the authors. This paper studies applicability of a deep classification method based on hierarchical representation of data obtained using decision tree ensembles.

2 The Proposed Classification Method

The decision tree divides the feature space into non-overlapping areas, corresponding to tree nodes, by means of a recursive procedure [5]. Given a data point decision tree performs a sequence of tests determining the way from the root to one of its leaves, that in turn defines the classification result. To ensure better statistical stability of the classifier, such methods for grouping the set of

trees as *boosting* [6] and *bagging* [7] are used. The output of a trees ensemble depends on the leaves that receive a certain object. In this context, Vens et al. [3] supposed that the set of tests performed for a certain data point could represent a better object view. Martyanov et al. [4] review different ways to use the internal structure of the random forest model [7] to construct efficient features. These ideas are further developed in this paper and a classification method based on hierarchical (deep) data representation is proposed. This method consists of the following stages:

1. Initial data representation is used to construct the ensemble of trees classifier.
2. Each tree in the ensemble defines a new feature whose value equals to the index of the leaf receiving the object in question.
3. New ensemble is constructed based on the derived features.
4. Steps 2 and 3 are repeated to ensure the required model depth.
5. The resulting feature descriptions are used to construct the classifier.

A possible advantage of the proposed method is that it can operate arbitrary initial feature descriptions. Therefore, using ones, that have been already successfully applied for computer vision purposes, one can improve both the classifier performance and the model quality. Deep model learning using this method can be based on widely known algorithms that are less complex compared to NN learning methods. The same is true for the use of a tuned model in the classification process, especially in case of cascaded classifiers enabling computation of feature values only when required.

Open-source C++ implementation of the proposed method utilizing random forest as a decision tree ensemble component was done [8]. Implementation allows an arbitrary initial feature description, however HOG [9] was used in all experiments described below.

3 Experimental Results

3.1 Vehicle Classification

To put the proposed classifier to the test the vehicle classification problem was solved. The MIT2 dataset [10] from [11] was used. It contains images of four classes: *sedan*, *van*, *taxi* and *car*. Because of the essential classes overlap only two-class problems were considered: *car* vs *van* and *sedan* vs *taxi*. The training set consist of 200 images while the test one includes 730. Sizes of objects in the dataset vary considerably with the average of 72×38 px. To properly apply different machine learning techniques, images were cropped and scaled to make all vehicles position and size identical.

It can be seen from the derived results (Table 1) that the proposed method has almost the same accuracy as (shallow) NN approaches: 97.5% and 97.75% correspondingly. According to Ma et al. [11], the best known result for the same data is 98.5% that is achieved by a Bayesian classifier applied to SIFT-like features. However this approach requires a complex parameter selection procedure.

In the absence of an open implementation nor a more detailed description it becomes even more tangled. Repeatability activities to obtain the results from [11] provided the accuracy of only 79.25% [12].

In case of less significant cross-class distinctive features the NN approaches showed a much lower classification accuracy than the proposed algorithm and the Bayesian classifier. The accuracy of convolutional NN on *sedans* vs *taxis* task is 36% compared to 94.24% and 96.06%, respectively (Table 1).

Table 1. Classification accuracy for MIT2 dataset

Classifier	*car vs van, %*	*sedan vs taxi, %*
Linear NN: fully-connected and softmax layer	97.75	0
Convolutional NN: convolutional and softmax layer	50.3	36.00
Bayesian classifier on SIFT-like features [11]	98.5	96.06
Bayesian classifier on SIFT-like features [12]	79.25	—
Proposed method: 3 levels, 500 trees of depth 2 at each	97.5	94.24

3.2 Image Classification with Large Number of Categories

To expect how the proposed method scales with the growth of classes count, experiments on CIFAR-100 dataset [13] were carried out. Dataset contains 60,000 32×32 px images. Each image belongs to one of the 100 classes that cover a wide range of visual objects. The whole set of images is divided into the learning (50,000 images) and test (10,000 images) parts.

Although the random forest, which is at the heart of the proposed method, is known as an algorithm that does not require careful selection of learning parameters, the optimal number of trees and their depth were determined on the basis of OOB accuracy [7] to find balance between model quality and size. Hyperparameters were selected in a greedy fashion: ones for the first layer were selected and fixed, then the second layer is explored.

For the first layer the OOB accuracy of 15.16% is obtained for 500 trees with the depth of 20. This is close enough to the observed maximum of 16.04% that corresponds to a much larger ensemble of 1000 trees. But this requires greater computational resources for learning and classification purposes and may also hinder learning of the second layer due to increased dimension of the new feature space. So 500 trees of depth 20 were fixed for the first layer setup. Achieved test accuracy for the one-layer model is 16.97%, while additional layer (1000 trees of depth 25) reduced it to 12.42%. It is much lower than the best known results of deep NN models for the same problem reaching 61–65% [14]. Such a gap may be explained by the nature of the solved problem, properties of the proposed algorithm and appropriateness of HOG description to which the proposed algorithm was applied.

Insufficient quality combined with uselessness of model depth increment for the considered problem indicates that the approach needs to be modified. The possible ways of improvement include the use of boosting instead of the random forest, replacement of the initial features or use of image patches of different size on different layers, to make connection between model layers and feature abstraction levels more explicit. Higher resolution images may be subject to quality improvement due to the use of keypoint descriptors by analogy with [4].

4 Conclusion

In this paper, a classification method based on hierarchical (deep) initial data representation was proposed. This method was tested by solving the vehicle classification problem and large scale image classification.

Acknowledgments. This work is supported by Russian Foundation for Basic Research (project No 14-07-31269) and has been done in the IT laboratory at CMC Department of Lobachevsky State University of Nizhni Novgorod.

References

1. Kustikova, V.D., Druzhkov, P.N.: A Survey of Deep Learning Methods and Software for Image Classification and Object Detection. In: Proc. of OGRW. (2014)
2. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. In: arXiv:1409.0575. (2014)
3. Vens, C., Costa, F.: Random Forest Based Feature Induction. In: Proc. of ICDM. pp. 744-753. (2011)
4. Martyanov, V., Polovinkin, A., Tuv, E.: Feature Learning for Image Classification with Code Book Based on Decision Tree Ensembles. In: Proc. of ICIM. pp. 188-191. (2013)
5. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees. Wadsworth & Brooks. (1984)
6. Freund, Y., Schapire, R.: A decision-theoretic generalization of online learning and an application to boosting. In: Comp. and Sys. Sciences. Vol. 55. pp. 119-139. (1997)
7. Breiman, L.: Random Forests. Machine Learning. Vol. 45. No. 1. pp. 5-32. (2001)
8. Deep random forest implementation in SWOD (Sliding Window Object Detection) library, `https://github.com/druzhkov-paul/swod/tree/deep_random_forest`
9. Dalal, N.: Finding people in images and videos. (2006)
10. Vehicle Classification in Surveillance (MIT2), `http://people.csail.mit.edu/xiaoxuma/proj/vehi_reco/index_vehi_reco.htm`
11. Ma, X., Grimson, W.E.L.: Edge-based rich representation for vehicle classification. In: Proc. of the 10th IEEE Int. Conf. on CV (ICCV05). Vol. 2. pp. 1185-1192. (2005)
12. Khanova, T.A., Kustikova, V.D.: One approach to solving the problem of vehicle classification. In: Proc. of HPC-2014. pp. 453-459. (2014) (In Russian)
13. Krizhevsky, A.: Learning Multiple Layers of Features from Tiny Images. (2009), `http://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf`
14. Classification datasets results, `http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html`

Families of Heron Digital Filters for Images Filtering

Ekaterina Ostheimer[1], Valeriy Labunets[2], and Filipp Myasnikov[2]

[1] Capricat LLC, 1340 S., Ocean Blvd., Suite 209, Pompano Beach, 33062 Florida, USA
katya@capricat.com,
[2] Ural Federal University, pr. Mira, 19, Yekaterinburg, 620002, Russian Federation
vlabunets05@yahoo.com, fsmyasnikov@gmail.com

Abstract. The basic idea behind this work is in extraction (estimation) of the uncorrupted image from the distorted or noised one. The idea is also referred to as the image denoising. Noise removal or noise reduction in an image can be done by linear or nonlinear filtering. The most popular linear technique is based on averaging (or meaning) linear operators. Usually, denoising via linear filters does not work sufficiently since both the noise and edges (in the image) contain high frequencies. Therefore, any practical denoising model has to be nonlinear. In this paper, we propose two new nonlinear data-dependent filters, namely, the generalized mean and median Heronian ones. These filters are based on the Heronian means and medians that are used for developing a new theoretical framework for image filtering. The main goal of the work is to show that new elaborated filters can be applied to solve problems of image filtering in a natural and effective manner.

Keywords: Nonlinear filters, generalized aggregation mean

1 Introduction

The basic idea of this work is in application of a systematic method to nonlinear filtering based on the Heronian averaging and median nonlinear operators [1–4]. The classical Heronian mean and median of two positive real numbers a and b have the following forms

$$\mathbf{MeanHeron}(a,b) = (\sqrt{aa} + \sqrt{ab} + \sqrt{bb})/3,$$

$$\mathbf{MedHeron}(a,b) = (\sqrt{aa}, \sqrt{ab}, \sqrt{bb}).$$

We are going to generalize and use these mean and median operators for constructing new classes of nonlinear digital filters. The general aim of this work is to clarify whether the filters based on such exotic meanings have any smoothing properties.

2 Generalized Heronian means and medians

Let $(x_1, x_2, \ldots, x_N)$ be an N-tuple of positive real numbers.

Definition 1. *The following generalized means and median*

$$\begin{aligned}
\mathbf{MeanHeron}_2^I(x_1,\ldots,x_N) &= \frac{1}{MH_2}\sum_i\sum_{\leqslant j}\sqrt{x_i x_j},\\
\mathbf{MeanHeron}_2^{II}(x_1,\ldots,x_N) &= \sqrt{\frac{1}{MH_2}\sum_i\sum_{\leqslant j}x_i x_j},\\
\mathbf{MedHeron}_2(x_1,\ldots,x_N) &= \mathbf{Med}\left[\left\{\sqrt{x_i x_j}\right\}_{i\leqslant j}\right] = \sqrt{\mathbf{Med}\left[\left\{x_i x_j\right\}_{i\leqslant j}\right]}
\end{aligned} \tag{1}$$

are called the Heronian means and median of the first and second kinds [1–3], respectively, where $MH_2 = N(N+1)/2 = \mathbf{MeanHeron_2}(1,1,\ldots,1)$.

Here, we want to generalize Definition 1 by summarizing up the k-th roots of all possible distinct products of k elements of $(x_1,\ldots,x_N)$ with repetition. The number of all such products is $C_{N+k-1}^k = MH_k$.. This determines the normalization factor and leads to the following definitions:

$$\begin{aligned}
\mathbf{MeanHeron}_k^I(x_1,\ldots,x_N) &= \frac{1}{MH_k}\sum_{i_1\leqslant}\sum_{i_2\leqslant}\cdots\sum_{\leqslant i_k}\sqrt[k]{x_{i_1}x_{i_2}\cdots x_{i_k}},\\
\mathbf{MeanHeron}_2^{II}(x_1,\ldots,x_N) &= \frac{1}{MH_k}\sqrt[k]{\sum_{i_1\leqslant}\sum_{i_2\leqslant}\cdots\sum_{\leqslant i_k}x_{i_1}x_{i_2}\cdots x_{i_k}}
\end{aligned} \tag{2}$$

for the generalized Heronian means and

$$\begin{aligned}
\mathbf{MedHeron}_k(x_1,\ldots,x_N) &= \mathbf{Med}\left[\left\{\sqrt[k]{x_{i_1}x_{i_2}\cdots x_{i_k}}\right\}_{i_1\leqslant i_2\leqslant\cdots\leqslant i_k}\right] =\\
&= \sqrt[k]{\mathbf{Med}\left[\left\{x_{i_1}x_{i_2}\cdots x_{i_k}\right\}_{i_1\leqslant i_2\leqslant\cdots\leqslant i_k}\right]}.
\end{aligned} \tag{3}$$

for the generalized Heronian median, where $MH_k = \mathbf{MedHeron_k}(1,1,\ldots,1)$.

Let us introduce the observation model and notion used throughout the paper. We consider noise images in the form $f(i,j) = s(i,j)) + \eta(i,j)$, where $s(i,j)$ is the original grey-level image and $\eta(i,j)$ denotes the noise introduced into $s(i,j)$ to produce the corrupted image $f(i,j)$. Here, $(i,j) \in \mathbf{Z}^2$ are 2D coordinates that represent the pixel location. The aim of image enhancement is to reduce the noise as much as possible or to find a method, which, for the given $s(i,j)$, derives an image $\widehat{s}(i,j))$ as close as possible to the original $s(i,j)$ subjected to a

suitable optimality criterion. In the standard linear and median 2D-filters with the square N-cellular window $\mathrm{M}(i,j)$ and located at (i,j), the mean and median replace the central pixel

$$\widehat{s}(i,j) = \underset{(m,n)\in M(i,j)}{\mathbf{Mean}}\left[f(m,n)\right], \quad \widehat{s}(i,j) = \underset{(m,n)\in M(i,j)}{\mathbf{Med}}\left[f(m,n)\right], \tag{4}$$

where $\widehat{s}(i,j)$ is the filtered image, $\{f(m,n)\}_{(m,n)\in \mathrm{M}(i,j)}$ is an image block of the fixed size $N = Q \times Q$ extracted from f by moving window $\mathrm{M}(i,j)$ at the position (i,j), and **Mean** and **Med** are the classical mean and median operators, where $Q = 2r+1$ is an odd integer. All pixels of this block are numbered by the following way: $(m,n) \to r$ has the following form $r = Q(m+1) + (n+1)$. For example, for the 9-cellular window of size $N = 3 \times 3 = 9$ we have $(-1,-1) \to 0, (-1,0) \to 1, (-1,1) \to 2, (0,-1) \to 3, (0,0) \to 4, (0,1) \to 5, (1,-1) \to 6, (1,0) \to 7, (1,1) \to 8$:

$$\{f(m,n)\}_{(m,n)\in \mathrm{M}(i,j)} = \begin{array}{|c|c|c|}\hline f(-1,-1) & f(-1,0) & f(-1,1) \\ \hline f(0,-1) & f(0,0) & f(0,1) \\ \hline f(1,-1) & f(1,0) & f(1,1) \\ \hline \end{array} \longrightarrow \begin{array}{|c|c|c|}\hline f^0 & f^1 & f^2 \\ \hline f^3 & f^4 & f^5 \\ \hline f^6 & f^7 & f^8 \\ \hline \end{array}.$$

3 Heronian mean and median filters

Now we modify the classical mean and median filters (4) in the following way:

$$\widehat{s}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{MeanHeron}_k^I}\left[f(m,n)\right] = \underset{r=1,2,\ldots,N}{\mathbf{MeanHeron}_k^I}\left[f_{(i,j)}^r\right] =$$
$$= \frac{1}{MH_k}\sum_{r_1\leqslant}\sum_{r_2\leqslant}\cdots\sum_{\leqslant r_k}\sqrt[k]{f_{(i,j)}^{r_1}, f_{(i,j)}^{r_2}, \ldots, f_{(i,j)}^{r_k}}, \tag{5}$$

$$\widehat{s}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{MeanHeron}_k^{II}}\left[f(m,n)\right] = \underset{r=1,2,\ldots,N}{\mathbf{MeanHeron}_k^{II}}\left[f_{(i,j)}^r\right] =$$
$$= \sqrt[k]{\frac{1}{MH_k}\sum_{r_1\leqslant}\sum_{r_2\leqslant}\cdots\sum_{\leqslant r_k} f_{(i,j)}^{r_1}, f_{(i,j)}^{r_2}, \ldots, f_{(i,j)}^{r_k}} \tag{6}$$

for the generalized Heronian meaning filers of the first and the second kinds, respectively, and

$$\underset{r=1,2,\ldots,N}{\mathbf{MeanHeron}_k^I}\left[f_{(i,j)}^r\right] = \underset{r=1,2,\ldots,N}{\mathbf{MeanHeron}_k^{II}}\left[f_{(i,j)}^r\right] =$$
$$= \mathbf{Med}\left[\left\{\sqrt[k]{f_{(i,j)}^{r_1}, f_{(i,j)}^{r_2}, \ldots, f_{(i,j)}^{r_k}}\right\}_{r_1\leqslant r_2\leqslant\cdots\leqslant r_k}\right] \tag{7}$$

for the generalized Heronian median filter.

4 Generalized Heronian aggregation

The aggregation problem [5, 6] consist in aggregating N-tuples of objects all belonging to a given set $\mathbf{D}$, into a single object of the same set $\mathbf{S}$, *i.e.*, $\mathbf{Agg} : \mathbf{S}^N \longrightarrow \mathbf{S}$. In the case of mathematical aggregation operator (AO) the set $\mathbf{S}$, is an interval of the real $\mathbf{S} = [0, 1] \subset \mathbf{R}$, or integer numbers $\mathbf{S} = [0, 255] \subset \mathbf{Z}$. In this setting, an AO is simply a function, which assigns a number y to any N-tuple of numbers $(x_1, x_2 \ldots, x_N)$: $y = \mathbf{Agg}(x_1, x_2, \ldots, x_N)$ that satisfies:

1. $\mathbf{Agg}(x) = x$.
2. $\mathbf{Agg}(a, a, \ldots, a) = a$.
   In particular, $\mathbf{Agg}(0, 0, \ldots, 0) = 0$ and $\mathbf{Agg}(1, 1, \ldots, 1) = 1$
   (or $\mathbf{Agg}(255, 255, \ldots, 255) = 255$).
3. $\min(x_1, x_2, \ldots, x_N) \leq \mathbf{Agg}(x_1, x_1, \ldots, x_N)) \leq \max(x_1, x_2, \ldots, x_N$.

Here $\min(x_1, x_2, \ldots, x_N)$ and $\max(x_1, x_2, \ldots, x_N$ are respectively the *minimum* and the *maximum* values among the elements of $(x_1, x_2 \ldots, x_N)$. All other properties may come in addition to this fundamental group. For example, if for every permutation $\forall \sigma \in S_N$ of $\{1, 2, \ldots, N\}$ the AO satisfies:

$$y = \mathbf{Agg}(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(N)}) = \mathbf{Agg}(x_1, x_2, \ldots, x_N),$$

then it is invariant (symmetric) with respect to the permutations of the elements of $(x_1, x_2, \ldots, x_N)$. In other words, as far as means are concerned, the *order* of the elements of $(x_1, x_2, \ldots, x_N)$ is - and must be – completely irrelevant.

We list below a few particular cases of means:

1. Arithmetic mean ($K(x) = x$): $\mathbf{Mean}(x_1, x_2, \ldots, x_N) = \frac{1}{N} \sum_{i=1}^{N} x_i$.
2. Geometric mean ($K(x) = \log(x)$): $\mathbf{Geo}(x_1, x_2, \ldots, x_N) = \sqrt[N]{\left(\prod_{i=1}^{N} x_i\right)}$.
3. Harmonic mean ($K(x) = x^{-1}$): $\mathbf{Harm}(x_1, x_2, \ldots, x_N) = \left(\frac{1}{N} \sum_{i=1}^{N} x_i^{-1}\right)^{-1}$.
4. One-parametric family quasi arithmetic (power or Hólder) means corresponding to the functions $K(x) = x^p$: $\mathbf{Hold}(x_1, x_2, \ldots, x_N) = \sqrt[p]{\left(\frac{1}{N} \sum_{i=1}^{N} x_i^p\right)}$.
   This family is particularly interesting, because it generalizes a group of common means, only by changing the value of p.

A very notable particular cases correspond to the logic functions $(\min, \max,$ **median**$)$: $y = \mathbf{Min}(x_1, \ldots, x_N)$, $y = \mathbf{Max}(x_1, \ldots, x_N)$, $y = \mathbf{Med}(x_1, \ldots, x_N)$.

When filters 5–7 are modified as follows:

$$\widehat{\mathbf{s}}(i, j) = \underset{(m,n) \in M(i,j)}{\mathbf{Agg}} \left[\mathbf{f}(m, n)\right], \tag{8}$$

we get the unique class of nonlinear aggregation filters [8–11].

In this work, we are going to use aggregation operator to the Heronian (extended) data. Let $(x_1, x_2, \ldots, x_N)$ be an N-tuple of positive real numbers.

Definition 2. *The following generalized aggregations*

$$\mathbf{HeronAgg}_2^I(x_1,\dots,x_N)=\mathbf{Agg}_{i\le j}\left\{\sqrt{x_ix_j}\right\}, \tag{9}$$

$$\mathbf{HeronAgg}_2^{II}(x_1,\dots,x_N)=\sqrt{\mathbf{Agg}_{i\le j}\left\{x_ix_j\right\}} \tag{10}$$

are called the Heronian aggregations of the first and second kinds, respectively.

Here, we want to generalize Definition 2 by summarizing up the k-th roots of all possible distinct products of k elements of $(x_1,\dots,x_N)$ with repetition. The number of all such products is $C_{N+k-1}^k=MH_k$. They form the Heronian (extended) data. This determines the following definitions:

$$\mathbf{HeronAgg}_k^I(x_1,\dots,x_N)=\mathbf{Agg}_{i_1\le i_2\le\dots\le i_k}\left\{x_{i_1}x_{i_2}\cdots x_{i_k}\right\}, \tag{11}$$

$$\mathbf{HeronAgg}_k^{II}(x_1,\dots,x_N)=\sqrt[k]{\mathbf{Agg}_{i_1\le i_2\le\dots\le i_k}\left\{x_{i_1}x_{i_2}\cdots x_{i_k}\right\}}. \tag{12}$$

5 Heronian aggregation filters

Now we modify the classical mean and median filters (4) in the following way:

$$\widehat{s}(i,j)=\underset{(m,n)\in M_{(i,j)}}{\mathbf{HeronAgg}_k^I}\left[f_{(i,j)}^{r_1},f_{(i,j)}^{r_2},\dots,f_{(i,j)}^{r_k}\right]=\underset{r=1,2,\dots,N}{\mathbf{HeronAgg}_k^I}\left[f_{(i,j)}^{r}\right]=$$
$$=\mathbf{Agg}_{r_1\le r_2\le\dots\le_k}\left\{\sqrt[k]{f_{(i,j)}^{r_1},f_{(i,j)}^{r_2},\dots,f_{(i,j)}^{r_k}}\right\}, \tag{13}$$

$$\widehat{s}(i,j)=\underset{(m,n)\in M_{(i,j)}}{\mathbf{HeronAgg}_k^{II}}\left[f_{(i,j)}^{r_1},f_{(i,j)}^{r_2},\dots,f_{(i,j)}^{r_k}\right]=\underset{r=1,2,\dots,N}{\mathbf{HeronAgg}_k^{II}}\left[f_{(i,j)}^{r}\right]=$$
$$=\sqrt[k]{\mathbf{Agg}_{r_1\le r_2\le\dots\le_k}\left\{f_{(i,j)}^{r_1},f_{(i,j)}^{r_2},\dots,f_{(i,j)}^{r_k}\right\}}, \tag{14}$$

for the generalized Heronian aggregating filters of the first and the second kinds, respectively. In particular case ($k=1$) we get the unique class of nonlinear aggregation filters [8, 9].

6 Experiments

Generalized aggregation Heronian filtering with $\mathbf{Agg}=\mathbf{Mean}$, $\mathbf{Med}$ has been applied to noised 256×256 gray level images "Dog" (Figures 1b, 2b). The denoised images are shown in Figures 1–2. All filters have very good denoising properties. This fact confirms that further investigation of these new filters is perspective. Particularly, very interesting is a question about the types of noises, for which such filters are optimal.

7 Conclusions

We suggested and developed a new theoretical framework for image filtering based the Heronian mean and median. The main goal of the work is to show that Heronian mean and median can be used to solve problems of image filtering in a natural and effective manner.

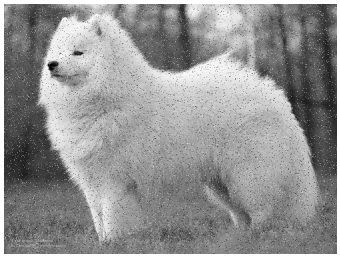
Acknowledgment. This work was supported by grants the RFBR Nos. 13-07-12168 and 13-07-00785.

References

1. Mitra, S.K.: Nonlinear Image Processing. Academic Press Series in Communications, Networking, and Multimedia, San Diego, New York, p.248, (2001)
2. Sykora, S.: Mathematical Means and Averages: Generalized Heronian means I. Stan's Library, Ed. S.Sykora, I (2009)
3. Sykora, S.: Generalized Heronian means II. Stan's Library, Ed. S.Sykora, II (2009)
4. Sykora, S.: Generalized Heronian means III. Stan's Library, Ed. S.Sykora, III (2009)
5. Mayor, G., Trillas, E.: On the representation of some Aggregation functions. In Proceeding of ISMVL, 20, 111–114 (1986)
6. Ovchinnikov, S.: On Robust Aggregation Procedures. Aggregation Operators for Fusion under Fuzziness. Bouchon-Meunier B. (eds.), 3–10 (1998)
7. Kolmogorov, A.: Sur la notion de la moyenne. Atti Accad. Naz. Lincei, 12, 388-391 (1930)
8. Labunets V. G.: Filters based on aggregation operators. Part 1. Aggregation Operators. 24th Int. Crimean Conference Microwave & Telecommunication Technology (CriMiCo2014), Sevastopol, Crimea, Russia, 24, 1239–1240 (2014)
9. Labunets, V.G., Gainanov, D.N., Ostheimer, E.: Filters based on aggregation operators. Part 2. The Kolmogorov filters. 24th Int. Crimean Conference Microwave & Telecommunication Technology (CriMiCo2014), Sevastopol, Crimea, Russia, 24, 1241–1242 (2014)
10. Labunets, V.G., Gainanov, D.N., Tarasov A.D., Ostheimer E.: Filters based on aggregation operators. Part 3. The Heron filters. 24th Int. Crimean Conference Microwave & Telecommunication Technology (CriMiCo2014), Sevastopol, Crimea, Russia, 24, 1243–1244 (2014)
11. Labunets, V.G., Gainanov, D.N., Arslanova, R.A., Ostheimer, E.: Filters based on aggregation operators. Part 4. Generalized vector median filters . 24th Int. Crimean Conference Microwave & Telecommunication Technology (CriMiCo2014), Sevastopol, Crimea, Russia, 24, 1245–1246 (2014)

Appendix. Figures

a) Original image



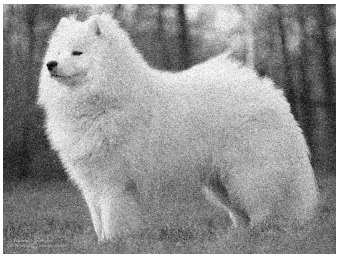
b) Noise image, $PSNR = 21.83$

c)
MeanHeron, $PSNR = 32.364$

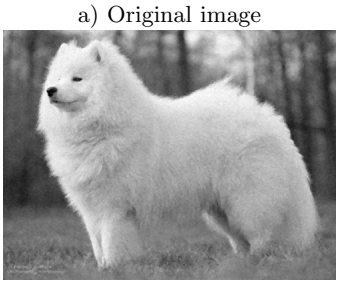
d)
MedHeron, $PSNR = 31.297$

Fig. 1. Original (a) and noise (b) images; noise: Salt-Pepper; denoised images (c)–(f)

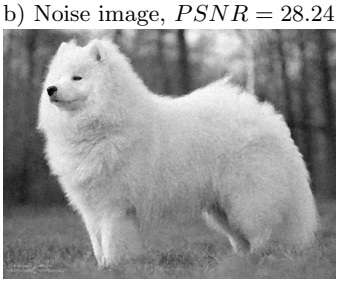
a) Original image



b) Noise image, $PSNR = 28.24$



c)
MeanHeron, $PSNR = 31.293$



d)
MedHeron, $PSNR = 29.531$

Fig. 2. Original (a) and noise (b) images; noise: Laplasian PDF; denoised images (c)-(f)

Imitation of Human Behavior in 3D-Shooter Game

Makarov Ilya, Tokmakov Mikhail, Tokmakova Lada

National Research University Higher School of Economics,
Department of Data Analysis and Artificial Intelligence
iamakarov@hse.ru
matokmakov@gmail.com
lrtokmakova@yandex.ru

Abstract. There are many algorithms for 3D-shooter game artificial intelligence that are based on automatic target recognition with the help of ray casting, script-based event systems and many others. In such models game agents have an advantage over human player because computers do not need time for enemy recognition and target aiming processes. The main goal of our research is to imitate a realistic enemy player under PC control, which acts similar to a human reaction. We consider such aspects of shooter in-game processes as visual recognition of players, the delay for the response on each game event and distributed strategies of decision making. The paper presents method of visual recognition of player's 3D-model with the use of geometry knowledge of the three-dimensional maze and target sighting with respect to human motor reflexes. The implementation is made in Unreal Engine 4, which provides large variety of methods for programming game agents.

Keywords: Game Artificial Intelligence, Neural Network, Visual Recognition, Sparse Network of Winnows, Boosting, Rule-based Systems.

1 Introduction

Entertainment industry provides a vast variety of different games, every time attempting to improve realism of virtual world. It involves modeling of an enemy under automatic computer control (so called, BOT) having reaction that is similar to human reaction. In fact, the user has to spend some time for enemy visual recognition, and adapt motor reflexes for target sighting of proper level. These circumstances together with the use of ray-casting algorithms allow an enemy to obtain immediate perfect target aim without even small time delays and sighting errors.

Multiple algorithms are presented to compensate BOT supremacy over human player, such as simulated time delay to fire back, bad accuracy during shooting process, almost infinite health level and infinite regeneration of human player and many other methods to retain balance between entertainment and game challenge in virtual world. We focus on creating game artificial intelligence agent for a 3-D shooter game imitating human-like behavior.

The article represents a new approach to the problem of creating realistic game agent for 3D-shooter. We try to find a solution to the inverse problem using machine and

neural learning algorithms to obtain maximum likelihood estimators of target correction parameters and enemy detection. We also make statistical experiment to find corresponding clusters of such delays relative to skill-level of human player.

Implementation is made using modern technologies of game agent modeling, which allow us to combine different schemes and algorithms into one common behavior tree algorithm, saving the properties of complementarity and interchangeability of the individual modules for decision making models and different visual recognition patterns.

The Blueprints visual scripting system in Unreal Engine 4 is a complete gameplay scripting system based on the concept of using a node-based interface to create gameplay elements from within Unreal Editor.

With the use of Blueprints, programmers can implement and modify any gameplay element:

- Gameplay conditions and rules.
- Developing diverse meshes and *Materials* or character customization.
- New camera perspectives and change dynamically the camera view.
- Items - weapons, pickups, med kits, triggers, and others.
- Environments - create procedurally-generated items.

Blueprints combine the best properties of previous AI models and propose sufficient interface to create complex models in short time.

Finally, we may say that developers frequently use existing models of computer player's behavior, which sometimes do not allow to achieve sufficient realism. In different games we may have texture issues, bad organized respawn, wanton shooting and the worst – disadvantage of human player in comparison with game AI due to widespread usage of engine algorithms in decision making algorithms. Our work is an attempt to form our own pattern of AI, which will combine human and machine visual perception and decision making models to implement human-like BOT.

Real-time game play has always been major setback for AI in video games. If the action continues arbitrary long, there will be no time to compute all possible actions and future states. In order to achieve diverse BOT difficulty levels a specific model should be constructed, providing computational part for:

- Confidence intervals for time delay and sighting accuracy using grade of visual recognition;
- Target sighting on fire recognition with adjusted parameters;
- Statistical model for the curves of the first (initial) and the second (correction) targeting accuracy and delays;
- Adaptation of shooting process with respect to recoil;
- Rule-based decision making system for priority actions with respect to a map of game events.

The first part of the paper is brief review of existing games. We analyze the history of BOT developing and snow minuses of automatically visual recognition. The second part of the article considers methods of image processing with the help of geometry

knowledge of three-dimensional maze. We present a statistical prediction model to obtain possible dangerous directions where an enemy appearance may occur more frequently and we organize some data structure to process these frequencies as priority goals for decision making module. The third part of our article describes basic algorithms of visual recognition and their applications to our problem relative to practical success in terms of time complexity and percent of successful enemy recognition. We deal with the description of samples' properties, and use neural network for frequencies, obtained from the geometry of a maze, to increase the possibilities for enemy recognition in dangerous zones. It does improve the quality of recognition in difficult places such as corners, box's covers, and does not affect for enemy recognition in front of a player, where the recognition method gives good results by itself. The next part of the paper deals with the decision making model and represents a combination of two mentioned above approaches into one AI algorithm. The article ends with the description of gaming process.

2 Brief Review of Existing 3D-shooters

The first "3-D action" was created in 1992 with the use of «Wolfenstein 3D» engine. The game has the same name, and visualization of 3D map and models was achieved by ray casting. This method emits one ray for each column of pixels, evaluates to see whether it has intersection with a wall and draws textures creating a one-dimensional Z-buffer. The game has widely popularized the genre on the PC and has found the basic run-and-gun type for many subsequent first-person shooter games (FPS games).

The game mentioned above was not as popular as its analog "Doom", which was developed in 1993. It was based on «Doom engine», which belongs to the same series of engines like «Wolfenstein 3D». In comparison with Wolfenstein 3D the Doom engine includes such features as non-perpendicular walls, full texture mapping of all surfaces, dynamic platforms and floors. In addition to ray casting the BSP tree technology created by Bruce Naylor was used. The lifelike environment was achieved by using the stereo sound system, which made it possible to quickly determine the direction and distance of a sound effect made by an enemy or occasional environment event. The player has to be on guard while he hears monsters' roars, and receives additional knowledge to find new areas in the form of sounds of some door which open remotely. Enemies can also become aware of the player's presence by hearing distant gunshots. There were also popular features of the Doom engine, such as the ability to create custom scenarios contributed significantly to the game's popularity.

It was the first step of FPS gameplay evolution. Nowadays the market of computer games is overcrowed. An artificial intelligence of many shooters has several shortcomings because developers concentrate more efforts on visual effects and entertainment, whereas BOT action can be not realistic. We shall now move to brief review of several modern games relative to artificial intelligence models and their properties in the context of setting goals for BOT.

One of the most popular FPS games, "Call of duty: Ghosts", was published in 2013 [1]. Despite the high-quality graphic and large variety of cinematic effects CoD series

continued modern tradition to publish AAA projects without sufficient testing procedures. One's virtual players can shoot the ceiling trying to kill enemies, which are on the floor above. Also, some adversaries can rest against the wall and not notice User in two steps. There is also a famous bug migrating in the CoD series: after the player's death reloading last checkpoint activates all enemy AIs without even appearing of player in the viewing angle. Some bugs appeared from wrong scripts, but many others involve low-level command block and determined algorithms, which do not cover all gameplay opportunities. One of the things that were highlighted during the Call of duty: Ghosts presentation was the feature that fish AI moves out of the way when user gets close. Actually, this feature is not even an innovation — it was nearly 20 years old!

The so-called "NextGen" notion for next generation of games is apparently relevant only for the graphical part of the games, but the vision of AI models is in creating tools for different games and simple developer kit but not for peculiar game properties.

"Counter-Strike: Global offensive" was developed in 2012. BOTs' interaction with outside world is not realistic in some cases, for example, enemies can jump from the second floor and their level of health will not change. However, to support the many community-created maps, CS:Source and CS:Global Offensive include an automatic mesh generation system. The first time you play a custom map with BOTs, the generation system will build a .nav file for that map. Depending on the size and complexity of the map, this process may go for a few minutes to a few hours.

The following steps occur during Navigation Mesh generation:

- Starting at a player spawn point, space for walking is sampled by "flood-filling" outwards from that spot, searching for adjacent walkable points
- Right-Angled Navigation Areas are built from the sampled data
- Hiding and Sniper spots are computed
- Encounter Spots and Approach Points are calculated (this can take awhile)
- Initial Encounter areas are computed

There are general mesh problems, for example, extraneous areas on stairs and ramps, railings, preventing excessive jumping, dealing with rotating doors.

BOTs will sometimes try to walk through the opened door brush from one nav area to another and get stuck.

Although mesh navigation had leading positions for a few years there were many issues with moving scene, extreme position cases, so a new concept for programming BOT action has appeared.

Summing up, we see that automatically enemy detection can lead to different types of errors in shooter games. It does not represent realism, for instance, identification of enemies through the walls and appearance of BOT in textures. The research aims to create own detection module in order to avoid mentioned above mistakes.

3 Geometry

We created BOT in Unreal Engine 4. The order of its actions depends on the objects encountered on its way. The appearance of the virtual player is a template shown in the picture:

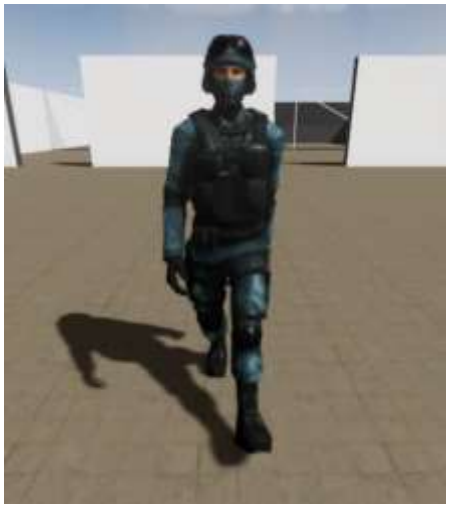


Fig. 1. Screenshot of BOT

We consider the one-story maze with the following objects:

- Walls
- Doorways
- Boxes
- Columns

The order of BOT's actions is a priority queue, which is denoted by S. When our virtual player faces the corners in the maze, S updates. All mentioned above objects have several corners. First of all, we calculate automatically mesh navigation to determine corners and separate possible fast scene changing. Then we process the queue considering a set of neighbors:

1. finding element with maximal priority
2. recalculating danger zones with respect to impact of neighbors by geometrical closeness with respect to mesh navigation
3. comparing the first K queue elements by using decision making module.

The constructed model allows us to separate geometrical combination of danger zones by Euclidian distance and different priorities in BOT reaction.

Let's look at an example illustrated in the picture:

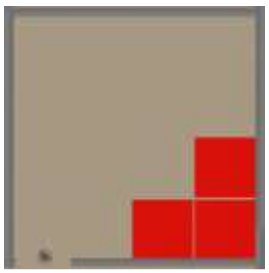


Fig. 2. Around the corner

We obtain the event “look around the corner” skipping visual recognition of the objects as if we already know them. We find out similar events with nearby distance from the current corner and include them in our calculations. The situation is separated by three simple steps. The first of them is the case in which BOT meets the nearest corner (the right side of the doorway). Then we process boxes as consequent corners in our model with respect to their priority and distance. The probability to turn on the angle X from 0 to 90 degrees with respect to the right wall is represented by density functions. A little circle represents BOT position.

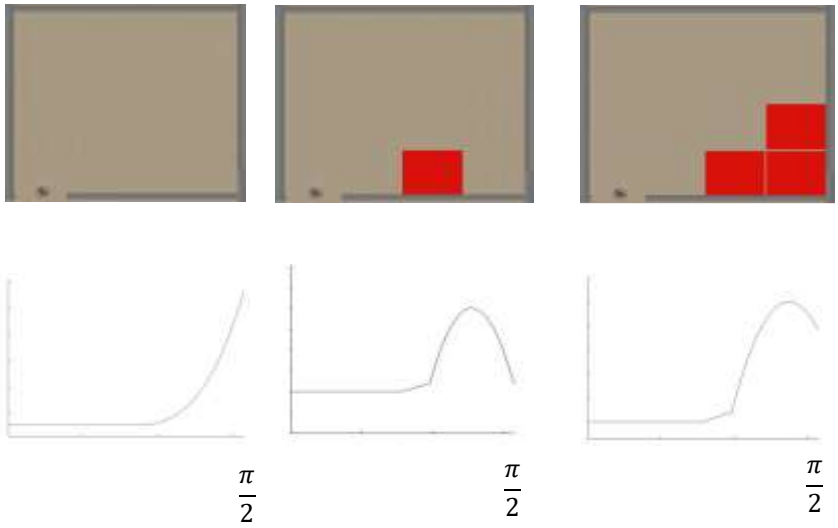


Fig. 3. Processing priority queue when looking around the corner

The obtained density functions represent the process of visual recognition of objects when we search for an enemy but have to spend some time on processing object’s shapes. Using geometry knowledge we reduce amount of time on visual recognition of objects using its statistical equivalent to find dangerous zones as equivalent of comparing visual images of objects and an enemy.

4 Visual Recognition

Visual recognition is the assignment of given pictures to a particular class of objects with the help of known methods of classification. We implement three standard steps of it:

- The converting of the source image to the initial idea;
- The extracting characteristic values of recognition;
- The classification of the image to a certain class with the use of the classifier.

In our work we consider three basic approaches to the visual recognition:

- *Heuristic methods* [2, 3]
    - *Full heuristic model*. Expert compiled a set of rules that describe the image of the object (built model), according to which detection is produced.
    - *Search for characteristic invariant attributes*. It is the description of characteristics of the object, which are invariant with respect to possible distortions (lighting changes, rotation, scaling).
- *Method of pattern matching*. It is the creation of a template for the image of the whole object or its characteristic features. The pattern can be a complicated structure and allows for different deformations and transformations. This approach allows implementation in real time.
- *Methods of learning by precedents.* The model is automatically built on the set of images of an object compiled beforehand from the possible input data of the system [4, 5, 6, 7].

The latter approach is the most interesting for us, because results of testing process show that such methods are more appropriate, so let's consider the last approach in detail. It consists of two units: converting the image into a feature vector and classification. The former means the most complete and informative representation of an image in the form of a numerical vector. The latter fact implies the verification of hypothesis accessories image to the class object images based on the observation that is the feature vector.

Denote a feature vector for $x \in X \subseteq R^n$ - the description of the object, which is output of the conversion unit. Class will be called a subset $K_y = \{x \in X, y^*(x) = y\}$ of the set X, $y \in Y \subseteq Z$ – the set of the class markers. We will view the case of binary classification (the image is an image of the desired object or the image is an image of anything else), consequently $Y = \{-1;\ 1\}$. $y^*(x)$ – the mapping, which is defined for all $x \in X$ and is specified the partition into subsets K_y. It is worth noting that classes K_y can overlap, because the feature vector is only a description of an object's attributes, and two different images can have the same specifications. The training set is pairs of precedents $T = \{(x_1, y_1), \dots, (x_l, y_l)\}: y^*(x_i) = y_i. i = \overline{1, l}.$

For the application of classification algorithms and pattern recognition we make the following hypothesis: the set $X \times Y$ is a probability space with probability measure P,

precedents $(x_1, y_1), \dots, (x_l, y_l)$ appear randomly and independently in accordance with the distribution P.

The task of classification is to build the function $F(x)$, which approximates the mapping $y^*(x)$ with the help of the training set.

Let's consider several methods of learning by precedents, which we tested in our work.

4.1 Bayesian Classification Methods

The first method is the principle of posteriori probability maximum [8]. It is based on following hypothesis:

- $X \times Y$ is the probability space with probability measure P, precedents $(x_1, y_1), \dots, (x_l, y_l)$ appear randomly and independently in accordance with the distribution P.
- We know density distributions of classes $p_y(x) = p(x|K_y), y \in Y$, which we will call likelihood functions.
- We know the probability of occurrence of each object from classes $P_y = P(K_y), y \in Y$, which we will call the priori probabilities.

Decision rule can be written in the following form:

$$F(x) = arg \max_{y \in Y} P(K_y|x) = arg \max_{y \in Y} p_y(x) P_y \quad (1)$$

The second method is called Bayesian networks [9]. This approach is based on the combination of the principle of posteriori probability maximum with graph theory. The idea is to create a graph with vertices corresponding to any component of the feature vector, and edges indicating a causal relationship.

4.2 Classical Neural Networks

The main idea is the serial conversion of a signal with the elementary functional elements (neurons), operating in parallel [10]. The basic principle of neural network configuration is to use optimization methods to minimize the mean square error. Also, the neuron networks are capable of retraining.

4.3 SVM - Support Vector Machine

The SVM algorithm constructs a linear separating surface (hyperplane), equidistant from the convex hulls of the classes, convex hull is based on the precedents [11]. If the hyperplane does not exist (classes are not linearly separable), sound conversion will apply for nonlinear classification, which projects X in the space of higher dimension, probably, infinite. The algorithm for constructing the classifier is reduced to the problem of quadratic programming. The solution of this task will be unique and the found extremum will be global.

4.4 SNoW - Sparse Network of Winnows

Sparse network of Winnows is a special type of neural networks [12]. The network consists of two (the number of possible classes) of linear neurons associated with the feature vector components. Geometrically, SNoW consists in two hyperplanes in the space of feature vectors. Vector belongs to the class, corresponding to which the hyperplane is the closest. The resulting separating surface is thus a hyperplane in the original space X.

4.5 Classifier Boosting

This approach combines primitive classifications into one stronger method. The main idea is an iterative minimization of a convex functional classification error. Boosting attempts to project the original data into a space where the classes are linearly separable.

4.6 Results

In our research we obtained the following results, using data from demorecords of multiplayer game sessions:

- SNoW and boosting distinguish the highest percentage of correct detections, about 94% [13, 14];
- SVM, SNoW and boosting provide high speed of recognition;
- SNoW and boosting have low level of the second kind error.

The correctness of visual recognition was manually checked. Thus, at the current level of experiment the most appropriate method for our problem is the combination of two methods: Sparse network of Winnows and Classifier boosting.

The picture below represents the result of visual recognition unite, using classifier boosting.

Fig. 4. Result of visual recognition

In this situation there are two objects, which can be recognized as enemies: the virtual player and its shadow. But the geometry unit defines that the latter is on the floor, so AI focuses only on the first one.

5 Decision Making Model

In the most basic games, there was simply a lookup table for what was currently happening and the appropriate best action. In the most complex cases, the computer would perform a short minimax search of the possible state space and return the best action [15]. The minimax search had to be of short depth since the game was occurring in real-time.

We use the blueprints visual scripting system in Unreal Engine 4 [16, 17]. The node-based interface allows us to create the maze with different objects (for instance, boxes, doors). Also, this concept gives an opportunity to make any appearance of BOT and to determine the direction of the eye with the help of a camera-element. The virtual player is the combination of several blueprints, each of which corresponds to the certain action: moving, shooting, behavior during shooting and search for enemies.

We focus on a rule-based system (RBS) [18, p.169] to identify simple situations and make a choice for a complex processes as a tuple of simple rules. RBS works with instructions and the interpreter. In our case, rules are oriented tree structures. Let's discuss the short list of game events:

1) Walking
    a) Bypassing and Search
    b) Moving toward some point in the maze
    c) Finding cover
2) Shooting
    a) Sighting
    b) Opening fire
    c) Correction relative to recoil and motor reflexes
3) Visual recognition
    a) Recognition of objects
    b) Recognition of dangerous zones
    c) Recognition of enemy

The simplest set of rules may be the following:

- $1a) \wedge 1c) \rightarrow 1a)$ (just bypassing)
- $1a) \wedge 3b) \rightarrow 2a) \wedge 1b)$ (accurate obstacle avoidance)
- $1a) \wedge 3c) \rightarrow 2a) \wedge 2b) \wedge 2c)$ (shooting after enemy recognition)

We give only a superficial representation of these rules, which are much more complicated when we go into details with respect to implementation in game engine.

Rule based system allow us to implement the neural network method when we have to choose another rule with the same premise, for example, if we have a rule:

$$1a) \wedge 3c) \rightarrow 1b) \wedge 1c) \quad (2)$$

This rule should be used, if a player has low health and his chance to win against enemy is almost zero, so he should hide and create an ambush for an enemy.

If we have time-limit and command should defend some place we should use the following rule:

$$1a) \wedge 3c) \rightarrow 1b) \wedge 2b) \quad (3)$$

So we will distract an enemy or sacrifice player to win some time for a command.

In a real game the player chooses different rules depending on his psychological temper, emotions or calculations, so it is really a good place to use learning techniques.

For each type of game we produce rules, which say whether or not player achieve his current goal and increase a part of command goal for this game. Of course command goal has greater priority, but in fact, in each situation the player decides without hesitation operating with current situation and parameters from visual recognition.

Formal Concept Analysis methods are used to find out a basis of rules (implications) when we have a non-empty intersection of premises. Machine Learning is used to find weight combination for rules, which have the same premises.

Model of neuron networks can be applied to univariate and multivariate analysis. In our work using neuron approach is appropriate to imitate personal and team player's goals. One player worries about its level of health; also, it is interested in winning. These two ideas characterize its behavior during the game; it means that the following situation is possible: BOT may decide to shoot and be with the low level of its health, because it will win. But, when we consider team game, BOT will not risk by it-self, it will defend, possible, shoot or run away. This case of game is complicated, but we can divide it into several simple cases, each of them is the same as for personal game. That is why neural networks can successfully reflect behavior of players in shooter.

6 Process of Gaming

Process of gaming starts with visual recognition. As far as BOT defines an enemy with certain threshold probability or after enemy fire identification, it starts to aim. Given that X is the horizontal rotation change speed, and Y is the vertical. The relative error depends on the angle between BOT's rotation movements and direction on the target, and the sign of this angle. It also depends on X-speed and Y-speed: the stronger the need to change the speed of rotation change, the greater the relative error will occur.

Let's consider two situations of possible targeting. In both examples visual recognition shows the place of the enemy. Also, we can see the rotation of BOT, which describes by means of the sum vector of X and Y. Black circles shows the approximate trajectory of aiming which is close to the human response to the detection of the adversary. The figure 5 represents the situation in which the direction to the target is opposite to the BOT rotation, therefore the relative error will be very high.

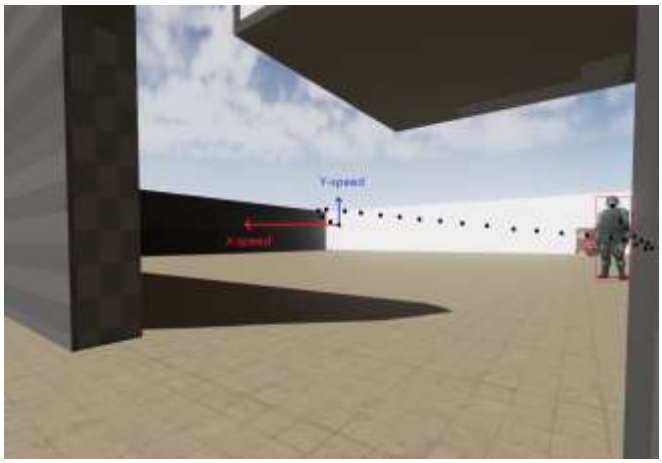


Fig. 5. Example 1

The figure 6 displays the case in which the direction to the goal and the rotation are the same; consequently, the relative error will be rather low.

Fig. 6. Example 2

So, we divide aiming by three phases:

- inertia during recognition,
- instant aiming,
- braking or correction near the goal,

and include them in our model.

7 Current Progress and Conclusions

We have stated game AI model and test it manually on demorecords. We will provide statistical parameters from experiment on one enemy at three-dimensional maze with over 500 different positions of players. In the current state we are verifying the parameters to identify the dangerous zones and we proceed with the comparison of the methods of visual recognition to apply for our model. RBS is now constructed automatically using Unreal Engine 4 and we recalculate bases of implications for every addition of new rules block. The system is flexible enough to extend to a new rule and to recalculate parameters of neural networks for a problem of rule choice with similar premises.

The following picture represents the current work of our AI:

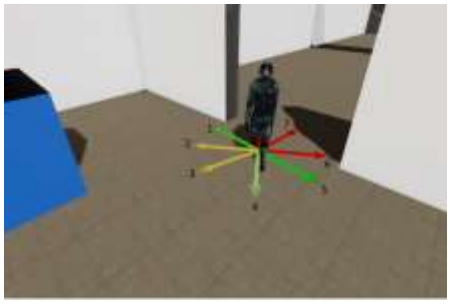


Fig. 7. Looking around the corner

The BOT AI can decide in which direction it will move on. Bright green arrows (1, 5) show that appropriate routes are the most probable, whereas red arrows (6, 7) correspond to the least probable paths. Yellow cursors (2, 3) display that such directions are possible, but the probability is not very high.

We aim to implement the blueprint, which will correspond to target sighting with adjustments. At the current progress we test the model of targeting with respect to dedicated case of one moving enemy at scene, but we will continue to verify this module with many enemies and partial recognition cases after completing results on the best visual recognition models.

Acknowledgements

The authors would like to thank the colleagues from Higher School of Economics Geoffrey Decrouez for great support during preparation process and Anton Konushin for well-timed advice on visual recognition methods.

References

1. Yanovich I. “Review of Call of Duty: Ghosts” (in Russian) http://kanobu.ru/articles/syuzhetnaya-kampaniya-call-of-duty-ghosts-367451/
2. Yang G. and Huang T. S., *“Human Face Detection in ComplexBackground”*. Pattern Recognition, vol. 27, no. 1, pp. 53-63, 1994.
3. Kotropoulos C. and Pitas I., “*Rule-Based Face Detection in Frontal Views”.* Proc. Int'l Conf. Acoustics, Speech and Signal Processing, vol. 4, pp. 2537-2540, 1997.
4. Pham T. V., Worring M., and Smeulders A. W. M. “*Face detection by aggregated bayesian network classifiers”*. Pattern Recognition Letters, 23(4):451-461, February 2002.
5. Henry A. Rowley, Shumeet Baluja, and Takeo Kanade. “*Human face detection in visual scenes”*. Technical Report CMU- CS-95-158R, School of Computer Science, Carnegie Mellon University, November 1995.
6. Roth D., Yang M.-H., and Ahuja N. “*A SNoW-based face detector”.* Advances in Neural Information Processing Systems 12 (NIPS 12), MIT Press, Cambridge, MA, pp. 855-861, 2000.
7. Viola P. and Jones M. “*Robust Real-time Object Detection”.* In Proc. 2nd Int'l Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing and Sampling, Vancouver, Canada, July 2001.
8. Vorontsov K.V. “Bayesian classification algorithms. Drafts of lectures” (in Russian) http://www.ccas.ru/voron/download/Bayes.pdf
9. Heckerman, D. *“A Tutorial on Learning with Bayesian Networks”.* In Jordan, M. (Ed.), Learning in Graphical Models, MIT Press, 1998.
10. Vezhnevec A. “ Popular neural network architecture” (in Russian) http://cgm.graphicon.ru/metodyi/populyarnyie_neyrosetevyie_arh itekturyi.html
11. Vapnik V. and Lerner A. J. “*Generalized portrait method for pattern recognition”.* Automation and Remote Control, vol. 24, no. 6, 1963.
12. Roth D. “*The SNoW Learning Architecture”.* Technical Report UIUCDCS-R-99-2102, UIUC Computer Science Department, 1999.
13. Ming-Hsuan Yang, David Kriegman, and Narendra Ahuja. “*Detecting Faces in Images: A Survey”.* IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 24, no. 1, pp. 34-58, 2002.
14. Viola P. and Jones M. *“Robust Real-time Object Detection”.* In Proc. 2nd Int'l Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing and Sampling, Vancouver, Canada, July 2001.
15. Szelinski R. “Computer Vision: Algorithms and Applications”. Springer, 2011.
16. https://www.unrealengine.com
17. http://unreal-engine4.ru
18. Champandard A.J. “AI Game Development: Synthetic Creatures with Learning and Reactive Behaviors”. New Riders Games, 2003.

The Text Network Analysis: What Does Strategic Documentation Tell Us About Regional Integration?

A.Murashov[1], O.Shmelev[2]

[1] Yaroslav-the-Wise Novgorod State University
(murashov.andrey@mail.ru)
[2] Kalashnikov Izhevsk State Technical University

Abstract. Values and attitudes towards the regional integration process of the Russian political elites are considered as an indication of what regional integration (RI) tends to be and how it evolves over time. This paper suggests how to systematically grasp and integrate elite's attitude into the analysis of RI by means of text network analysis. The text network analysis allows one to visualize the meanings and agendas present within political manifests which are supposed to reflect values and attitudes towards RI of the local political elite.

Keywords. igraph, political elite attitude, R, regional integration, regional strategy, system of indicators, text mining, text network analysis (TNA)

1 Introduction

This paper is a part of PhD thesis aimed at constructing a so-called System of Indicators of Regional Integration in Russia. Values and attitudes towards the regional integration process of the Russian political elites are considered as an indication of what regional integration tends to be and how it evolves over time. One of the shortcomings of conventional approaches is insufficient and unsystematic consideration of political elite's attitude towards regional integration and decision-making in this field. The question how to systematically grasp and integrate elite's attitude into analysis remains open.

In comparative politics measuring the attitudes of the political elite is often undertaken either by expert surveys or by the analysis of political manifests [De Lombaerde P. et al., 2011]. Our interest lies in researching the political manifests – a regional strategy reflects opinion of a local authority, like a party manifest directly reflects opinion of a political party.

Research questions are those related to monitoring and analyzing regional integration process. How do regions cooperate? What forms of integration emerge within the selected regions? What problems / challenges does integration impose? Which industries are mainly affected by integration process? How do regions choose their region-partners? What is beyond the choice?

2 Background

From theoretical perspective this paper is supposed to contribute to the investigation of values and attitudes towards the regional integration process that are represented in political manifests. This topic is covered in particular by:

(1) Comparative Manifesto Project (CMP) maintained by Manifesto Research Group. Their purpose is to discover party stances by quantifying their statements and messages to their electorate, method used is quantitative content-analysis [CMP, 2014];

(2) Leontief Center's Study of Russian Regions' Strategies aimed at, among other things, building a classification of regional strategies based on their content, method used is expert review and content-analysis [Zhikharevich B. et al., 2013];

(3) Philippe De Lombaerde from United Nations University, Institute on Comparative Regional Integration Studies (UNU-CRIS) and his team who employing multi-disciplinary approach in developing quantitative and qualitative tools to monitor regional integration process [De Lombaerde P. et al., 2011].

3 Method

From methodological perspective this paper applies an approach which combines two methods - comparative text-mining and graph analysis – "text network analysis". The text network analysis allows one to visualize the meanings and agendas present within political manifests. This approach outputs a graph of relations between key terms where each node represents a term and edges express logical associations between terms.

Putting it in a general scenario of social networks, the terms are taken as people and the segments of text as groups on LinkedIn or Facebook, and the term-document matrix can then be taken as the group membership of people. Several notions of co-occurrence have been used in the literature to group words [Saeedeh M. et al., 2010]: document-wise/sentence-wise /window-wise/syntax-wise co-occurrence. We build a network of terms based on their co-occurrence in the same text segments (paragraphs) extracted from the documents in the course of expert review. There is an edge between two terms if they appear in the same text segments (paragraphs). The weight of an edge is its frequency [Batagelj V. et al. 2002, Polanco X., 2006]. Such a network (or conceptual map [Chernyak E. et al., 2014]) visualizes logical associations between concepts presented in the political manifests.

1. Establish text corpus and transform it

Data to analyze is regional strategies of socio-economic development as a central and most capacious source of information about political elite's views on regional integration process. We are interested in 6 Russian regions situated alongside the Moscow – St Petersburg transport corridor: Moscow, Moscow region, Tver' region, Novgorod region, Leningrad region, St Petersburg. Their strategies are studied. There may exist a wide range of other official documents on regional integration but unfor-

tunately we are not able to cover all of them, so we decided to limit our sample by regional strategies only.

Using Atlas.ti (qualitative data analysis software) we establish text corpus and retrieve those text segments (paragraphs) from the regions' strategies which refer to regional integration process, refine it in a specific way then (lemmatization, filter stopwords, punctuation and numbers removing, etc.).

2. Explore text corpus (igraph & tm packages)

Text network analysis is performed with R [Yanchang Zhao, 2012], specifically, with packages {igraph} and {tm} (provides functions for text mining). We build a document-term matrix, after that, it is transformed into a term-term adjacency matrix, where the rows and columns represent terms, and every entry is the number of concurrences of two terms, after that, frequent words and their associations (fastgreedy.community) are found from the matrix.

3. Visualization

Finally, we visualize the result by means of {igraph} package in R environment: (1) plot the graph to show the relationship between frequent terms (graph.adjacency, layout = layout.fruchterman.reingold, delete.edges), (2) dendrogram (dendPlot).

4 Results

First we review general graph statistics. Snapshot of the network metrics is in the table following (tab.1). Volume of the strategies varies from 396 vertices for Moscow region to 800 vertices for Tver' region. Function assortativity.degree uses vertex degree (minus one) as vertex values. The coefficient throughout the corpora is negative suggesting that the connected vertices tend to have different degrees. Centralization is a general method for calculating graph-level centrality score based on node-level centrality measure. Novgorod region's strategy is that one having most centralized structure (centralization degree of 68% of its theoretical maximum). We also arrive at the conclusion that there is a substantial amount of centralization in the Moscow region's strategy. In general, the power of individual terms varies rather substantially, and this means that, overall, positional advantages are rather unequally distributed in each strategy. The global version of clustering coefficient (function transitivity) indicates that the degree to which nodes in a graph tend to cluster together is relatively low. This makes sense since we removed from the graphs singular edges for the sake of simplicity (here we refer to a parameter n which is discussed below). Fastgreedy algorithm identifies from 6 to 10 communities in the graphs with moderate modularity. As we can see from the table 1 the graphs are quite similar in terms of their mathematical conception. Much more insightful and interesting results come from analysis of the networks' content.

Table 1. Graphs' key metrics[1]

Parameter	{igraph} function	Strategy					
		SP	M	LO	NO	TO	MO
Number of vertices	vcount	737	463	490	491	800	396
Number of edges	ecount	5039	2311	2517	2913	7296	1897
Assortativity	assortativity.degree	-0.25	-0.34	-0.32	-0.24	-0.31	-0.29
Transitivity	transitivity	0.19	0.22	0.16	0.21	0.22	0.16
Average path length	average.path.length	2.54	2.64	2.48	2.43	2.52	2.58
Graph density	graph.density	0.019	0.022	0.021	0.024	0.023	0.024
Centralization Degree	centralization.degree	0.49	0.41	0.52	0.68	0.48	0.68
Centralization Closeness	centralization.closeness	0.54	0.48	0.53	0.67	0.50	0.61
Centralization Betweenness	centralization.betweenness	0.30	0.30	0.27	0.55	0.19	0.39
Eigenvector Centrality Scores	centralization.evcent	0.92	0.91	0.92	0.92	0.91	0.92
Diameter	diameter	13	10	13	13	10	14
Number of communities (best split)	fastgreedy.community	6	6	10	8	8	8
Modularity (best split)	fastgreedy.community	0.40	0.49	0.35	0.38	0.32	0.38

To demonstrate some examples for applying the strategies to study regional integration the graphs following are built (fig.1). They are based on the strategy of St Petersburg. The graph (fig.1,a) is crowded with many vertices and edges, it represents most of the ideas we can find in the strategy. To simplify the graph we remove insignificant terms. With function delete.edges, we remove edges which have weight less than a certain value. To do it in our experiment we introduce a parameter n referring to a number of text segments (paragraphs) where a certain term appears. After removing edges, some vertices become isolated and are also removed. The produced graph is on fig.1,b. The interpretation is that we exclude from the scope of analysis most rare and random concepts.

Let us set n equal to 8. The resulting graph on fig.2,a is crowded with many vertices and edges, we can interpret it at some extent but we need to get more precise picture. We identify vertices whose removal increases the number of connected components in the graph. They are: city, petersburg, development, etc. To simplify the graph and find relationship between terms beyond the selected keywords, we remove major articulation points (or alternatively those terms whose removal, we expect, will lead to a result we are looking for) so that the layout is rearrange and new concepts and links between them are revealed. We see that some of the articulation points are not necessarily meaningful but just the highly frequent words carrying less meaning than those with a moderate or low frequency and are thus not very valuable to explore.

[1] SP = St Petersburg, M = Moscow, LO = Leningrad region, NO = Novgorod region, TO = Tver' region, MO = Moscow region

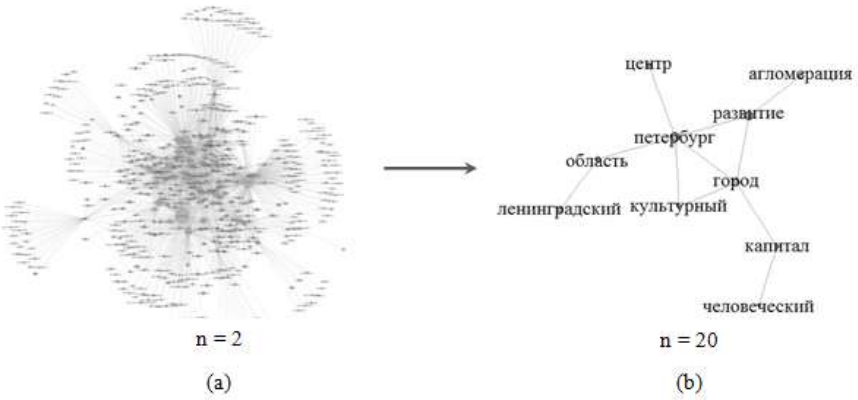


Fig. 1. Example of graph evolution (a – initial graph; b – truncated graph)

Next, we try to detect communities in the graph. Graph community structure is calculated with the fastgreedy algorithm [Kincharova A., 2013]. The nodes that cluster together (communities) are shown with the same color on fig. 2, indicating contextual proximity of the terms used. The communities tell us that the local authorities focus quite heavily on patterns of spatial development , unique role of St Petersburg and its attractiveness for migrants, close association between the City and Leningrad region, etc.

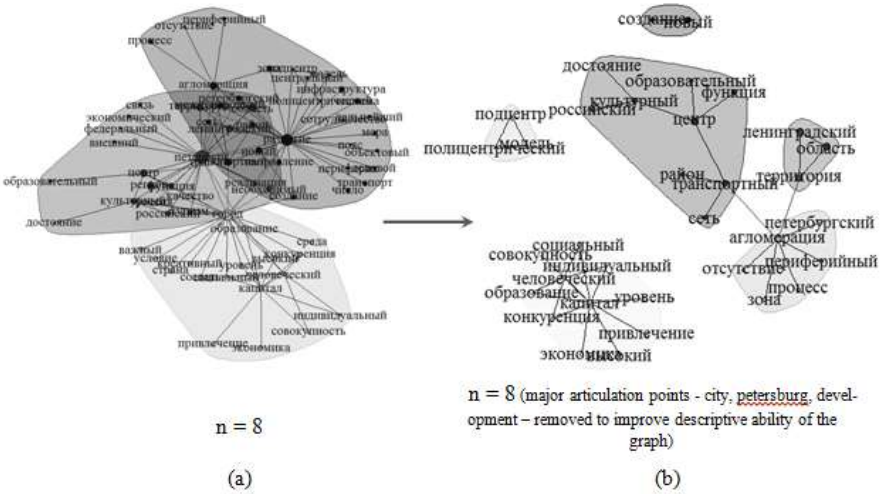


Fig. 2. Graph improvement by managing articulation points (a – initial graph; b – refined graph)

We can also have a further look at which terms colocations are most frequent in each strategy (fig.3). Parameter n tells us how many times the plotted collocations appear. Parameter n is a lower bound of the frequency, that is, collocation «Moscow –

St Petersburg» appears not less than 33 times in different text segments within Tver' region's strategy. Each strategy mentions those regions more frequently which are supposed to be their main partners.

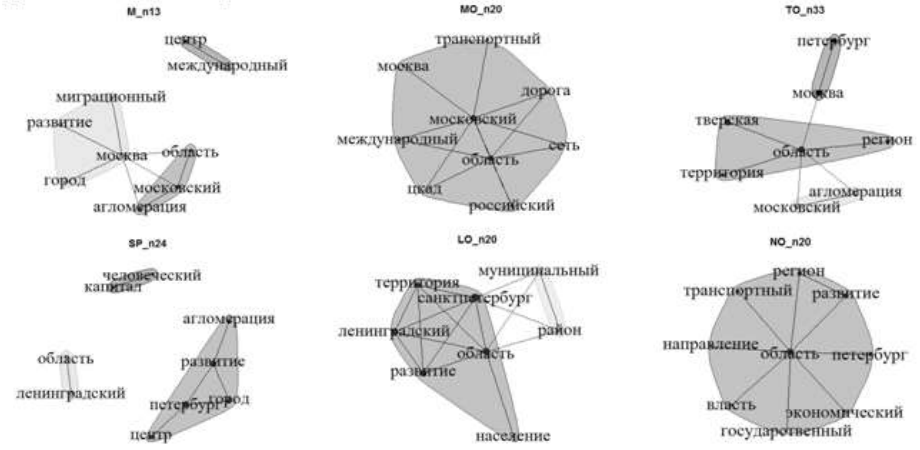


Fig. 3. Most frequent terms colocations

Plots for Moscow region and Leningrad region suggest that the development of the two regions is strongly connected to two largest Russian cities that they surround. For instance, the Leningrad region strategy suggests strong spatial planning on territories adjoined to St Petersburg. The Moscow region strategy repeatedly highlights importance of Moscow and its agglomeration.

5 Concluding remarks

Given the work-in-progress style of our study, it has a number of weaknesses; among them is a lack of the comparative perspective. A solution might be to build a network of text segments (paragraphs) based on the terms shared by them (two-mode network). Alternatively, we would probably arrive at some interesting insights if could put a network on another one to see joint term clusters and terms-outliers [Ermakov A. et al., 2002]. We appreciate the more sophisticated models use phrases or n-grams instead of words and this also could be a possibility to improve our analysis. These basic tools implemented here are only the beginning of possibilities for applications of TNA. It is worth noting that we were not meant to provide an extension of the base technology of computer-supported TNA but suggest an example of its practical implementation in social science.

The text-mining approach combined with graph theory appears to be a valuable method for extracting elite's attitude towards regional integration process from public strategic documentation (allows us to access a large amount of textual material, regional analysis may provide interesting input for text network analysis, etc.). Modeling the data using this method provided us with the specific insights on what local authorities really focus on and how the strategies differ from each other.

References

1. Batagelj V., Mrvar A., Zaversnik M. Network analysis of texts. URL: http://nl.ijs.si/isjt02/zbornik/sdjt02-24bbatagelj.pdf (25.03.2015).
2. Chernyak E., Morenko E., Mirkin B. Conceptual Maps: Construction Over a Text Collection and Analysis. Analysis of Images, Social Networks and Texts Communications in Computer and Information Science Volume №436, 2014, pp.163-168.
3. De Lombaerde P. et al. The Regional Integration Manual: Quantitative and Qualitative Methods. Routledge, London, 2011.
4. Ermakov A., Pleshko V. Informatization and information security of enforcement officials. XI International scientific conference. Conference proceedings, Moscow, 2002, pp. 343-347.
5. Kincharova A. Application of community detection algorithms for sociological investigation of blogs: results of a piloting study. URL: www.hse.ru/data/2013/06/10/1283702757/dzh.pdf (25.03.2015).
6. Manifesto Project Database. URL: https://manifestoproject.wzb.eu/ (25.03.2015).
7. Polanco X., San Juan E. Text data network analysis using graph approach. Vicente P. Guerrero-Bote. I International Conference on Multidisciplinary Information Sciences and Technology, Oct 2006, Merida, Spain. Open Institute of Knowledge, vol. 2, pp.586-592. URL: https://hal.archives-ouvertes.fr/hal-00165964 (25.03.2015).
8. Saeedeh M. et al. A Comparative Study of Word Co-occurrence for Term Clustering in Language Model-based Sentence Retrieval. Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL, pp. 325–328, Los Angeles, California, June 2010. URL: http://www.aclweb.org/anthology/N10-1046 (25.03.2015).
9. Yanchang Zhao. R and Data Mining: Examples and Case Studies. Academic Press, Elsevier, 2012.
10. Zhikharevich B., Zhunda N., Rusetskaya O. Proclaimed and actual priorities of regional and local authorities: approaches to reveal and compare // The Region: Economics and Sociology, 2013, №2, pp. 108 – 132.

A Nonlinear Dimensionality Reduction Using Combined Approach to Feature Space Decomposition

Myasnikov E.V.

Samara State Aerospace University,
Image Processing Systems Institute of the Russian Academy of Science
Samara, Russia
mevg@geosamara.ru

Abstract. In this paper we propose a new combined approach to feature space decomposition to improve the efficiency of the nonlinear dimensionality reduction method. The approach performs the decomposition of the original multidimensional space, taking into account the configuration of objects in the target low-dimensional space. The proposed approach is compared to the approach using hierarchical clustering in the original space and to the approach based on the decomposition of the target space using KD-Tree.

Keywords: dimensionality reduction; multidimensional scaling; Sammon's mapping; hierarchical clustering; KD-Tree

1 Introduction

Dimensionality reduction methods, operating on the principle of preserving the pairwise distances between objects can be used as a means to display multidimensional data in scientific research and in production activities in a number of areas: biology, genetics, sociology, economics, etc. In modern information systems, such methods can be used to create navigation systems for multimedia databases, as well as in interface design to virtual directories. In the field of image analysis and processing nonlinear dimensionality reduction techniques have been applied not only for research but also to solve a number of applied problems: creation of automated systems for image segmentation, thematic mapping from satellite imagery and others.

One of the most well-known dimensionality reduction methods is a method [9], which minimizes the following error of multidimensional data representation

$$\varepsilon = \frac{1}{\sum\limits_{\substack{o_i, o_j \in O \\ i<j}} d(o_i, o_j)} \cdot \sum_{\substack{o_i, o_j \in O \\ i<j}} \frac{\left(d(o_i, o_j) - d^*(o_i, o_j)\right)^2}{d(o_i, o_j)} \qquad (1)$$

where o_i is an object from a set O, $d(o_i, o_j)$ is the distance between objects in the original space, $d^*(o_i, o_j)$ is the distance between objects in the target space.

Iterative procedure that minimizes the error (1) may be based on the following recurrence relations for the coordinates $\mathbf{y}_i$ in the target low-dimensional space (gradient descent):

$$\mathbf{y}_i(t+1) = \mathbf{y}_i(t) + m \cdot \sum_{o_j \in O} \varsigma(o_i, o_j) \tag{2}$$

where (m can be calculated once):

$$m = \frac{2 \cdot \alpha}{\sum\limits_{\substack{o_i, o_j \in O \\ i<j}} d(o_i, o_j)}$$

$$\varsigma(o_i, o_j) = \frac{d(o_i, o_j) - d^*(o_i, o_j)}{d(o_i, o_j) \cdot d^*(o_i, o_j)} \cdot \left(\mathbf{y}_i - \mathbf{y}_j\right) \tag{3}$$

It is worth noting that the performance of the nonlinear mapping has been improved by extending multidimensional data representation error (1) with Bregman divergences [11, 12].

Unfortunately, a significant drawback of the nonlinear dimensionality reduction techniques, operating according to the iteration scheme (similar to that shown above) is the high computational complexity (O(K^2) per iteration, where $K = |O|$ is the number of objects). For this reason a number of techniques with a reduced computational complexity have been proposed [3, 4, 5, 6, 7].

One of the most effective ways to reduce the computational complexity is the hierarchical decomposition of space. After performing such decomposition objects can be considered not individually, but in groups, that allows to speed up the iterative optimization process. So, if all the objects of the original set O are divided into groups $s_j \in S$, then (2) takes the form

$$\mathbf{y}_i(t+1) = \mathbf{y}_i(t) + m \cdot \sum_{s_j \in S} \tilde{\varsigma}(o_i, s_j) \tag{4}$$

$$\tilde{\varsigma}(o_i, s_j) = |s_j| \cdot \frac{d(o_i, s_j) - d^*(o_i, s_j)}{d(o_i, s_j) \cdot d^*(o_i, s_j)} \cdot \left(\mathbf{y}_i - \mathbf{y}_{s_j}\right) \tag{5}$$

where $d(o_i, s_j)$ is the distance from the object to the center of the group s_j in the original space, $d^*(o_i, c_j)$ is the distance from the object to the center of the group in the target space, $\mathbf{y}_{s_j}$ is the coordinates of the center of the group in the target space.

Such decomposition allows to reduce the computational complexity to O(*LK*) per iteration, where $L = |S|$ is the number of groups.

It is obvious that the expression (4) allows to approximate (2) with a certain accuracy that depends on how well objects are divided into groups. For this reason, there is a question about how to better carry out such decomposition. It is understood that the decomposition can be performed in both the source and target space. Approach based on decomposition of the original space, is implemented in a nonlinear method for dimensionality reduction of data using reference nodes [5] proposed by the author of this article. Approach based on decomposition of the target space has been implemented, for example, in [1] (regular decomposition) and [8] (hierarchical decomposition) to solve the problem of drawing graphs to approximate the forces acting on the vertices of the graph.

In this paper we propose a new method that uses the combined approach to feature space decomposition to improve the efficiency of the nonlinear dimensionality reduction method. The proposed method is based on the nonlinear method for dimensionality reduction of data using reference nodes [5] that performs the decomposition of the original space, and complemented with the additional control on the configuration of objects in the target space. The proposed method is compared to the method based on the decomposition of the original space, and to the method based on the decomposition of the target space, both in terms of the quality of the mapping, as well as in terms of the operating time.

The paper is organized as follows. The next section is devoted to the description of the known approaches used in the present study, and consists of two subsections. The third section is devoted to the description of the proposed method. The fourth section presents the results of numerical experiments. At the end of the paper the conclusion is given.

2 Brief Description of the Methods Used in the Study

2.1 Nonlinear Method for Dimensionality Reduction of Data Using Reference Nodes

Approach based on the decomposition of the original space, is implemented in a nonlinear method for dimensionality reduction of data using reference nodes [5]. This method consists of four steps, briefly described below.

The inputs to the method are the feature vectors describing the objects in the original multidimensional space. The outputs of the method are the feature vectors describing the objects in the target low-dimensional space.

Step 1: Construction of a hierarchy of clusters. At the first stage of the method we make a hierarchical partition of the initial set of objects into clusters (subsets of objects with similar characteristics in the original feature space). When we refer to the hierarchy of clusters we mean tree-like data structure, the root of which is a cluster of

top-level, and each vertex-cluster contains either subclusters or the objects of the original set.

Step 2: Initialization of the coordinates of objects in the target space. Initialization of the coordinates of objects in the target space can be performed in various ways, e.g., using random values, the results of the PCA method, etc.

Step 3: Construction of the lists of reference nodes. At the third stage, for each object of the original set the lists of reference nodes are built. The reference node of the object o refers to a certain object $o_i \neq o$ or group of objects $\{o_i\}_{i=1..N}$ that possesses close characteristics in the original multidimensional space and considered as a single object.

Let o be an object for which we need to form a list S of the reference nodes and let C be an arbitrary cluster of the hierarchy. By using the hierarchy of clusters obtained at the first stage the construction of the reference nodes list may be implemented using the following recursive algorithm.

1. If the cluster C satisfies the decomposition criterion with respect to the considered object and subclusters $C_i \in C, i = 1..N$ are contained in cluster C, then one must apply this recursive algorithm to each of the subclusters $C_1, C_2, .., C_N$.
2. If the cluster C satisfies the decomposition criterion with respect to the considered object and objects $o_i \in C, i = 1..N$ are contained in cluster C, then one must:
   (a) add all objects that close to the considered object to the list of reference nodes S;
   (b) distinguish the set V of objects situated at a distance from the considered object as an incomplete cluster;
   (c) in the case when this set V is nonempty, add it to the list of reference nodes.
3. If the cluster C doesn't satisfy the decomposition criterion with respect to the considered object then add it to the list of reference nodes S.

Decomposition criterion used in this paper is slightly different from the criterion used in [5] and restricts (by the value δ) the estimation of the angle at which the hypersphere of the cluster is observed from the object o in the original space.

Figure 1 illustrates this process. The cluster, which is observed from an object o at an angle α will be divided into three clusters if $\alpha \geq \delta$. The cluster, which is observed at an angle β will be considered as a whole if $\beta < \delta$.

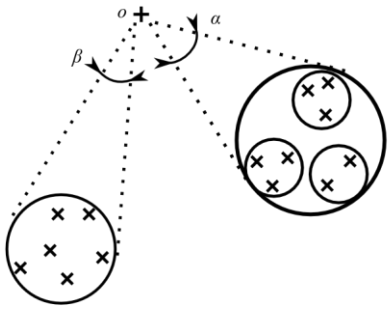


Fig. 1. Illustration to the construction of a list of reference nodes

Step 4: Iterative optimization procedure. At the final stage of the method iterative procedure is performed, which allows to clarify the position of objects in the target low-dimensional space. The work of the iterative procedure is based on (4), (5).

2.2 Nonlinear Dimensionality Reduction Method Using KD-Tree for the target space decomposition

In contrast to the approach discussed above decomposition of the target space inevitably leads to the need for periodic renewal or adjustment of the structure by which the decomposition is performed. The reason for this consists in changing the coordinates of objects in the target space as you perform the optimization process.

In this paper, KD-trees are used for the decomposition of the target space. Tree construction is performed at each iteration of the optimization process using the following recursive procedure.

1. Calculation of the characteristics of the current tree node (average coordinates in the source and target space, boundaries of the node in the target space).
2. In the case when the node contains only one object, then exit.
3. Splitting the node into the two child nodes using the perpendicular to the most elongated boundary so that the number of objects in the child nodes is approximately the same.
4. Perform this procedure for the newly formed child nodes.

Constructed tree is used to perform the next iteration of the optimization process. For each object, the sum (4) is calculated by traversing the tree as follows.

1. If the decomposition criterion is not satisfied and the objects contained in the current node of the tree can be considered as a group, then the next term in (4) is calculated using (5).
2. Otherwise decomposition criterion is satisfied, and both subtrees (child nodes) traversed similarly.

A decomposition criterion used in the tree traversal is similar to the above, and restricts (by the value δ) the estimation of the angle at which the minimum bounding rectangle of the node is observed from the position of the object o in the target space.

Figure 2 illustrates this process. Different decomposition levels are shown using different lines: the first level is shown by the solid line, the second level is shown by the dash-dot line, the third level is shown by the dashed line. A group of four objects (minimum bounding rectangle is shown by dots) will be devided into two groups if the angle β at which the minimum bounding rectangle is observed from the position of the object o is greater than the decomposition angle δ.

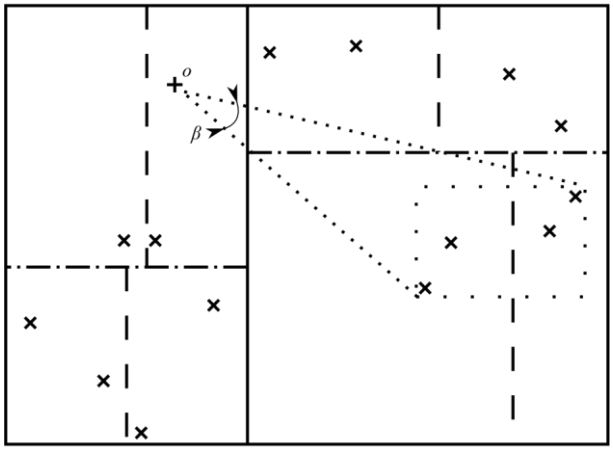


Fig. 2. Illustration of decomposition in the target space using KD-Tree

It is worth noting that this algorithm is executed at each iteration for each object, so to improve the efficiency the constructed tree is traversed not recursively, but iteratively using special pointers.

Note that some additional experimental results on comparison of methods described in this section can be found in [6].

3 Proposed Method

As noted above the expression (4) approximates (2) with some accuracy. The error of approximation for some object o and the set s of individual objects can be represented using (5) as follows:

$$\rho(o,s) = \left\| \sum_{o_j \in s} \frac{d(o,o_j) - d^*(o,o_j)}{d(o,o_j) \cdot d^*(o,o_j)} \cdot (\mathbf{y} - \mathbf{y}_j) - \tilde{\boldsymbol{\varsigma}}(o,s) \right\|. \quad (6)$$

Unfortunately, the use of the described above approaches can cause significant (in some cases unbounded) approximation errors as in the first case the decomposition criterion takes into account only the characteristics of nodes in the original space and in the second case criterion takes into account only the characteristics of nodes in the target space. The way out of this situation may be a combination of the original space decomposition with the additional control on the configuration of objects in the target space.

In this paper, for this purpose a nonlinear method for dimensionality reduction of data using reference nodes is complemented by the preservation and analysis of the boundaries of nodes in the target space. The proposed method consists of the same steps as the base method [5]:

1. Construction of a hierarchy of clusters
2. Initialization of objects coordinates in the target space
3. Construction of reference nodes lists
4. Iterative optimization procedure

At the first step the difference between the proposed method and the base one is in the structure of nodes in the cluster tree. In the proposed method each node contains information about its boundaries in the target low-dimensional space. The second and the third steps are the same as in the base method. At the fourth step iterative optimization procedure is different in the proposed method.

At the each step of the optimization process for each object $o_i \in O$ with the corresponding reference nodes list S_i it recalculates the coordinates $\mathbf{y}_i$ in the target low-dimensional space. To do this according to (4) it iterates through the list of reference nodes S_i and calculates the approximated term for a given node $s_j \in S_i$ using the following recursive algorithm.

1. If the given node s_j is the object $o_j \in O$ of the initial set of objects, then the exact value of the corresponding term is calculated using (3).
2. If the decomposition criterion is not satisfied for given node s_j in the target low-dimensional space and the objects contained in the given node can be considered as a group, then the approximated term is calculated using (5).
3. Otherwise the decomposition criterion is satisfied for the given node s_j in the target low-dimensional space, and all subtrees of the given node in the cluster tree traversed similarly using this algorithm.

4 Experimental Study

In this paper all the studied methods have been implemented in C++ language in the integrated development environment Borland Turbo C++ 2006 Explorer. The PC

based on Intel Core i5-3470 CPU 3.2 GHz has been used in the experiments. Corel Image Features Data Set (http://archive.ics.uci.edu/ml/databases/CorelFeatures/CorelFeatures.data.html) has been used in the experiments as an initial dataset. This dataset contains a set of features, calculated from the digital images from the Corel image collection (http://corel.digitalriver.com/). In particular, the following feature sets have been used.

Feature set 1 contains color moments [10]. Three features were calculated for each color component: mean, standard deviation, and skewness. The dimensionality of the feature space is 9.

Feature set 2 contains color histograms [13] constructed in the HSV color space. Color space was divided into 8 ranges of H and 4 ranges of S. The dimensionality of the feature space is 32.

Feature set 3 contains texture features based on co-occurrence matrices [2]. Four co-occurrence features (second angular moment, contrast, inverse difference moment, and entropy) were computed in four directions (horizontal, vertical, and two diagonal). The dimensionality of the feature space is 16.

Fragments containing features information for the required number of images has been selected from these feature sets. Then for sets 1 and 3, the identity component has been added to the feature information and normalization has been performed.

To evaluate effectiveness of the methods the value of multidimensional data representation error (1) has been calculated and the average per iteration execution time of the optimization procedure has been measured.

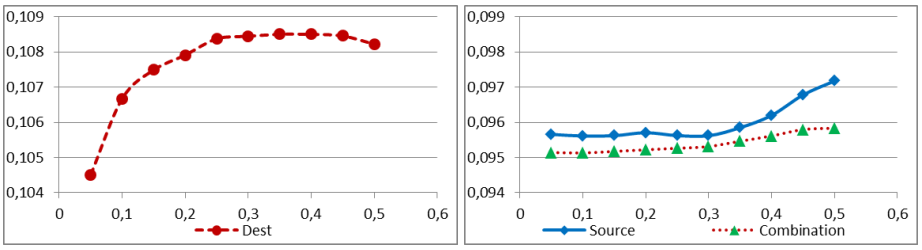
Work of the methods stopped when the rate of decreasing the error slowed down (i.e. the relative decrease in error for ten iterations does not exceed 0.05). In all cases, the dimension of the target space has been set equal to two (two-dimensional data mapping). Initialization is performed using the PCA method.

Some results for the feature set 1 are shown in Fig. 3-6.

Fig. 3 and 4 show the dependence of the qualitative and temporal characteristics on the decomposition angle δ at which the algorithm moves to the child nodes of the corresponding hierarchical structure (cluster tree or KD-tree).

As can be seen from these results, the large values of the angle δ leads to the expected deterioration of the mapping quality due to a coarser approximation, which is reflected in the higher values of the multidimensional data representation error ε (see Fig. 3). The time it takes to perform a single iteration, decreases with increasing angle δ (see Fig. 4), due to the smaller number of processed nodes of the corresponding hierarchical structure during the construction of approximations.

It should be noted that the approach using the original space decomposition provides less value of multidimensional data representation error ε than the approach using the decomposition of the target space.



(a) (b)

Fig. 3. Dependence of the multidimensional data representation error ε on the decomposition angle δ (in fractions of π radian)

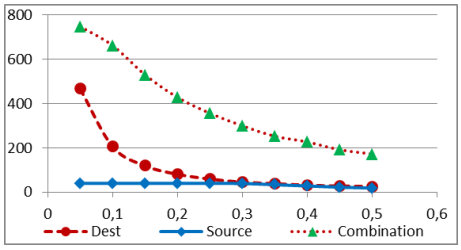


Fig. 4. Dependence of the average execution time per iteration (in ms) on the decomposition angle δ (in fractions of π radian)

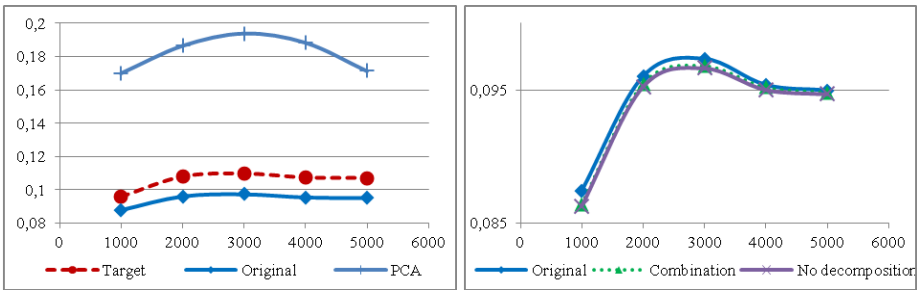
At the same time the average execution time per iteration is much smaller when using the approach with the decomposition of the original space in the range $\delta \in [0.05\ \pi, 0.2\ \pi]$.

The use of the combined approach allows slightly reduce the multidimensional data representation error in comparison with the approach based on the decomposition of the original space. At the same time the average time per iteration using a combined approach is significantly larger than when using other considered approaches.

Dependencies of quality and temporal characteristics on the number of objects shown in the figures 5 and 6 confirm said above. The approach using the original space decomposition provides less value of multidimensional data representation error ε than the approach using the decomposition of the target space (fig. 5a). The quality of mapping, measured by the multidimensional data representation error ε, formed using decomposition in the original feature space is only slightly inferior to the combined approach (fig. 5b). The quality of mapping formed using combined approach is virtually identical to the quality obtained with the base method without decomposition.

At the same time approach, using the decomposition of the original space, allows us to construct mapping much faster than the other considered approaches (Fig. 6).

Figure 5a also shows the multidimensional data representation error obtained after initialization of objects coordinates in the target low-dimensional space using the PCA.



(a) (b)

Fig. 5. Dependence of the multidimensional data representation error ε on the number of objects

Note that the experiments performed on other datasets described above, show similar results.

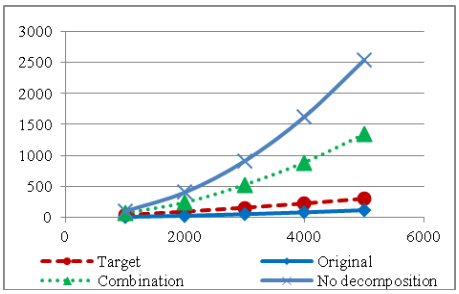


Fig. 6. Dependence of the average execution time per iteration (in ms) on the number of objects

5 Conclusion

In this paper, we conducted a study of different approaches to the decomposition of space in terms of quality and speed of the nonlinear dimensionality reduction method. The study showed that for the presented data sets decomposition of the original space is more preferable than the decomposition of target space in terms of both quality and time. Combining the decomposition of the original space with the decomposition of target space has allowed slightly improve quality.

Acknowledgments. This work was financially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years and by the Russian Foundation for Basic Research, project № 15-07-01164 -a.

6 References

1. Fruchterman, T., Reingold, E.: Graph Drawing by Force-directed Placement. Software – Practice and Experience, vol. 21, no. 11, pp. 1129-1164 (1991)
2. Haralick, R.M., Shanmugam, K., Dinstein, I.: Texture Features for Image Classification. IEEE Trans. on Sys. Man. and Cyb. SMC-3(6) (1973)
3. Lee, R.C.T., Slagle, J.R., Blum, H. A.: Triangulation Method for the Sequential Mapping of Points from N-Space to Two-Space. IEEE Transactions on Computers, vol. 26, no. 3, pp. 288-292 (1977)
4. Morrison, A., Ross, G., Chalmers, M.: Fast Multidimensional Scaling through Sampling, Springs and Interpolation. Information Visualization, vol. 2, pp.68-77 (2003)
5. Myasnikov, E.V.: A Nonlinear Method for Dimensionality Reduction of Data Using Reference Nodes. Pattern Recognition and Image Analysis, vol. 22, no. 2, pp. 337–345 (2012)
6. Myasnikov, E.V.: The Choice of a Method for Feature Space Decomposition for Non-Linear Dimensionality Reduction. Computer optics, vol. 38, no. 4, pp. 790-797 (2014)
7. Pekalska, E., de Ridder, D., Duin, R.P.W., Kraaijveld, M.A.: A New Method of Generalizing Sammon Mapping with Application to Algorithm Speed-up. Proc. ASCI'99, 5th Annual Conf. of the Advanced School for Computing and Imaging, pp. 221-228. Heijen, Netherlands (1999)
8. Quigley, A., Eades, P.: FADE: Graph Drawing, Clustering, and Visual Abstraction. Proceedings of the 8th International Symposium on Graph Drawing, pp. 197–210 (2001)
9. Sammon, J.W. Jr.: A Nonlinear Mapping for Data Structure Analysis. IEEE Transactions on Computers, vol. C-18, no. 5, pp. 401-409 (1969)
10. Stricker, M., Orengo, M.: Similarity of color images. In Proc. SPIE Conf. on Vis. Commun. and Image Proc (1995)
11. Sun, J., Crowe, M., Fyfe, C.: Extending metric multidimensional scaling with Bregman divergences. Pattern Recognition, 44 (5), pp. 1137–1154 (2011)
12. Sun, J., Fyfe, C., Crowe, M.: Extending Sammon mapping with Bregman divergences. Information Sciences (2011)
13. Swain, M., Ballard, D.: Color indexing. International Journal of Computer Vision, 7(1) (1991)

Studies of Anthropometrical Features using Machine Learning Approach

The Long Nguyen[1], Thu Huong Nguyen[1], and Aleksei Zhukov[2]

[1] Irkutsk State Technical University, Irkutsk, Russia
{thelongit88,thuhuongyb}@gmail.com
[2] Irkutsk State University, Irkutsk, Russia
zhukovalex13@gmail.com

Abstract. In this article we propose the novel approach to measure anthropometrical features such as height, width of shoulder, circumference of the chest, hip and waist. The sub-pixel processing and convex hull technique are used to efficiently measure the features from 2d image. The SVM technique is used to classify men and women based on measured features. The results of real data processing are presented.

Keywords: anthropometrical features, SVM, image processing, body size, support vector machine

1 Introduction

Development of efficient methods of image recognition is an important field of computer sciences. The theory of such methods uses the machine learning methods enabling an automatic scenes analysis.

Recently, automatic detection and feature extraction of the human bodies is widely used in many fields such as non-contact measurements of body size [1], the construction of 3D models of humans [2,3], [12], the analysis of human action [4] and pose estimation [5].

In this paper we propose a system of non-contact anthropometrical features extraction based on the analysis of digital images of humans. Background subtraction is used to detect contours data. Edge detection operators are employed for contour detection (silhouette). This approach combines with the algorithm of face recognition, background subtraction, the detection of skin color and contour analysis. The convexity hull defects are also used for anthropometrical features extraction, then the support vector machine (SVM) will be applied for gender classification.

Cluster analysis methods allow to analyze the space of feature vectors which obtained sizes to find out the clusters corresponding to the most characteristic anthropometric features of people, that can help to develop recommendations for the clothing industry, and also can be used in other fields of natural science, such as axiology.

The novelty and motivation of our approach is usage of the image processing techniques for efficient measurement of anthropometric features for further

classification with machine learning techniques. As an application male/female classification task was examined.

2 Related Work

All detection methods of body parts can be divided into two main categories: model based and learning methods. Model based approaches are the "top-down" methods, which are used to obtain preliminary information, e.g. about the shapes of the human body in various poses [6]. In machine learning, one apply the principle of learning data to extract useful knowledge from the data.

In [7] an effective approach for silhouette detections is presented. It is used to represent the contour curves of the form of the human body in the form of 8-connected chain code Freeman [8]. The classic algorithm for background subtraction and the Canny edge detector are used to get the silhouette. Then the contour data is divided into a number of segments. Thus, special rules for measuring the differences between the directions of the segments are used for feature extraction points. This approach has several disadvantages including high sensitivity noise.

The contour of body is also divided into the parts in [9]. The convexity and curvature boundaries of each segment of the contour are used to define the body parts.

In [6] the approach for body parts segmentation in noisy silhouette images was proposed. The weighted radial histogram of distances and directions are used as features. Authors also used Hidden Markov Model (HMM) to model the silhouette as a sequence of body parts. A general model is trained using shape context features extraction from the labeled synthetic data.

In [10] authors use the segmentation method based on sub-pixel data processing and relatively large regions or segments. It also detects a significant upper and lower extremities from these segments, identifies potential head and torso positions at the same time. Both modules combine these parts from partial images configuration of the body, by applying global constraints to recover full body configurations.

3 Anthropometric Data Extraction

In the first part of the approach we detect the silhouette using the background subtraction method. The morphological operations including erosion, dilation, opening, and closing are performed to reduce noise and to smooth the silhouette contour. A convex hull is created around the silhouette frame, and convexity defects are used as the features for analysis. Individually these features are the start and end convexity defect points and convexity defect locations.

To find the convex hull of a 2D set points, we use the Sklansky's algorithm [14] that has $\mathcal{O}(N \log N)$ complexity.

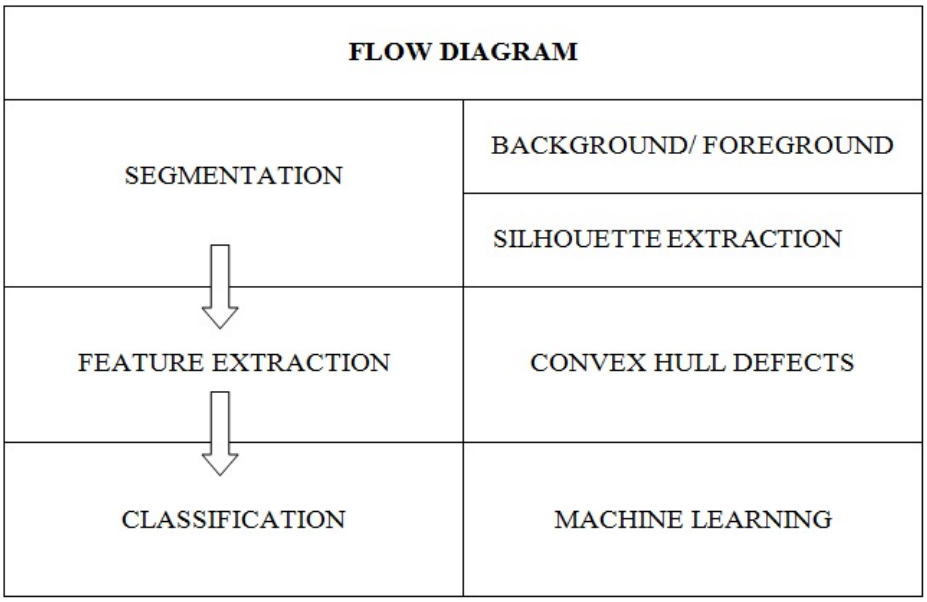


Fig. 1. Flowchart of machine learning with anthropological features extraction.

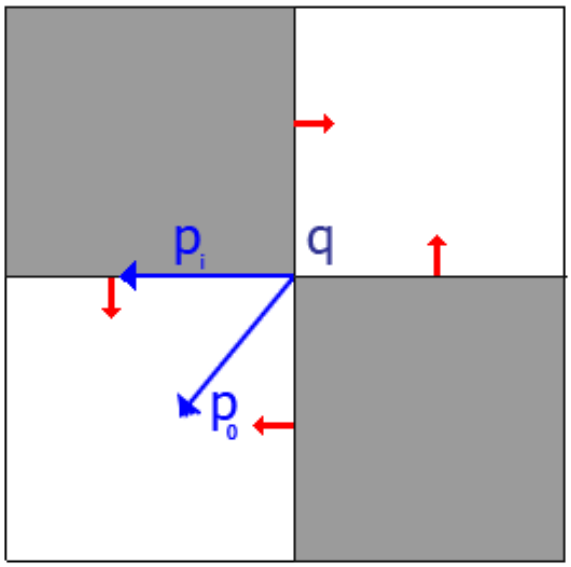
Background Subtraction Algorithm: The method is based on a comparison between the two images, namely foreground (FG) and background images (BG). The scene of background image is obtained when there is no object motion [9,10]. Call $FG(x, y)$ is the intensity values of pixel with coordinates (x, y) in the foreground image, belonging to the interval [0.255]. $BG(x, y)$ is the intensity values of pixel with coordinates (x, y) of the background image. A pixel with coordinates (x, y) in the foreground image of the dominant component if it satisfies: $|FG(x, y) - BG(x, y)| > T$

Where, $T(x, y)$ is the threshold value, which enables initializing by the value was determined. Pixels with label 1 are an object if $|FG(x, y) - BG(x, y)| > T$ or not object (value 0) if $|FG(x, y) - BG(x, y)| < T$.

Sub-pixel Processing: To achieve better accuracy sub-pixel processing we employed sub-pixel processing. Conventional image processing is performed in units of 1 pixel, while the sub-pixel processing method performs position detection in units down to 0.01 pixels. This enables high accuracy position detection, expanding the application range to precise part location and dimension measurement.

As described in [18] sub-pixel procedure iterates to find the sub-pixel accurate location of corners, as shown on the fig. 2.

Sub-pixel accurate corner locator is based on the observation that every vector from the center q to a point p located within a neighborhood of q is orthogonal



.

Fig. 2. Sub-pixel corner location principle illustration (Red arrows mean gradient direction).

to the image gradient at p subject to image and measurement noise. Consider the expression:

$$\epsilon_i = {DI_{p_i}}^T \cdot (q - p_i)$$

where DI_{p_i} is an image gradient at one of the points p_i in a neighborhood of q. The value of q is to be found so that ϵ_i is minimized. A system of equations may be set up with ϵ_i set to zero:

$$\sum_i (DI_{p_i} \cdot {DI_{p_i}}^T) - \sum_i (DI_{p_i} \cdot {DI_{p_i}}^T \cdot p_i)$$

where the gradients are summed within a neighborhood of q. Calling the first gradient term G and the second gradient term b gives:

$$q = G^{-1} \cdot b$$

.

The algorithm sets the center of the neighborhood window at this new center q and then iterates until the center stays within a set threshold. In our work we use OpenCV library to perform sub-pixel corner detection [18].

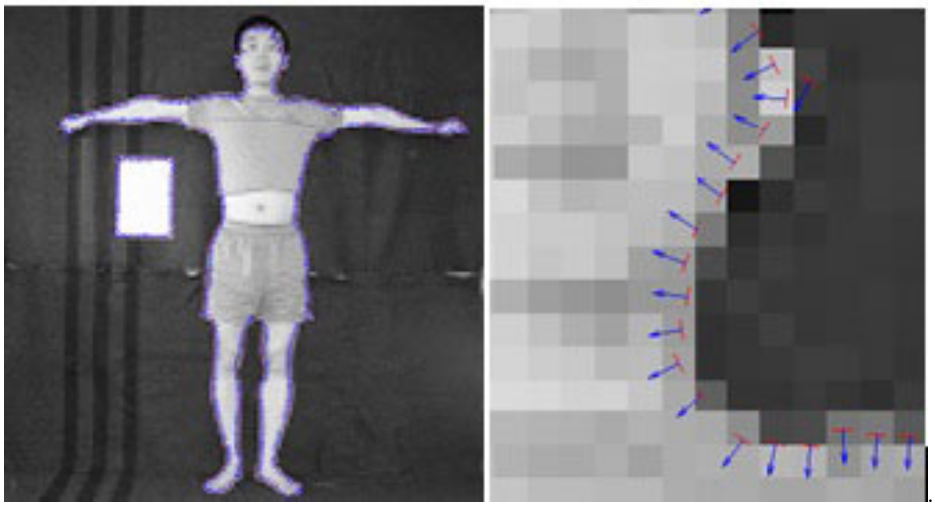


Fig. 3. Sub-pixel processing result.

In our approach sub-pixel helps us to find corners. It uses the dot product technique to refine corners. The function works iteratively, refining the corners until the termination criteria is reached. Most sub-pixel algorithms require a good estimate of the location of the feature. Otherwise, the algorithms may be attracted to the noise instead of desired features.

Features Extraction based on Convexity Hull Defects:

In our approach the human body is described by convexity defect triangles. The bodies are represented by triangles with three coordinates called the convexity defect start $(x_{ds}; y_{ds})$, defect end $(x_{de}; y_{de})$, and defect position points $(x_{dp}; y_{dp})$, labeled as P_1,P_2 and P_3 respectively.

We applied the convex hull method to extract contours and obtained a lot of convexity defects, including areas with very small depth, even a value of 0 - these are not areas containing features to be extracted. So, we determined the area of interest which contains the parts of the body is area that has $depth > 50$. That depth value was obtained empirically. Thus we got 5 convex regions correspond to conditions. Then we determined of the human body, which contains three parts of chest, waist and hips. We continue applying the convex hull to locate the waist.

Once we got the coordinates of the points determined, we perform calculation of the distances between points in pixels, and finally converted the measurements into cm. The convexity defect of triangle is determined based on: $(x_{ds}; y_{ds})$, $(x_{de}; y_{de})$, and $(x_{dp}; y_{dp})$. A convexity defect is presented wherever the contour of the object is away from the convex hull drawn around the same contour. Convexity defect gives the set of values for every defect in the form of vector.

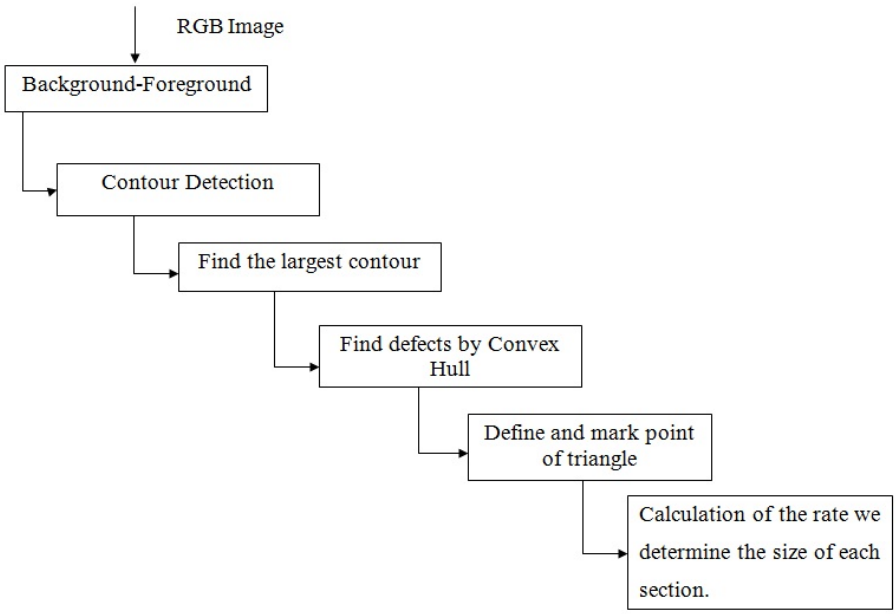


Fig. 4. Flowchart feature extraction based on convex hull.

This vector contains the start and end point of the line of defect in the convex hull. These points indicate indices of the coordinate points of the contour. These points can be easily retrieved by using start and end indices of the defect formed from the contour vector. Convexity defect also includes index of the depth point in the contour and its depth value from the line. In fact, the person may be represented by many triangles point defects, it is piecewise convex. However, in this approach, we are interested in two triangles have the biggest area - It obviously corresponds to leg-armpit-arm triangle, and it includes the location of the parts we need to calculate interest: chest, waist, hips.

Finally, we obtain the coordinates of the points on the body.

Therefore in this paper we propose a simpler and cheaper system comparing with other systems, see e.g. [16]. In our experiments we used single digital camera and A4 sheet ($210 \times 297mm$) for calibration. Source images for the method must be captured in special way: with given background and calibration sheet, human must stand straight with arms stretched.Three dimensions (chest circumference, waist circumference, hip circumference) were selected because of their relevance to clothing sizing, and human classification which was the main purpose of our system. Table 1 shows some results of measurements(from 50 measurements of people in the database) sizes of human body using convex hull method comparing with manual method.

Table 1. Results of measurement sizes of human body using convex hull method.

BODY SIZES	***Manual method***	***Convex hull method***
Chest	87.98 cm	88.12 cm
Waist	67.95 cm	68.05 cm
Hip	90.52 cm	91.68 cm
Chest	88.64 cm	89.02 cm
Waist	66.61 cm	67.13 cm
Hip	93.58 cm	94.93 cm
Chest	87.22 cm	88.15 cm
Waist	67.19 cm	67.96 cm
Hip	89.36 cm	89.01 cm
Chest	88.16 cm	89.46 cm
Waist	65.64 cm	66.42 cm
Hip	92.17 cm	93.02 cm
Chest	90.96 cm	91.23 cm
Waist	71.44 cm	70.82 cm
Hip	93.56 cm	94.19 cm
Chest	86.94 cm	87.12 cm
Waist	67.68 cm	68.12 cm
Hip	85.28 cm	84.56 cm

Measurements errors are mostly caused by camera resolution or non tight clothing and noises of environment near by the object. In our case, we used a basis phone camera of model Samsung Galaxy S4 with resolution 13 Mega Pixel. We recommended using high-quality resolution camera with flash opened during the time capture photos and people should wear tight clothes body to reduce maximum noises as measurements errors. In addition, we have also performed averaging over several measurements to reduce measurement and calibration errors.

The circumferential measures were generated by approximating the shape of the respective body part. For example, neck circumference was approximated with the elliptical shape. The major and minor axes lengths were obtained from the front and side views. The chest circumference was determined by approximating the shape as a combination of a rectangle and an ellipse, using the method with formula mentioned in [16].

4 Anthropometric Data Analysis

Base on method anthropometric features extraction which mentioned in previous section, we collected sizes of human body parts. We have a train set contains sizes of 50 people sizes (25 men and 25 women) and test set from 18 people (10 men and 8 women). Each measurement contains 3 feature for each object: chest, waist and hip circumferences.

4.1 Classification of Men/Women using Support Vector Machine(SVM)

To solve the problems of classification we used well-known machine learning method – support vector machine (SVM), proposed by Vladimir Vapnik [11]. We choose SVM as a one of the most effective algorithm which have many real world applications [17].

Our goal is use a support vector machine for gender classification based on anthropometric data. As a features we use three human body parameters: chest, waist, hip circumferences.

The LibSVM library [15] was used as Support Vector Machine implementation. The following main parameters of SVM were chosen. Radial basis function with $\gamma = 0.333$ parameter was used as a kernel function. SVM model parameters Gamma (γ) and cost (C) were obtained empirically. Gamma parameter is needed for all types of kernels except linear. Constant C is an regularization term in the Lagrange formulation. We will use the supplied parameter ranges (C - cost, γ - gamma), using the train set. The range to gamma parameter is between 0.000001 and 0.1. For cost parameter the range is from 0.1 until 10. It's important to understanding the influence of this two parameters, because the accuracy of an SVM model is largely dependent on the selection them. For example, if C is too large, we have a high penalty for non separable points and we may store many support vectors and over-fit. If it is too small, we may have an under-fitting. The results of this algorithm is shown in Fig.5. Obtained test classification error for current dataset is 20%.

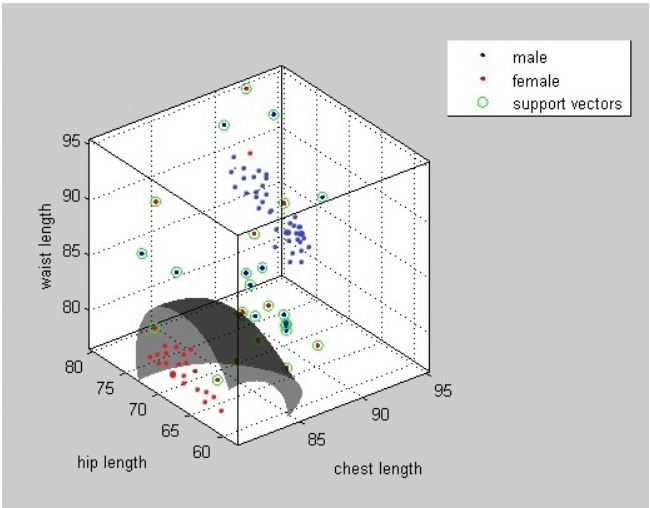


Fig. 5. The result of classification by SVM. Blue dots: men, red dots: women, green dots: support vectors.

5 Conclusion

Human classification is an useful application for many future scenarios of human-computer interaction. Our approach is presented in this paper describes the classification of people based on the anthropometrical features using machine learning approach. Proposed approach allows to extract form images a number of anthropometrical features, including the length of arms, chest width, shoulder width, hips, leg length. In this article we propose solution of male/female classification task by image based on support vector machine. Obtained test error is sufficiently large but in feature work we are going to examine more classifiers to choose best one and also use bigger datasets to reduce misclassification rate. Based on these results, we hope to build solution that can help to solve tasks of the clothing industry.

References

1. Lin, Y.-L., Wang, M.-J.J.: Automatic Feature Extraction from Front and Side Images. In: International conference on Industrial Engineering and Engineering Management, pp. 1949–1953. IEEE, Singapore (2008)
2. Lin, Y.-L., Wang, M.-J.J.: Constructing 3D Human Model from 2D Images. In: International conference on Industrial Engineering and Engineering Management, pp. 1902–1906. IEEE, Xiemen (2010)
3. Lin, Y.-L., Wang, M.-J.J.: Constructing 3D Human Model from Front and Side Images. Expert Systems with Applications, Vol. 39, No. 5, 5012–5018 (2012)
4. Rahman, S.A., Cho, S.-Y., Leung, M.K.H.: Recognizing Human Actions by Analyzing Negative Spaces. IET Computer Vision, Vol. 6, No. 3, 197–213 (2012)
5. Pickup, D., Sun, X.-F., Rosin, P.L., Martin, R.R., Cheng, Z.-Q., Lian, Z., Aono, M., BenHamza, A., Bronstein, A., Bronstein, M., Bu, S., Castellani, U., Cheng, S., Garro, V., Giachetti, A., Godil, A., Han, J., Johan, H., Lai, L., Li, B., Li, C., Li, H., Litman, R., Liu, X., Liu, Z., Lu, Y., Tatsuma, A., Ye, J.: SHREC'14 Track: Shape Retrieval of Non-Rigid 3D Human Models. In: Proc. 7th of Eurographics Workshop on 3D Object Retrieval, pp. 101–110. Eurographics, Strasbourg (2014)
6. Barnard, M., Matilainen, M., Heikkila, J.: Body part segmentation of noisy human silhouette images. In: Multimedia and Expo, IEEE International Conference on, pp. 1189–1192. IEEE, Hannover(2008)
7. Jiang, L., Yao, J., Li, B., Fang, F., Zhang, Q., Meng, M.Q.-H.: Automatic Body Feature Extraction from Front and Side Images. Journal of Software Engineering and Applications, Vol. 5, No. 12, 94–100 (2012)
8. Freeman, H.: On the Encoding of Arbitrary Geometric Configuration. IRE Transactions on Electronics Computers, Vol. EC-10, No. 2, 264–268 (1961)
9. Mittal, A., Zhao, L., Davis, L.S.: Human Body Pose Estimation using Silhouette Shape Analysis. In: Advanced Video and Signal Based Surveillance, 2003. Proceedings. IEEE Conference on, pp. 263–270. IEEE, USA (2003)
10. Mori, G., Ren, X., Efros, A.A., Malik, J.: Recovering Human Body Configurations: Combining Segmentation and Recognition. In: Computer Vision and Pattern Recognition. Proceedings of the 2004 IEEE Computer Society Conference on, Vol. 2, pp. 326–333. IEEE, Washington (2004)

11. Cortes, C., Vapnik, V.: Support-Vector Networks. Machine Learning, Vol. 20, No. 3, 273–297 (1995)
12. Han, E.: 3D Body-Scanning to Help Online Shoppers find the Perfect Clothes fit. The Sydney Morning Herald. National Newspaper (Australia) (2015)
13. MacQueen, J. B.: Some Methods for Classification and Analysis of Multivariate Observations. In: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability 1, pp. 281–297. University of California Press, USA (2009)
14. Sklansky, J.: Finding the Convex Hull of a Simple Polygon. Pattern Recognition Letters, Vol. 1, No. 2, 79–83 (1982)
15. Chang, C.-C., Jen, C.: LIBSVM - A Library for Support Vector Machines, `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
16. Kohlschetter, T.: Human Body Modelling by Development of the Automatic Landmarking Algorithm. Technical Report No. DCSE/TR-2012-11 (2012)
17. Wang, Lipo, ed. Support Vector Machines: theory and applications. Vol. 177. Springer Science & Business Media (2005).
18. OpenCV 2.4.11.0 documentation, `http://docs.opencv.org/modules/imgproc/doc/feature_detection.html`

An Approach to Multi-Domain Data Model Development Based on the Model-Driven Architecture and Ontologies

Denis A. Nikiforov, Igor G. Lisikh, Ruslan L. Sivakov

Centre of Information Technology, Ekaterinburg, Russia

Denis.Nikiforov,Igor.Lisyih,Ruslan.Sivakov@centre-it.com

Abstract. To date, there are many diverse data representation technologies (EDIFACT, XML, JSON, CSV, relational model, NoSQL). Transition to new technologies or the integration of information systems based on different technological stacks is a complex and expensive process. Platform-independent models take an important role in this process. The structure of such a model is described in this article. However, given the data model has been created at the junction of different domains, it may be not enough. In such case, a one more step of abstraction and a movement to the computation-independent model is required. The authors propose to create it in an ontological form.

Keywords: ontology, model-driven architecture, platform-independent model, computation-independent model, data model

1 Introduction

When developers create complex, heterogeneous, distributed information systems, they face a number of questions: which technologies for data storage and transfer to choose, how to ensure data-model consistency across different participants of information exchange, and how to simplify future maintenance of the system under development. In order to assist the developers of such information systems, the OMG consortium has developed the model-driven architecture [1]. This architecture considers an information system as a set of models and the development process is transformation of some models into others.

The architecture is not a technical specification. It describes only basic principles. Specifically, it describes platform-dependent, platform-independent, computation-independent models without governing the structures thereof. In other words, developers choose existing or create new modeling languages (metamodels) for each specific information system. The NIEM specification [2] is an example of such a metamodel. It describes a platform-independent metamodel and a platform-dependent metamodel as well as the rules for transforming instances of the former into instances of the latter. We have also developed a similar platform-independent metamodel as well as the rules for transforming its instances into platform-dependent models (an XML schema and an ER model).

As a rule, the platform-independent model solves a considerable number of problems related to the development and maintenance of an information system without the need to provide the third level of abstraction (in the computation-independent model). However, if information exchange is sufficiently complex and covers multiple domains, then the computation-independent model is required.

The purpose hereof is to describe the problems occurring in the data-modeling process that could indicate the need for the computation-independent model. We also show that such a model is actually an ontology. We hope that our experience will help the developers of complex, heterogeneous, distributed information systems to take correct architectural decisions. We will try to answer the question whether and why ontologies are needed for data modeling.

2 The Platform-Dependent Data Model

Let us consider the following example. Some organization produces shrimps at its aquafarm and sells these to another organization. If the organizations reside in different countries, then this process additionally involves customs, sanitary and veterinary, transport and other control authorities. All the stages of this process are carried out together with the storage, transfer, and processing of data on the consignor, consignee, forwarder, transport vehicle, cargo, and other objects (Fig. 1).

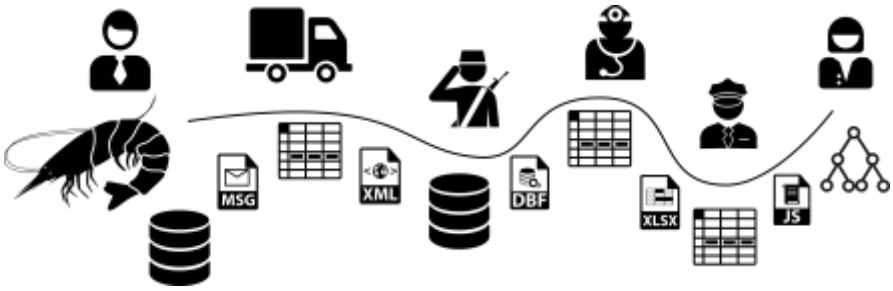


Fig. 1. The process of delivering the commodity and its accompanying data

The participants process data on the same objects. However, they can use different technologies for data storage (RDBMS, NoSQL, Excel) and transfer (XML, JSON, CSV). It means that, in the general case, every participant can have its own platform-dependent data models.

For example, an applicant can file a customs declaration in electronic form in the XML format, whereas a customs authority can store the same data in a relational DB (Fig. 2). In this case, they need two platform-dependent data models: an XML schema and an ER model, respectively.

In order to enable the customs authority to extract the data from the XML message and save the data in its DB, another model mapping XML elements to DB fields is required, and such a model always exists. Even if operators manually input details into the DB, this model exists in their minds or as text instructions. In the general case, the

number of mapping models is proportional to the square of the number of exchanging information systems. Obviously, the integration of such heterogeneous information systems with different data storage, transfer and mapping models is rather labor-consuming and error-prone. All these models will have to be changed in case a data structure changes.

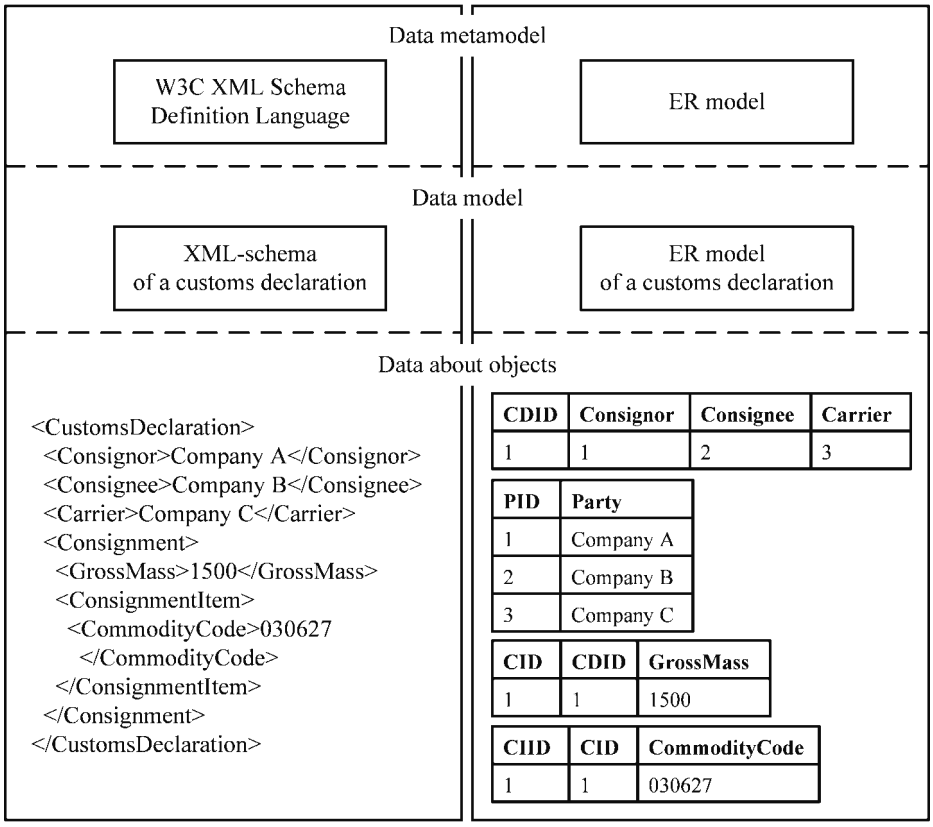


Fig. 2. The example of platform-dependent models and metamodels

3 The Platform-Independent Data Model

Platform-dependent models described above have rather different forms but, on the whole, consider more or less the same data. The development and maintenance of all these models can be significantly simplified if we abstract away from data-representation differences between different platforms and create a single platform-independent model. If one needs to modify a data structure, corrections will have to be made only in this model and then platform-dependent models and mapping models will be automatically generated on its basis. Such an approach is described in the model-driven architecture [1]. An example of the platform-independent model is shown in Fig. 3.

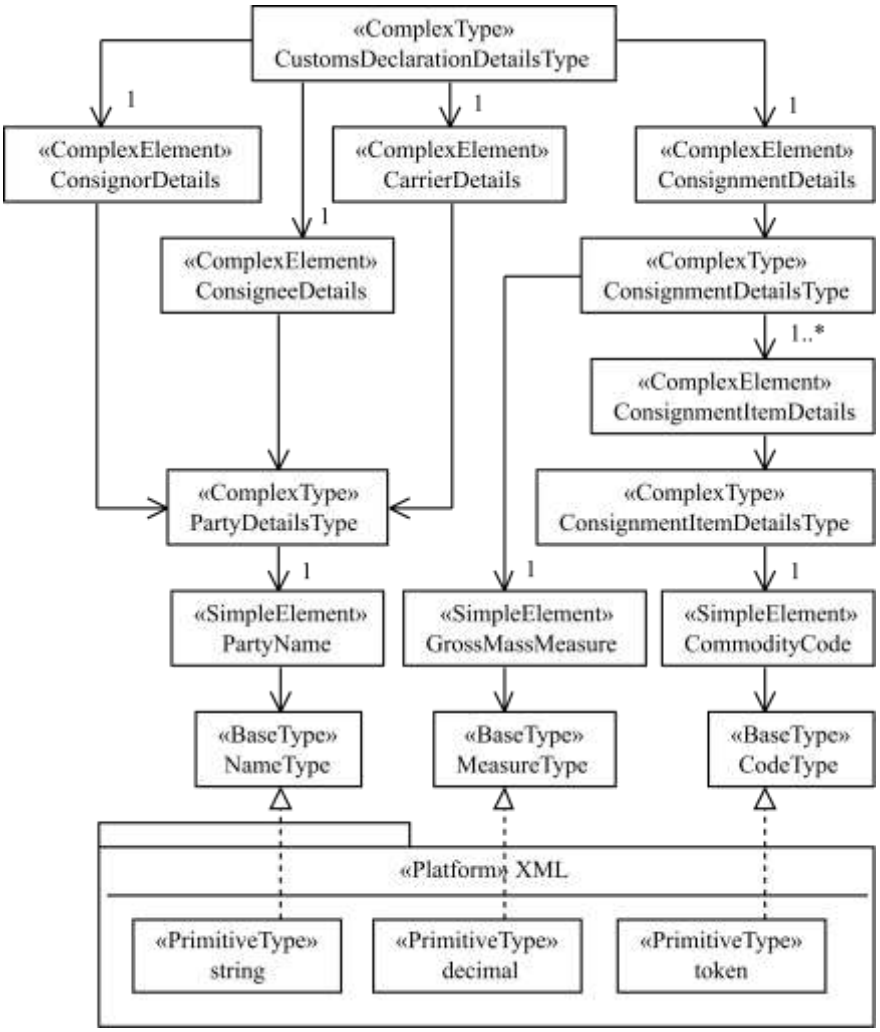


Fig. 3. The example of the platform-independent data model

In our approach, a UML [3] profile is used to develop platform-independent models (Fig. 4). The profile is based on ISO/IEC 11179 [4] and contains the following stereotypes.

Classifier is any object of a data model.

Namespace is the container for logically-related objects with unique names.

Namespace Subset is the subset of related objects of the Namespace.

Platform is the set of the Primitive Types of a certain platform.

Primitive Type is the primitive data type of a certain platform.

Base Type is the data type used to abstract away from the platform. It is to be implemented by exactly one primitive data type for every platform.

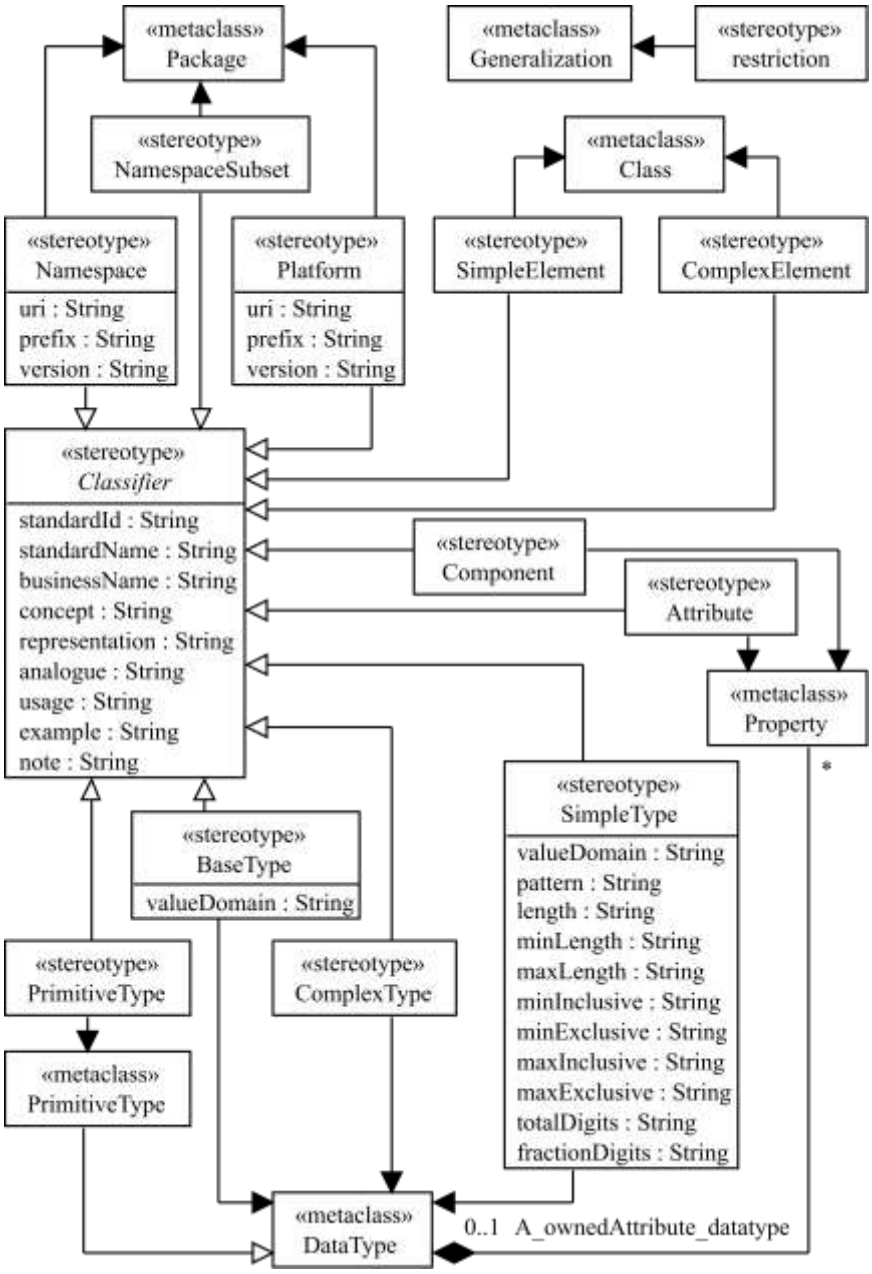


Fig. 4. The profile of the platform-independent data metamodel

Element is the data unit with the designated definition, identifier, representation, and permitted values.

Simple Element is the data element without properties.

Complex Element is the data element with properties.

Simple Type is the set of permitted values of the Simple Element.

Complex Type is the set of the Element's properties that are also Elements, in their turn.

Component is the Element's property.

Attribute is the context characteristic of the data type.

There are two types of inheritance: extension and restriction.

Every data element and type is described as a separate independent entity. This allows generating maximally-normalized platform-dependent models: XML schemas of the "Garden of Eden" pattern and relational models with the relations in the 6th normal form. If necessary, one can generate less normalized models. In other words, such a platform-independent model contains enough information to generate platform-dependent models on its basis with any required characteristics. Further, the metamodel allows describing, using the OCL language [5], business rules that can be converted into SQL or XPath expressions. Model transformation is implemented in [6] using the QVTo language [7], which detailed description falls beyond the scope hereof.

On the one hand, data elements in the described platform-independent model are abstracted away from specific data-modeling languages. They are not XML elements, not entities, not relations but some syntax-neutral data units. On the other hand, the data elements can be re-used in different data structures. This means that the same properties of real-world objects are described using the same data element in all particular data sets (documents, messages). One can conclude that, in fact, delinking of data elements from specific data-representation languages and from their contexts makes the described model conceptual (Fig. 5). The next section shows why this is not the case, and what difficulties it can cause.

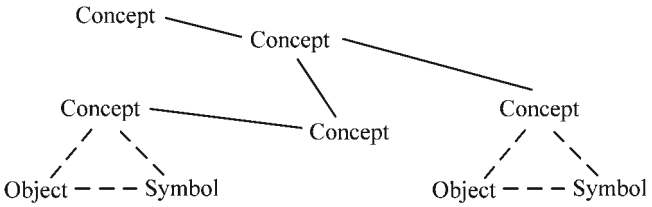


Fig. 5. A conceptual model

4 The Computation-Independent Data Model

The platform-independent data metamodel described above can considerably simplify the development and maintenance of information systems. However, if we attempt to create a single model for different participants (Fig. 1), then the following problems will be encountered.

Firstly, the participants can use different dictionaries and code lists for semantically identical data elements. For instance, a customs authority can code commodities based on the Harmonized Commodity Description and Coding System, whereas a sanitary

and veterinary authority can use different classification codes of the products under control. Kinds of submitted documents for which every participant has its own code list is another example. Such semantically identical data elements can be combined in two ways: either to unify code lists or to accompany codes with references to the used code lists.

Secondly, the participants can structure details of the same objects in different ways. For example, customs authority needs to know a commodity code, price, intended use of goods, packaging kind, and mass. Transport control authority is not interested in price and intended use of goods. Sanitary and veterinary control authority checks date of production, best before date, and age and taxon information of shrimps (Fig. 6). All these appropriate authorities keep under control transportation of one and the same object, i.e. shrimps, but we had to create separate data types for each kind of authority. It increases the size of the model and complicates its maintenance. We can try to harmonize these types by one of the three methods: extension (general characteristics are put in a base data type), restriction (all possible characteristics are put to a base data type, and usage of these characteristics in derived data types is restricted), or composition (each characteristic is described as independent object wherein composite data types refer to these global characteristics only).

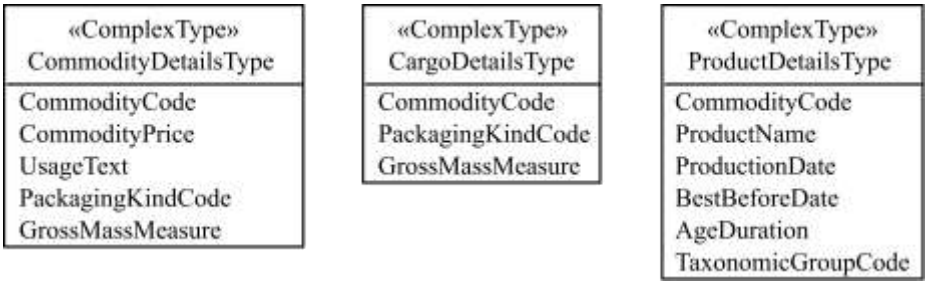


Fig. 6. Different details of one and the same object (shrimps)

Thirdly, the participants use different terminologies. For example, a customs authority mainly uses the term “commodity”, a transport authority uses the term “cargo”, and a sanitary and veterinary authority uses the term “product” to designate the shrimps (Fig. 1). In spite of this, the same object is controlled by all the authorities, which means that only one term should correspond to this object in a single data model. The question is: which of the above terms? Some sources use these terms as synonyms. Other sources state that “a commodity is a product of labor made for sale.” On the other hand, for example, a land lot can be a commodity without being a product of labor. In its turn, a cargo can or cannot be a commodity or a product, and so on.

Although the first and the second problems can be solved within the mentioned platform-independent model, they imply that this model is not as universal as was described above. It can contain several different types and elements in order to represent details of the same object. This fact complicates the integration of information systems and the maintenance of the model.

The third problem explicitly indicates a fundamental shortcoming of the described model. Despite abstracting away from particular data-representation forms (relational model, XML schema, et al.), our model still depends on the usage context, on the domain.

This is connected with the fact that the model describes particular sets of details of real-world objects (particular documents, messages), which can be different for different participants, rather than the objects themselves. A single model can be developed only if the context, documents, messages are disregarded and real-world objects are modeled. The goal of the process described in Fig. 1 is not to transfer the documents but to change the statuses, properties, and relations of the real-world objects. Documents are only a tool to reach the process's goal. If we abstract away from the documents, a single computation-independent model for all the participants can be obtained.

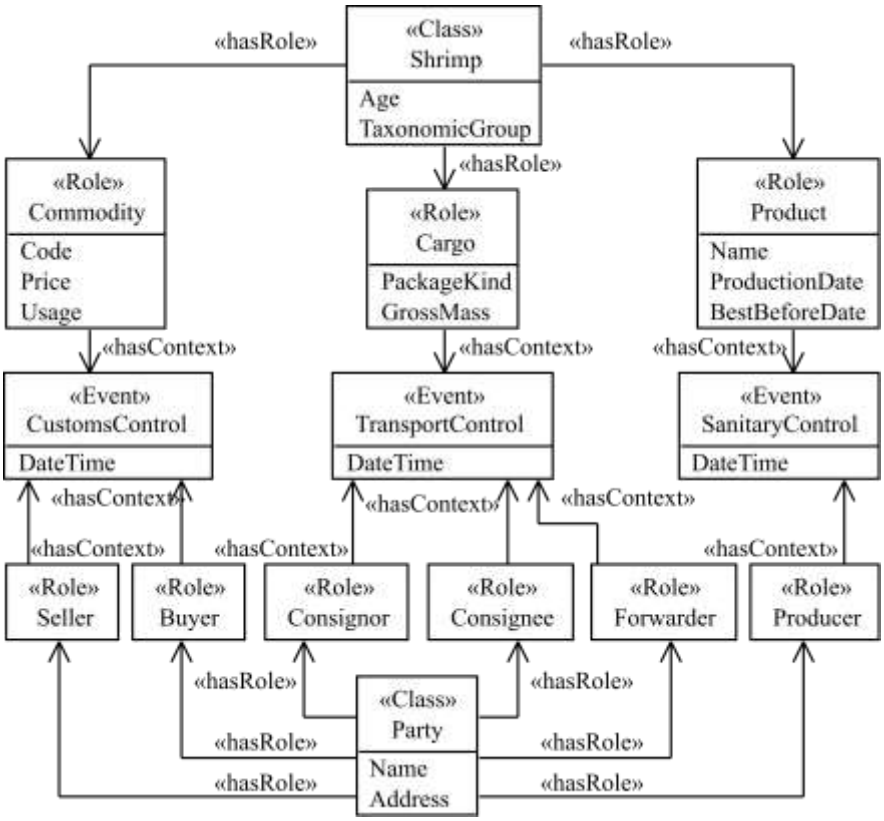


Fig. 7. The example of harmonization of different object' details by means of ontology

Upon reaching the ontological level of modeling, we can conclude that a commodity, a cargo, and a product are not the entity of the object (which data we transfer in the process) but the temporal roles that this object can play. A commodity is the role of the object that has a seller and a buyer. A cargo is the role of the object that has a consignor,

a consignee, and a forwarder. A product is the role of the object that has a producer and a consumer. If the difficulties described in this section occur in the data-modeling process, this means that the capacity of the used modeling language is insufficient and an ontology is required.

An example of such ontology is presented in Fig. 7.With our approach; the ontology can include concepts of three types: classes, roles and events. Class describes an intrinsic essence of the object. Role describes temporary characteristics or relations of the object which appear only in a particular context [8, 9]. Context can be presented by a class (for example, an employee is a role of a man appearing in the context of organization) or by an event. Some authors extract one more kind of a context, i.e. *process*. [10]. However, we view processes as compound events [11]. The computation-independent model, firstly, does not have duplicate elements and types, and, secondly, does not change significantly with appearance of new messages or by change of existing documents or messages.

5 Similar Studies

As early as in the 1970-ies, the need for conceptual data modeling was understood, which resulted in the development of ER models and IDEF1X. It is interesting that the conceptual level of the ER model contains minimum information on which basis a more particular and detailed logical model is further created. In other words, the conceptual model is considered as the first approximation to a logical and then physical model. However, in IDEF1X, a conceptual model is needed to integrate several external and internal data models. Moreover, the conceptual model contains a single, integral, sufficiently detailed representation of data rather than the minimum information or the first approximation to other models.

Our approach based on the model-driven architecture is very similar to ER and IDEF1X: it also has 3 levels but slightly different ones. The first level is a computation-independent model describing real-world objects, the relations and properties thereof. In fact, this is a conceptual model but more detailed than the one of the ER model (in addition to entities and relations, it also contains properties) and less detailed than the one of IDEF1X (it does not indicate data types). The second level is a platform-independent model describing particular structures of documents, messages that contain data of real-world objects. This model is something between external and conceptual models as approached by IDEF1X. The third level is a platform-dependent model that is similar to the logical ER model but can also be non-relational. Another difference between our approach and the ER model and IDEF1X is that transformations of platform-independent models into platform-dependent ones are completely formalized and automated. Finally, our platform-independent model describes not only a data schema but also business rules in the OCL language.

NIEM is another analog of our approach [2]. Similarly, it is based on the model-driven architecture, its specification considers platform-independent and platform-dependent models, transformation of one model into the other is described using the QVTo language, and the metamodel is implemented as a UML profile. There are no

other similarities, however. Firstly, our approach does not have a profile for platform-dependent models. There is no need for it as such a model is fully automatically generated from the platform-independent model. If a platform-dependent model with other characteristics needs to be generated, then the transformation itself should be changed. Secondly, the platform-independent model of NIEM contains several kinds of complex types (objects, roles, associations, et al.). In our approach, such a division is done on the level of the computation-independent model that NIEM lacks. In contrast to our approach, NIEM does not allow describing business rules using the OCL language and transforming these into XPath expressions.

CCTS is another analog of our approach [12]. Similarly, this one is based on ISO/IEC 11179 [4]. Data is modeled on two levels: core components and business-information entities that, for the purposes of this discussion, correspond to the computation-independent and the platform-independent models used in our approach. The difference is that, in CCTS, basic core components, association core components (and the corresponding business-information entities) cannot be reused. For this reason, only partially-normalized platform-dependent models can be generated from such a platform-independent model (the "Garden of Eden" and the 6th normal form are impossible). The second distinction is that new aggregate core components (and aggregate business information entities) cannot be defined in CCTS by extension or restriction; only qualification can be used that is not supported by any platform that the authors hereof are aware of. The third distinction is that, despite the fact that CCTS allows describing business rules in a data model, business rules are described using the XPath language and supported only in XML schemas but, for example, cannot be transformed into SQL expressions. In our platform-independent model, business rules are described using the OCL language and can potentially be transformed into expressions in the language of any platform.

Finally, Semantic Web gains rather significant popularity recently [13]. RDF and OWL allow to describe real-world objects conceptually and computation-independently. Moreover, several working groups developed data integration standards (ISO 15926, ISO 21127), intended on active usage of these technologies. Some authors propose to use RDF as a universal language for data exchange. [14], [15]. Other authors allow usage of XML or relational models for data exchange, but they propose to transform XML schemas and ER models to ontology [16], [17]. The latter, in turn, can be used for data integration.

Our approach goes in the opposite direction. Firstly, we consider that it is not always possible to be limited to usage of RDF and triplestore in data transfer and storage, sometimes XML schemas or relational models are required. Secondly, computation-independent model describing real-world objects must be primal. A data modeler has to develop platform-independent model describing documents and messages on basis of it. This platform-independent model can be transformed into platform-dependent models (XML schemas, ER models). We found no one article with description of transformation of OWL model to ER model or XML schema. Only backward transformations are available. However, for us, this direct transformation is of the most interest.

6 Conclusion and Follow-up Studies

Data modelers (including authors of this article) often face an issue of motivation for the use of one or another data modeling approach: ontologies, model-driven architecture, IDEF1X, ER. In this article, we tried to correlate these approaches. We also tried to specify problems which can testify a necessity of use of model-driven architecture and ontologies.

Multi-level data models including the conceptual model were developed as early as in the 1970-ies. At large, neither ontologies nor the model-driven architecture introduces anything fundamentally new into this field. These are only steps on the way to generalize, unify existing ideas. Ontologies play exactly the same role as the classical conceptual ER models or IDEF1X in the data-modeling process. In the first case, the ontology is the first approximation to a more particular, detailed, contextual data model that describes the sets of details of real-world objects represented by documents, messages rather than the objects themselves. In the second case, the ontology is a sufficiently detailed data model common for several participants.

If only data-exchange schemas (documents, messages) or particular data-storage schemas need to be designed, then XML schemas, relational models, or other logical models are sufficient, there is no need to use ontologies. If there is a need to unify all these particular data schemas, then we should abstract away from the sets of objects' details and start modeling the real-world objects themselves, and for this purpose, ontologies can be used. The possibilities of the languages intended for data modeling on a logical (rather than conceptual) level are rather limited when there is a need to unify data structures using different code lists, slightly different sets of details, and different terminologies. Such languages allow modeling only particular sets of details of objects rather than the real-world objects themselves.

It should also be noted that the developers of data models based on the model-driven architecture, as a rule, restrict themselves to platform-independent and platform-dependent models thus omitting computation-independent models. The present paper attempts to show that platform-dependent models can be well dispensed with as these can be automatically generated from platform-independent models. However, if the developed model is rather complex, covers multiple domains, the computation-independent model (describing real-world objects rather than particular sets of details thereof) can considerably simplify the development and maintenance of information systems. Furthermore, the computation-independent data model is an ontology.

The metamodel presented in this article has no essential novelty. As already stated, similar metamodels were discussed in [2], [12]. As a rule, these metamodels have only one target architecture. Our approach has a benefit: from one platform-independent model we form not only XML schemas but ER models as well. Secondly, we describe not only data schema but also business rules, and we do it in platform-independent OCL, not in a language of target architecture. Thirdly, we predicate that in some situations one platform-independent model is not enough, and a computation-independent model is required.

Currently, a metamodel is used for data modeling in several domains: customs control, transport control, technical regulation, sanitary and veterinary control, and others.

Our approach to data modeling can be used with other types of B2G or G2G interaction. Developed platform-independent model has around 2,000 elements and types. Around 100 exchange structures were developed on the basis of this model. XML schemas and ER models are formed automatically from this model. Unfortunately, transformation rules [6] are too voluminous to present them in this article. It will be described in details in next articles.

At this date, development of the computation-independent data model is in the initial stage. In next articles, we shall describe in details the model formation rules and rules for creation of platform-independent model on its basis. We shall also try to evaluate at what extent the computation-independent model will help to improve the quality of the platform-independent model.

7 References

1. Miller, J., Mukerji, J.: MDA Guide Version 1.0.1 (2003). http://www.omg.org/cgi-bin/doc?omg/03-06-01
2. Object Management Group: UML Profile for NIEM, version 1.0 (2014)
3. Object Management Group: OMG Unified Modeling Language, version 2.4.1 (2011)
4. ISO: ISO/IEC 11179-1:2004. Information technology – Metadata registries – Part 1: Framework (2004)
5. Object Management Group: Object Constraint Language, version 2.4 (2014)
6. Nikiforov, D.A.: UML Model to XML Schema 1.1 Transformation. doi:10.5281/zenodo.16151 (2013)
7. Object Management Group: Meta Object Facility (MOF) 2.0 Query/View/Transformation Specification, version 1.1 (2011)
8. Pradel, M., Henriksson, J., Aßmann, U.: A good role model for ontologies: Collaborations. In: International Workshop on Semantic-Based Software Development (2007)
9. Henriksson, J., Pradel, M., Zschaler, S., and Pan, J. Z.: Ontology Design and Reuse with Conceptual Roles. In Proceedings of the 2nd International Conference on Web Reasoning and Rule Systems, RR ’08, pp. 104–118, Berlin, Heidelberg. Springer-Verlag (2008)
10. Mizoguchi, R., Kozaki, K., and Kitamura, Y.: Ontological Analyses of Roles. FedCSIS, pp. 489-496 (2012)
11. Partridge, C.: Business Objects: Re-Engineering for Re-Use, section 8.3.1.4. Butterworth Heinemann (1996)
12. UN/CEFACT: Core Components Technical Specification, version 3.0 (2009)
13. W3C: Semantic Web. http://www.w3.org/standards/semanticweb/
14. Gorshkov, S. “Business Semantics” practice of application integration using Semantic Web technologies. In international data science conference on Analysis of Images, Social Networks, and Texts (2013)
15. Booth, D., Dowling, C., Fry, C. E., Huff, S., Mandel, J.: RDF as a Universal Healthcare Exchange Language. In Semantic Technology and Business Conference San Francisco, CA (2013)
16. Bedini, I., Matheus, C., Patel-Schneider, P., Boran, A., Nguyen B.: Transforming XML schema to OWL using patterns. ICSC 2011 – 5th IEEE International Conference on Semantic Computing, Palo Alto, United States. pp.1-8 (2011)
17. Myroshnichenko, I., Murphy, M. C. Mapping ER Schemas to OWL Ontologies. In ICSC '09. IEEE International Conference on Semantic Computing (2009)

Implementation of Image Processing Algorithms on the Graphics Processing Units

Natalia Papulovskaya, Kirill Breslavskiy, and Valentin Kashitsin

Department of Information Technologies of the Ural Federal University,
620002, 19 Mira street, Ekaterinburg, Russia
pani28@yandex.ru,

Abstract. The paper describes features of the multithreaded algorithms implementation on contemporary CPU and GPU. The features of access of a graphics processing unit (GPU) memory are reviewed. The bottlenecks have been identified, in which there is a loss of speed in image processing. Recommendations are made for optimization of algorithms for processing image of various size. Examples of implementation of the algorithms are given in the software and hardware architecture CUDA, which is well suited for a wide range of applications with high parallelism.

Keywords: Image processing, Graphics processing unit, Parallel computing, CUDA

1 Introduction

Restricted computing power is one of the problems for solving many tasks. Especially, it appears in processing huge arrays or real-time problems. Processor performance growth of processors has always been associated with increasing the frequency and number of transistors. However, according to the Moore's Law [1], the lost of power is because the laws of physics cannot be changed. Namely, increasing the frequency and amount of transistors can not be infinite. Recent ten years have been devoted to search for other solution: despite projected increase in frequency, there were optimizations of structure and growth of cores quantity. But cores quantity growth is not an absolute solution. New opportunities arise for parallel computing after introducing General Purpose Computing on Graphics Processing Units (GPGPU) [2] and Compute Unified Device Architecture (CUDA) by NVIDIA company [3,4,5].

Usually, the Central Processing Unit (CPU) has only 4 or 8 computing threads. But the Graphics Processing Unit (GPU) has hundreds and thousands computing threads. It provides significant acceleration for algorithms with high parallelism degree. Nevertheless, sometimes the CPU wins in competition with the GPU on well-paralleled tasks.

Mostly, image processing algorithms are easy for parallelization. However, under certain conditions, an algorithm implementation on the central processing unit (CPU) is faster. For proper use of GPU, it is necessary to identity its bottlenecks and describe capabilities of computing resources in tasks of image processing and analysis.

2 Differences between the CPU and GPU

Any program running on CPU uses the operating memory, processor cache, and processor itself. Being run on GPU, the working process of a program is more complex: the algorithm and data should be previously uploaded to GPU before use. The majority of graphics adapters are implemented with the PCI-Express bus, which has limited bandwidth. This is one of the reasons why the central processing unit can perform calculations faster. Obviously, the value of the data and memory uploading time is the crucial parameter for calculation speed.

The CUDA memory model includes the host DRAM memory (regular operating memory), device DRAM (graphics memory), and registered and shared memory that are located on every multiprocessor of the graphics chip. The CUDA threads have access to all kinds of the GPU memory, which differ by scope, speed, and size (Table 1). The registered memory is used for local variables. Other variables that exceed the registered memory are placed in the local memory, which is suited outside of the chip and has low access speed. The shared memory is used for interaction between the threads. All threads have read/write access to the global memory. The scope for global memory is the whole application, and contents of the global memory doesn't change while starting different cores. Moreover, the central processor has also access to the global memory; this is why the global memory is used for data exchange between the CPU and GPU.

Types of memory available for the CUDA applications

Type	Scope	Speed	Size	Applying
Registered	Thread	High	16384 registers per SM	Local variables
Local	Thread	Low	Up to global memory size	Local variables, exceeding registered memory
Shared	Block	High	16 Kb per SM	Threads inter operation
Global	Application	Low	Up to 4 Gb	Data storage with CPU

Thus, the most data operations and exchange with the CPU are implemented by the global memory, which has low speed, as seen from Table 1. The memory delay problem in the CPU is solved by caching and code predicting. But the GPU goes in the other way: in the case of waiting the data access in some thread, the video chip attaches to another thread that has already got all necessary data. Moreover, the video memory mostly has wider bandwidth than the common memory. Instead of active cache memory used by the CPU, the graphics cards have only 128-256 Kb of the cache memory, which is applied to bandwidth extension and of reduction delays.

Multithreading is supported in the GPU on the hardware level. Vendors achieve instant switching between threads (by 1 clock cycle) and support up to 1024 threads on every core.

Large number of threads in the CPU is not advisable due to significant time loss while switching, which may consume up to several hundreds of clock-cycles. In addition, one CPU core is able to process only 1 or 2 threads simultaneously.

3 Benchmarks

3.1 Baseline

Performance tests were done on image processing algorithms software implementation. Four different resolutions of image (50x50, 150x150, 300x300, and 512x512) were used with the quality of JPEG Quality 25 and JPEG Quality 100. As a reference image, the well-known file Lena.jpg [6], was used. Lenna or Lena is the name given to a standard test image widely used in the field of image processing since 1973.

The images were processed by application of Sarcis [7] developed by the Department of Information Technologies of the Ural Federal University. This program is designed for image processing of color and gray scale images and can be used for calculation both on the CPU and GPU. For the tests, four algorithms were chosen: the weighted linear filter, reverse gradient filter, edge detection Canny's filter, and morphology processing filter erosion. An example of application of the Canny's filter is shown in Fig.1. Workstation configurations

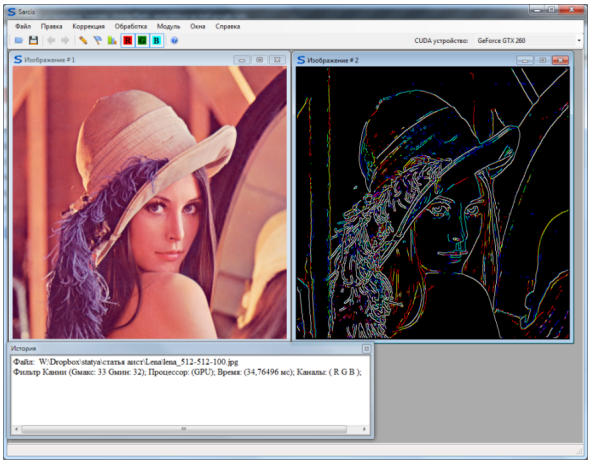


Fig. 1. Image processing by the Sarcis program

used for benchmarks:

CPU Intel Core i5-2450 (4 threads, 2.5Ghz), GPU NVIDIA GT630M (96 CUDA cores, 800Mhz);

CPU Intel Core i7-860 (8 threads, 2.8Ghz), GPU NVIDIA GTX260 (216 CUDA cores, 576Mhz);
CPU Intel Core i5-3470 (4 threads, 3.2Ghz), GPU NVIDIA GTX760Ti (1344 CUDA cores, 915Mhz).

3.2 Theoretical description of filtering algorithms

Weighted linear filtering. The linear smoothing filters are good ones for removing the Gaussian noise and, also, the other types of noise. A linear filter is implemented using the weighted sum of the pixels in the successive windows. Typically, the same pattern of weights is used in each window; this means that the linear filter is spatially invariant and can be implemented using a convolution mask. If different filter weights are used for different parts of the image (but the filter is still implemented as a weighted sum), then the linear filter is spatially varied [8].

One of the simplest linear filters is implemented by a local averaging operation where the value of each pixel is replaced by the average of all the values in the local neighborhood. The weighted linear filtering mask is given by the mask $M = \frac{1}{16}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$. The main strategy is to set the largest weight at the central pixel and inversely proportional to distance values to other pixels:

Reverse gradient filter. The idea of the reverse gradient filter is in choosing the mask weights. The greater bright difference between the central and the next point, the less weight is prescribed to the next point.

The mean gradient module value calculated for each aperture brightness level is:

$$H^G(l) = \frac{1}{h(l)} \sum_{m=1}^{h(l)} G_l^m(i,j)$$

Where $H^G(l)$ is the mean brightness gradient function value for l; $h(l)$ is the number of points with brightness l; $G_l(i,j)$ is the gradient value in the mth point with brightness l at position (i,j). The Laplace operator $\bigtriangledown^2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$ is used to calculate the gradient value

Canny filter. The Canny filter is considered to be the optimal choice for any task, which requires edge extraction or edge enhancement of elements with different shapes and characteristics. This is due to its easy implementation and ability to be adapted to different circumstances [9].

There are the following three criteria for the algorithm:

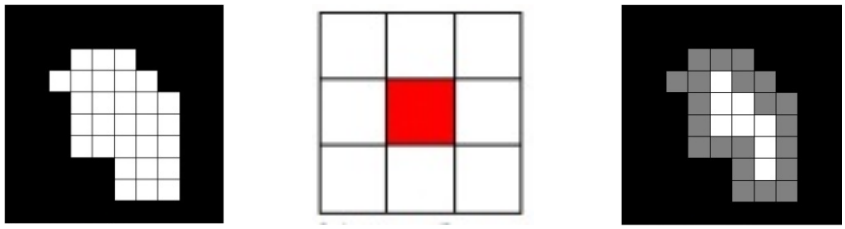
- smart detection (increasing the signal/noise ratio);
- good localization (correct definition of boundary position);
- only one response to the border.

The Canny filter differs from other classic edge detecting filters because it considers directionality of the element and permits personalization of the filter behavior by choosing the value of parameters of the algorithm. This provides changes in direction of the filter polarization and sensibility of the edge extraction.

The algorithm runs in 5 separate steps [10].

1. Smoothing: blurring of the image to remove noise.
2. Finding gradients: the edges should be marked where the gradients of the image have large magnitudes.
3. Non-maximum suppression: only local maxima should be marked as edges.
4. Using the double thresholds: potential edges are determined by using the thresholds.
5. Edge tracking by hysteresis: the final edges are determined by suppressing all edges that are not connected to a strongly determined edge.

Morphological Image Processing. The morphological image processing is a collection of non-linear operations related to the shape or morphology the image features. Erosion of images is generally used for getting rid of the image of random insertions. The idea is that the inclusions are eliminated by blurring, while large and, thus, more visually important regions remain.



a) b) c)

Fig. 2. Morphological image processing example; a) source image; b) kernel; c) erosion result

Erosion (morphological narrowing) is a convolution of the image or its area selected with some kernel. The kernel may be of arbitrary shape and size. Here, in calculating the convolution, only one leading position is selected in core, which is aligned with the current pixel. In many cases, the core is selected as a square or circle with a leading position at the center (Fig. 2b). The core can be regarded as a pattern or a mask.

Applications of dilation is reduced to the passage pattern throughout the image and use some operator to search for a local minimum intensities of the pixels in the image, which covers the template. The gray color filled pixels appear black due to erosion (Fig. 2c).

3.3 Benchmark results

Weighted linear filtering. As seen from graphs in Fig. 3, the images with size 50×50 are processed for the same time by both the CPU and GPU, excluding NVIDIA GTX260, which works a little better. The larger images are processed faster by the GPU.

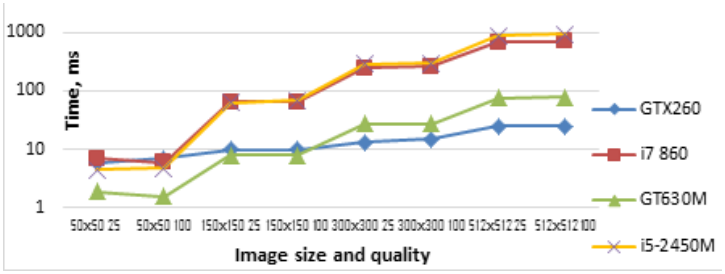


Fig. 3. Weighted linear filter benchmarks

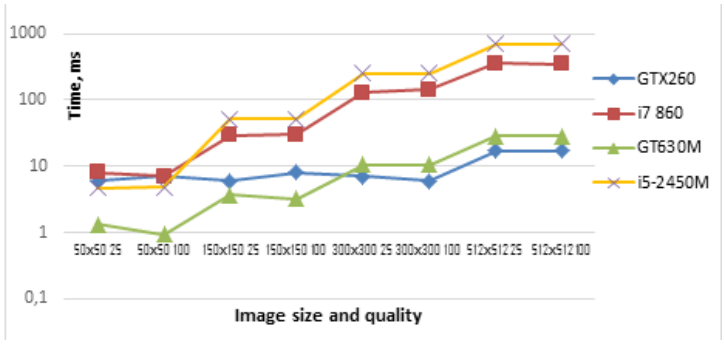


Fig. 4. Reverse gradient filter benchmarks

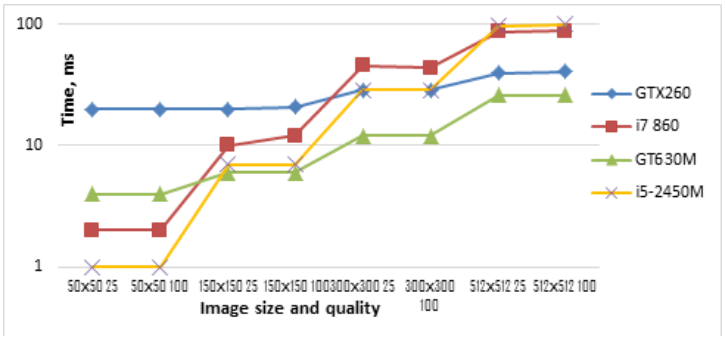


Fig. 5. Canny filter benchmarks

Reverse gradient filter. As seen from graphs in Fig. 4, application of the reverse gradient filter gives comparable results w.r.t. the weighted linear filter.

Canny filter. It is obvious from Fig. 5 that the CPU runs faster on images up to 150×150 quality 25, but the GPU NVIDIA GT630M takes advantage in times on larger images. The GPU NVIDIA GTX260 works better when the CPU starts from images 300×300 regardless of image quality.

Morphological Image Processing. Analysis of graphs in Fig. 6 shows that in using the morphological processing similar to Canny's filter, the GPU NVIDIA GT630M starts process faster than CPU from images size 50×50; and the GPU NVIDIA GTX260 is faster from size 300×300.

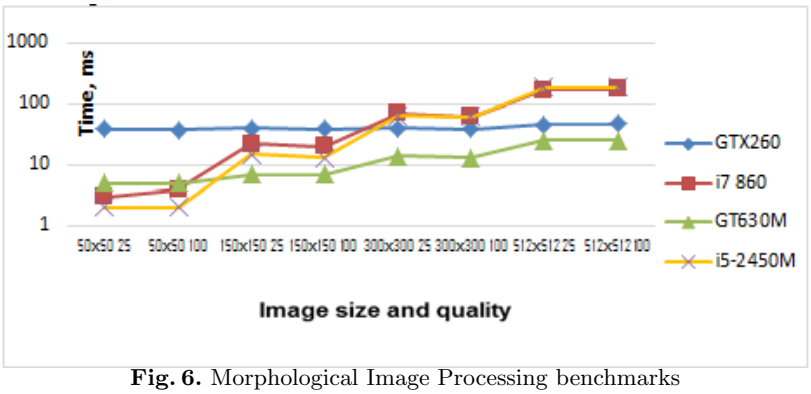


Fig. 6. Morphological Image Processing benchmarks

4 Conclusion

Results of this research show a rise in the efficiency of the GPU computing while image size grows. Nowadays, the images with resolution less than 1024×768 can be found rarely. From resolutions 300×300, the GPU starts to process images faster than the CPU; so, it is advisable to use image processing algorithms applying the GPU computing. The CPU computing is suitable in small image processing or hardware systems without the GPU.

References

1. Moore's Law, `http://www.mooreslaw.org/` csl783/canny.pdf (2009)
2. GPGPU. General-Purpose Computation on Graphics Hardware, `http://gpgpu.org`.
3. Halfhill, T. R.: Parallel processing with CUDA. Micriprocessor report, `http://www.nvidia.com/docs/IO/55972/220401-Reprint.pdf` (2008)
4. NVIDIA. CUDA in Action, `http://www.nvidia.com/object/cuda-home-new.html`
5. NVIDIA CUDA: Compute Unified Device Architecture, NVIDIA Corp. (2007).
6. The Lenna Story, `http://www.cs.cmu.edu/~chuck/lennapg/lenna.shtml`
7. Dorosinsky, L.G., Kruglov, V.N., Papulovskaya, N.V., Chiryshev, A.V.: CUDA technology in digital image processing, Ekateringurg, 192 p.,(2011)

8. Jain, R., Kasturi, R., Schunck, B.G.: Machine vision, http://www.cse.usf.edu/~r1k/MachineVisionBook/MachineVision.files/MachineVision-Chapter4.pdf, 118-122
9. Pirotti, F., Vettore, A.: Canny filter applications and implementation with grass, http://www.academia.edu/313972/
10. Canny Edge Detection, http://www.cse.iitd.ernet.in/~pkalra/

Am I Really Happy When I Write "Happy" in My Post?

Pavel Shashkin and Alexander Porshnev

National Research University Higher School of Economics
p-sh@live.ru, aporshnev@hse.ru

Abstract. Posts published on the Internet could serve as a valuable source of information regarding emotion. Recommendation systems, stock market forecast and other areas are likely to benefit from the advancement in mood classification. To deal with this task, researchers commonly rely on preassembled lexicons of emotional words. In this paper we discuss the possibility of extracting emotion-specific words from user-annotated blog entries. The study is based on analysis of the collection from 14800 Live Journal posts containing the "Current mood" tag, specified by the author. The analysis findings and possible applications are discussed.

Keywords: *sentiment analysis, user-annotated data, computational linguistics, emotional states, psycholinguistics*

Introduction

Over the last few years, a considerable amount of work has been done to reduce the overload of user-generated web content [1]. The possibility of grouping data in accordance with sentiment is repeatedly discussed in recent investigations [2, 3]. The system capable of extracting emotions inherent in the text is likely to assist both human-computer and human-human interactions, and help in various tasks. For example, automatic analysis of emotions could be used in some applications, such as: recommendation systems (personal emotions expressed during evaluation could be taken into account), monitoring of psychological user states (customer satisfaction or diagnostics of potential illness), business intelligence (evaluation of the emotional tone of comments circulating about one's company can be used to improve financial decisions).

Online diaries provide researchers with extremely diverse and manifold data. Blog entries are rich in deeply personal and subjective content. Unlike other corpora used in sentiment analysis, "Current Mood" is a text attribute directly specified by the author at the time of writing, rather than by some independent annotator. We expect that analysis of user-annotated data could provide new information about words people use to express their emotional states.

Related work

Over the past few decades there have been several projects devoted to analysis of emotions in the Internet posts. For example, a research project for measuring emotions is "Pulse of a nation" is based on analysis of Twitter messages from September 2006 to August 2009 [4]. In their research, Mislove and coauthors tried to find places where life is sweet, people are happier, and to reveal the unhappiest time of a day. Although, we were unable to find scientific articles, the authors of the "Pulse of a nation" project took part in several TV programs and published the results in newspapers and periodicals (including The Wall Street Journal and The New York Times).

To measure emotions in each tweet, Mislove and coauthors used the ANEW word list [5]. The methodology of emotion analysis was to calculate a sentiment score as a ration of the amount of positive messages to that of negative messages. A message is regarded as positive if it has at least one positive word and as negative if it has at least one negative word (the same message can be both negative and positive) [6].

The project focused on the expression of happiness in social media was developed by a group of researchers from the University of Vermont. They tried to measure happiness in Twitter posts [7].

First of all, Dodds and his coauthors conducted a survey using Amazon Mechanical Turks to obtain happiness evaluations of over 10,000 individual words, representing a tenfold size improvement over similar existing word sets (chosen by frequency of usage in collected samples of nearly 4.6 billion expressions posted over a 33 month span). The created words list contains ranks of their relation to happiness. For example, the top happiness words in their rank are laughter (rank=1) and happiness (rank=2). Next, they created on-line service hedonometer.org, which provides real time happiness analytics based on analysis of frequencies of the words from the list. It is worth mentioning that this service also has a rank for the word "birthday", so the expression "Happy Birthday" is not excluded from analysis.

Lansdall-Welfare, Lampos, and Cristianini tried to measure several emotions in twitter posts by counting the frequency of emotion-related words in each text published on a given day [8]. They also use a lexical approach and base their analytics on the word lists extracted from the WordNet Affect ontology [9].

After this pre-processing Lansdall-Welfare, Lampos, & Cristianini compiled four word lists containind 146 anger words, 92 fear words, 224 joy words and 115 sadness words. The evaluation of emotions in tweets was based on counting the amount of tweets containing each word from the compiled list. Lansdall-Welfare, Lampos, & Cristianini say they do not expect the high frequency of the word 'happy' to necessarily signify a happier mood in the population, as this can be due to expressions of greeting, like "Happy Birthday". Although they do not filter this and similar expressions in their analysis.

We can conclude that the projects running analysis of emotions and moods in social networks usually use the lexicon methodology based on expert-annotated words lists.

Application of the lexicon approach based on expert or naïve rating of emotions in the Internet posts can be supported by the findings made by Gill, Gergle, French and Oberlander. They examined the ability of naive raters of emotion to detect one of the eight emotional categories by asking participants to read 50 and 200 word samples of a real blog text and evaluate whether this message expresses one of the eight emotions: anticipation, acceptance, sadness, disgust, anger, fear, surprise, joy or being neutral [10]. Comparing the results of evaluation by expert raters and naive experts allowed the conclusion that rater agreement increased with longer texts, and was high for ratings of joy, disgust, anger and anticipation, but low for acceptance and 'neutral' texts.

Although raters show agreement in annotation of emotions, we can raise a question about its validity from the psychological perspective. We are not sure that all people express their emotions in a straightforward way, using words closest to the chosen mood category.

An interesting study of emotion in the context of a computer-mediated environment was conducted by Hancock, Landrigan, & Silver [11]. They organized an experimental study, in which some of the participants were asked to express either positive (happy) or negative (unhappy) emotions during a chat conversation, without explicitly describing their (projected) emotional state. Even though, their chat partners did not know about their instructions and their emotional state, they could accurately perceive their interlocutor's emotions. Linguistic analysis showed that the authors portraying positive emotions used a greater number of exclamation marks and more words overall. The participants portraying negative emotions used an increased number of affective words, words expressing negative feeling, and negations.

In this study the people understand emotions of their partner even if these emotions were not explicitly expressed. This raises a question: could we extend the lists of emotional words by analyzing data annotated with the current mood of an author?

Analysis of text semantics, therefore, can provide information about user emotions and we expect analysis of user-annotated data from LiveJournal to help extend the existing words lists related to emotions.

Data collection

We used DuckDuckGo[1] search engine in conjunction with "GoogleScraper"[2] Python module to make a list of English-speaking LiveJournal users who have at least once used the "Current Mood" functionality. The list of obtained URLs is passed down to the web crawler hosted on "import.io"[3] platform. Each visited page is parsed to extract user messages and links to other LiveJournal blogs to be added to crawling query (e.g. from the comment section). Data collecting

[1] `https://duckduckgo.com/`
[2] `https://github.com/NikolaiT/GoogleScraper`
[3] `https://import.io/`

continues until the specified maximum page depth is reached.

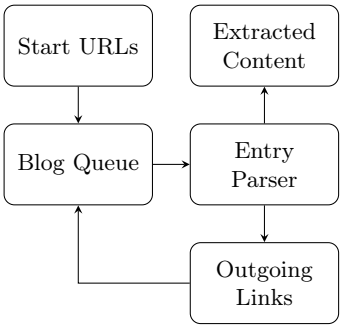


Fig. 1. Data collection system architecture

Data-set highlights

For each message in visited blogs we extract a web address, title, text content and mood tag (usually accompanied by "Current mood:", "Feeling rather:" or just "Mood:"). Although the presence of subjective content in a title or web address is questionable, they are needed to identify continuous entries (eg. stories divided into series of posts). Blog posts, especially the ones with a fair amount of text, are not as frequent as, say, twitter posts. For that reason we do not use time stamps and rarely present geolocation data. The acquired dataset contains 14,800 documents tagged with 800 unique mood labels. 6% of the labels were responsible for 60% of data entries (Figure 2). Average text length is 420 words. An approximate post count for the average author is 5 messages (Figure 3). The most popular mood tags are: "accomplished", "cheerful", "tired" and "amused".

Pre-Processing

The initial step is to clean data from invalid entries (non-Latin or comprising only media content). The dictionary is then reduced by transforming everything to lowercase, stemming words, removing punctuation, stopwords and numbers. If we find negations, like “don’t”, “didn’t” or “not”, the subsequent token is replaced by not_token. For example, “they didn’t come” includes three tokens: “they”, “didn’t”, “come”. We also keep negations as we can expect negative moods negotiations to carry some additional information. URLs are shortened to their respective domains and repeating letters (more than three) are reduced to three. Words and numbers representing time or date are replaced with "time_date".

Fig. 2. Label frequencies

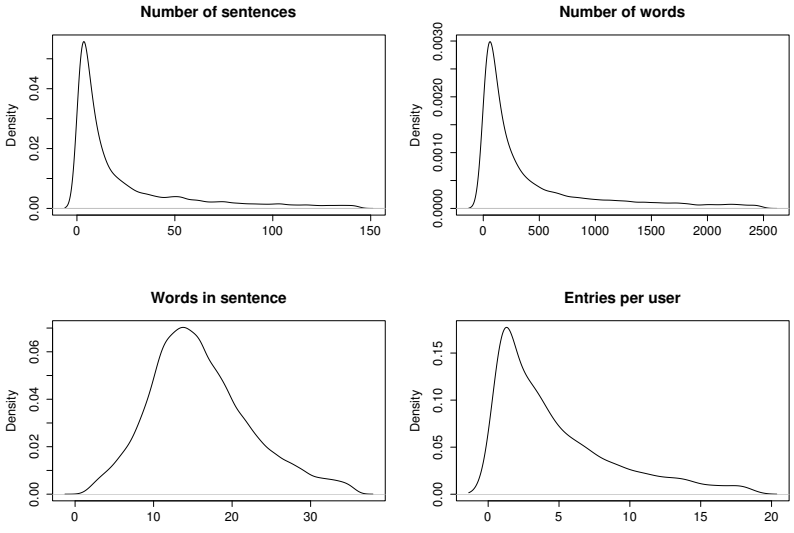


Fig. 3. Empirical distribution density for text statistics (sentences detected with openNLP, outliers removed)

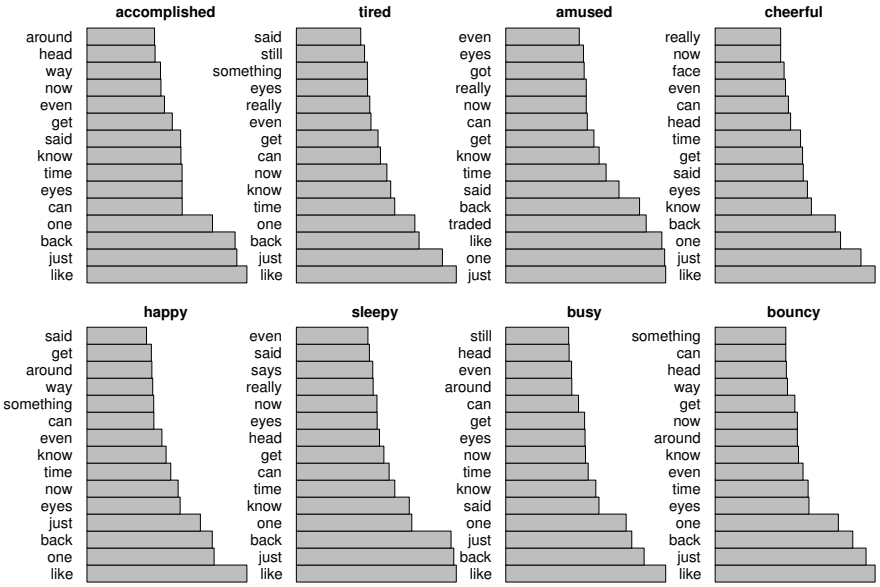


Fig. 4. Word frequencies distribution for popular mood labels

After pre-processing the portion of non-sparse terms doubled. Overall dimensionality of feature space was reduced by more than 3 times.

Fifteen highest word frequencies for 8 most used mood labels are very similar and do not provide any evidence that people use emotional words to mark their emotions (Figure 4). We can see that words highly associated with a mood are not included in the list with top 20 frequencies. For example, it is not often that messages tagged “Happy” contain “happy” in their body. The list of top 20 words does not contain many words from emotional lists. Words “like”, “one”, “back” are not put on the list of 10,000 words related to “happy” according to a Hedonometrics survey [7]. They are not included in the list of Affective Norms for English Words either [5].

The TF-IDF coefficient frequently used for document classification can provide more focused information about semantics of each emotion categories. To calculate TF-IDF, we joined all documents of the category into one document. First, the calculated TF-IDF allowed us to find most of the names used in posts. The words with the highest TF-IDF scores were “leo", "maes", "vampir", "jare","sandi","gaara", and "roger". After including the names in a list of stopwords, we received almost the same situation as with calculation of term frequencies (see Table 1).

accomplished		tired		cheerful		sleepy	
hand	505.7	hand	222.4	artwork	173.3	ghost	342.0
said	447.1	rift	216.7	hand	167.7	hand	170.2
eye	416.0	say	191.8	said	154.5	margin	153.3
back	389.8	f*ck	170.9	head	151.1	head	151.1
head	375.5	eye	156.9	eye	142.8	back	144.5
smile	368.7	back	155.6	smile	135.7	color	142.3
look	360.8	look	149.5	face	135.2	say	140.1
say	355.2	head	148.6	back	119.3	eye	130.8
robot	353.5	doesnt	145.5	look	117.9	superhero	130.0
mule	338.5	said	145.2	lip	110.8	said	125.4

amused		happy		busy		bouncy	
trade	692.7	array	279.9	dev	263.9	sampl	120.7
prize	379.7	hand	164.9	alt	178.5	hand	115.1
ward	143.9	eye	164.7	hand	148.6	introspect	106.3
claim	135.6	back	137.1	said	147.0	head	91.3
vote	128.4	lip	127.4	border	133.4	back	89.4
said	91.6	smile	125.5	back	118.3	said	84.8
hand	86.6	head	122.0	head	114.8	charact	80.6
bill	86.0	knew	122.0	eye	114.7	lip	76.5
materia	79.7	realis	121.2	knew	101.3	look	74.6
back	69.4	face	119.1	multi	101.1	kiss	74.4

Table 1. Words with highest TF-IDF score for eight most popular mood labels (with proper nouns removed)

Next, we introduced the TF-ICF coefficient. In order to identify important group-specific words, the term frequencies $TFij$ for word i in group j are multiplied by:

$$log \frac{||D||}{\sum_{j=1}^{||D||} \frac{TF_{ij}}{\max_{t \in T_j} TF_{tj}}} \tag{1}$$

where $||D||$ is the number of document groups and T_j - unique words in document group j.

The results produced by this transformation are listed in Figure 5 (apart from persons, locations and brands on top of the list) and provide more information regarding sentiment. For example, the word "finally" has a high value in documents tagged with "accomplished" or "tired". Although, some of these results are relatively counter-intuitive or even contradictory (e.g. "bed" is present in "accomplished", "bouncy", "cheerful", "busy", but absent in "sleepy").

This suggests that the distance between documents written in different emotional states could be shorter than that between documents written in the same emotional state by different authors. To test this hypothesis, we filtered documents by author and then, using vector representation of documents, we calcu-

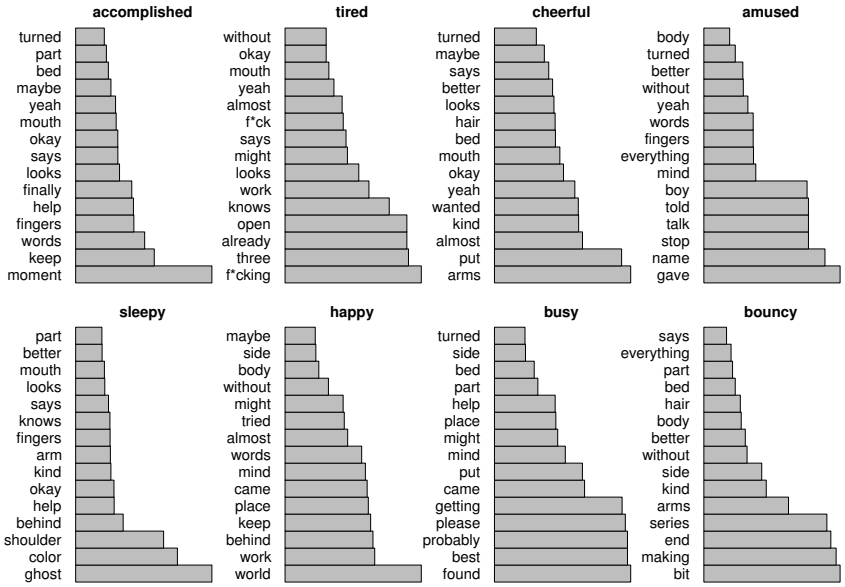


Fig. 5. Most important words according to TF ICF (with proper nouns removed)

lated cosine similarity between every pair of documents. The same procedure was carried out for documents filtered by current mood tag.

The only pair of tags "nervous" and "accomplished" has the distance between mood labels shorter than the average distance between different authors. This is probably because they carry a lot of objective content, which should have been filtered at earlier stages. The previously mentioned self-containing states of mind fall within the same group of labels, whose distances do not exceed the global average.

The vector model, therefore, contained enough information to distinguish emotions and what we needed was to find an approach to extracting words with maximum information. To solve this task, we used the Mutual Information feature selection algorithm [12].

Application of the mutual information feature selection algorithm showed that the word “happy” provided relevant information about the mood of an author. However, the top twelve terms for the “happy” category only contained two emotional words included in the Hedonometics or ANEW list (“happy” and “wonder”).

We saw that, according to mutual information feature selection, many of the categories were determined by the terms not included in emotional words lists. Then we checked whether or not category name synonyms obtained from WordNet were frequently encountered in the documents labeled with the same mood [13]. Most of the time mood was not specified in the text in an obvious

accomplished		**tired**		**cheerful**		**sleepy**	
eye	0.0084	three	0.0014	yes	0.0005	found	0.0009
head	0.0080	got	0.0011	feel	0.0004	name	0.0007
pull	0.0079	home	0.0010	still	0.0004	probabl	0.0007
turn	0.0074	day	0.0010	like	0.0004	knew	0.0007
smile	0.0072	lot	0.0009	way	0.0003	side	0.0006
arm	0.0071	tell	0.0009	right	0.0003	bad	0.0006
first	0.0070	realli	0.0009	see	0.0003	tri	0.0006
pair	0.0069	far	0.0008	guy	0.0003	rate	0.0006
behind	0.0069	think	0.0008	think	0.0003	time	0.0005
hand	0.0068	let	0.0008	girl	0.0003	walk	0.0005
away	0.0067	time	0.0008	hold	0.0002	mayb	0.0005
side	0.0063	part	0.0008	just	0.0002	find	0.0005

amused		**happy**		**busy**		**bouncy**	
knew	0.0008	happi	0.0020	thank	0.0012	your	0.0006
still	0.0008	dont	0.0006	set	0.0011	far	0.0006
way	0.0007	found	0.0005	pleas	0.0010	someon	0.0005
week	0.0007	ive	0.0005	use	0.0010	feel	0.0004
cant	0.0007	wonder	0.0004	everi	0.0010	ask	0.0004
pleas	0.0007	thank	0.0004	ask	0.0009	sinc	0.0003
feel	0.0007	wait	0.0004	man	0.0009	talk	0.0003
far	0.0006	also	0.0004	ill	0.0007	word	0.0003
will	0.0006	home	0.0004	leav	0.0007	dont	0.0003
tri	0.0005	one	0.0004	found	0.0007	guy	0.0003
found	0.0005	isnt	0.0003	comment	0.0006	way	0.0003
day	0.0005	day	0.0003	tag	0.0006	see	0.0003

Table 2. Most important words according to mutual information feature selection algorithm

way. Only 14 of 50 popular moods or their synonyms are frequently encountered in a text tagged with the same mood: crazy (stressed, crazy, sick), curious (good, curious, sore), depressed (depressed, hopeful, artistic), ecstatic (hopeful, ecstatic, productive), hopeful (hopeful, crazy, sad), pissed off (pissed off, nervous, okay), sad (sad, frustrated, pissed), sick (confused, crazy, sick), sleepy (stressed, sleepy, sick), sore (sad, ecstatic, sore), stressed (stressed, depressed, curious).

Surprised by such results, we tried to analyze the document using words from the Hedonometrics list. Analysis of frequencies of top twelve words from the Hedonometrics list in texts written in different moods showed that these words have the most common usage in emotional states different from “happy” (Table 3). Only one word “successful” is used more frequently by authors who tagged their message with the current mood “happy”.

	accomplished	tired	cheerful	sleepy	amused	happy	busy	bouncy
laughter	1.37	1.45	1.28	1.16	2.22	1.54	1.61	2.25
love	23.97	25.32	29.18	20.92	27.27	27.29	26.46	30.04
happy	7.28	6.74	8.90	7.98	7.83	11.95	7.71	12.14
laugh	10.83	12.03	12.43	7.28	9.79	7.48	8.88	9.01
excellent	0.50	0.33	0.56	0.54	0.26	0.46	0.54	0.38
joy	0.89	0.66	1.52	0.77	1.17	1.31	0.45	1.25
successful	0.64	1.06	0.72	0.62	0.65	1.39	0.45	0.13
win	1.42	1.19	1.20	0.85	3.26	0.92	1.97	0.75
rainbow	0.56	0.20	0.32	0.31	0.26	0.08	0.09	0.50
smile	27.29	25.45	28.86	21.23	23.10	25.67	20.63	24.53
won	0.73	0.60	1.20	0.46	0.91	0.39	0.90	1.25
pleasure	1.59	1.26	1.60	1.24	1.04	1.46	1.97	3.13
celebration	1.14	1.12	1.92	0.54	1.30	0.69	1.17	0.75

Table 3. Word frequencies $\times 10^5$

Conclusion

Analysis of user-annotated blog messages showed that connections between emotions and their linguistic expression could not necessarily be straightforward as is usually expected by compilers of emotional words lists. The most frequent words in each mood category are not included in the list of emotional terms. Application of TF-IDF and the calculated TF-ICF coefficient did not change the situation. Words with the highest scores continue not to be included in popular lists used for mood analysis. Application of the Mutual Information feature selection algorithm allowed us to find the most important words in each category, but only few of them are included in popular lists of emotional words. We can confirm that. according to the mutual information coefficient, the word "happy" has high discriminative power, while other words from the Hedonometrics list were not as successful.

People show a high ability to evaluate emotions of other persons even in a computer-mediated environment, although the way we can understand other people's emotions still raises questions. On the one hand, the ability to understand emotions also exists in situations where emotions are not explicitly expressed; on the other hand, our analysis showed a paradoxical situation when the terms used for evaluation of emotions are not among the top 20 frequent or discriminative words for each of mood categories. These facts raise a question about psychological validity of straightforward techniques for measuring emotions.

In our further research we plan to move in two different directions. One is to compare results of emotion analysis by applying the classical lexical approach with two dictionaries (ANEW and Hedonometrcs) and Naïve Bayes algorithm using the probabilities calculated in the current research. The other direction is to test agreement between naïve or expert annotators and authors of mood

labels. We also intend to develop more sophisticated procedures to filter objective content and detect invalid entries, establish a meaningful connection between content and label and further extend our database to improve validity of our study.

References

1. Agichtein, E., Castillo, C., Donato, D., Gionis, A., Mishne, G.: Finding high-quality content in social media. In: Proceedings of the 2008 International Conference on Web Search and Data Mining, ACM (2008) 183–194
2. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: Sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10. EMNLP '02, Stroudsburg, PA, USA, Association for Computational Linguistics (2002) 79–86
3. Yu, L.C., Wu, J.L., Chang, P.C., Chu, H.S.: Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. Know.-Based Syst. (March 2013) 89–97
4. Mislove, A., Lehmann, S., Ahn, Y.Y., Onnela, J.P., Rosenquist, J.: Pulse of the Nation: U.S. Mood Throughout the Day inferred from Twitter (2010)
5. Bradley, M.M., Lang, P.J.: Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical report, Technical Report C-1, The Center for Research in Psychophysiology, University of Florida (1999)
6. O'Connor, B., Balasubramanyan, R., Routledge, B.R., Smith, N.A.: From tweets to polls: Linking text sentiment to public opinion time series. In: Proceedings of the International AAAI Conference on Weblogs and Social Media. (2010) 122–129
7. Dodds, P.S., Harris, K.D., Kloumann, I.M., Bliss, C.A., Danforth, C.M.: Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. PloS one (2011)
8. Lansdall-Welfare, T., Lampos, V., Cristianini, N.: Effects of the Recession on Public Mood in the UK. In: Proceedings of the 21st international conference companion on World Wide Web, ACM (2012) 1221–1226
9. Strapparava, C., Valitutti, A., Stock, O.: The affective weight of lexicon. In: Proceedings of the Fifth International Conference on Language Resources and Evaluation. (2006) 423–426
10. Gill, A.J., Gergle, D., French, R.M., Oberlander, J.: Emotion rating from short blog texts. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2008) 1121–1124
11. Hancock, J.T., Landrigan, C., Silver, C.: Expressing emotion in text-based communication. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM (2007) 929–932
12. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to information retrieval. Volume 1. Cambridge university press Cambridge (2008)
13. Miller, G.A.: Wordnet: A lexical database for english. Commun. ACM (November 1995) 39–41

Study of the Mass Center Motion of the Left Ventricle Area in Echocardiographic Videos

Porshnev S.V.[1], Mukhtarov A.A.[1], Bobkova A.O.[1], Zyuzin V.V.[1], and Bobkov V.V.[2]

[1] Ural Federal University
named after First President of Russia B.N. Yeltsin,
Ekaterinburg, Russia (620002, Ekaterinburg, Mira Str., 19)
sergey_porshnev@mail.ru,andrew443209993@yandex.ru,
iconismo@gmail.com,zvvzuzin@gmail.com
[2] Ural State University of Economics and Ural Institute of Business,
Ekaterinburg, Russia, (620144, Ekaterinburg, 8 th of March Str. 62)
btow@yandex.ru

Abstract. The study of the mass center (CM) motion of the left ventricle (LV) area in echocardiographic videos is presented. CM of any phase of the cardiac cycle is inside an ellipse. An area of the ellipse is not more than 2% of the area of LV in systole. The criterion to identify correct and incorrect forms of contours are proposed for automatic contouring of LV in the frames of a video sequence.

Keywords: left ventricle, contouring, echocardiographic images (echocardiography), image processing.

1 Introduction

Echocardiography is one of the most wide spread ways of not invasive heart muscle deseases diagnosis, which most commonly uses ultrasound (US) images of the apical four-chamber heart projection. Sequence analysis of ultrasound images allows to analyze the dynamics of the heart muscle. Of particular interest to cardiologists is left ventricle (LV), since most various diseases and pathologies of the heart can change their state.

To assess the state of the LV, cardiologist puts contour on each frame of ultrasound sequence, which limits the region of the LV. As a rule, the cardiologist makes it manually or by semiautomatic modes. Using the contouring, they calculate the geometric dimensions of the LV (systolic and diastolic volumes), ejection fraction, wall contractility, etc. The obtained quantitative indicators provide reliable assessment of the heart muscle condition.

Analysis of the experience of cardiac ultrasound shows that in most cases the LV boundary is conducted subjectively and depends on the skill of the physician who performs the processing of ultrasound images. In this context, the task of developing algorithms for automated delineation of the LV on echocardiography

images, that eliminate the element of subjectivity and increase the processing speed of ultrasound images of the heart, is relevant.

In this article the results of the kinematics analysis of mass center on the ultrasound images are discussed and the identification approach of inaccurate LV contour built in automatic mode are submitted.

2 Statement of the Problem

Let's consider the features of contouring procedures LV on the ultrasound image. Typical source frame of an ultrasound image is presented in figure 1. Figure 1 shows that the images are low-contrast. There is a variety of artifacts due to the presence of the papillary muscle in the heart tissue. The figure 1 also presents the expert LV contour. The expert had the right border of the LV, ignoring existing artifact in the image.

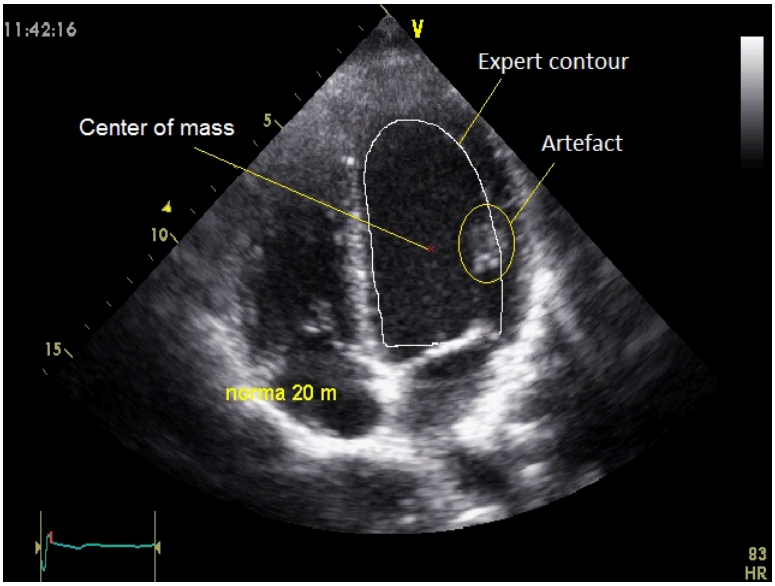


Fig. 1. The ultrasound frame with expert LV contour.

It should be noted that the presence of this artifact is critical for automatic algorithm. This problem is illustrated in figure 2, the results of the automatic delineation LV from the image are shown in figure 1.

Consequently, the automatic contouring process should have a pre-treatment procedure to remove speckle noise, artifacts and to enhance the contrast of ultrasound images. These algorithms are considered in [1], [4], [7] [8]. Analysis of the experience of their application shows that they really can improve the contrast

Fig. 2. The LV area is obtained from ultrasound image.

of the ultrasound images and in some cases remove artifacts. However, this can be done not for every patient.

Also the works [2], [3] should be particularly noted. For LV wall motion of the heart the optical flow algorithm [5] was used and analysis of LV shape changes over time. Work [9] describes the isolation area of the LV on echocardiography image. However, in these works the criteria to evaluate the correctness of contouring are not given.

Comparative analysis of the LV contour built by the expert and the contour built in automatic mode (see. Fig. 2), allows us to conclude that the CM coordinates of these contours are significantly different from each other. Thus, we can assume that the CM coordinate values are an informative parameter. This parameter allows to distinguish the contours of regular and irregular shapes (Fig. 3).

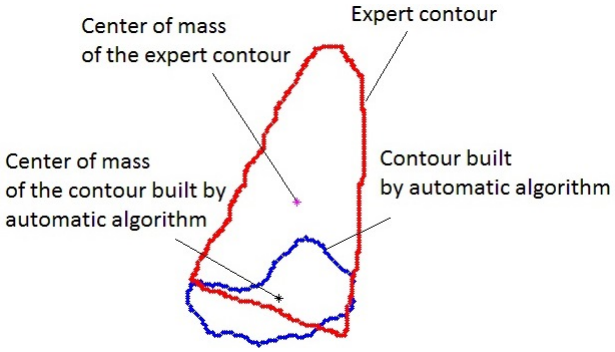


Fig. 3. Expert and contouring with the center of mass.

To test this hypothesis, the CM motion of expert contours was investigated. The results are presented in the next section.

3 The study of the CM LV motion.

In the course of the research ultrasound records of 16 patients were used, the total number of frames was 320. At each frame LV contours were built by experts. Figure 4 shows the coordinates of the CM expert LV contours of the heart. Figure 4 shows that the position of a CM is inside the minimal ellipse [6].

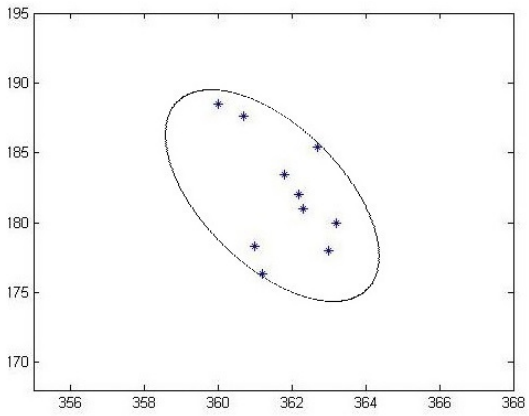


Fig. 4. CM of expert contours for one patient and ellipse constructed by CM points.

Similar calculations were performed for the remaining patients. Next, for each patient area of the respective ellipses was calculated. This area is compared with the area of LV in systole. Calculated coefficients are presented in table 1.

Table 1. The ratio of the ellipse area covering CM LV to the area of expert contour in systole.

Patient	Coefficient (expert data)	Coefficient (data obtained from automatic algorithm)	Patient	Coefficient (expert data)	Coefficient (data obtained from automatic algorithm)
B	0.0086	0.0424	K	0.0137	0.5185
C	0.0113	0.1180	L	0.0117	0.0237
D	0.0168	0.0110	N	0.0163	0.0349
E	0.0138	0.0117	O	0.0161	0.0189
F	0.0095	0.0136	R	0.0200	0.0235
G	0.0143	0.0674	T	0.0254	0.0205
H	0.0173	0.0533	V	0.0167	0.0981
I	0.0168	0.0987	X	0.0193	0.2188

Table 1 shows that the ellipse area encompassing LV CM does not exceed 2% of the left ventricular in systole. The results of similar calculations for contours constructed automatically, some of which have an irregular shape, are presented in table 1. (The automatic algorithm is considered in [10]).The maximum value of the ratio of the ellipse area, covered CM LV, to expert contour area in LV systole is 51%. Thus, the position of the CM allow to automatically identify the correct construction of the contour, using the following step:

1. Calculate the CM coordinates of the each video frame.
2. Construct a minimum ellipse which includes the CM of all contours.
3. Calculate the area of the ellipse.
4. Calculate the ratio of the areas of the ellipse to the area of LV in systole.
5. If the ratio is greater than 2% then apply the clustering procedure CM [11].
6. Identify the contours of the irregular shape which CM are assigned to a remote cluster.

Figure 5 shows the CM arrangement examples in the case, where the shape of one contour is incorrect. Figure 5 shows that CM for contours with regular shape are grouped in a certain region, while the contours of irregular shape are located at a remote distance from the grouping area.

It should be noted that the errors of second type (the contour is wrong, but CM belongs to the ellipse) were not found.

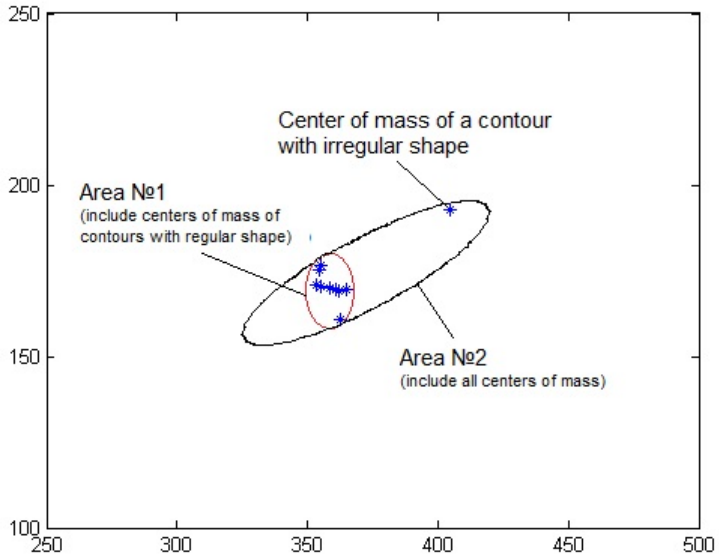


Fig. 5. Location of CM for patient X, one circuit is built wrong.

4 Conclusion

The CM for regular shape countours are grouped within the ellipse area. The CM for irregular shape countours are located out of this area. Methods of points clustering can solve this problem. An approach based on an analysis of the CM LV contours location, allowing to identify the contours of an irregular shape, is given in this work. The approach will be embeded in automatic algorithm of contouring the LV area.

References

1. Amerbaev, V., Kalnej, S., Rychagov, M., Frolova, G.: Restoration of medical ultrasound images based on the effective deconvolution of scan data. Medical equipment (3), 9–12 (2004)
2. Bosnjak, A., Colmenares, L., Montilla, G., Venezuela, V.: Spatial and temporal estimation of left ventricle wall from ultrasound images using optical flow algorithm. Computing in Cardiology (CinC) pp. 573–576 (2012)
3. Comaniciu, D., Zhou, X.S., Krishnan, S.: Robust real-time myocardial border tracking for echocardiography: an information fusion approach. Medical Imaging, IEEE Transactions on 23(7), 849–860 (2004)
4. Gonzalez, R., Wood, R.: Digital image processing, vol. 1072 (2005)
5. Lai, S.H., Vemuri, B.C.: Reliable and efficient computation of optical flow. International Journal of Computer Vision 29(2), 87–105 (1998)
6. Loa, C.F.V.: Using the ellipse to fit and enclose data points, `http://www.cs.cornell.edu/cv/otherpdf/ellipse.pdf`
7. Porshnev, S., Zyuzin, V., Bobkov, V., Bobkova, A.: Analysis of methods for removing noise and artifacts on echocardiographic images. In: The 11th International Conference: "PATTERN RECOGNITION and IMAGE ANALYSIS: NEW INFORMATION TECHNOLOGIES". pp. 525–528. No. 8-1 (2013)
8. Priorov, A., Hrjashhev, V., Sladkov, M.: Improving the quality of ultrasound medical images. Medical equipment (4), 11–14 (2008)
9. Varlamov, A., Makarova, E.: An automatic selection method of objects on heart's ultrasound images. Algorithms, methods and data processing systems 17, 49–54 (2011)
10. Yatchenko, A.: Numerical methods of analysis and heart's image processing. Ph.D. thesis, Lomonosov Moscow State University (2013)
11. Zagoruiko, N.: Applied methods of data analysis and knowledge (1999)

Accuracy Analysis of Estimation of 2–D Flow Profile in Conduits by Results of Multipath Flow Measurements

Mikhail Ronkin and Aleksey Kalmykov

Ural Federal University, pr. Mira, 19, Yekaterinburg, 620002,
Russian Federation
z_sm@mail.ru

Abstract. The paper presents a model of distorted velocity distribution of a flow in conduit, which passes through its bent section. The proposed model is properly analyzed. An algorithm for recovering the pipe profile by multipath ultrasonic measurements in the presence of *a priori* information is proposed. Numerical simulations with the proposed algorithm are performed as well.

Keywords: flow measurements, flow profile, multipath flowmeter

1 Introduction

In practice, it is important to know the set of values of the stream head of a flow in each section of the pipeline. The flow head is one of the main characteristics of the conduit. When the flow passes through a bent section (such as elbow or valve, etc.), its head gets lost. In practice, losses of the head are determined by experiments [1]. The loss of head is connected with the flow profile distortion. There is a symmetric flow profile in a long straight section of pipeline (Fig. 1a,b). However, when flow passes a bent section (such as elbow, valve, etc.), the flow profile, *i.e.*, the distribution of velocity in the cross section of the pipe, becomes distorted (Fig. 1 c,d). Thus, the profile of flow characterizes a bent section of the pipe [1].

One way to determine the flow profile is to recover it by ultrasonic multipath flow measurements. The ultrasonic technology allows measurement of the velocity distribution in a number of planes between two transducers/receivers of acoustic waves as it is shown in Fig. 2a. The distortion of a flow profile can be recovered by approximation of such measured values in a sufficient number of planes [2].

The task of ultrasonic measurement of the flow rate is the radar task of measurement of propagation time in an investigated media. The emitted wave accelerates when moving in the direction of a flow and slows down in the opposite direction (Fig. 2b). The velocity of a flow is calculated as a difference of measured times of propagation:

$$t_{12} = \frac{L}{c + v_l \cos\alpha}, \quad t_{21} = \frac{L}{c - v_l \cos\alpha}, \tag{1}$$

$$v_l = \frac{\Delta t c^2}{2L\cos\alpha}, \quad v_z = \frac{\Delta t c^2}{2L}\tan\alpha, \tag{2}$$

where t_{12} is the time of propagation in the flow direction; t_{21} is the time of propagation in the opposite direction; c is the speed of acoustic wave; α is the angle between direction of the flow and direction of the wave propagation; L is the length of the wave path in the investigated media; v_l is the projection of the flow velocity on the wave path; v_z is the velocity in z–direction of the flow [3].

The velocity profile is always distributed not uniformly through the cross–section area of a pipe. This means that velocity in the central region is usually higher than near the wall. The measured value of velocity v_l is the average of this distribution in the direction of the path. Measured value usually differs from average velocity through the whole cross section v_c:

$$v_l = \int_0^L v(l)dl, \; Q = \int_s v(x,y)dS, \tag{3}$$

$$k = \frac{Q}{\pi R^2 v_l}, \tag{4}$$

where $v(l)$ is the distribution of velocity in the direction of the wave propagation; Q is the flow rate; $v(x,y)$ is the distribution of velocity through the whole cross–section of the conduit; R is the radius of the pipe; S is the cross–section area of the pipe; k is the meter factor, which is connected with the measured and actual flow rate [3].

Equation (4) is valid for any symmetric flow. However, in the case of asymmetric propagation of the flow wave patch, the flow rate would be calculated with error. The type of function $v(x,y)$ for asymmetric flow does not have general analytical expression. The function of flow distribution strongly depends on the conduit configuration, and its characteristics (such as material, roughness, temperature of the controlled media, etc.). The error of the calculated flow rate for a single path flow meter can achieve a 10% and be even larger [3].

The meter factor can be calibrated in laboratory conditions. Such calibration is generally performed on a long straight section of a conduit where a symmetry flow profile takes place [3]. In the case of multipath flow meter, all asymmetric features are assumed to be present for calibration of a symmetric flow. In - practice, it is necessary to have more than 4 paths to consider the meter factor as 1 with error less than 0.1% [4, 5]. However, the solution of the task of recovering a flow profile with such accurate flow measurements is an open issue.

For some types of flow profiles, the analytical expression has been obtained by Salami [6]. Based on his works, some authors proposed analysis of configuration of the multipath measurements and suggested some interpretations of analytical profiles [4–8]. The solution of the profile recovering problem by ultrasonic flow measurements was only proposed on the basis of ultrasonic tomography or by Abel's integration transform, which doesn't take into account asymmetry features of flow [2, 9].

In this work, a technique is suggested for recovering the flow profiles using the multipath measurements and *a priori* information about discretionary function, which can be determined for each type of a bent section of the conduit.

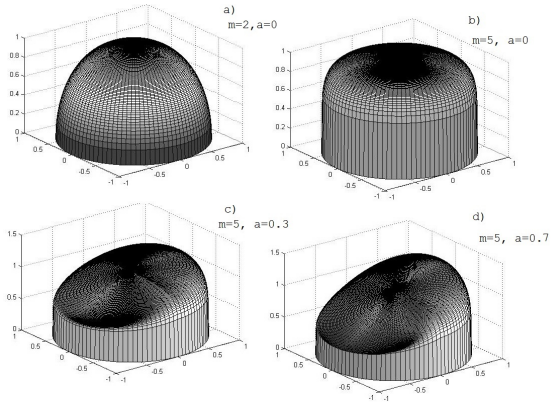


Fig. 1. Flow profiles for a) m=2, a=0; b) m=5, a=0;) m=5, a=0.3; d) m=5, a=0.7

2 Model of flow profile

The model of flow profile, which have been proposed by authors [6, 7] can be considered as a particular case of a flow profile that has passed through a conduit section, whose specified characteristic determines the distortion of the velocity distribution. The proposed model and generalized approach can be analyzed regarding to the accuracy of the profile recovering.

In general, the flow rate could be calculated as an integral of velocity distribution over the cross-section area as

$$Q = \int_{s(z)} v_z(x,y)dS = \int_{x_1=-R}^{x_2=R} \left[\int_{y_1=-\sqrt{(R^2-x^2)}}^{y_2=\sqrt{(R^2-x^2)}} v_z(x,y)dy \right] dx, \qquad (5)$$

Thus

$$Q = \int_{-R}^{R} v_z(x)dx = \frac{c^2 \tan(\alpha)}{2} \int_{-R}^{R} \Delta t(x)dx, \qquad (6)$$

where $\Delta t(x)$ is the distribution of difference values of measured propagation time in the flow direction and in the opposite one versus the coordinate. In theory of numerical integration, it is proposed to substitute the integral by sum with specified weight, which may be chosen, for instance, by criteria of optimal distribution of nodes locations and corresponding coefficients (the Gauss quadrature

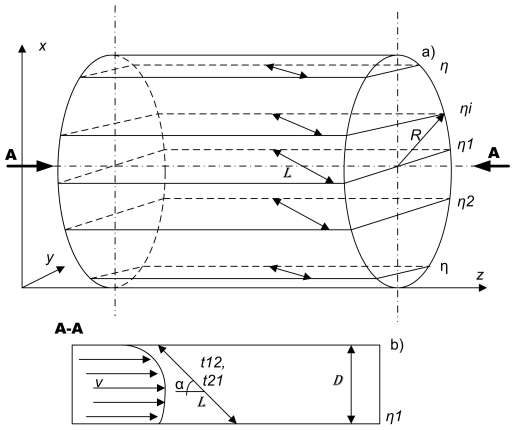


Fig. 2. Schematic draw picture a) multipath flowmeter and b) the center plane cut, η_i is the ith plane

method). The solution is presented in Table 1.

$$\xi = \frac{x}{R}\int_{-R}^{R} \Delta(x)dx \;\rightarrow\; R\int_{-1}^{1} \Delta(\xi)d\xi = R\sum_{i}^{n} \lambda_i \Delta t(\xi), \tag{7}$$

where ξ is the normalized coordinate of the measurement plane, λ_i is the coefficient of weight function (determined by solution of the Gauss quadrature task for number of measurements); $\Delta t(\xi_i), i = 1, 2 \ldots n$ are differences of propagation times in planes with coordinate ξ_i; n is the number of planes. In the case of a

Table 1. Solution of Gauss quadrature task [10]

	n=2	n=3		n=4		n=5		
λ_i	± 0,5773	0	± 0,7746	± 0.3400	± 0.8611	0	± 0.5385	± 0.9062
ξ_i	1	0.8889	0.5556	0, 6521	0,3479	0.5689	0.2369	0.4786

multipath flow meter, the flow rate can be calculated in accordance with (6) as

$$Q = k\frac{c^2 R\ tan(\alpha)}{2}\sum_{i}^{n} \lambda_i \Delta t(\xi_i) = k\sum_{i=1}^{n} \lambda_i v_z(\xi_i), \tag{8}$$

where the meter factor k can be considered as 1 if there are more than 4 measured planes [4]. There are many profiles proposed by Salami [6], which have been verified by experiments. For instance, authors [2] noted that profile of flow that passes through a single elbow could be expressed as follows:

$$\frac{v_z(r,\varphi)}{v_0} = \sin\left(\frac{\pi}{2}(1-r)^{1/m}\right) + \alpha\sin\left(\pi(1-r)^{1/2}\right)\exp(-0.2\varphi)\sin(\varphi), \tag{9}$$

where $v_z(r,\varphi)$ is the velocity distribution in cylindrical coordinates; r is the radius; φ is the angle; v_0 is the velocity of flow at the center of cross-section ($v_z(0,0) = v_0$); α is the velocity of flow at the center of cross-section; m is the coefficient of symmetric flow profile, which characterizes the flow profile depending on velocity if asymmetric coefficient is equal to zero. In the original work [6], the author proposed values $m = 5$, $\alpha = 0.3$. The first part of the right side of equation (4) corresponds to the symmetric part of the flow, and the second part corresponds to asymmetry distortion with coefficient α [6]. Without asymmetry distortion, the flow rate is calculated according to the equations (8)-(9) as

$$Q = \int_S \sin\left(\frac{\pi}{2}(1-r)^{\frac{1}{m}}\right) dS = \int_0^{2\pi}\int_{-R}^{R} \sin\left(\frac{\pi}{2}(1-r)^{\frac{1}{m}}\right) r dr d\varphi, \tag{10}$$

$$Q = kR\sum_{i=1}^{n} \lambda_i v_z(\xi_i). \tag{11}$$

For symmetric model of flow (10, 11) calibration $m = f(k) = f(Q/v_0)$ could be performed. Normalized profiles are shown in Fig. 1 a,b. It can be noted that profile with *m=2* corresponds to the laminar flow mode, and cannot be considered in model. Profile with *m=5* corresponds to the turbulent mode [2]. It is well known that the profile of a flow changes toward symmetry distribution with distance from a section of the conduit, which caused distortion [3]. So, it may be supposed that the influence of the second part of equation (9) decreases as asymmetric coefficient tends to zero with increasing distance from the distorted section of the conduit. Based on the condition of constant flow ($Q(z) = const$) it should be noted that m is also changing with distance from the section of the conduit that caused the distortion.

It may be proposed that type of the second part of the right side of equation (9) depends on the type of section of the conduit, which caused distortion. Such expressions can be defined theoretically or from experiments. In this work, type of the profile (9), which corresponds to signal elbow [2], is considered. Profiles of flow for $m = 5, \alpha = 0.3$ *and* $\alpha = 0.7$ are shown in Fig. 2c,d. It has been mentioned above that in each measurement plane, the value of velocity is the integral over the direction of acoustic wave propagation. The normalized measured velocity in the plane with coordinate x_i corresponds to equation (9) can be expressed as

$$\frac{v_z(x_i)}{v_0} =$$

$$= \int_{-A}^{+A} \left[\sin\left(\frac{\pi}{2}(1-(x_i^2+y^2)^{1/2}\right)^{1/m}\right] + \left[\alpha \sin\left(\pi\left(1-(x^2+y^2)^{1/2}\right)^{1/2}\right)\right] \times$$

$$\times \exp(-0.2\arctan\left(\frac{y}{x_i}\right) \sin\left(\arctan\left(\frac{y}{x_i}\right)\right) dy, \tag{12}$$

where $A = \sqrt{R^2 - x_i^2}$. Denote:

$$F_1\left(x_i, m\right) = \int_{-A}^{+A} \left[\sin\left(\frac{\pi}{2}(1 - (x_i^2 + y^2)^{1/2}\right)^{1/m}\right] dy, \tag{13}$$

$$F_2\left(x_i\right) = \left[\alpha \sin\left(\pi\left(1 - (x^2 + y^2)^{1/2}\right)^{1/2}\right)\right] \times$$

$$\times \exp(-0.2 \arctan\left(\frac{y}{x_i}\right) \sin\left(\arctan\left(\frac{y}{x_i}\right)\right) dy. \tag{14}$$

Consequently,

$$\frac{v_z\left(x_i\right)}{v_0} = F_1\left(x_i, m\right) + \alpha F_2\left(x_i\right), \tag{15}$$

where $F_1(x_i, m)$ is the function, which characterizes symmetric part of flow, $F_2(x_i)$ is the function, which characterizes an asymmetric part of flow when it passes through a bent section. Expression of the flow rate corresponding to (15) is

$$Q = \int_{-R}^{R} v_z(x)dx = v_0 \int_{-R}^{R} \left[F_1(x_i, m) + \alpha F_2(x_i)\right] dx \approx kR \sum_{i=1}^{n} \lambda_i v_z(\xi_i). \tag{16}$$

In accordance with (15), the flow rate can be expressed as

$$Q \approx v_0 R \sum_{i}^{n} \lambda_i \left[F_1(\xi_i, m) + \alpha F_2(\xi_i)\right]. \tag{17}$$

In equation (17), the meter factor is not used because since it is included in the calibration dependence for $F_1(x_i, m)$. From equation (17) in accordance with (11) and (16), the symmetric flow rate normalized on v_0 can be expressed as

$$\frac{Q_c}{v_0} = R\left[k \sum_{i=1}^{n} \lambda_i v_z\left(\xi_i\right) - \alpha v_0 \sum_{i=1}^{n} \lambda_i F_2(\xi_i 2)\right] = f^{-1}(m), \tag{18}$$

where $f^{-1}(m) = Q_c/v_0$ denotes the calibration relation in coordinates $f(Q_c/v_0)$. Such calibration can be implemented because of symmetry of function $F_1(x_i, m)$ and symmetry of profile Q_c. Hence, if we know the flow rate under symmetric conditions and radius of the pipe, we can deduce v_0.

There are three unknown parameters in equation (17): v_0 is the value of velocity at the center of the pipe cross-section; m is the coefficient of symmetric part of flow; α is the asymmetric coefficient. Here, it is assumed that type of function $F_2(x_i)$ is known as a function of distorted section of the conduit. Relation $m = f(F_1(x, m)) = f(Q/v_c)$ assumed to be known from preliminary calibrations on a long straight pipeline. The determination of parameters named above requires solution of the system of equations for each parameter. The position of planes (x_i) is supposed to be chosen here in accordance with Table 1

and from symmetry of $F_1(x_i, m)$. The solution of the system mentioned above has the following expression:

$$\begin{cases} \alpha = [(v_z(x_3) - v_z(x_1))\, F_1(x_1, m)]/[F_2(x_2)v_z(x_1) - v_z(x_3)F_2(x_1)], \\ v_0 = [F_2(x_3)v_z(x_1) - v_z(x_3)F_2(x_1)]/[F_1(x_1, m)\,(F_2(x_1) - F_2(x_3))], \\ \sum\limits_{i=1}^{n} \lambda_i F_1(x_i, m) = \Big[k \sum\limits_{i=1}^{n} \lambda_i v_z(x_i) - \alpha v_0 \sum\limits_{i=1}^{n} \lambda_i F_2(x_i)\Big]/v_0 = f^{-1}(m). \end{cases} \tag{19}$$

Here, the first and the second equations are calculated depending on F_1 and hence on m. However, the multiplier αv_0 may be calculated without the third equation:

$$\alpha v_0 = \frac{v_z(x_3) - v_z(x_1)}{F_2(x_1) - F_2(x_3)}. \tag{20}$$

In the third equation, the meter factor can be considered as 1 if more than 4 measurement planes are used. Consequently, in the third equation m is determined from the calibration relation. The solution of system (19) allows one to calculate velocity distribution in accordance with (15) for distorted flow profile in cross-section of the conduit.

3 Results of Modeling

In the Matlab software, simulation of the algorithm described above is carried out for function of distorted profile (9). The first stage of modeling requires the number of measurement planes and their positions. In our study, *4* planes ($k \approx 1$ in this case) have been chosen according to the Table 1. For the chosen model of distortion relation, functions $m = f(Q/v_0)$ have been calibrated, shown on Fig. 3.

$$x_1 = 0.34R,\ x_2 = 0.86R,\ x_3 = -0.34R,\ x_4 = -0.86R,$$

$$\lambda_1 = 0.652,\ \lambda_2 = 0.347,\ \lambda_3 = 0.652,\ \lambda_4 = 0.347.$$

In Figure 4 function $F_2(x)$ for the chosen type of distortion section of the conduit is shown. In the model of profile (9), coefficients were selected as $m = 5,\ \alpha = 0.3,\ v_0 = 1$, along with the measured values of velocities:

$$v_z(x_1) = 1.9;\ v_z(x2) = 1;\ v_z(x3) = 1.4;\ v_z(x_4) = 0.5.$$

The calculated value αv_0 accordingly to (20) $\alpha v_0 = 0.3$. The calculated value of the flow rate (considering that $k = 1$) is $Q = 2.7307$.

The relation for m (Fig.3) has been found from the third equation of system (9). For calibration, the value $m = 2.83$ has been determined. The difference of the measured and the preliminary chosen m value may be explained by not sufficient accuracy of the assumption that $k \approx 1$ or with influence of F_1 on F_2, which is not taken into account in the proposed model.

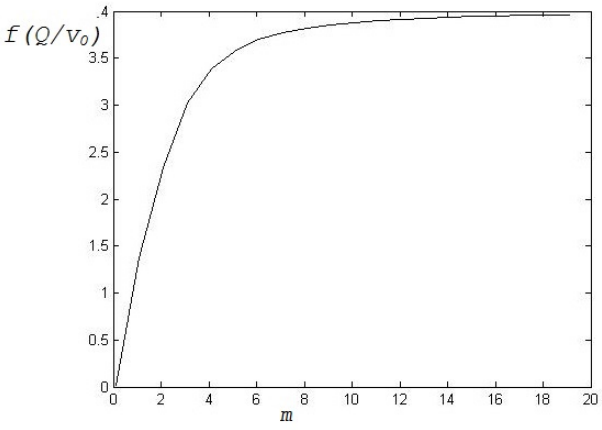


Fig. 3. - Relation $m = f(Q/v_0)$

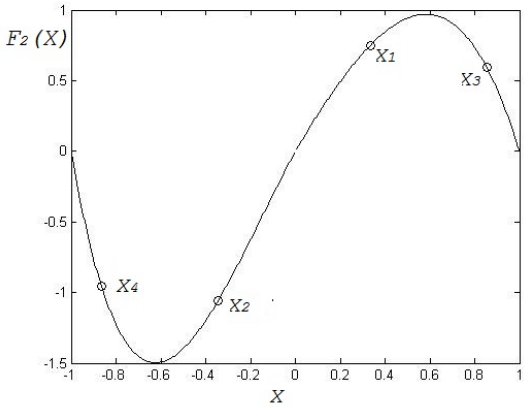


Fig. 4. Function $F_2(x)$ for the chosen type of distortion section of the conduit

The calculated values of $F_1(m, x_i)$, in accordance with (13) and the determined m, are

$$F_1(m, x_1) = F_1(m, x_3) = 1.51;\ F_1(m, x_2) = F_1(m, x_4) = 1.1.$$

From the estimated value of m $F_1(m, x_i)$ and by calculation of equations (20) $\alpha = 0.26, v_0 = 1.16$. The error of determination α, v_0 obviously connected with accuracy of estimation m.

The error of recovering the profile of a flow (9) with the calculated values of parameters and set values estimated as

$$\left(1 - \frac{Q_{teor}}{Q_{calc}}\right) 100\% = 6.7\%, \tag{21}$$

where Q_{teor} is the flow rate calculated form the selected values m, α, v_0, and Q_{calc} is the flow rate calculated from estimated values of parameters. The relation of error versus the asymmetry coefficient and $m = 5,\ 7,\ 10$ shown in Fig. 5. The

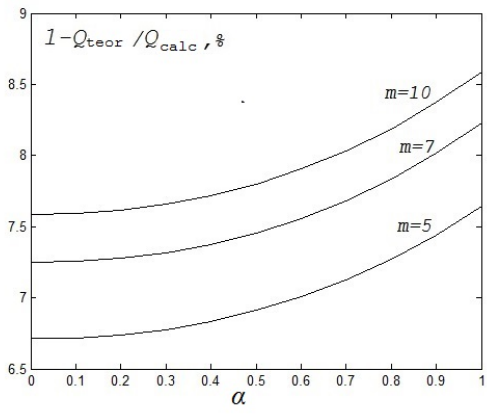


Fig. 5. Relation of error depending on asymmetry coefficient with $m = 5,\ 7,\ 10$ and $v_0 = 1$

accuracy analysis shows that the error dependence has a constant part that corresponds to symmetry flow distribution $m = f(Q/v_0)$. Such constant value of the error can be decreased by correction of calibration relationship $m = f(Q/v_0)$. The result with such calibration for each value m is shown in Fig. 6. In the whole, such behavior of the error can be caused by the insufficient accuracy of the model assumptions. Particularly, it may be supposed that the asymmetry part of $F_2(x)$ in (13) is influenced by the symmetry part of $F_1(x, m)$.

The general algorithm of flow profile recovering on the basis of multipath ultrasonic flow rate measurements and *a priori* information about distorting function of the conduit section has the following stages:

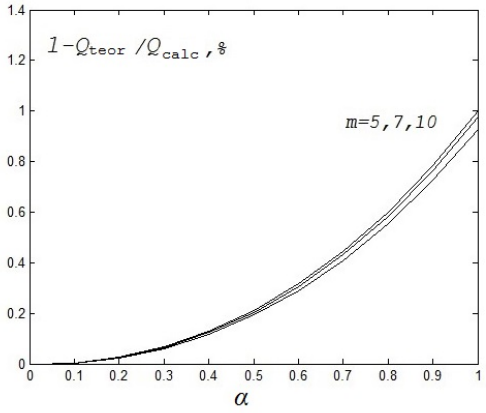


Fig. 6. Relation of error with corrected $m = f(Q/v_0)$ depending on asymmetry coefficient with $m = 5,\ 7,\ 10$ *and* $v_0 = 1$.

1. selection of the function $F_2\,(x, m)$ type;
2. carrying out calibration for symmetric flow $m = F(Q)$;
3. measurements of flow velocities in planes for each x_i and determination the flow rate by equation (16);
4. calculation value αv_0 from (20);
5. determination of m in agree with (18) and calibration of the relationship;
6. evaluation of correction relation $m = F(Q)$ for decreasing symmetric flow error ($\alpha = 0$);
7. calculation of values $F_1(m, x_i)$, α and v_0 by equations (19);
8. calculation of velocity distribution and flow profile.

4 Conclusion

The new model of behavior of velocity distribution in conduit and the algorithm for flow profile recovering based on it are proposed in the paper. The model generalizes analytical expressions for distorted flow profiles in the conduit, which were proposed by Salami. In contrast to [6], here, some coefficients of the model has been generalized and their physical interpretation provided.

The algorithm for profile recovering on the basis of multipath ultrasonic measurements and *a priori* information about distorting section of conduit is proposed. The model can be used for any type of profile with theoretical or experimental description.

The numerical simulation for generalized expression of the function for profile after single elbow has been carried out [2]. The results of simulation have been presented in the paper. The flow profile after single elbow, measured by ultrasonic in 4 planes is considered. The error is estimated as the relation between the flow

rate calculated with theoretical set of parameters and parameters that were determined by the proposed algorithm. The obtained accuracy gives less than 1% error.

The reached that the model can be considered as reliable. In further developments of the model, the influence of symmetry part of flow on asymmetry part will be investigated. Moreover, investigation of number of measurements and orientation of planes influence on accuracy should be done.

References

1. Idelchik, I.E.: Handbook of hydraulic resistance. Mashinostroenie, Moscow (1992) (in Russian)
2. Rychagov, M., Tereshenko, S.: Multipath flowrate measurements of symmetric and asymmetric flows. Invers Problems, 16, 495–504 (2000)
3. Kremlevskij, P.P.: Flowmeters and counters of the amount of substance. Handbook. Politeh, Saint-Peterburg (2004)
4. Brown, G.J., Augenstein, D.R., and Cousins T.: An 8-patch ultrasonic master meter for oil custody transfer. XVIII IMEKO WORLD CONGRESS, 17–22 (2006)
5. Rychagov, M.N.: Ultrazvukovye izmereniya potokov v mnogoploskostnyh izmeritel'nyh mod-uljah. Akusticheskij zhurnal 44 (6), 829–836 (1998) (in Russian)
6. Salami, L.A.: Application of a computer to asymmetric flow measurement in circular pipes. Trans. Inst. MC. 6(4), 197–206 (1984)
7. Moore, P., Brown and J., Stimpsom, B.: Ultrasonic transit–time flowmeters modeled with theoretical velocity profiles: methodology. Meas. Sci. Technol., 11, 1802–1811 (2000)
8. Zheng, D., Zhao, D. and Mei, J.: Improved numerical investigation method for flowrate of ultrasonic flowmeter based on Gauss quadrature for non-ideal flow fields. Flow measurement and Instrumentation. 41, 28-25 (2000)
9. Lui, J., Wang, B., Cui, Y. and Wang, H.: Ultrasonic tomographic velocimeter for visualization of axial flow fields in pipes. Flow measurement and instrumentation, 41, 57-66 (2000)
10. Krylov, V.I.: An approximate calculation of integrals. Nauka, Moscow (1967)

Semi-automated Integration of Legacy Systems Using Linked Data

Ilya Semerhanov and Dmitry Mouromtsev[1]

ITMO University, Saint-Petersburg 197101, Russia,
ailab@mail.ifmo.ru

Abstract. A lot of valuable data is stored in standalone legacy systems inside enterprise infrastructure across different domains. It was always a big challenge to integrate such systems with each other even on structural level, but with the recent development of Semantic Web technologies its now clear that integration on semantic level could be achieved and data from different sources could be used more effectively. Its now becoming a trend to open data to the public according to Linked Data principals, but there is no common workflow that could be used with legacy systems. In this paper we propose our solution for semi-automated integration of legacy systems using Linked data principals. We analyze the motivation and current state of the art in the domain, present our method and algorithms for data extraction, transformation, mapping and interlinking. Finally we show our own implementation of the proposed solution and discuss the results.

Keywords: Semantic Web, Linked Data, Database integration, Ontology, Resource Description Framework

1 Introduction

Data integration nowadays is a big and challenging problem, due to the fact that a lot of business and personal data is stored in large amount of different computer systems: Customer Relationship Management systems, Supply Chain Management systems, Enterprise Resource Planning systems and etc. Also information is stored in web based applications such as intranet portals, blogs, personal web pages and so on. The amount of such data is growing rapidly every year and because of this industry is in need of a solution for accessing and managing distributed data. Data integration approaches can be separated in two different groups: consolidation approaches and virtualization approaches. Approaches from the first group are meant to physically transfer data from distributed data sources into one unified storage. On other hand approaches from the second group concentrate on providing a virtual way of accessing the data without any physical relocation. In this paper we understand integration according to the approach from the second group, as a unified virtual interface for accessing distributed heterogeneous data sources. But we can go with classification even further and separate virtual data integration approach into several classes: structure based and semantic based integration. The structural integration is based

on an idea that full integration can be achieved by providing a mapping between different structural parts of integrated systems. A lot of approaches and methods were developed in past years with this idea and the most popular that are used now are usually based on such technologies as Web Services, for example Service Oriented Architecture. For data representation general approach in such a case is to use a markup language, for example XML, and for data transferring - some transport network protocol, like, for example, HTTP. Thus we can say that structure based integration methods are mostly used for representing and transferring data from different systems in a common way and providing a solution for mapping one system structure to another system structure. Semantic data integration, on other hand, is dealing not only with structural mapping, but also trying to define equivalency in semantic meaning of distributed integrated data. This kind of integration, compared to the structural integration, is operating several levels above in DIKW pyramid, using not only simple data structures for integration, but also context of its usage and knowledge of how to use it [1]. This approach can bring several benefits, like for example it can increase level of automatic decision making and provide means for better analytical data search. In other words semantic integration methods can make integration tools more intelligent and help them deal more effectively with big amount of data.

This paper consists of 6 sections. Section 2 explains our motivation in this work and why we started the research. Section 3 focus on the state of the art it the area and related work. Section 4 covers our proposed solution for the problem, including method, algorithms for data integration and also implementation of it as a tool. Sections 5 and 6 concludes the paper and focus on discussion of results and future work.

2 Motivation

It is an indisputable fact that currently most of the applications and computer systems in different domains use traditional technologies such as relational databases for data storage and retrieving. From the perspective of Semantic Web they could be called legacy systems, or in other words outdated computer systems. Applying semantic integration approach to the legacy systems could take a big amount of the effort and manual work, and as a result it could be an obstacle for introduction of such approach, regardless the benefits that it could bring. There are already several solutions for mapping relational databases to Resource Description Framework (RDF) datasets available, as well as there is also a World Wide Web Consortium (W3C) R2RML language specification that intends to standardize the approaches for such mapping [2]. Nevertheless we are convinced that they all miss the unified methodology for data retrieval from different types of relational databases and publication on the Web. The level of automation in such tools is still no enough, so they require a lot of manual work through whole data integration life cycle. In this paper we present our own view on the problem and a solution for semi-automated integration of legacy systems. This solution will include the methodology and algorithms for data extraction and

transformation, as well as a prototype of a tool, that is used by us for evaluation of the results.

3 Related Work

Semantic data integration approaches are discussed for a long time by professionals in data integration domain and we can already say that the most promising solutions are based on metadata usage. Some brief survey of the current state of the art in the domain was presented by Klaus R. Dittrich and Patrick Ziegler [3]. One of the directions of development of metadata approach is closely linked to the idea to use ontology models for data integration and system interoperability. This idea is not really a new one and was already presented for example in the paper [4]. Usually there are 3 main approaches for ontology based integration: single ontology approach, multi ontology approach or hybrid approach. The idea behind single ontology is based on the assumption that its possible to create single ontology that could describe all entities and concepts in integrated systems. This is a straightforward solution that could be applied only if all systems works in one domain and it was implemented in SIMS [5]. In second approach, instead of one ontology each integrated system provide its own ontology. Compared to first solution in this case its easier to integrate new system, because ontologies could be developed independently for all integrated systems, like for example in OBSERVER [6]. But without common vocabulary its problematic to interlink concepts in different ontologies. Thats why there is a third solution that uses both ideas: for describing systems it uses several single ontologies; for interlinking them with each other global ontology is used as a vocabulary. But still there are a lot unresolved problems in this approach, for example its still not clear how to get single ontology for every system automatically and how to map them automatically with each other with the use of global ontology. The solutions for relational database mapping to RDF data sets are described in a W3C RDB2RDF Incubator Group survey [7] and in survey [8]. In this surveys several tools and techniques for data transformation are described and could be separated in three classes: direct mapping solutions, semi-automated direct mapping solutions and solutions that use domain semantic for extracting the data from relational database. Although some of them are already a mature software they still lack solutions for stable automated ontology mappings, duplicates discovery and other features that could make data integration process more friendly.

4 Implementation

4.1 Semantic integration with ontology

According to Tomas Gruber definition, ontology is an explicit specification of conceptualization [9]. It can describe concepts and relationships that can exist in one domain and basically it is a set of concepts, relations and functions that

connects them with each other:

$$O = \{T, R, F\}, where: \tag{1}$$

O – ontology of the domain;
T - set of concepts in described domain;
R - set of relationships in the domain;
F - set of functions that connects concepts and relationships inside one domain.

By using it for data integration concepts and relationships could be described in every integrated subsystem as well as in the whole subject domain. In order to do that every object that has some valuable data in the integrated systems should be described with the use of semantic metadata, in terms of one general ontology of subject domain in which all systems work. Metadata is information presented in special format that describes content of objects, it also could be called data, that describes data. In formal form we can express it as:

$$M_i = T_i \vee E_i, where: \tag{2}$$

M_i - metadata of object i;
T_i - set of concepts connected to object i;
E_i - set of concept instances in ontology.
In other words ontology of subject domain plays a role of coordinate system for all integrated applications. There are a lot of benefits of using ontology for semantic data integration:

- Instead of simple structure mapping we have a relationship mapping that describes how to use data. In other word we are moving from data level to information and knowledge level in DIKW pyramid.
- Ontology could be parsed automatically and it could provide ways for automatic decision making with the use of description logic.
- Ontology can be easily extended with new concepts and relations, thus we can integrate new subsystems without additional effort.
- One of the big advantages is that ontology can have several levels. Low level ontology could describe every individual system, middle level ontology could describe the set of integrated systems in one subject domain and high level ontology could describe integrated systems even between different domains.

In order to use such approach for integration, data that is originally stored in distributed legacy systems should be transformed to semantic friendly format first and then ontology for describing this data and systems should be created. There are numbers of methodologies for manual ontology creating, for example IDEF5, TOVE, METHONTOLOGY [10, 11]. However, there is no standard approach for semi-automatic ontology generation from legacy systems that should be later integrated, therefore we present in this paper our own solution for that.

4.2 Linked data approach

One of the good things behind ontology based semantic integration is that it can be used to integrate not only structured data, but also semi-structured and unstructured. Using Linked Data principals as the basis for integration gives us a possibility to apply the same techniques for integrating any kind of data. Linked Data describes a method of publishing data on the web so that it can be interlinked and become more useful. It is based on W3C standard technologies, such as HTTP, RDF, Web Ontology Language (OWL), and according to Tom Heath and Christian Bizer this is one of the best solutions for publishing interlinked data on the web [12]. By defining relationships between data, in the way that it can be parsed by the computers, Linked Data gives a possibility to aggregate the information from distributed data sources, create new relations, visualize connections and also extend it by connecting to other external resources. In this approach it is also possible to use ontology and Ontology Web Language for describing relations between objects in order to achieve semantic integration goal. In other words Linked Data will play a role of unified virtual interface for accessing data, stored in distributed legacy systems. The big challenge of this research was a legacy systems data extraction, transformation and load problem. In order to overcome it we developed our own method and several algorithms for ETL (Extract, Transform, Load) procedure. As the overwhelming majority of legacy systems, such as Enterprise Resource Planning systems (ERP), Customer Relationship Management systems (CRM) and custom applications, make use of relational databases to store the data we focused on this kind of data storages.

4.3 Method

In legacy computer systems, that use relational databases, data is stored in tables, where rows represent entity instance and columns represent attribute values, describing each instance. To achieve integration with Linked Data principals data should be extracted from tables and published on the Web in appropriate format. Usually in one infrastructure there are several legacy systems, that store different kind if data, but in one subject domain. That allows us to use domain ontology for describing common relations in all systems in the domain. Furthermore some general relations are independent from the domain and could be used in any computer system. We intend to combine legacy system data model, domain ontology and upper ontology in order to automatically extract data from distributed relational databases and publish it as integrated Linked Data on the Web. For this purpose we developed a method, that is based on the IDEF5 method for ontology modeling, but instead of manual ontology creation it provide steps for semi-automated ontology generation from data storage structure. This method consists of four main steps:

1. Individual RDF structure. Extracting information about data model structure of every integrated legacy system and transforming it to RDF model;
2. Common RDF structure. Combining extracted RDF models in a common RDF model;

3. Common global ontology. Creation of global ontology model on the basis of upper ontology, subject domain ontology and extracted common RDF model of legacy systems;
4. Integrated data ontology. Extracting data from distributed data sources and presenting it as semantic metadata with the use of common global ontology and automated decision making tools.

According to the method on the first step there is a primary translation of legacy system data model to the RDF model. In case of relational databases table names transforms to RDF classes, table fields to RDF properties. Then, on the second step extracted RDF models automatically combined into one common RDF model. On this step we also provide a solution for searching for similar properties and classes and providing relations between them. On the next step generated RDF model should be enriched by different concepts and relations from subject domain ontology and upper ontology, such as Friend Of A Friend ontology (FOAF). This step is not automated and should be done manually with the use of ontology editors. On the last step, based on the common global ontology, data is extracted from distributed legacy systems and described with semantic metadata. This metadata is stored together with common global ontology in OWL format and could be published on the web as Linked Data.

Provided method make use of semantic relations, described by combined ontology, between objects in integrated systems across the domain, for automated data extraction and interlinking with other data. The benefit from this approach is that there is no need for manual interlinking between extracted data, instead semantics relations between objects will be used for this purpose. Semantic relations, on other hand, will be extracted automatically from initial data model and then extended manually by the domain experts.

4.4 Algorithms

Within the proposed method we also provide several supporting algorithms that should be used for legacy systems integration:

- Algorithm for automated common RDF model extraction from legacy systems data model;
- Algorithm for data extraction from distributed legacy systems with the use of common global ontology and publishing it on the Web as Linked Data.

We also used similarity analysis procedure in our method for comparing entities in integrated systems.

Common RDF model extraction algorithm was developed in order to extract data model from each of integrated systems as RDF, and combine it together as unified RDF data model. Given there are two legacy systems LS_1 and LS_2 in one subject domain, that use relational databases for storing the data. In this case the algorithm goal is to transform LS_1 and LS_2 database structure

into common RDF model. In relational databases for description of its structure database schema is used. This schema defines the tables, fields, relationships, views and a lot of other elements, however this is only enough for integration on structural level. In order to achieve semantic integration domain ontology should be used and in this case it will add to the extracted model relations between concepts in the subject domain. The input data for proposed algorithm is database schema of each integrated systems and domain ontology.

Legacy system LS_1 is using database schema S_1 and system LS_2 – database schema S_2:

$$S_1 = \{Tb_1, \ldots, Tb_n\}, S_2 = \{Tb_1, \ldots, Tb_k\}, where: \quad (3)$$

S_1, S_2 - database schemes,
Tb_n , Tb_k - schema tables.

$$Tb_1 = \{At_1, \ldots, At_i\}, where: \quad (4)$$

At_i – table attributes.

The algorithm consists of five steps:

1. Structure mapping. Sequential mapping of S_1 and S_2 into RDF format.

$$Tb_{\mathrm{n}} \rightarrow T_{\mathrm{m}}, Tb_{\mathrm{k}} \rightarrow T_{\mathrm{m}}, At_{\mathrm{i}} \rightarrow A_{\mathrm{i}}$$

   where T_m - RDF classes, A_i - RDF properties.
2. Automatic relations creation. Providing semantic properties P_j , by automatic similarity analysis of database structure. Analysis is based on several measures: data types similarity, database names similarity, string similarity.
3. Enrichment. Subject domain ontology and upper ontology import with the use of OWL property *owl:import*.
4. Manual relations creation. Editing of extracted model with the ontology editor. Manual creation of relations between concepts from upper and domain level ontologies and objects in extracted model.
5. Output of created common RDF model into file or in RDF triple store.

The output is a common RDF model that describe objects and their relations in legacy systems, as well as concepts and their relations inside subject domain.

Data extraction algorithm was designed for data extraction from distributed legacy systems with the use of previously generated common RDF model, and for later publishing on the Web. The input data for this algorithm is common RDF data model and list of tables from which data should be extracted. Suppose in input there are i number of tables with data, then:

$$Tb_i = \{V_1, \ldots, V_n\}, where: \quad (5)$$

Tb_i - table database i;
V_n - table entry.

The algorithm consists of five steps:

1. Import of common global model. Import of extracted earlier common global RDF model.
2. Entries extraction. Extraction of every V_n entry from table Tb_i in each integrated database.
3. Similarity analysis. Similarity analysis of extracted entries with each other. If there is a match, one of the tree semantic properties should be applied: *skos:closeMatch*, *skos:narrowMatch*, *skos:exactMatch*.
4. Reasoning. Creation of new semantic properties by logical reasoning, that is working because of description logic, which were imported from common global model.
5. Output of created integrated data ontology in OWL format in a file or in RDF store.

The result of the algorithm is global meta model that contains objects and its relations in the subject domain and also interlinked data in RDF format.

Similarity analysis was done in both algorithms in the method. In common RDF model extraction algorithm it is used for automated creation of relations between objects in different integrated systems. In data extraction algorithm it is used for comparing extracted string values.

For automated creation of relations between objects in different systems during common model extraction, we propose to use complex method of comparison by several parameters:

- string comparison of elements names and descriptions, for example attribute names in tables of relational database.
- comparison by data type. Different relational databases use different data types, however in the research we created common mapping between all relational databases data types and XML Schema data types.
- corpus-based comparison of elements. By using subject domain ontology as thesaurus, we can find relations between objects during extraction.

For string similarity comparison there are several approaches available, for example Jaccard similarity coefficient, Tanimoto similarity coefficient, Levenshtein distance or Sorensen-Dice similarity coefficient [13]. For all approaches we propose to use it not against a single character in a string or a whole line, but against a string, separated in intersecting N-grams. N-gram is a contiguous sequence of N items from a given sequence of text and it can be of arbitrary length. We propose to select length dynamical, based on the length of the initial string and use w-shingling for tokenizing it in N-grams. For example for Sorensen-Dice similarity coefficient, that originally looks like:

$$p = \frac{2 \cdot |X \wedge Y|}{|X| + |Y|}, where: \tag{6}$$

p – similarity coefficient,
X, Y – number of characters in string X and Y.
The solution with N-grams and Sorensen-Dice coefficient should look like:

$$p = \frac{2 \cdot |Ngrams(X) \wedge Ngrams(Y)|}{|Ngrams(X)| + |Ngrams(Y)|}, where: \tag{7}$$

p – similarity coefficient,
Ngrams(X) – function calculating N-grams sets from string X.

Every parameter, that is used for complex comparison has its own weight. In the end every compared element could be described as a set of parameter similarity coefficient and its weight:

$$sim(S, E) = \{p_1 w_1, \ldots, p_k w_k\}, where: \tag{8}$$

sim(S,E) – final similarity between element S and E in integrated system,
p_k – similarity coefficient of parameter k,
w_k – weight of parameter k.
In real implementation of the algorithms the final similarity value should be computed as a leaner function or with some more effective approach, for example linear regression [14].

4.5 Semantic data integration tool

For practical implementation and testing of our method and algorithms we developed a prototype of semantic data integration tool. The high-level architecture of the tool, called R2RMapper, is illustrated on figure 1. The tool consists of four main modules: R2RMapperCore, R2RMapperWi, R2RMapperBatch and MatchingDB. R2RMapperCore is the main module that is used for data extraction from distributed legacy systems. It can work like a standalone library or inside R2RMapper tool. Communication with integrated data sources is achieved by web services or by direct JDBC connection. R2RMapperWi module is a web frontend that gives access for the user to the main features. R2RMapperBatch is a module for scheduling of different tasks, for example nightly synchronization with integrated systems. MatchingDB is a Redis based memory caching mechanism that caches different information, such as similarity analysis results, in memory during ontology extraction. The extracted OWL ontology is stored as RDF in Jena TDB storage from where it is published by Jena Fuseki server on the web as a linked data cloud. This linked data cloud works as a virtual interface, with which client systems or users can access data in distributed legacy systems by SPARQL queries. As the linked data is backed up by the real ontology with the sets of concepts and relations, all benefits of ontology based integration are provided.

For testing of the tool we used a CentOS 6.2 server with 16GB RAM and Intel i5 processor 3.10 GHz. We executed it against Oracle 11g database with 50 000 entries. The results of the execution are presented on figure 2. Due to use

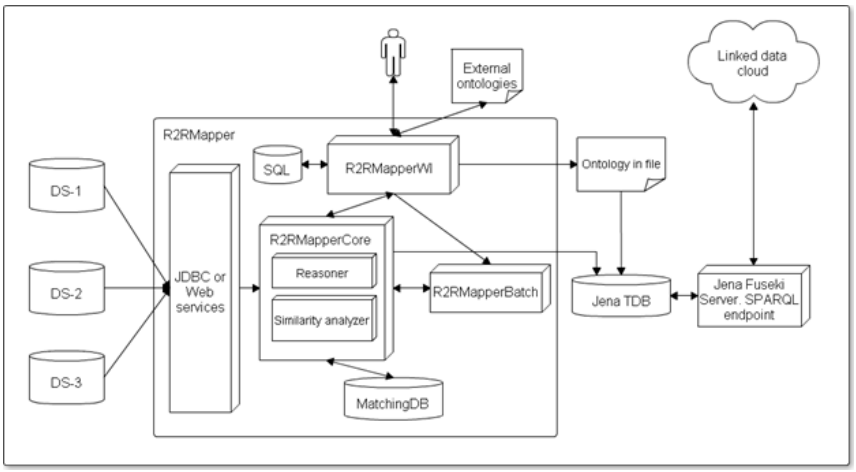


Fig. 1. R2RMapper architecture

of Redis as a caching mechanism we managed to decrease time for extraction of data from relational databases two times.

As a result of integration, user can work with integrated data through one endpoint and more effectively. For example there is a possibility to do an analytical semantic search query which will also include in search results related information about the required resource. Thus we can say that by using such approach users can work not only with the integrated data, but with integrated semantic knowledge of how to use this data.

5 Discussion

Today providing of an easy access to the big amount of stored data is becoming a trend, no matter whether it is inside one organisation infrastructure or distributed between different organisations and domains. It is equally important for big enterprise companies, open communities and research centers to open the data and make it available for other internal systems, external clients and of course users. In the research we were developing a unified method that could be applied to every domain for data extraction, transformation, publication and integration. The main goal is to automate the process as much as possible by using hybrid multilevel ontology approach, natural language processing and machine learning for automated decision making during integration. Although the method and its implementation is an already working solution it still has a room for improvement, for example in the area of semi-structured and unstructured data extraction. We also attempt to make global ontology extraction procedure more convenient for the users and domain experts.

As one of the first big practical use cases of presented method we intend to develop a platform for publishing of open science data from universities as

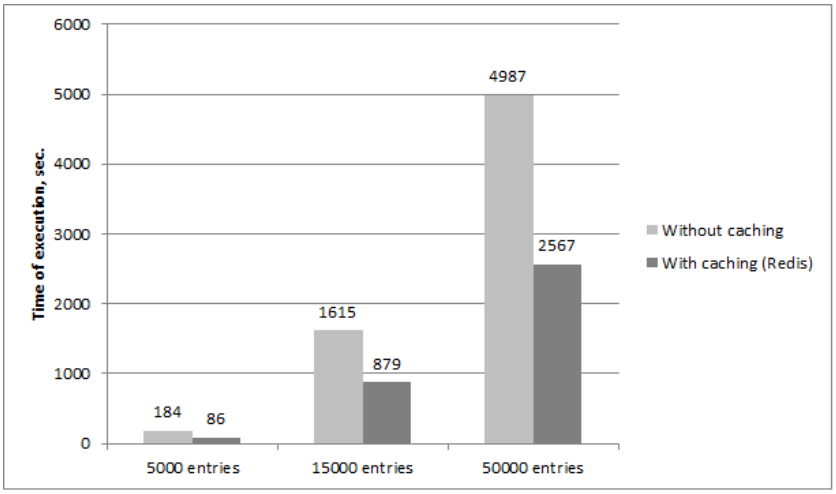


Fig. 2. Time of data extraction with R2RMapper tool

Linked Data. The first candidates for opening data by this platform are ITMO University and Leipzig University.

6 Conclusions and future work

In the presented paper we discussed semantic data integration approach, based on ontology usage and linked data technology. We showed the main benefits of this idea and presented our solution for providing distributed data in a semantic friendly linked data format. We also presented a developed prototype of a semantic data integration tool, which implements our solution. Our method currently works only for the structured data, but we are working on an extension of this method in the direction of working with semi-structured data. In future work we plan to research possibility of applying our solution also for unstructured data in order to parse it and represent as Linked Data.

In collaboration with AKSW [1] group from Leipzig we intend to improve the solution and use it in a generic platform for open science data integration, knowledge acquisition and collaboration beteen universities. Together with legacy system integration using ETL procedures this platform will be used for integration of existing RDF datasets by the use of SPARQL federated queries. This topic will be discussed in our future papers.

[1] http://aksw.org/

References

1. Rowley J. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2):163–180, 2007.
2. R2RML: RDB to RDF mapping language. http://www.w3.org/TR/r2rml/. Accessed: 2015-03-20.
3. Dittrich K. R. Ziegler P. Three decades of data integration - all problems solved? *World Computer Congress - IFIP*, pages 3–12, 2004.
4. Visser U. Stuckenschmidt . Wache H., Vogele . Ontology-based integration of information - a survey of existing approaches. *IJCAI-01 proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, pages 4–10, aug 2001.
5. Yigal Arens, Chun-Nan Hsu, and Craig A. Knoblock. Readings in agents. chapter Query Processing in the SIMS Information Mediator, pages 82–90. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998.
6. Eduardo Mena, Arantza Illarramendi, Vipul Kashyap, and AmitP. Sheth. Observer: An approach for query processing in global information systems based on interoperation across pre-existing ontologies. *Distributed and Parallel Databases*, 8(2):223–271, 2000.
7. Sebastian Hellmann Kingsley Idehen Ted Thibodeau Jr Sren Auer Juan Sequeda Ahmed Ezzat Satya S. Sahoo, Wolfgang Halb. A survey of current approaches for mapping of relational databases to RDF. Technical report, W3C RDB2RDF Incubator Group, jan 2009.
8. Catherine Faron-Zucke Franck Michel, Johan Montagnat. A survey of RDB to RDF translation approaches and tools. Under review in Semantic Web Journal, 2014.
9. Gruber T. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.
10. Perakath C. Benjamin, Christopher P. Menzel, Richard J. Mayer, Florence Fillion, Michael T. Futrell, Paula S. deWitte, and Madhavi Lingineni. *IDEF5 Method Report*, September 1994.
11. Uschold M. King M. Towards a methodology for building ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing.*, 1995.
12. Bizer C Heath T. Linked data: Evolving the web into a global data space (1st edition). *Synthesis Lectures on the Semantic Web: Theory and Technology*, pages 1–136, 2011.
13. Safa'a I. Hajeer. Comparison on the effectiveness of different statistical similarity measures. *International Journal of Computer Applications*, (8), 2012.
14. Anhai Doan, Jayant Madhavan, Pedro Domingos, and Alon Halevy. Learning to map between ontologies on the semantic web. *WWW-02 Proceedings of the 11th international conference on World Wide Web*, pages 662–673, 2002.

Adaptive Regularization Algorithm Paired with Image Segmentation

Tatiana Serezhnikova[1,2]

[1] Krasovsky Institute of Mathematics and Mechanics UB RAS
[2] Ural Federal University, Ekaterinburg, Russia
sti@imm.uran.ru

Abstract. This paper presents the subsequent improvements of our adaptive regularization algorithm. The current version of the algorithm contains an additional stabilizer. This stabilizer includes the coefficient, which utilizes prior information about segmentation of an image to be processed. We demonstrate that our adaptive regularization algorithm can exploit only the observed image segmentation for both the adaptive coefficient calculation and successful restoration of the original image segmentation.

Keywords: image restoration, Tikhonov regularization, adaptation, image segmentation

1 Introduction

In the model experiments, see [1-4], we demonstrated, that our adaptive regularization algorithm can successfully restore and provide more deblurred and denoisy intensity of the original model image. In this paper, we improve our algorithm, which is based on the results from [5-7], by an additional stabilizer. This stabilizer includes the coefficient, which utilizes some information about segmentation of an image to be restored.

Most artificial objects are made of surfaces, resulting in images with several subimages (or a single subimage) together with their boundary lines.

These facts motivate us to use ideas from image segmentation in order to improve our algorithm. In the sequel, we describe the stabilizer coefficient construction, using ideas from image segmentation.

Image segmentation algorithms are based on properties of intensity values: discontinuity and similarity, see [8-11]. Starting with the observed object intensity, setting $\beta(x, y) \equiv 0$ in our adaptive algorithm, we successfully apply the original approach for both sub-images and their boundaries. In the next iterations, the adaptive algorithm utilizes this initial result for calculation of the coefficient β.

Thus, we consider image segmentation both as result and tool in every image restoration.

We prove numerically that the proposed algorithm can reveal the latent image segmentation of observed images.

The paper is organized as follows. In Section 2, we describe regularizing functionals and present our adaptive technique. In Section 3, we describe the calculation process of our stabilizer coefficient. The last section concludes the paper.

2 Mathematical Model for Image Deblurring

We consider the following two-dimensional Fredholm integral equation of the first kind:

$$Au \equiv \int_0^1 \int_0^1 K(x-\xi, y-\eta)u(x,y)dx, dy = f(\xi,\eta). \tag{1}$$

In image restoration, the estimation of u from the observation of f is referred to as the two-dimensional image deblurring problem.

We construct and develop the novel technique for numeric solution of equation (1). Abstract methods with convergence analysis of regularization for this problem are presented in [3-7].

The foundation of the regularization method is given by

$$\min\left\{||A_h u - f_\delta||^2_{L_2} + \alpha\left(||u||^2_{L_2} + J(u)\right) : \ u \in U\right\}, J(u) = \int_D |\nabla u|\, dx, \tag{2}$$

where ∇u denotes the gradient of smooth function u, $(u \in W_1^1(D))$, $J(u)$ is the total variation of the function u on D. The practical implementation of this method requires minimization of the functional in (2).

The novelty was that we proposed to add the version of the iterative technique, containing additional parameters $\beta_{i,j}\,, \beta_{i,j} \geq 0\,$, see [1-4]:

$$\mathbf{u}^k = \arg\min\left\{\Phi_N^\alpha(\mathbf{u}) + \sum_{i,j} \beta_{i,j}\,(u_{i,j} - u_{i,j}^{k-1})^2 \ : \mathbf{u} \in R^N\right\}. \tag{3}$$

Here, $N = n^2$, $\Phi_N^\alpha(\mathbf{u})$ is the discrete form of the functional in (2).

$\sum\limits_{i,j} \beta_{i,j}\,(u_{i,j} - u_{i,j}^{k-1})^2$ is a discrete form for the additional stabilizer in (3).

We use the iterative subgradient method in order to compute $\mathbf{u}^k$ defined in (3), see details in [1-4].

3 Stabilizer coefficient selection

3.1 Image Segmentation: Motivation

The previous version of the algorithm (3) used the constant β in every mesh point of discrete models. The low resulting accuracy of this two-dimensional model led us to change the way of $\{\beta_{i,j}\}$ selection.

As we said above, most man-made objects are made of surfaces, resulting in images with a single image or some sub-images. The quality of the images restoration features the quality of sub-images restoration together with sub-image boundary lines restoration.

In [1], we proposed the more general form of the stabilizer I^{β}:

$$I^{\beta} = \int\limits_{Q} \beta(x,y)[u(x,y) - u^{k}(x,y)]^{2}\, dxdy, \tag{4}$$

we used

$$\begin{cases} \beta(x,y) = u_{\max}, & \text{for} \quad (x,y) \in Q, \\ \beta(x,y) = 0, & \text{for} \quad (\mathrm{x},\mathrm{y}) \in \Pi/\mathrm{Q}, \quad \text{where} \end{cases} \tag{5}$$

$$u_{\max} = \max\{u_{\text{true}}(x,y)\,, (x,y) \in \mathit{\Pi} = [0,1] \times [0,1]\}, \tag{6}$$

$$(x,y) \in Q \Leftrightarrow u_{\text{true}}(x,y) = u_{\max}. \tag{7}$$

So, in [1] our regularization algorithm stabilizer includes the coefficient β, which uses some information about true image segmentation, see (6) and (7). This facts motivates us to exclude true image parameters. We approach this purpose using image segmentation ideas.

3.2 Image Segmentation as Tool

In this paper we propose to begin calculations with the parameter $\beta \equiv 0$.

So, in this case, we use standard regularization method:

$$\min\left\{||A_h u - f_\delta||^2_{L_2} + \alpha\left(||u||^2_{L_2} + J(u)\right) :\ u \in U\right\}, J(u) = \int\limits_{D} |\nabla u|\, dx, \tag{8}$$

where ∇u denotes the gradient of smooth function u, $J(u)$ is the total variation of the function u on D.

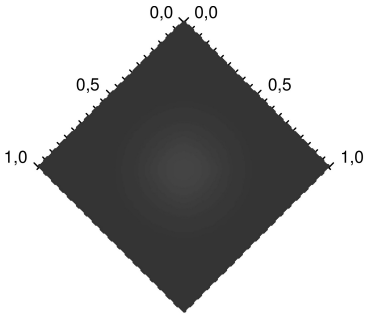
After obtaining the solution of problem (8), we get approximate $\tilde{u}^0(x,y)\,, Q^0\,, \mathit{\Gamma}_{Q^0}$, and we use them for the $\beta_{i,j}$ construction:

$$\beta_{i,j} \in \{o; c\},\ c = max\{\tilde{u}^0(x,y), \quad (x,y) \in Q^0 \cup \mathit{\Gamma}_{Q^0}\}. \tag{9}$$

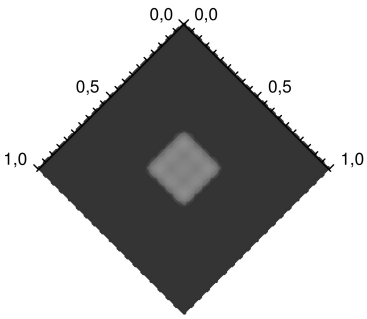
3.3 Image Segmentation as Result

Now, using β calculated in (9), the algorithm computes the final image restoration including its final image segmentation.

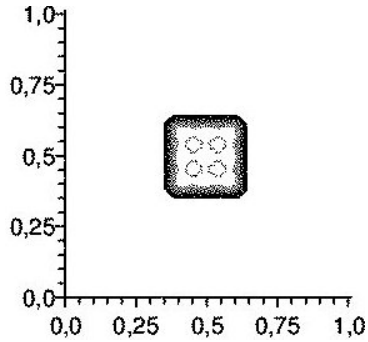
Numerical experiments have confirmed our ideas. Thus, we see in Figure 1:
(a) the observed image; (b) the restoration image;
(c) the central image segment restoration, $Q^0 \cup \Gamma_{Q^0}$, calculated with $\beta \equiv 0$;
(b) the central image segment restoration, $\tilde{Q} \cup \tilde{\Gamma}_{\tilde{Q}}$, calculated with adapted β.



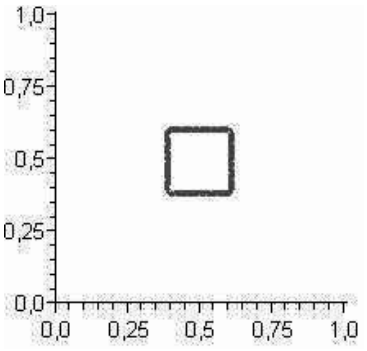
a) The observed image



b) The recovered image



c) The model segment restoration for $\beta(x, y) \equiv 0$



d) The model segment restoration for the adapted stabilizer coefficient $\beta(x, y)$

Fig. 1. Model experiments

It is clear, that we can write the final image segmentation of the recovered image in the form:

$$\Pi = \{\Pi/\tilde{Q}\} \cup \tilde{Q}, \tag{10}$$

were $\tilde{Q} \cup \tilde{\Gamma}_{\tilde{Q}}$ we see in Figure 1.(d), $\Pi = [0, 1] \times [0, 1]$.

3.4 Conclusion

In this paper, we have described and verified the image segmentation approach for the construction of stabilizer coefficient β.

Starting with the observed object intensity and setting the stabilizer coefficient $\beta \equiv 0$ in the regularization algorithm, we successfully apply the initial approach for both the sub-image and the sub-image boundary, Q^0, Γ_{Q^0}.

After that, we use calculated results for construction of the adaptive stabilizer coefficient β. At this step the image segmentation is a tool.

Using the calculated coefficient β, the algorithm computes the final recovered image. We obtain a good localization of segments, $\tilde{Q}$, $\Gamma_{\tilde{Q}}$, and we see a more deblurred final image segmentation.

As a result, in this paper, we have described and verified that our adaptive algorithm does not need a priori information. For successful image restoration and restoration of image segmentation, our adaptive algorithm can utilize the image of an observed object only.

We plan the following modifications and experiments:

- inclusion of the solution estimate for equation $u(x, y, \Gamma_Q(x, y)) = c$ to the estimate list;
- validation of β^k inclusion in (4), where k is the iteration number;
- processing of thin lines and narrow segments.

Acknowledgments The author expresses special gratitude to Prof. V.V.Vasin from the Institute Mathematics and Mechanics of Ural Branch of RAS. The author would like to thank the colleagues from AIST Program and Organizing Committees for their helpful advice and guidance in the paper preparations. This work was supported by the Program of UB RAS "Mathematical models, algorithms, high-performance computational and information technologies and applications", project 15-7-1-13.

References

[1] T.I. Serezhnikova. An algorithm based on the special method of the regularizationand the adaptation for improving the quality of image restorations.Univer.J.Comp.Math.,V.2(1) 11–16 (2014)

[2] T.I. Serezhnikova. Image restoration algorithm based on regularization and adaptation.In.: AIST 2014, CCIS 436, Springer, 213–221 (2014)

[3] V.V. Vasin and T.I. Serezhnikova. A regularizing algorithm for approximation of a nonsmooth solution of Fredholm integral equations of the first kind. J. VICH. TECH., V. 15(2), 15–23 (2010)

[4] V.V. Vasin and T.I. Serezhnikova. Two steps method for approximation of a nonsmooth solutions and nosy image reconstructions. J. Avt. Tel., No.2, 126–135 (2004)

[5] V.V. Vasin. Proximal algorithm with projection in problems of the convex programming.Ural Sci. Center AN SSSR, Instit. Mathem. I Mekhan., Sverdlovsk, pp. 47 (1982)

[6] A.N. Tikhonov and V. Arsenin. Solutions of Ill-Posed Problems. Wiley, New York (1977)

[7] C.R. Vogel. Computational methods for inverse problems. SIAM, Philadelphia (2002)

[8] R.C. Gonzalez, R.E. Woods. Digital image processing,3-ed ed. Prentice Hall,New Jersey(2008)

[9] X.Chen, M.K. Ng, C. Zhang. Non-Lipshitz l_p- regularization and box constrained model for image reconstruction. IEEE Tr. Im. Proc. V.21(12), 4709–4721 (2012)

[10] P. Arbelaer, M. Maire, Ch. Eowlkes, J. Malik. Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence. 33(5), 898-916 (2011)

[11] J.Canny A computational approach to edge detection. IEEE Trans. Pattern Analysis and Machine Intelligence. V.8(6), 679-698 (1986)

Algorithm of Interferometric Coherence Estimation for Synthetic Aperture Radar Image Pair

Andrey Sosnovsky and Victor Kobernichenko

Ural Federal University, pr. Mira, 19, Yekaterinburg, 620002,
Russian Federation
sav83@e1.r

Abstract. Interferometric coherence is an important indicator of reliability for interferograms obtained by interferometric synthetic aperture radar (Interferometric SAR, InSAR). Areas with low coherence values are unsuitable for interferometric data processing. Also, it may be used as a classification parameter for various coverage types. Coherence magnitude can be calculated as an absolute value of the correlation coefficient between two complex SAR images with averaging in a local window. The problem in coherence estimation is in its dependence on phase slope caused by relief topography (topographic phase). A method for suppression of the topographic phase influence is proposed, based on the spatial phase derivation.
Keywords: Synthetic aperture radar images, InSAR systems, Coherence estimation

1 Introduction

Interferometric data processing for extraction of information about the Earth terrain and its changes becomes one of the general guidelines in the development of contemporary space-based radar systems together with the implementation modes of ultra-high spatial resolution (1-3 meters) and full–polarimetric processing. The method of space-based radar interferometry implies a joint processing of the phase fields obtained by simultaneous scattering of the terrain with two antennas or by non-simultaneous scattering with one antenna moved by two different parallel orbits [1, 4]. This method combines high accuracy of the phase measurements with high resolution of the synthetic aperture radars (SAR) technology. Technology of the differential SAR interferometry (InSAR) makes possible to get maps of the elevation changes between the radar passages. The stages of the interferometric processing are implemented in specialized software systems for the remote sensing data processing such as the SARscape, IMAGINE Radar Mapping, Photomod Radar, RadarTools. So we can talk about Information technology (IT) of the digital elevation models (DEM) generation by remote sensing data.

However, for practical application of these technologies, one has to overcome a number of significant problems. Two general problems of interferometry are

the temporal and spatial decorrelation of the received data and the problem of phase unwrapping, *i.e.* a recovery of the absolute phase information from the relative phase, which is wrapped into $[-\pi, \pi]$-interval. Another important field of investigation in radar interferometry is selection of the most efficient processing algorithms and obtaining the experimental estimates of the generated DEM accuracy. This work is devoted to the first of the mentioned problems, *i.e.* to the data decorrelation (summary temporal and spatial) and its estimation.

2 Interferometric coherence

Interferometric coherence is an important indicator of suitability of the data scene obtained by a radar remote sensing system for the further processing and solving the final problem, *i.e.* generation of digital elevation model or terrain changes map. The coherence factor is calculated as the absolute value of the correlation coefficient between samples of two complex radar images (single-look data complex, SLC) got in the local windows

$$\hat{\gamma_0} = |\hat{\rho_0}| = \frac{\Sigma \dot{z}_1(m,n) \cdot \bar{z}_2(m,n)}{\sqrt{\Sigma |\dot{z}_1(m,n)|^2 \cdot |\bar{z}_2(m,n)|^2}}, \tag{1}$$

where $z_{1(2)}(m,n)$ is the SLC samples ($\bar{z}_{1(2)}(m,n)$ are complex-conjugate samples) [2–5], $\hat{\gamma_0}$ takes values in interval $[0,1]$, near-zero values correspond to areas of high or full decorrelation, which are not suitable for interferometric data processing. The values higher than 0.5 mean good data correlation.

However, this approach entails some problems because, in fact, a random variable is estimated here, but not a random process. So, any phase gradients caused by both natural topography variability and by point-of-view geometry (remote sensing radar systems have a side-scattering configuration) lead to the degradation of the estimate (1). Its value depends on the slope and tends towards the value $|\rho_{12}(N)|_{\rho=0}$, *i.e.* a bias of the estimate for independent Gaussian values of the correlation coefficient (N is the number of samples) [3], which in practice takes the value about 0.1–0.3. Thus, coherence loses its properties as measure of the quality of the interferogram, its value becomes dependent on relation between topographic and fluctuation components of the phase.

3 Differential phase coherence estimate

To eliminate the effect described above, the following modification of the coherence estimate can be offered taking into account influence of the topographic component. Modified coherence evaluates not the samples $z_1(m,n)$ and $z_2(m,n)$ of the SLC–image pair, but the following values:

$$\dot{w}_1(m,n) = \dot{z}_1(m,n) \cdot \bar{z}_1(m+1,n), \;\; \dot{w}_2(m,n) = \dot{z}_2(m,n) \cdot \bar{z}_2(m+1,n), \tag{2}$$

where the new phase values of the $w_1(m,n)$ and $w_1(m,n)$ will characterize the slope of the topographic phase in the direction of growth of the m^{th} picture coordinate. So, this operation performs the phase derivation along one coordinate.

Similarly, one can use the gradients along the n^{th} coordinate. Expression for the coherence estimation using values (2) looks as follows:

$$\hat{\gamma_1} = \sqrt{\frac{\Sigma \dot{w}_1(m,n) \cdot \bar{w}_2(m,n)}{\sqrt{\Sigma |\dot{w}_1(m,n)|^2 \cdot |\bar{w}_2(m,n)|^2}}}. \tag{3}$$

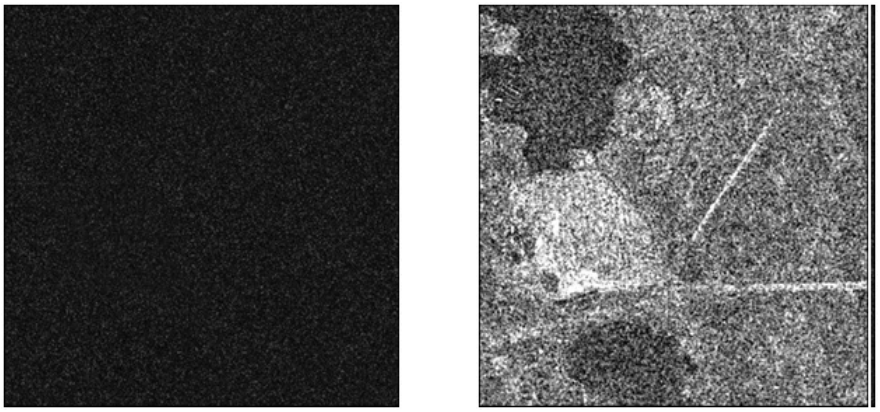
Although this estimate is non-Gaussian, it is wealthy, and it can be shown that it is insensitive to the linear phase trend.

4 Experimental results

Test now the work of the coherence estimation algorithms for RADARSAT–1 data (wavelength 56 mm). Figure 1 presents a fragment of the radar image of an area with surfaces of different reflectance including surfaces with volume scattering and the water surface, which has generally low coherence. Signal differential phase for the given fragment has a slope in the horizontal direction of the order of 0.3 radians per sample (due scattering geometry). The coherence maps of the fragment were constructed using conventional estimate $\hat{\gamma_0}$ and modified estimate $\hat{\gamma_1}$ with the sampling size 11×11 are presented in Figs. 2a, 2b.

Fig. 1. A RADARSAT-1 radar image scene

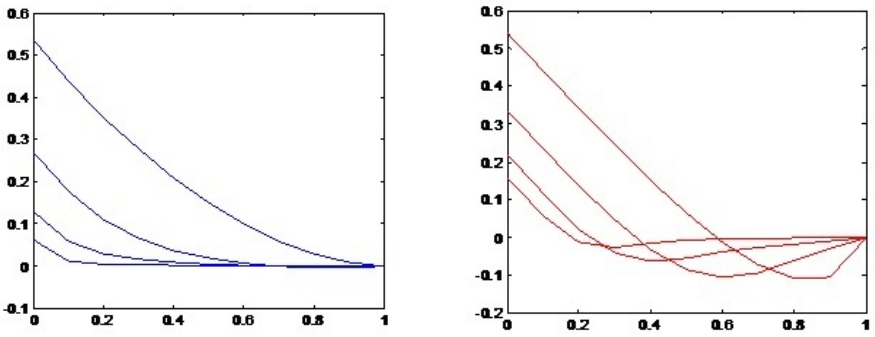
One can see that the first map is degraded since different surfaces give the same low coherence values regardless on the surface type. The map obtained



a) Traditional estimate $\hat{\gamma_0}$ b) Modified estimate $\hat{\gamma_1}$

Fig. 2. Coherence map calculated

using modified estimate (Fig. 2b) has a good sensitivity to the surface type. However, the estimate has a larger bias at low values than the $\hat{\gamma_0}$, and, so, it requires increasing the sample value towards to $\hat{\gamma_0}$ (Fig. 3a, 3b). A quantitative accuracy assessment for the scene is not available because of poor reference DEM for this territory.



a) Traditional estimate $\hat{\gamma_0}$ b) Modified estimate $\hat{\gamma_1}$

Fig. 3. Coherence estimates biases for different sample sizes (N=2, 5, 10, 25)

5 Conclusion

A modified method for estimation of spatial coherence of the interferometric radar images pairs is developed. The method consists in calculation of the cor-

relation between pairs of complex multiplications of neighbour elements. The research is implemented on radar images RADARSAT–1. The result shows that the modification allows one to solve the problem of estimate degradation under the differential phase trend, which always takes place in side-scattering radar systems.

Acknowledgment The work was supported by the RFBR grants Nos. 13-07-12168, 13-07-00785.

References

1. Elizavetin,I. V., Ksenofontov, E. A.: Resultaty experimentalnogo issledovaniya vozmozhnosti pretsizionnogo izmereniya reliefa Zemli interferentsionnym metodom po dannym kosmicheskogo RSA [The results of experimental research of precious Earth relief measurement by interferometric method with space-based SAR]. Issledovaniya Zemli iz kosmosa, 1, 75–90 (1996) (in Russian)
2. Touzi, R., Lopes, A., Bruniquel,J., Vachon, P.W.: Coherence Estimation for SAR Imagery. IEEE Trans. Geosci. Remote Sensing, 37, 135–149 (1999)
3. Lopez-Martinez, C., Pottier, E.: Coherence Estimation in Synthetic Aperture Radar Data based on Speckle Noise Modeling. Applied Optics, 46(4), 544–558 (2007)
4. Kobernichenko, V. G., Sosnovsky, A. V.: Particular Qualities of Digital Elevation Maps Generation in Interferometric SAR Technology. SPIIRAS Proceedings, 5(28), 194–208 (2013)
5. Kobernichenko, V. G., Sosnovsky,A. V.: Coherence Estimation in Interferometric SAR Data Processing. 2014 24th Int. Crimean Conference Microwave and Telecommunication Technology (CriMiCo–2014). 7–13 September, Sevastopol, Crimea, Russia, 24, 1153-1154 (2014)

Irregularity as a Quantitative Assessment of Font Drawing and Its Effect on the Reading Speed

Dmitry Tarasov[1], Alexander Sergeev[1,2]

[1]Ural Federal University, Ekaterinburg, RUSSIA
[2]Institute of Industrial Ecology, UB of RAS, Ekaterinburg, RUSSIA
datarasov@yandex.ru

Abstract. It is proposed to use irregularity, the scale invariant index based on the ideas of fractal geometry to assess the spatial features of font drawings. The index is sensitive to the shape of characters in the font, which affects text legibility. Preliminary results have shown promising application of the proposed index for classifying fonts by reading speeds.

Keywords. Font; Fractal; Scale invariance; Legibility; Typeface

1 Introduction

Research in the fields of legibility and readability are maintained for over a hundred years. They are particularly important for the development of textual materials intended for readers with emerging reading skills. A significant place in these studies takes fonts. Many researchers have investigated the clarity, legibility, readability of different fonts, the influence of serifs, the influence of the pattern and spatial characteristics of the font on the understanding and memorising the content of the text and some other factors. The obtained results are contradictory. So far there is no consensus on what fonts features and how affect the reading process. This is largely due to the lack of an objective index, which could describe the typeface, and allows comparing different fonts.

Artemov [1] proposed to divide the concepts of visibility and readability of the font. Readability is influenced by reader's physiological characteristics. Visibility depends on the quality of font drawing and vision features of the person. Differences in typeface readability investigated in [2-4]. Some fonts are marked as the most readable. The superiority of some small book fonts connected to their shapes and drawings is demonstrated. Thick font reads faster. At the same time, respondents preferred the other fonts. Similar results were obtained in [5]. Studies have shown the presence of subjective preferences of readers, as well as an objective difference in readability of fonts with different shapes. The review [6] analysed the various features of fonts with respect to their readability, but also contains a large number of different, often conflicting, views on the impact of serifs, size and font style for readability. Results of study [7] compares the readability of some common fonts by testing the reading speed

of texts in Russian. A higher reading speed for serif fonts is demonstrated. However, no explicit font characteristics affecting readability are identified. The work [8] provides an overview of the situation of modern typography of textbooks and considers contradictions of the current state with the font design to the rules of the current technical regulations. Lots of researchers consider serif fonts more legible and it is because of their serifs which add more information to the eyes [9] and enhance the legibility of a text by helping the readers to distinguish the letters and words more easily [10]. Results in [11], [12] indicated serif fonts are believed to be read faster due to their invisible horizontal line made by serifs. Results of study [13] is against the prominence of serif fonts. The space between letters in serif fonts is slightly reduced due to the ornaments that they have. Consequently, as mentioned in [14], serifs act as visual noise when the readers' eyes attempt to detect the letters and words. The reduction of the space leads to other problems: One is a problem of *crowding* which is hindering of letter recognition when a letter is flanked by other letters (cited in [15]) and the other is that letter position coding may be hindered which decreases the ability of word recognition [13]. The results of studies [15], [16] showed out equal legibility and perception between serif typefaces and sans serif ones.

Thus, almost equal numbers of studies showed advantages and disadvantages of serifs, as well as a preference of other features of text. The preferences of specific font features and font size are highly dispersed, too. It can be suggested that legibility is more sensitive to some combinations of spatial features of text. No special type font is suggested to use. The point to pay attention to is the familiarity of the subjects with special typefaces and subjects' preferences. The aim of this work is to find the way to assess the spatial features of font drawings by using an objective scale invariant index.

2 Approach

An assessment of the visual characteristics of fonts represents certain difficulties associated with the difference in approaches to the understanding of what is a set of visual characteristics and what criteria should be used in their assessment. The similarity of some graphic elements of letters in font and the letters themselves, as well as the font as a whole, suggests the possibility of using the ideas of fractal geometry to make the assessment. A special case of the fractal dimension d is expressed by well-known formula that combines the number of objects n, with which the measurement is taken, and the geometric size of the object a:

$$d = \log n : \log a^{-1}. \tag{1}$$

Mandelbrot showed [17] that for fractal sets the expression relating the length of the perimeter of the object P and its area S is performed:

$$P^{1/d} : S^{1/2} = \text{const}, \tag{2}$$

which implies that $S \sim P^{2/d}$.

The fractal dimension can be understood as the degree of filling of the space by irregularly distributed substance. Thus, in either family of flat figures (like font), geometrically similar but having different linear dimensions, the ratio of the length of the shapes border to the square root of its area is a number that is completely determined by the general form for the family. The equivalence of different linear extensions in many cases is very useful [17]. The relation between abris' length (perimeter *P*) of the character or set of characters in the font and its area (*S*) can be used as a unique font index. Considering the font as a coherent geometric set, by analogy with the way proposed in [18], it is possible to apply the definition of compactness of the set *C* (3), circularity coefficient C_c (4) and irregularity C_n (5) which is proposed to use as such unique font index.

$$C = P^2 : S. \tag{3}$$

$$C_c = 4\pi S : P^2, \tag{4}$$

$$C_n = C_c^{-1} = \mathrm{C} : 4\pi = P^2 : 4\pi S \tag{5}$$

Vector graphics software having an intrinsic macro language based on *VBA* can helps to solve the task of index calculation. In the present work a public macro *CurveInfo* for *CorelDraw* package is used. The macro calculates perimeter (in mm) and area (in mm^2) of a coherent vector object.

Fig. 1. Set of font letters and its division

As a representation of a font the full set of 66 uppercase and lowercase letters (for Russian language) of each font is used (see Fig. 1). To obtain information about the perimeter of a particular letter the perimeter of the external abris of a letter (P_{out}) must be added to (if available) of the internal perimeter of the letter space (P_{in}). The perimeter of the full set of 66 letters (*P*) equals the sum of the perimeters of all letters (6). To obtain an area of a letter it must subtract the internal area of the letter space (S_{in}) from the general area bounded by the outer abris of the letter (S_{out}). The area of the full set of 66 letters (*S*) is equal to the sum of the areas of all the letters (7).

$$P = \Sigma\,(P_{out} + P_{in}) \tag{6}$$

$$S = \Sigma\,(S_{out} - S_{in}) \tag{7}$$

3 Results and discussion

For the measurement 21 fonts (straight light drawing, sizes 12 and 18 pt) were selected: 9 sans serif fonts, 11 serif fonts and a script font. For the selected set of fonts the described procedures and formulas were applied. By the distribution of the values of irregularity C_n for groups of serif and sans-serif fonts the mapping is undertaken.

Also, 5 fonts among the full set are used to test their reading speeds. Participants of the experiment are 10 students. They read text samples with different font layout and count the reading speed. The correlation between reading speeds and irregularities is assessed. The results of calculations by formula (5) and reading speed measures for 5 fonts are given in Table 1. As it can be seen from Table 1, a irregularity has almost constant values for each font. It indicates the objective nature of the scale invariant index proposed that is useful for research. A small variation of the index for some fonts can be explained by features of font scaling. Figure 2 shows the distribution of irregularity for groups of serif and sans-serif fonts by peer review. Serif fonts are read slightly faster than sans serif ones, on average. Script fonts have a low reading speed but extremely high irregularity. Figure 3 shows the distribution of reading speeds and its dependence from irregularity for 5 selected fonts, regression and confidence intervals. Statistical analysis reveals strong negative correlation between reading speed and irregularity (correlation coefficient –0,69, $p<0,05$).

Table 1. Irregularities and reading speeds.

No	Font	Feature	Reading speed, chars per sec	C_n, 12 pt	C_n, 18 pt
1	Vanta	sans-serif		305	305
2	Text Book	sans-serif		418	418
3	Comic Sans	sans-serif		439	416
4	Verdana	sans-serif		459	464
5	Avantgard	sans-serif		470	470
6	Arial	sans-serif	43,2	481	469
7	Century Gothic	sans-serif		575	575
8	Futuris	sans-serif	35,5	605	605
9	Futura	sans-serif		825	825
10	Bruskovaya	serif		472	472
11	Antiqua	serif		585	585
12	Mysl Narrow	serif		653	653
13	Times New Roman	serif	33,6	675	655
14	Baskerville	serif		682	682
15	Journal	serif		703	704
16	Book Antiqua	serif		714	714
17	Baltica	serif		778	792
18	Classic Russian	serif	31,4	796	796
19	Garamond	serif		875	875
20	Literaturnaya	serif		880	802
21	*Art Script*	script	28,5	1717	1651

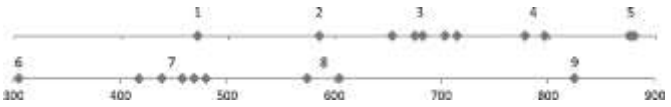


Fig. 2. Distribution of irregularity for groups (numbers) of serif (above) and sans-serif (below) fonts by peer review

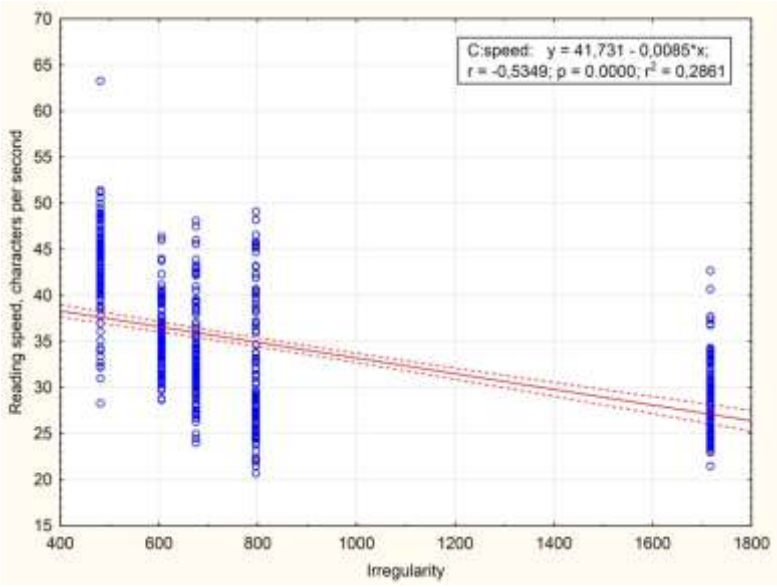


Fig. 3. Distribution of reading speeds and its dependence from irregularity for 5 fonts, regression and confidence intervals (drops).

4 Conclusion

Although there are massive bodies of analysis considering typography and font features, there is no agreement among researchers regarding legibility factors in print. One of the most complicated issue is accounting for the effect of font drawing on legibility. The work offers a solution to this problem. The scale invariant index to assess the spatial features of font drawing is proposed. Accounting for the index in research of reading (e.g. [19]) might help to identify predictors of reading speed, as well as quality of assimilation not only for paper, but also for electronic texts. In any case, the scale invariance of the proposed index allows to accumulate experimental results without any subsequent processing or conversion, which is convenient.

5 References

1. Artemov V.A. Technographic analisys of summarized letters of new alphabet. Writing and revolution. 1933, 1, 58–76.
2. Ushakova M.N. The new typeface for newspapers. Polygraphic manufacturing. 1952, 4, 22–23.
3. Ushakova M.N. The new typeface for narrative literature. Polygraphic manufacturing. 1952, 11, 26–28.
4. Ushakova M.N. The new typeface for large circulation narrative literature. Polygraphic manufacturing. 1959, 11, 26–27.
5. Alexandrova N.A., Chiminova V.G. New typefaces for newspapers. Polygraphic manufacturing. 1962, 5, 57–61.
6. Tokar O.V., Zilbergleit M.A., Petrova L.I. Legibility of typefaces. Proceedings of the higher educational institutions. Problems of printing and publishing. 2004, 2. 79–92..
7. Tokar O.V., Zilbergleit M.A., Litunov S.N. Font legibility assessment on the sample of official document. Omsk scientific herald. 2009, 2(80). 246–249.
8. Tarbeev A.V. Font design of modern textbook. Proceedings of the higher educational institutions. Problems of printing and publishing. 2004, 4. 93–102.
9. Geske J. (1996). Legibility of Sans Serif Type for Use as Body Copy in Computer Mediated Communication. Presented at Annual Meeting of the Association for Education in Journalism and Mass Communication 1996
10. McCarthy M.S. & Mothersbaugh D.L. (2002). Effects of typographic factors in advertising-based persuasion: A general model and initial empirical tests. Psychology &Marketing, 19(7-8), 663–691.
11. Romney C. (2006). Improving the visual appeal of classroom handouts. JALT2005 Conference Proceeding. Tokyo: JALT. Retrieved May 21, 2012, from http://jalt-publications.org/archive/proceedings/2005/E121.pdf
12. Arditi A. & Cho J. (2005). Serifs and font legibility. Vision Research, 45, 2926–2933.
13. Perea M. & Moret-Tatay C., Gómez P. (2011). The effects of inter letter spacing in visual word recognition. Acta Pschologica, 137, 345–351.
14. Woods R. J., Davis K., Scharff L.V.F. (2005). Effects of typeface and font size on legibility for children. American Journal of Psychological Research, 1(1), 86–102.
15. O'Brien B.A., Mansfield J.S., Legge G.E. (2005). The effect of print size on reading speed in dyslexia. J Res Read, 28(3), 332–349.
16. Paterson D. G. & Tinker M. A. (1932). Studies of typographical factors influencing speed of reading. X. Style of type face. Journal of Applied Psychology, 16(6), 605–613.
17. Mandelbrot B. Fractal geometry of nature. Moscow, Institute of computer studies, 2002, p.656.
18. Alexanina M.G., Karnatsky A.Y. Comparison of the spatial characteristics of the fields of sea ice and cloud according to the visible channel AVHRR / NOAA on the example of the Okhotsk Sea / The Sixth All-Russian Open Annual Conference "Modern Problems of Remote Sensing of the Earth from Space". Moscow, IKI RAS 10-14 nov 2008, 299–302.
19. Tarasov D.A. & Sergeev A.P. The leading as a factor of readability: development of the methodology for educational use. Procedia – Social and Behavioral Sciences, Elsevier, v.106, 2013. 2914–2920.

Fast Full-Search Motion Estimation Method Based On Fast Fourier Transform Algorithm

Elena I. Zakharenko, Evgeniy A. Altman

Omsk State Transport University
{zaxarenko.elena, altmanea}@gmail.com

Abstract. Motion estimation (ME**)** is used extensively in video codecs based on MPEG-4 standards to remove interframe redundancy. Motion estimation is based on the block matching method which evaluates block mismatch by the sum of squared differences (SSD) measure. Winograd's Fourier transform is applied and the redundancy of the overlapped area computation among reference blocks is eliminated in order to reduce the computational amount of the ME. When the block size is $N \times N$ and the number of reference blocks in a search window is the same as the current block, this method reduces the computational amount (additions and multiplications) by 58 % of the straightforward approach for $N = 8$ and to 81 % for $N = 16$ without degrading motion tracking capability. The proposed fast full-search ME method enables more accurate motion estimation in comparison to conventional fast ME methods, thus it can be applied in video systems.

Keywords: motion estimation, fast Fourier transform, correlation, convolution theorem, full-search, video encoding.

1 Introduction

The popularity of video as a mean of data representation and transmission is increasing. Hence the requirements for a quality and size of video are growing. High visual quality of video is provided by coding. In 1960s the motion estimation (ME) and compensation were proposed to improve the efficiency of video coding [1].

The current frame is divided into non-overlapping blocks. For each block of the current frame the most similar block of the reference frame within the limited search area is found. The criterion of the similarity of the two blocks is called a metric comparison of the two blocks. The position of the block, for which an extremum of metric is founded, determines the coordinates of the motion vector of the current block.

The full search algorithm is the most accurate method of the block ME, i.e. the proportion of true motion vectors found is the highest [2]. The current block is compared to all candidate blocks within the restricted search area in order to find the best match.

This ME algorithm requires a lot of computing resources. Therefore, a lot of alternative fast motion estimation algorithms were developed. In 1981 T. Koga and other authors proposed a three-step search algorithm (TTS) [3].

The disadvantage of fast search methods is finding a local extremum of a function of the difference of two blocks. Consequently motion estimation degrades by half degradation in some sequences compared to brute-force and visual quality of video degrades as well [4].

2 The Criterion To Compare Blocks

The standards of video coding do not regulate the choice of criterion for matching two blocks (metric).

One of the most popular metrics is the sum of square difference (SSD):

$$SSD(i, j) = \sum_{y=0}^{N_h-1} \sum_{x=0}^{N_w-1} (B(x, y) - S(x+i, y+j))^2, \quad (1)$$

where i, j – the coordinates of the motion vector of the current block, $i \in (-V_w/2; V_w/2)$, $j \in (-V_h/2; V_h/2)$, where $V_w \times V_h$ – size of the area which can be is the upper left corner of the title block on the reference frame; x, y – coordinates of the current block B; $N_w \times N_h$ – block size B; S – reference area of size $S_w \times S_h$, where $S_w = N_w + V_w$, $S_h = N_h + V_h$; B and S – luminance images in color format YUV.

Inside the search area size $S_w \times S_h$ is the minimum value of SSD criterion for the current block B, which determines the coordinates of the motion vector in order.

SSD can be calculated through fewer number of operations by decomposition into three components [5]:

$$\sum_{y=0}^{N_h-1} \sum_{x=0}^{N_w-1} B^2(x, y) - \quad (2)$$

$$-\sum_{y=0}^{N_h-1} \sum_{x=0}^{N_w-1} B(x, y) S(x+i, y+j) + \quad (3)$$

$$+\sum_{y=0}^{N_h-1} \sum_{x=0}^{N_w-1} S^2(x+i, y+j), \quad (4)$$

In [5, 6] for the (3) computation were used Fast Fourier Transform (FFT). We propose to replace this algorithm by other fast transforms: Winograd algorithm and the number-theoretic transform of Farm (NTT).

3 Research Results

A programming model for block motion estimation using SSD metrics was implemented in Matlab in order to analyze the effectiveness of the fast Fourier transform algorithm for computing the block matching criterion.

Expression (4) is calculated using the algorithm described in [2].

To analyze the effectiveness of the algorithms of fast Fourier transform (FFT) for motion estimation were considered Cooley-Tukey and Winograd algorithms. We selected *B* block 16 × 16 and 8 × 8 pixels, a reference area *S* is twice as much as *B* block, i.e. is 32 × 32 and 16 × 16, respectively.

Operating time of algorithm is significant for processing and analysis of video, it depends on its computational complexity. The number of arithmetic operations (multiplications (×) and additions (+)) is significant when the complexity of the method for motion estimation is measured. Therefore, this criterion was chosen as the metric calculation efficiency of SSD.

Table 1 shows the results for a block size of 16 × 16 pixels and the reference area *S* 32 × 32 dots.

Table 1. The computational complexity of algorithms SSD when $S = 2B = 32$ for one of the current block

Algorithms	Real arithmetic operations								In total	The complexity, %
	expression (2)		expression (3)		expression (4)		in total			
	×	+	×	+	×	+	×	+		
SSD according to (1), FS	–	–	–	–	–	–	65 536	130 816	196 352	100
SSD according to (1), TTS	–	–	–	–	–	–	2 560	5 110	7 670	4
SSD according to (2), (3) and (4)	256	255	65 536	65 280	65 536	65 280	131 328	130 815	262 143	134
Winograd FFT	0	0	7 203	25 923	1 426	2 115	8 629	28 038	36 667	19
Cooley-Tukey FFT	0	0	23 552	33 792	1 426	2 115	24 978	35 907	60 885	31

The results for block 8×8 and the search area 16×16 similar to presented in Table 1. Winograd algorithms reduces the computational complexity of the ME algorithm by 58 % for block size 8×8.

Having analyzed the results shown in Tables 1, it is obvious that the choice of convolution algorithm has a significant impact on the computational complexity of motion estimation.

For the hardware implementation of a significant impact on the area occupied on the chip, providing multiplication. As can be seen from Table 1 application of the Winograd algorithms for calculating SSD reduces number of the real multiplications.

The algorithms examined operate with complex numbers in order to get the result. For identifying more efficient ways to compute the two-dimensional convolution the number-theoretic transform of Farm (NTT) was considered. This algorithm is based on modular arithmetic and uses only integer real numbers.

It is hard to assess the computational efficiency by the number of arithmetic operations of NTT relative to considered algorithms, because besides addition and multiplication there are operations modulo Farm. In this case the criterion of comparison of these methods is the estimated performance. This criterion evaluates the time required

to calculate the coordinates of the moving blocks with each of the FFT algorithms considered in the developed model. The results of Matlab simulation are presented in Table 2.

Table 2. The estimated performance of the algorithms SSD when $S=2B=16$ for one of the current block

Algorithms	Time, ms
Winograd FFT	5,5
Cooley-Tukey FFT	6,7
NTT	6,1

As shown in Table 2 motion estimation method based on block matching using Winograd algorithm has the lowest complexity.

4 Conclusions

A new fast full search algorithm for block motion estimation based on convolution theorem and Winograd's Fourier transform is presented. Proposed method reduces the computational complexity of the ME algorithm by 58 % the size of block 8 × 8, and – 81 % the size of block 16 × 16 pixels for the search area, exceeding twice the size of the current block. The developed motion estimation algorithm of full search can be used in the environment with a high demand for the quality of video, and the computational complexity of the algorithm is not critical. The further direction of research in this area will be the analysis of fast algorithm for computing two-dimensional convolution through its decomposition into several convolutions shorter length.

References

1. Haskell B.G., Limb J.O.: Predictive video encoding using measured subjective velocity. U.S. Patent No. 3,632,865 (1972)
2. Lee C.-H., Chen L.-H.: A Fast Motion Estimation Algorithm Based on the Block Sum Pyramid. J. IEEE Trans. Image Processing, vol. 6, pp. 1587–1591 (1997)
3. Koga T., Iinuma K., Hirano A., Iijima Y., Ishiguro T.: Motion compensated interframe coding for video conferencing. In: Proc. Nat. Telecommun. Conf., New Orleans, LA, pp. G5.3.1–5.3.5 (1981)
4. Cheung C. H., Po L. M.: A novel cross-diamond search algorithm for fast block motion estimation. J. IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 12, pp. 1168–1177 (2002)
5. Kilthau S. L., Drew M. S., Moller T. Full search content independent block matching based on the fast Fourier transform. IEEE 1CIP, I, pp. 669–672 (2002)
6. Zubarev Iu.B., Dvorkovich V.P., Nechepaev V.V., Sokolov A.Iu.: Metody analiza i kompensatsii dvizheniia v dinamicheskikh izobrazheniiakh. J. Elektrosviaz', no. 11, pp. 15–21 (1998)

Development of Trained Algorithm Detection of Fires for Multispectral Systems Remote Monitoring

Sergey Zraenko, Margarita Mymrina, and Vladislav Ganzha

Ural Federal University, Mira str., 19, Yekaterinburg, 620002,
Russian Federation
z_sm@mail.ru

Abstract. The paper describes research of a trained threshold algorithm for analysis of the fire situation using multispectral remote sensing data. The algorithm is based on application of sub-satellite information about the fire on the territory covered by a satellite image. Thresholds are selected for each of the spectral bands for the determination of known fire threshold and then are used to select all thermal anomalies in the picture. According to study of the algorithm the informative spectral channels are defined for detecting fires. Combination of the results from the multiple channels spectroradiometer MODIS for increasing the probability of correct detection is demonstrated.

Keywords: Forest fires, MODIS spectroradiometer, MOD14 algorithm, trained threshold algorithm, spectral channels, probability of fire detection

1 Introduction

Data of the earth remote sensing (ERS) is widely used for forest fire detection. In this case, pictures of Imaging Spectroradiometer MODIS [1], which is a part of equipment of American Terra and Aqua satellites, are intensively used now. An operative receiving of information from MODIS is carried out by personal receiving ground stations, one of which (UniScan-24 [2]) is set at the Ural Federal University.

The MODIS data are formed in 36 spectral bands and characterized by a wide swath (2300 km) and high frequencyof udate (up to several times a day). However, these images have low spatial resolution (1 km in the most of the infrared spectral channels) that does not allow one to detect reliably fires in a fraction of a hectare under various observation conditions (surface fire, lack of fire intensity, poor weather conditions, *etc.*).

2 Problem formulation and development of trained threshold algorithm

Complexity of solving the problem of detection of the forest fires using remote sensing data may be shown by presence of a large number of Russian and foreign Internet services. This information about the fires differs not only from the actual, but, also, among the sources. The evidence of this is confirmed [3] by analysis of the fire in State Nature Reserve "Denezhkin Kamen" (Fig. 1) in 2010 according to the data of several online fire detection and monitoring services.

Information from these services is based on the MODIS sensor data that are processed by the MOD14 fire detection algorithm and its modifications [4].

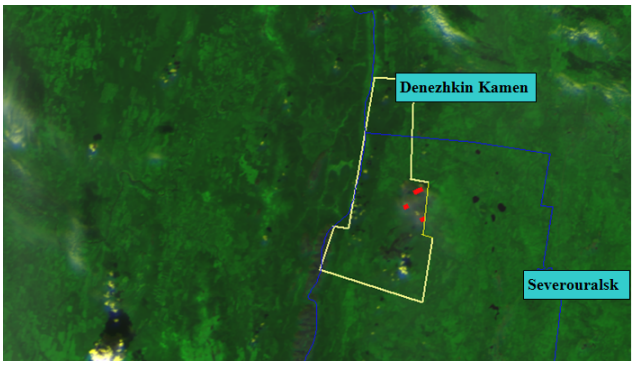


Fig. 1. Daily snapshot MODIS 31.07.2010 combined with mask fires and the boundary of the reserve "Denezhkin Kamen"

This algorithm uses the temperature data of the pixels at two infrared spectral channels: 4 μm (21 channel) and 11 μm (channel 31). The brightness temperature of a "hot" pixel in channel 21 should be higher 312K by day and 305K by night. In addition, if the temperature difference of a pixel between the 21 and 31 channels and the difference between channel 21 and the environment temperature are larger, the probability that the pixel belongs to the fire is higher. Pre-image is also evaluated under presence of clouds using masks clouds and global clouds; and the pixels belonging to the solar flare are eliminated. As a result, the mask of fires is built that contains the so-called "thermal points", *i.e.*, the pixels exceeding the preset temperature thresholds.

It should be noticed that this algorithm is currently the basis for a number of specialized software tools for processing remote sensing data, such as the program product ScanEx Image Processor [5]. In the fact, efficiency of the MOD14 is insufficient in areas with high forest vegetation, and it can detect only relatively large fires. This factor is caused by fixed values of temperature thresholds that do not take into account specific conditions of the observation.

To increase the likelihood of a correct detection of temperature anomalies, we have proposed [6] to train our threshold algorithm on the basis of sub-satellite information about the current fire. Selection of the threshold level was made using channel 21 of the spectroradiometer MODIS for pictures dated by July 28 (night) and July 29 (day), 2010. The training was based on the known fire in the reserve "Denezhkin Kamen". For this purpose, a part of the Sverdlovsk region was cut and converted to a range of brightness from 0 to 255, and after that the detection thresholds were determined. It is also supposed (Fig. 2) that there is the area free from fire adjacent with another seat of fire, the area of 1 pixel size (gray area) separated from the fire at the distance about 0.5 – 1.5 pixels is the boundary region part, which can be occupied with the fire (white square) round the seat of fire (a dark oval). In our case (Fig. 2), when the seat of the fire is in the pixel with coordinates (X = 217, Y = 193), this means that "hot" pixels are absent at the points X = 215–219 when Y = 191 and Y = 195 and at the points Y = 191–195 when X = 215 and X = 219.

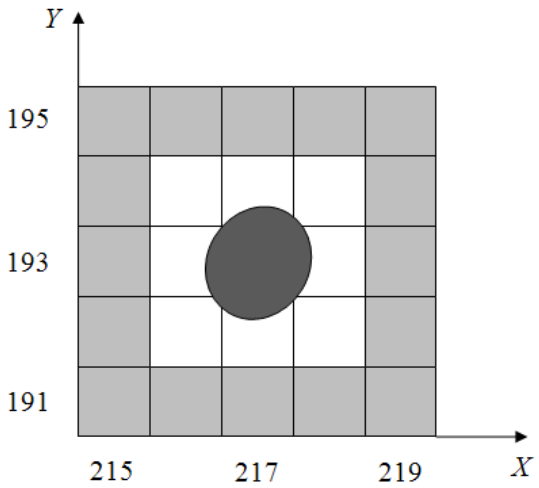


Fig. 2. The definition of adjacent area to the fire

The threshold in spectral channel 21 (I_{POR}) is defined as the average of the luminance values of pixels from the known fire hearth (I_{OI}) and the adjacent region (I_{SI})

$$\begin{aligned} I_{\mathrm{POR}} &= (I_{\mathrm{OImin}} + I_{\mathrm{SImax}})/2, \\ I_{\mathrm{OImin}} &= \min(I_{\mathrm{OI}}), \\ I_{\mathrm{SImax}} &= \max(I_{\mathrm{SI}}). \end{aligned} \tag{1}$$

Here, I_{OImin} is the minimum brightness of a pixel of the source area with known fire, I_{SImax} is the maximum brightness of a pixel of the adjacent area. In this case, ($I = 1$) are pixels (correlated with the fire) in the image (I) exceeding this

threshold

$$I = \begin{cases} 0, & if\ I < I_{POR}\,, \\ 1, & if\ I \geq I_{POR}\,. \end{cases} \tag{2}$$

3 Research of the trained algorithm

Study of the proposed algorithm was performed using MODIS imagery territory of the Sverdlovsk region 28–29 July, 2010 and 01.08.2012. The actual fire situation was determined by the summary of the Ministry of Emergency situations posted the Internet. In the night picture, the threshold of detection of fire in the reserve "Denezhkin Kamen" in channel 21 MODIS (in relative brightness values 0–255) was determined in the range $B_{21NIGHT} = 59$–69. After selecting the middle value $B_{21NIGHT} = 64$, the mask of detected fires (Fig. 3) was constructed. The fire figures (black dots) are visible in the reserve "Denezhkin Kamen", Garinsky, and Sosvinsky counties.

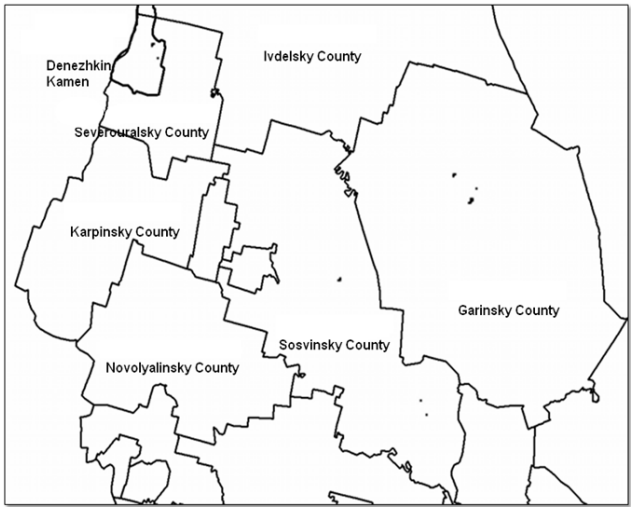


Fig. 3. Mask fires of the trained algorithm for the night snapshot, 28.07.2010; the black dots are selected thermal points

Table 1 presents information about known fires and the results of detection of the thermal points obtained (using the night MODIS image) by the proposed algorithm and the MOD14 algorithm with selected (from the condition of maximum detection of existing fires) value $T_{21NIGHT} = 300$K.

The chosen threshold value $T_{21NIGHT} = 300\text{K} = 270\text{C}$ was lower than the value $T_{21NIGHT} = 305$K, which is considered as the boundary for the recognition of the pixel to be the thermal points. It is quite realistic to suppose that this was

the night temperature of crowns of trees in the Northern part of the Sverdlovsk region. Note that the proposed algorithm has demonstrated its efficiency, but it showed slightly worse results compared to the MOD14 with selected thresholds in channels 21 and 31.

Table 1. Detection of fire by the night snapshot 28.07.2010; the number of thermal points in each snapshot is given in parentheses

County	Actual fires	MOD14 (selected thresholds)	Trained algorithm
reserve "Denezhkin Kamen"	1	3 (4; 3; 1)	2 (3; 1)
Severouralsky	11	2 (1; 1)	no
Sosvinsky	15	6 (2; 3; 4; 2; 2; 12)	3 (3; 1; 1)
Ivdelsky	9	1 (1)	no
Garinsky	10	10 (1; 1; 1; 2; 1; 1; 2; 3; 2; 7)	3 (6; 3; 1)

When applying the proposed methodology on a daily snapshot, it took to use as a training standard the fire in Garinsky county because the creeping fire in the reserve "Denezhkin Kamen" failed to be detected due to its masking by the canopy of trees. In this case, the average value of the threshold $B_{21DAY} = 43$ for channel 21 fires in Sosvinsky, Garinsky, Krasnoturinsky, Serovsky, Kachkanarsky, and Karpinsky counties were detected.

In this experiment, the threshold value $T_{21DAY} = 320$K was selected by the condition of maximum detection of existing fires and was slightly above the recommended ($T_{21DAY} = 312$K) to reduce the number of false thermal points. Analysis (Table 2) has determined that the false "hot" pixels were detected in Kachkanarsky county, and the fires in the reserve "Denezhkin Kamen", Severouralsky, and Ivdelsky counties were skipped. According to the above data, the trained algorithm showed approximately the same results as MOD14 with thresholds chosen to the highlight of known fires. In general, the results of a daily snapshot for both algorithms were slightly worse than for the night ones because of the influence of surface temperature and possible specular reflections from water surfaces and edges of the clouds.

Verification of the trained algorithm was carried out using the daily MODIS image of the Sverdlovsk region dated 01.08.2012. In this case, the fire at Krasnouralsky county was selected as "training". The comparison of the results obtained here for the trained algorithm with the actual number of fires, as well as, with the MOD14 algorithm with a threshold chosen for the situation in 2010 is given in Table 3.

The obtained result demonstrates the effectiveness of the proposed algorithm (selected only one false fire) compared to the MOD14 algorithm where the threshold was selected for the snapshot of the same territory made two years earlier (with a precision of 3 days), where lots of false fires were found.

Table 2. Detection of fires on a daily snapshot 29.07.2010; the number of thermal points in each snapshot is given in parentheses

County	Actual fires	MOD14 (selected thresholds)	Trained algorithm
reserve "Denezhkin Kamen"	1	no	no
Severouralsky	12	no	no
Sosvinsky	15	2 (1; 1)	1 (1)
Ivdelsky	9	2 (1; 1)	no
Garinsky	11	6 (1; 2; 2; 2; 3; 5)	4 (5; 2; 1; 1)
Karpinsky	7	1 (1)	2 (3; 1)
Krasnoturinsky	3	no	2 (8; 1)
Serovsky	1	2 (1; 1)	1 (11)
Kachkanarsky	1	1 (12)	2 (15; 12)

Table 3. Detection of fires on a daily snapshot 01.08.2012

County	Actual fires	MOD14 (the thresholds selected for the 29.07.2010 snapshot)	Trained algorithm
Krasnouralsky	1	1	1
Kachkanarsky	no	2	1
Sosvinsky	no	1	no
Ivdelsky	no	7	no
Severouralsky	no	1	no
Krasnoturinsky	no	1	no
Karpinsky	no	2	no

Further improvement of accuracy of fires detection using MODIS data can be done by integration of the detection results obtained by processing a series of images (especially day and night) received at the same or next day. However, at the time of the next shooting, some fires can be extinguished, and some new fires may appear. Shooting conditions can also change. Therefore, the most expedient way is to extract information about the fire according to the other spectral channels of a single snapshot. Thus, parameters of the signal passing through the atmosphere from a thermopoint are various for different spectral channels (depend on a condition of the atmosphere, temperature of a thermopoint and temperature sensitivity of the channel). So we conducted research of temperature sensitivity of different channels.

In contrast to our previous publications [3], [6], this paper shows the results that were received using pictures 28.07.2010 (night) and 29.07.2010 (day) with the subsequent association of results in spectral channels. Results for the MODIS channels that are informative to find thermal anomalies are presented in Table 4. The appearance of the sign "–" for channel 24 means that it is uninformative during the day because it could not allocate any points of real fires. In addition,

the reserve "Denezhkin Kamen" is a part of Severouralsky county, and the table presents aggregated information about fires in these area.

Table 4. Results of study of the spectral channels reliability

The channel number		20	21	22	23	24	25
County	Actual fires (night/ day)	Fires found (night/day)					
Severouralsky (reserve "Denezhkin Kamen")	11/12	2/0	2/0	2/0	2/0	2/–	2/0
Sosvinsky	15/15	6/0	3/0	3/0	4/0	6/–	5/0
Ivdelsky	9/9	0/6	0/0	0/0	0/0	13/–	0/0
Garinsky	10/11	3/3	3/2	3/3	3/3	3/–	3/2
Karpinsky	0/7	0/0	0/0	0/0	0/0	0/–	0/0
Krasnoturinsky	0/3	0/2	0/0	0/0	0/0	0/–	0/0
Serovsky	0/1	0/0	0/1	0/1	0/1	0/–	0/1
Kachkanarsky	0/1	0/1	0/1	0/2	0/1	0/–	0/0
Taborinsky	0/0	0/0	0/0	0/0	0/0	0/–	0/2
The channel reliability night/day (P_{RN}/P_{RD})		0.244/ 0.203	0.178/ 0.068	0.178/ 0.085	0.200/ 0.085	0.444/–	0.222/ 0.051

Analysis of the obtained data allows one to determine the probability of a false detection of the thermal points as the ratio of the total area of the false "hot" pixels to the area of the territorial county where they are found. The area of the MODIS pixel in the selected channels is 1 km^2, so, the probability of false detection of a fire by daylight shots in Kachkanarsky and Taborinsky districts was equal to $3 \cdot 10^{-3}$ and $2 \cdot 10^{-4}$, respectively. Similarly using the night shot of four non-existent fires were found (13 discovered and reported 9) at Ivdelsky district. So, the probability of false detection of fire is $2 \cdot 10^{-4}$.

The probability of correct detection of the thermal point (P_R) in the spectral channel (the reliability of the channel) was determined as the ratio of the total (around snapshot) number of the detected fires in the channel to their actual number. So, the actual number of fires on territories at the night snapshot 28.07.2010 was 45, and number of thermal points found at the channel 20 is 11. As for the day shot, 59 points are actual fires and 12 are detected. So, the reliability of the channel 20 at the night is $P_{RN} = 0.244$, and at the day its $P_{RD} = 0.203$. Similarly, reliability of the remaining (21,...,25) channels (Table 4) was calculated.

Combining the results of several spectral bands, it is possible to was estimate the probability of correct fires detection based on remote sensing data and application of the proposed algorithm. For example, if one of the pixels on the night

image of the Garinsky district exceeded trained 20,...,25 channels thresholds, the probability that it belongs to the fire is

$$P_{RN} = 1 - \prod_{k=20}^{25} (1 - P_{RNk}) = 0.823. \quad (3)$$

Using information of only one channel, you can get the reliability of the fire detection not more than $P_{RN24} = 0.444$. If the pixel at the day image of the Garinsky district exceeded the threshold of 20,...,23, and 25 channels, the probability that it is thermal point is $P_{RD} = 0.410$. And finally, if it is the same pixel on both daylight and night shots, the probability that it belongs to the fire is

$$P_R = 1 - (1 - P_{RN})(1 - P_{RD}) = 0.896. \quad (4)$$

Comparing our algorithm with well-known ones, for example [7], [8], we should say that in them information about statistical characteristics of fire and background pixels is used to define a threshold of detection. But we suggest the approach that doesn't assume any initial allocation of potentially hot pixels for definition of distribution function of probability density for brightness of fire-fighters and the pixels surrounding them. It is based on the fact that detection of thresholds is performed at spectral channels on the basis of allocation of the known fire in the picture, according to the subsatellite information. In this case, it is possible to exclude an uncertainty in temperature of a surface and, partially, in the atmosphere parameters.

The conducted experimental study of our algorithm, which is based on the existing fire, allows one to conclude that the results received with its help aren't worse than ones of the well-known algorithm MOD14. Similar results are received in papers [7], [8] that allows one to judge efficiency of our algorithm in comparison with ones offered there. We plan to conduct further researches using more extensive actual material for receiving adequate quantitative estimates. Combination of our results with ones received by the mentioned algorithms will allow increasing probability of the forest fires correct detection.

The further increase of probability of the correct fire detection according to MODIS is possible by exception of reflections from stationary natural (water) and anthropogenous (industrial and civil constructions) objects, and, also, from nonstationary objects (edges of clouds).

4 Conclusion

The paper presents description of the trained threshold algorithm that uses the true ground information about one of the fires and builds an effective procedures for analysis of the fire situation on the basis of the territory shot obtained by the spectroradiometer MODIS. Study of the proposed algorithm was implemented for the forest fire detection using data by 21 spectral channels of one night and

two daylight MODIS images. Good coincidence of the results of the fires detection with the ground true data and results of work of the well-known MOD14 algorithm was obtained. Informativity of the MODIS spectral channels for forest fire detection using the trained algorithm was analyzed. The example shows that approximately double increase of the probability of correct detection of fires (up to 0.896) is possible by combining the results of spectral bands 20,...,25 of the night and day shots.

5 Acknowledgment

The work was partially supported by the RFBR grants Nos 13-07-12168, and 13-07-00785.

References

1. MODIS specifications. NASA. http://modis.gsfc.nasa.gov/about/specifications.php
2. UniScan-ground station for X–band data reception from polar-orbiting remote sensing satellites. RnD Center "ScanEx". http://www.scanex.com/en/stations
3. Zraenko, S. M., Mymrina, M. A., Ganzha, V. V.: Obuchaemyi algoritm obnaruzheniya pozharov po dannym distantsionnogo zondirovaniya [Learning algorithm of fire detection from remote sensing data]. Zhurnal radioelektroniki: elektronnyi zhurnal, 11 (2014), URL: http://jre.cplire.ru/jre/nov14/8/text.pdf
4. Algorithm Technical Background Document. MODIS Fire Products. http://modis.gsfc.nasa.gov/data/atbd/atbd-mod14.pdf
5. SCANEX IMAGE PROCESSOR. In-depth Processing of Remotely Sensed Data and Creation of Thematic Products. RnD Center "ScanEx". http://www.scanex.com/en/software
6. Zraenko, S. M., Mymrina, M. A., Ganzha, V. V.: Application of remote sensing data for forest fires detection with the use of a trained thresholding algorithm / Proc. 24th International Crimean Conference Microwave and Telecommunication Technology (CriMiCo–2014), 2, 1151–1152 (2014).
7. Ying, Li, Vodacek, A., Kremens, R. L., e.a.: A Hybrid Contextual Approach to Wildland Fire Detection Using Multispectral Imagery. IEEE Transactions on Geoscience and Remote Sensing, 43 (9), 2115–2126 (2005).
8. Chin, S., Liew, A., Lim, L., Keong, K.: New Methods of Active Fire Detection Using MODIS Data. Proc. 26th Asian Conference on Remote Sensing and 2nd Asian Space Conference, 1, 411–415 (2005).

О методах исследования цифровых копий художественных произведений для определения их индивидуальных особенностей

Славных В.А., Сергеев А.П., Филимонов В.В.

УрФУ имени первого Президента России Б.Н. Ельцина
s.nika@list.ru, alexanderpsergeev@gmail.com, fvv1408@list.ru

Аннотация. Работа посвящена обзору отечественных исследований и патентных разработок в области определения индивидуальных особенностей художественного произведения. Систематизированы характерные признаки авторских стилей, выделяемые различными исследователями и представлены методы определения указанных признаков. Обзор опубликованных работ показал, что данная тема недостаточно представлена в отечественной литературе. Большинство исследований посвящено решению задачи идентификации автора.

Ключевые слова: цифровое изображение, атрибуция, классификация, идентификация автора, признаки авторского стиля, экспертиза.

1 Введение

Индивидуальное своеобразие художника формируется в процессе его профессиональной деятельности и выражается в неповторимости его произведения. Среди множества критериев творчества одним из главных является способность отражения индивидуальных позиций и передача мыслей и отношений к реальности.

Установление авторства произведений искусства является актуальной проблемой. Во-первых, она возникает во время аукционных торгов, при определении стоимости произведения искусства. Во-вторых, при работе с музейными коллекциями. В-третьих, существует большой пласт художественных произведений, авторство которых еще не установлено. Актуальным также остается вопрос о большом количестве копий произведений искусства.

Важной задачей является систематизация произведений искусства и организация цифровых баз данных музейных коллекций, архивов и библиотек в Интернете. Подобная деятельность необходима для поиска новой информации о ранее неизвестном произведении и сличения с цифровой копией уже атрибутированных работ предполагаемого автора. Деятельность, направленную на выявление соответствия произведения искусства заявленным атрибутам называют «экспертизой».

Экспертиза произведений искусства решает задачи, которые мы предлагаем разделить на технико-технологические и искусствоведческие.

Технико-технологические задачи - установление состояния сохранности объекта и предполагаемая стоимость реставрации, техники и технологии исполнения представленной работы, ее возраста и рыночной стоимости.

Технико-технологическая экспертиза использует современные физические, химические и физико-химические методы анализа материалов, из которых состоит картина, иногда требующие изъятия пробы. Полное технологическое исследование произведений искусства имеет необходимость использования, технически хорошо оснащенной лаборатории, высококвалифицированного персонала, а также наличие сравнительного материала, достаточно высокую стоимость и по этой причине не всегда целесообразно.

Искусствоведческие задачи - определение авторства художественного произведения, принадлежности художественному направлению, стилю или школе, поиск влияний и связей между художниками, культурной или исторической ценности произведения.

Искусствоведческая экспертиза использует сравнительно-исторический, формальный, иконологический, семиотический методы анализа формы, композиции, технических приемов.

К экспертизе привлекаются специалисты в соответствующей области деятельности. Оценка экспертной комиссии зависит от квалификации и знаний каждого члена группы. Только специалисты могут обеспечить полноту исследования и надежность выводов. В связи с чем, важной задачей представляется изучение и развитие новых компьютерных методов бесконтактного исследования произведений искусства для целей атрибуции.

Цель работы: обзор исследований и патентных разработок в области атрибуции произведений искусства по их цифровым копиям.

2 Отечественные работы по исследованию цифровых копий художественных произведений

Мурашов Д.М., Березин А.В., Иванова Е.Ю. [1-2] (Москва, Вычислительный центр им. А.А. Дородницына РАН, Государственный Исторический музей) предложили описание фактуры картин на основе характеристик хребтов полутоновых изображений мазков и локальных признаков структурного тензора на изображении фрагмента. «Гистограмма формируется по значениям составляющих градиента уровней полутонов в точках, расположенных на хребтах информативных фрагментов» [1].

Размер кисти художника (толщина волоса, ширина кисти) вычисляется на основе значения модуля волнового вектора, который определяет пространственный период волны в направлении ее распространения. Данные для исследования формируется по информативным фрагментам изображений без предварительной сегментации отдельных мазков кисти. Для сравнения информативных фрагментов применяется теоретико-информационная мера различия на основе дивергенции Кульбака-Лейблера - меры удаленности друг от друга двух вероятностных распределений. Данные исследования показали, что значения исследуемого признака фрагментов картин разных авторов различаются в пять раз больше, чем значения признака на картинах одного автора. Описанная методика была протестирована на изображениях портретов, написанных в XVIII–XIX вв. [1].

Иваненко Ю.М., и др. [3] выдвинули гипотезу: живописное произведение обладает внутренней динамикой, которая характеризуется соотношением детерминированности и хаотичности. Исследователи выявили, что полутоновые изображения отличаются конечностью размерности аттрактора. На основе отсчетов

яркости реального изображения были выбраны специальные переменные для построения многомерного фазового пространства, в котором производился анализ динамики исследуемого изображения путем вычисления интегральной корреляционной функции аттрактора. Исследователи установили, что значения размерности аттрактора отражают вид изображения: «полиграфическая копия» произведения искусства, фотографическое изображение лица, фрагменты скульптурного изображения, зафиксированные с помощью видеокамеры, модельные изображения различных объектов, сгенерированные на компьютере, в соответствии с разработанной шкалой сложности изображений. Определенная зависимость размерности аттрактора от числа фазовых переменных позволяет произвести идентификацию индивидуального творческого почерка художника [3].

Новиков Г.И. [4] (Институт физиологии им. И.П. Павлова РАН) предложил для идентификации автора художественных произведений использовать колориметрические измерения «контраста цветового насыщения», (противопоставление ярких, насыщенных и блеклых, затемненных цветов), зависящего от пропорций и равновесности композиции, присущего конкретным авторам. Исследователь предложил вычислять соотношение основных и дополнительных цветов на основе сопоставления цвета картины в разных точках по отношению к нейтральному серому фону. Вывод был основан на рассмотрении колориметрических измерений объектов русской и западноевропейской живописи второй половины XIX и начала XX веков.

Костюков Ю.В. [5] (Краснодарский краевой художественный музей им. Ф.А. Коваленко) предлагает осуществлять отбор характерных признаков для исследования художественного произведения, статистическими методами, которые позволят выявить тесноту связи между признаками и выделить наиболее идентификационно значимые. Он предложил устанавливать факт наличия взаимосвязи между признаками и ее значимости на основе критерия хи – квадрат. Описанная методика была протестирована при анализе декорировки столовой и чайной посуды завода Гарднера конца XVIII – последней трети XIX вв. В результате исследования Костюков Ю.В. установил, что на изделиях завода Гарднера наиболее тесно связаны между собой признаки «тема» и «форма» декора. Зная значение одного из них, можно с определенной вероятностью предсказать значение другого [5].

3 Отечественные патенты на изобретения

Хачатрян А.Х. [6] (ООО «Научно-исследовательский институт экспертизы произведений искусств и архитектуры») описывает три запатентованных технологии бесконтактного неразрушающего анализа произведений искусства:

- «Способ бесконтактного неразрушающего анализа художественных полотен и других произведений искусства по их цифровым копиям для определения индивидуального рисунка рельефа и углублений поверхностных и внутренних слоев» [7].
- «Способ опосредованного неразрушающего исследования художественных полотен, монументального и других произведений искусства для выявления их скрытых индивидуальных особенностей (варианты), способ определения подлинности и/или авторства художественного или других произведений искусства» [8].

- «Способ защиты художественного полотна и других произведений искусства от подделки, а также определения их авторства» [9].

Цифровая копия обрабатывается в графическом редакторе (увеличение масштаба, затемнение до появления углублений и контуров, осветление до более четкого проявления рельефа, обработка при помощи инструмента «контурная резкость» до получения максимально четких контуров). При этом каждая отдельно взятая цифровая копия художественного произведения разделяется на слои и сохраняется в отдельные файлы, как часть исследуемого произведения искусства. Каждая из них обрабатывается отдельно; количество слоев конкретного произведения зависит от индивидуального характера каждого из них, а каждый изученный слой рассматривается как отдельное исследование. Далее эксперт зрительно анализирует полученные копии художественного полотна [6].

По мнению авторов, предложенные способы выявляют скрытые индивидуальные особенности композиции и характер художественного произведения, что позволяет определить их подлинность, послойно рассмотреть картины разных художников, обнаружить тайные знаки, раскрывающие замысел и обнаружить пустоты, полости и скрытые предметы в художественных произведениях [6].

Кастальская-Бороздина Н.К. [10] (Свято-Троице Сергиева Лавра) предлагает способ идентификации художественных произведений через сопоставление наборов зашифрованных данных фрагментов живописи по соответствующим цветам палитры произведения. Исследователь утверждает, что у каждого автора существует неповторимый набор исходных «полу(четверть) тоновых шкал колеров» [10] ожидаемых в любой его работе. Кастальская-Бороздина Н.К. предлагает идентификацию исследуемого объекта осуществлять на основе сравнения «выявленных уникальных признаков произведения живописи и оригинала по графическим зависимостям коэффициента гашения обертонов от суммарного колера, состоящего из красного, зеленого и синего цветов, вычисленным по самым темным, самым светлым и срединным колерам с учетом естественной аппаратной погрешности» [10]. Описанный способ был опробован на цифровых копиях таких художников как, Айвазовский И., Левитан И., Кустодиев Б., Поленов В.

Шольцен В. [11] предлагает способ определения автора художественного произведения по ее цифровой копии путем сравнения характеризующих признаков и деталей характеризующих признаков (точек, линий, групп точек или линий, узоров), полученных преобразованием Хафа.

4 Характерные признаки авторских стилей

Таким образом, на основе проведенного обзора можно выделить следующие характерные признаки авторских стилей и способы их вычисления, используемые различными исследователями.

Фактура картины:

Определяемая, гистограммой направлений хребтов полутоновых изображений мазков, формируемой на основе вторых частных производных от функции, описывающей полутоновый рельеф изображения по пространственным координатам и гистограммой локальной ориентации окрестностей на изображении фрагмента картины, формируемой на основе первых частных производных от функции, описывающей полутоновый рельеф изображения по пространственным координатам [2].

Получаемая, после обработки в графическом редакторе Adobe Photoshop или FS Viewer путем масштабирования, затемнения, осветления, увеличения контурной

резкости (настройка параметров эффекта резкости, радиуса пикселей, и порога уровней), яркости и контрастности. [8]

Получаемая, после обработки в графическом редакторе путем масштабирования, затемнения, осветления, увеличения контурной резкости (настройка параметров эффекта резкости, радиуса пикселей, и порога уровней) [9].

Размер кисти, использовавшейся художником – локальное волновое число, получаемое из изображения при использовании преобразования Рисса [1].

Выразительные графические средства (точки, группы точек, линии, группы линий, узоры), определяемые методом Хафа. Дополнительно устанавливаются опорные признаки, путем масштабирования, растяжения или сжатия, изменения или изгибания линии, изменения угла между двумя линиями соответствующего характеризующего признака [11].

Цветовой контраст:

Яркость изображения, определяемая зависимостью размерности аттрактора от числа фазовых переменных [3].

Контраст цветового распространения, зависящий от пропорций и равновесности композиций автора. Сопоставление цветов картины в разных точках по отношению к нейтральному серому фону. Количественные расчеты производятся исходя из установленных И. Гёте простых числовых выражений. Контраст цветового распространения подчинен гармонии равновесия отношения светлоты и дополнительных цветов. [4]

Цвет, получаемый после обработки в графическом редакторе Adobe Photoshop или FS Viewer путем масштабирования, затемнения, осветления [8].

Цветовые пятна, определяемые графической зависимостью коэффициента гашения обертонов и суммарного колера, вычисленным по самым темным, самым светлым и серединным колерам с учетом естественной аппаратной погрешности [10].

5 Выводы

В результате проведенного обзора были выявлены характерные признаки авторских стилей, выделяемые различными исследователями для целей атрибуции художественных произведений. Все признаки можно разделить на три основные группы: фактура картины, цветовой контраст, выразительные графические средства.

Рассмотренные в настоящей работе исследования в области анализа цифровых копий произведений искусства направлены на определение авторства художественного произведения. Однако в области экспертизы художественных произведений остается еще целый ряд задач, которые можно решать: классификация стилей, жанров и периода создания, установление взаимовлияний и связей между художниками. Дальнейшие исследования могут быть направлены на развитие и апробацию бесконтактных методов анализа художественных произведений, которые недостаточно представлены в отечественной литературе, а также изучение опыта зарубежных исследователей.

В представленных работах не предложено описание технологии получения и критерии оценки качества, предъявляемые к цифровым копиям художественных произведений, в свиязи с чем, возникает вопрос об объективности результатов исследования изображений, имеющихся в сети Интернет.

Очевидно, что для успешного использования цифровых бесконтактных методов анализа художественных произведений необходимо наличие обширных баз оригинальных атрибутированных работ, а так же баз вычисленных данных для сравнения.

Список литературы

1. Мурашов Д.М., Березин А.В., Иванова Е.Ю. Формирование признакового описания фактуры картин // Машинное обучение и анализ данных. 2013. Т. 1, № 6. С. 779–786.
2. Мурашов Д.М., Березин А.В., Иванова Е.Ю. Сравнение изображений картин по информативным фрагментам // Машинное обучение и анализ данных. 2014. Т. 1. № 8. С. 941–948.
3. Иваненко Ю.М., Макулов В.Б., Раскин А.Г., Трифонов М.И. Математическая формализация художественной ценности произведения искусства, их голографический копий и компьютерных изображений // Информационный бюллетень РФФИ, 5 (1997) Науки о человеке и обществе.
4. Новиков Г.И.Колориметрические исследования в живописи // Материалы IX Международной конференции «Прикладная оптика-2010», 18–22 октября 2010, Санкт-Петербург, С. 236–240.
5. Костюков Ю.В. Возможности применения вероятностно-статистического метода для оценки значимости признаков в идентификации произведений искусства // Теория и практика актуальных исследований (Материалы Международной научно-практической конференции, 17 апреля 2012).
6. Хачатрян А.Х. Исследование произведений монументального и изобразительного искусства с помощью нового изобретения, его роль и важность в мировой науке // Гуманитарные, социально-экономические и общественные науки. 2013, № 4. С. 461–470.
7. Хачатрян А.Х., Симонян З.А., Хачатрян В.А. Способ бесконтактного неразрушающего анализа художественных полотен и других произведений искусства по их цифровым копиям для определения индивидуального рисунка рельефа и углублений поверхностных и внутренних слоев // Патент России № 2012157043, 10.07.2014.
8. Хачатрян А.Х., Симонян З.А., Хачатрян В.А. Способ опосредованного неразрушающего исследования художественных полотен, монументального или других произведений искусства для выявления их скрытых индивидуальных особенностей (варианты), способ определения подлинности и/или авторства художественного полотна, монументального или других произведений искусства // Патент России № 2012138269, 20.03.2014.
9. Хачатрян А.Х., Симонян З.А., Хачатрян В.А. Способ анализа художественных полотен и других произведений искусства для определения их авторства и/или защиты от подделки // Патент России № 2012118006, 10.11.2013.
10. Кастальская-Бороздина Н.К. Способ идентификации произведений живописи на предмет их авторства // Патент России № 2333613, 10.09.2008.
11. Шольцен В. Устройство и способ идентификации автора произведения искусства // Патент России № 2011149157, 27.06.2013.

About The Methods of Research Digital Copies Works of Art to Determine Their Specific Features

Viktoriya Slavnykh, Alexander Sergeev, Viktor Filimonov

Ural Federal University named after the first President of Russia Boris Yeltsin,
s.nika@list.ru, alexanderpsergeev@gmail.com, fvv1408@list.ru

Abstract. The work is devoted to the review of Russian studies and patents to determine the specific features of works of art. The article provides a classification the characteristic features of the author's style works of painting and are described methods for determining specified characteristics.

Keywords: digital image, image features, classification, attribution of paintings.

Эволюционирующие онтологии в аспекте управления темпоральными или изменяющимися фактами

А. А. Демидов

Институт программных систем им. А.К. Айламазяна РАН,
alex@dem.botik.ru

Аннотация В работе предлагается алгебраический подход к построению онтологий, способных к эволюции под влиянием новых фактов и обладающих внутренними механизмами валидации. Для этой цели строится формальная модель взаимодействия объектов и выясняются ограничения на операции с объектами, накладываемые такой моделью. Затем в контексте формальной модели определяются основные понятия модели представления знаний: концепты, экземпляры, свойства и отношения. При этом формальные ограничения переносятся в модель представления знаний естественным образом.

Ключевые слова: эволюционирующие онтологии, представление знаний, формальные системы.

Введение

Онтология — это система понятий и утверждений об этих понятиях, на основе которых можно строить классы, объекты, отношения, функции и теории. В общем случае онтологии содержат концепты (понятия или классы), экземпляры (индивиды или объекты), свойства концептов и экземпляров (атрибуты или роли), отношения между концептами или экземплярами, а также дополнительные ограничения, определяемые аксиомами и правилами вывода. Сформулировано множество определений понятия онтологии, наиболее точным из них представляется следующее: «Онтология — это формальная теория, ограничивающая возможные концептуализации» [1]. Поскольку мы хотели бы также хранить в онтологии отдельные экземпляры, данное определение требует обобщения: «Онтология — это формальная система, взятая вместе с её интерпретацией».

Эволюция онтологии — это процесс последовательной адаптации онтологии к происходящим изменениям и непротиворечивого распространения этих изменений [2]. Это нетривиальный процесс, поскольку изменение в одной части онтологии может привести к возникновению противоречия в других её частях, информация различных источников может оказаться неполной или противоречивой. В силу сложности учёта большого количества факторов инженер по знаниям оказывается не в состоянии учесть все побочные

эффекты вносимых изменений, поэтому необходима реализация специальных механизмов, отвечающих за поддержание согласованности онтологии в процессе её актуализации.

Можно выделить два основных подхода к представлению онтологий: логический и объектный [3]. Логический подход связан с представлением онтологии как формальной системы со своим синтаксисом, аксиомами и правилами вывода, он находится в русле классического искусственного интеллекта, изучающего способы представления знаний. Объектный подход предполагает представление онтологии в виде графа, состоящего из классов, объектов и связей между ними, он удобнее для реализации и чаще используется в прикладных разработках, в рамках него обычно создаются сверхбольшие ресурсы, используемые в широких предметных областях: различного рода словари, таксономии, рубрикаторы и тезаурусы.

Обладая лучшей наглядностью и удобством, объектный подход имеет серьёзный недостаток: он не предоставляет никаких внутренних механизмов контроля непротиворечивости — в отличие от формальных систем, имеющих средства контроля полноты и непротиворечивости. Поэтому при объектном подходе инженер по знаниям должен определить систему внешних ограничений, описывающих допущения предметной области. Но проблема в том, что предметная область сама зависит от вновь прибывающих фактов, она всё время меняется. Поэтому система внешних ограничений также должна всё время меняться. Но тогда возникает проблема контроля непротиворечивости изменений уже этой системы ограничений — получается замкнутый круг.

1 Постановка задачи

Несложно показать, что онтологии, создаваемые в рамках обоих подходов, могут быть преобразованы от одного вида к другому путём перехода к отношениям: каждая непротиворечивая формальная теория имеет модель в теоретико-множественном представлении, с другой стороны, каждый объект может быть представлен как набор записей реляционной базы данных. Поэтому корректней говорить о различных представлениях одной и той же онтологии — логическом и объектном. Вопрос состоит в том, какой возможный механизм контроля в объектном представлении может соответствовать концепции противоречивости в логическом представлении.

2 Состояние исследований

Большинство работ по эволюции онтологии, известных в настоящее время, описывают различные эвристические подходы — их обзор дан в работе Ф. Заблиса и др. [4]. Как правило, в этих работах предлагается некий сложный эвристический алгоритм, берущий на себя рутинную часть работы по согласованию изменений, и предполагается участие инженера по знаниям, выполняющего интеллектуальную часть работы. Другие системы — такие

как EvoPat, RUL — реализуют возможность автоматического разрешения конфликтов, но они действуют на основе жесткой системы внешних ограничений, заданных априори при создании онтологии.

Более перспективно выглядят методы эволюции онтологий, обоснованные теоретически. В работе П. Хаазе и Л. Стояновича [5] выделяется три вида непротиворечивости: структурная (синтаксическая), логическая (семантическая) и определяемая пользователем (внешняя). Большинство динамических онтологий обеспечивают только синтаксическую непротиворечивость на основе соответствия языку описания онтологии или схеме данных. В работе не предлагается решения, обеспечивающего логическую непротиворечивость, однако приводятся эффективные алгоритмы локализации и устранения логических противоречий с помощью исключения конфликтующих аксиом.

В работах Т. Шарренбаха и др. [6,7] конфликты аксиом не запрещаются, поэтому их исключение из онтологии не требуется. Вместо исключения аксиом (явных знаний) авторы предлагают объявлять недействительными конкретные выводы (неявные знания), которые вызывают противоречие в онтологии. Так, оба заключения A и $\neg A$ могут оказаться выводимы, невыводимым останется только противоречие $A\&\neg A$. Это приводит к неоднозначности онтологии.

Задача поиска механизма контроля онтологии в объектном представлении, который соответствовал бы концепции противоречивости в логическом представлении, в явном виде ранее не ставилась.

3 Мир как динамическая система

3.1 Формальная логика и клеточные автоматы

Машина Тьюринга как универсальный вычислитель способна реализовать как систему продукций аксиоматизируемой формальной логической системы, так и систему продукций клеточного автомата. Отсюда следует, что выразительные способности тьюринг-полных клеточных автоматов достаточны для представления логических ограничений на структуру знаний.

Далее мы будем рассматривать явления предметной области в терминах динамических систем — как изменения состояния динамической системы, эволюционирующей с течением времени. Покажем, что клеточный автомат также может рассматриваться как динамическая система, введём в этой модели понятие объектов, опишем их взаимодействия. Поскольку возможностей тьюринг-полного клеточного автомата достаточно для моделирования произвольного процесса эволюции, то найденные закономерности будут справедливы и в общем случае.

3.2 Динамические системы

В самом общем смысле динамическая система есть тройка (T, X, ϕ), где T — аддитивный моноид, X — множество, а ϕ — функция

$$\phi : T \times X \to X, \tag{1}$$

такая что

$$\begin{aligned} \phi(0, x) &= x, \\ \phi(t_2, \phi(t_1, x)) &= \phi(t_1 + t_2, x), \end{aligned}$$

где T — множество неотрицательных вещественных чисел $\mathbb{R}^+$ (непрерывное время), либо натуральных чисел с нулём $\mathbb{N}$ (дискретное время) [8,9].

Функция $\phi(t, x)$ называется оператором эволюции динамической системы, она ставит в соответствие каждой точке множества X единственный образ, зависящий от переменной t, называемой параметром эволюции. Множество X называется фазовым пространством (или пространством состояний), в котором переменная x определяет начальное состояние системы.

3.3 Клеточные автоматы

Пусть G — группа, а A — множество. Тогда (см. [10])

Определение 1. *Клеточный автомат над группой G и алфавитом A есть отображение $\tau : A^G \to A^G$, обладающее свойством: существует конечное подмножество $S \subset G$ и отображение $\mu : A^S \to A$, такие что*

$$\tau(x)(g) = \mu((g^{-1}x)|_S) \tag{2}$$

для всех $x \in A^G$ и $g \in G$, где запись $(g^{-1}x)|_S$ означает сужение конфигурации $g^{-1}x$ на множество S.

Путём последовательных итераций клеточного автомата $\tau : A^G \to A^G$ можно получить дискретную динамическую систему. Это означает, что конфигурация может пониматься эволюционирующей во времени в соответствии с τ: если $x \in A^G$ — конфигурация в момент времени $t \in \mathbb{N}$, то $\tau(x)$ — конфигурация в момент времени $t + 1$. Суперпозиция $\tau^t = \tau \circ \tau \circ \cdots \circ \tau$ итераций по параметру t даёт оператор эволюции динамической системы во времени

$$\phi(t, x) = \tau^t(x). \tag{3}$$

3.4 Объекты в конфигурационном пространстве

Неформально говоря, объектом будем считать глайдер в конфигурационном пространстве. Глайдеры — это хорошо известные стабильные динамические структуры (паттерны) в клеточном пространстве игры Конвея

«Жизнь». Глайдер (солитон) является возмущением в активной среде, распространяющимся в этой среде с постоянной скоростью и сохраняющим свою целостность в течение некоторого времени.

Отображение $\mu : A^S \to A$ в формуле (2) определяет систему окрестностей $O(g) = g \cup Sg$ для каждого $g \in G$. Окрестностью множества $O(g)$ является объединение

$$O(O(g)) = \bigcup_{q \in O(g)} \big(q \cup Sq\big). \tag{4}$$

Последовательные итерации этой формулы задают систему окрестностей $\mathcal{P}$, определяющую предбазу топологии на G

$$\mathcal{P}(t, g) = \{O^t(g) : t \in \mathbb{N}, g \in G\}. \tag{5}$$

Если для $g_t, g_0 \in G$ выполняется условие $g_t \in O^t(g_0)$, то будем говорить, что элемент g_t причинно зависит от g_0 на интервале времени t.

Определение 2. *Объектом O назовём динамическую область конфигурационного пространства $O(g_i) \in \mathcal{P}$, где g_i пробегает последовательность причинно зависимых элементов $g_0, \ldots, g_t$, такую что все конфигурации $A^{O(g_i)}$ равны между собой.*

4 Модель взаимодействия объектов

4.1 Взаимодействие как эволюция состояния

Будем рассматривать изменение объектов — их взаимное превращение в результате взаимодействия в некоторой области конфигурационного пространства G с заданной на этом пространстве конфигурацией $x \in A^G$, эволюционирующей под действием моноида времени T.

Нас будет интересовать только конечный результат взаимодействия после того, как объекты разойдутся в конфигурационном пространстве на достаточное расстояние, чтобы не оказывать влияния друг на друга.

Обозначим через Ψ множество конфигураций всех возможных объектов

$$\Psi = \{A^O : O \in \mathcal{P}\}. \tag{6}$$

Тогда процесс изменения конфигурации $\psi = \psi_1 \otimes \ldots \otimes \psi_n$ системы взаимодействующих объектов можно записать в привычном виде

$$\psi' = U\psi, \tag{7}$$

где $U = U(t, x)$ — оператор эволюции состояния $x \in A^G$ за время $t \in T$, а $\psi \in \Psi$ — функция состояния системы объектов.

4.2 Взаимодействие как преобразование объектов

Произвольным образом разобьём исходную динамическую систему на две подсистемы — ψ_a и ψ_b так, чтобы $\psi_a\psi_b = \psi$ (для удобства будем опускать знак операции «$\otimes$»). Взаимодействие систем ψ_b и ψ_a тогда может быть задано с помощью операции «$\odot$» в виде

$$\begin{cases} \psi'_a = \psi_a \odot \psi_b \\ \psi'_b = \psi_b \odot \psi_a, \end{cases} \tag{8}$$

где первое равенство описывает трансформацию ψ_a в ψ'_a с помощью системы ψ_b, а второе — обратное влияние системы ψ_a на ψ_b.

Модель взаимодействия тогда вполне определяется двумя операциями

$$\otimes\colon \Psi^2 \to \Psi, \quad \text{«из чего состоит» в пространстве,} \tag{9}$$

$$\odot\colon \Psi^2 \to \Psi, \quad \text{«как изменяется» во времени.} \tag{10}$$

Операция «$\otimes$» ассоциативна, поскольку свойство «из чего состоит», очевидно, удовлетворяет тождеству $(\psi_a \otimes \psi_b) \otimes \psi_c = \psi_a \otimes (\psi_b \otimes \psi_c)$. Но не коммутативна — порядок аргументов важен, поскольку задаёт расположение объектов в пространстве. Свойства операции «$\odot$» будут установлены далее (очевидна лишь некоммутативность).

Правила преобразования конфигураций могут быть заданы функциями на множестве Ψ. Иначе говоря, существует инъекция

$$\xi\colon \Psi \to \mathcal{F}(\Psi), \tag{11}$$

позволяющая отождествить конфигурации объектов Ψ с элементами симметрической полугруппы $\mathcal{F}(\Psi)$ преобразований над ними.

4.3 Связь двух операций «$\otimes$» и «$\odot$»

Теорема 1. *Инъекция $\xi\colon \Psi \to \mathcal{F}(\Psi)$ множества конфигураций объектов в симметрическую полугруппу преобразований является гомоморфизмом относительно операции «$\odot$», если эта операция действует в пространстве T, связанном с пространством G преобразованием масштаба.*

Доказательство. Зададим начальное разбиение системы $\psi \in \Psi$ в пространстве G так, чтобы она состояла из трёх подсистем следующим образом: $\psi_a\psi_b\psi_c = \psi$. Будем попарно объединять эти подсистемы и описывать взаимодействие такого объединения с оставшейся частью полной системы.

Поскольку результат операции «$\otimes$» зависит от порядка аргументов, при объединении необходимо следить за тем, в какой системе координат производится данное действие. Для удобства введём унарную операцию $\neg\colon \Psi \to \Psi$ поворота системы координат, связанную с операцией «$\otimes$» тождеством

$$\neg(\psi_a\psi_b) = \neg\psi_b\neg\psi_a. \tag{12}$$

Объединим подсистемы $\psi_a\psi_b = \psi_{ab}$, тогда правила трансформации динамической системы $\psi_a\psi_b\psi_c = \psi_{ab}\psi_c$ запишутся в виде

$$\begin{cases} \psi'_{ab} = \psi_{ab} \odot \psi_c = \psi_c(\psi_{ab}) \\ \psi'_c = \neg(\neg\psi_c \odot \neg\psi_{ab}) = \neg\psi_{\neg b\neg a}(\psi_{\neg c}). \end{cases} \tag{13}$$

Во втором уравнении появляется операция обращения аргументов, поскольку относительно подсистемы ψ_c, на которую действует объединённая подсистема, подсистемы ψ_a и ψ_b расположены в порядке ψ_b, ψ_a — именно в таком порядке и происходит действие. Операция обращения всего выражения отвечает за возврат системы координат в исходное положение.

Теперь объединим подсистемы $\psi_b\psi_c = \psi_{bc}$, тогда правила трансформации динамической системы $\psi_a\psi_b\psi_c = \psi_a\psi_{bc}$ запишутся в виде

$$\begin{cases} \psi'_a = \psi_a \odot \psi_{bc} = \psi_{bc}(\psi_a) \\ \psi'_{bc} = \neg(\neg\psi_{bc} \odot \neg\psi_a) = \neg\psi_{\neg a}(\psi_{\neg c\neg b}). \end{cases} \tag{14}$$

Во втором уравнении также появляется операция обращения аргументов по аналогичной причине: подсистема ψ_a действует сперва на ближайшую к ней подсистему ψ_b, и только затем — на ψ_c. Операция обращения всего выражения отвечает за возврат системы координат в исходное положение.

Поскольку $\psi'_{ab}\psi'_c = \psi'_a\psi'_{bc} = \psi' = U\psi$, то можно приравнять правила трансформации первой и второй систем уравнений:

$$((\psi_a \otimes \psi_b) \odot \psi_c) \otimes\neg(\neg\psi_c \odot (\neg\psi_b \otimes \neg\psi_a)) = \tag{15}$$

$$= (\psi_a \odot (\psi_b \otimes \psi_c)) \quad \otimes\neg((\neg\psi_c \otimes \neg\psi_b) \odot \neg\psi_a). \tag{16}$$

Равенство выполняется, если и только если существует унарная операция $\propto\colon \Psi \to \Psi$, связывающая операции «$\odot$» и «$\otimes$» тождеством

$$\propto(\psi_a\psi_b) = \psi_a \odot \psi_b, \tag{17}$$

— тогда равенство можно преобразовать к очевидной форме:

$$\propto((\psi_a \otimes \psi_b) \otimes \psi_c) \otimes\neg \propto(\neg\psi_c \otimes (\neg\psi_b \otimes \neg\psi_a)) = \tag{18}$$

$$= \propto(\psi_a \otimes (\psi_b \otimes \psi_c)) \quad \otimes\neg \propto((\neg\psi_c \otimes \neg\psi_b) \otimes \neg\psi_a). \tag{19}$$

Унарная операция «$\propto$» существенно ограничивает возможный вид операции «$\odot$», в частности, должно соблюдаться условие

$$(\forall\psi_x, \psi_y \in \Psi) \quad \psi_x\psi_y = \psi_a \;\to\; \psi_x \odot \psi_y = \propto \psi_a, \tag{20}$$

которое означает, что все решения уравнения $x \otimes y = \psi_a$ с двумя неизвестными также являются решениями уравнения $x \odot y = \propto \psi_a$. В этом смысле операции «$\otimes$» и «$\odot$» эквивалентны с точностью до преобразования масштаба, определяемого операцией «$\propto$». При этом операция «$\odot$» оказывается, вообще говоря, неассоциативной:

$$(\psi_a \odot \psi_b) \odot \psi_c \;\neq\; \psi_a \odot (\psi_b \odot \psi_c), \tag{21}$$

$$\propto (\propto (\psi_a\psi_b)\psi_c) \neq \propto (\psi_a \propto (\psi_b\psi_c)). \tag{22}$$

Введём формальное обозначение $\propto^{-1}\propto \psi_a = \psi_a$ и перепишем тождество (17) в виде

$$\propto (\psi_a\psi_b) = \propto^{-1}\propto \psi_a \odot \propto^{-1}\propto \psi_b, \tag{23}$$

обозначим операцию «$\odot$» вместе с преобразованием «$\propto^{-1}$» обратного масштабирования её аргументов через $\tilde{\odot}\colon \Psi^2 \to \Psi$, тогда станет очевиден изоморфизм

$$\propto (\psi_a \otimes \psi_b) = \propto \psi_a \tilde{\odot} \propto \psi_b. \tag{24}$$

Также, после замены переменных $\psi_a \to \propto \psi_a$ и $\psi_b \to \propto \psi_b$ в тождестве (17) и перехода к обозначению $\tilde{\otimes}\colon \Psi^2 \to \Psi$, объединяющему операцию «$\otimes$» вместе с преобразованием «$\propto$» прямого масштабирования её аргументов, выявляется изоморфизм

$$\propto (\psi_a \tilde{\otimes} \psi_b) = \propto \psi_a \odot \propto \psi_b. \tag{25}$$

Изоморфизмы (24) и (25) связывают пространства T и G преобразованием масштаба, определяемом операцией (17).

Операция «$\otimes$» — ассоциативна, поэтому полугруппа $\langle \Psi, \otimes \rangle$ вкладывается в симметрическую полугруппу $\mathcal{F}(\Psi)$ трансформаций множества Ψ. В силу изоморфизма (24) операция «$\odot$» тождественна операции «$\otimes$», если все её аргументы переведены в пространство T преобразованием «$\propto^{-1}$». Следовательно, инъекция $\xi\colon \Psi \to \mathcal{F}(\Psi)$ (11) является однозначным гомоморфизмом относительно операции «$\tilde{\odot}$».

4.4 Спецификация модели взаимодействия объектов

Таким образом, построенная модель включает:

- алгебру $\langle \Psi, \otimes, \odot \rangle$ динамических систем (6) и (9);
- симметрическую полугруппу $\mathcal{F}(\Psi)$ преобразований (11);
- инъективный гомоморфизм $\xi\colon \Psi \to \mathcal{F}(\Psi)$ (теорема 1);
- преобразование масштаба $\propto (\psi_a\psi_b) = \psi_a \odot \psi_b$ (17),

где операции «$\otimes$» и «$\odot$» связаны равенством

$$\psi_a \otimes \psi_b = (\propto^{-1} \psi_a) \odot (\propto^{-1} \psi_b). \tag{26}$$

5 Модель представления знаний

Построенная модель взаимодействия объектов сводится к подполугруппе $\mathcal{J}$ симметрической полугруппы $\mathcal{J} \subset \mathcal{F}(\Psi)$ преобразований множества Ψ, которая равномощна этому множеству $|\Psi| = |\mathcal{J}| < |2^{\Psi}|$. Такое сокращение числа допустимых преобразований указывает на существование естественных модельных ограничений, что может быть использовано для целей контроля онтологии, однако число степеней свободы всё ещё очень велико.

Определение 3. *Если по отношению к динамической системе ψ_a подмножество систем $\Psi_B \subset \Psi$ является эквивалентным в смысле тождества*

$$(\forall \psi_b \in \Psi_B)(\psi_a \odot \psi_b = \psi'_a),$$

то множество Ψ_B будем называть классом со свойством $\psi_a \to \psi'_a$.

Совокупность всех свойств объекта ψ_b полностью определяет преобразование $\psi_b \colon \Psi \to \Psi$ из полугруппы $\mathcal{J}$. Совокупность всех свойств класса также определяет преобразование, частично определённое на множестве Ψ, которое является пересечением преобразований всех объектов данного класса

$$\psi_B = \bigcap_{\psi_b \in \Psi_B} \psi_b. \tag{27}$$

В эволюционирующей онтологии ни один объект не определён окончательно, под влиянием новых фактов отношения достраиваются и перестраиваются. Поэтому между объектами и классами нет принципиальной разницы — те и другие являются частично определёнными преобразованиями.

В прикладных задачах построения онтологий часто приходится иметь дело не с функциями, а с отношениями. Можно показать, что такой переход правомерен, если в теореме 1 перейти к гомоморфизму $\xi \colon \Psi \to 2^{\Psi}$. В силу однозначного гомоморфизма $\xi \colon \Psi \to 2^{\Psi}$ операция «$\otimes$» является суперпозицией отношений соответствующих объектов. С другой стороны, она выражается через операцию «$\odot$» с помощью унарной операции «$\propto^{-1}$» преобразования масштаба (26). Тождество

$$\psi_b \circ \psi_a = \psi_a \otimes \psi_b = (\propto^{-1} \psi_a) \odot (\propto^{-1} \psi_b), \tag{28}$$

где слева стоят отношения, а справа — соответствующие им объекты, является мощным средством обеспечения непротиворечивости онтологии, для чего даже не требуется просматривать весь ресурс знаний.

Для организации хранения знаний можно предложить следующую базовую структуру из двух таблиц.

Таблица 1. Структура хранения знаний (с валидацией)

head	
id	Идентификатор
label	Наименование

body	
subj	Subject ($\to$ head.id)
obj	Object ($\to$ head.id)
prop	Result *subj* $\odot$ *obj* ($\to$ head.id)
test	Result *subj* $\otimes$ *obj* ($\to$ head.id)

На начальном этапе все таблицы пусты. В процессе работы каждое новое понятие фиксируется в таблице **head**, а каждое его свойство попадает в таблицу **body**. Пока функция $\propto^{-1}$ не определена, поле body.test не заполняется

— это можно сделать позже на основе значения body.prop. Классы можно получать динамически как пересечение нескольких отношений, однако из соображений быстродействия имеет смысл запустить автономный процесс, который бы создавал и удалял классы на основе устойчивых корреляций между объектами. Этот же процесс мог бы выполнять трудоёмкую задачу вычисления полей body.test пока быстрое вычисление с использованием функции $\propto^{-1}$ не станет возможным.

Тождество (28) позволяет вычислять значения функции $\propto^{-1}: \Psi \to \Psi$ динамически по мере добавления новых фактов: поскольку все объекты являются отношениями, кажется, что их суперпозиция тоже должна задавать объект; это почти так — если отношения предварительно масштабировать функцией $\propto^{-1}$, а результат масштабировать обратно с помощью $\propto$. Для всех пар $\psi_a \circ \psi_b$ данная функция должна быть одной и той же — если так не получается, то можно констатировать наличие в онтологии противоречия.

6 Заключение

В работе предложен алгебраический подход к построению онтологий, способных к эволюции под влиянием новых фактов и обладающих внутренними механизмами валидации. Для этой цели установлено соответствие между двумя основными подходами к представлению онтологий: логическим и объектным. В результате этого в объектном представлении удалось отыскать такой механизм контроля непротиворечивости онтологии, который соответствует концепции противоречивости в логическом представлении — на основе функции $\propto^{-1}$. В терминах алгебраического подхода определены основные понятия модели представления знаний: концепты, экземпляры, свойства и отношения. При таком подходе формальные ограничения переносятся в модель представления знаний естественным образом.

Найденное решение является новым, его можно использовать для построения больших прикладных онтологий, не пренебрегая средствами контроля непротиворечивости. Что, в свою очередь, открывает возможности к организации процесса эволюции такой онтологии в автоматическом режиме, что до сих пор было невозможным.

Построенная алгебра с двумя операциями — суперпозицией отображений «$\otimes$» и неассоциативной операцией «$\odot$» — является программной алгеброй. Программные алгебры интенсивно исследуются в настоящее время в контексте построения высокопроизводительных параллельных вычислительных систем [11].

Результаты данной работы планируется использовать при построении ресурса знаний большого объёма для системы извлечения информации из текстов ИСИДА-Т [12].

Благодарности. Работа выполнена в рамках НИР «Моделирование модально–временного аспекта описания ситуаций в задаче извлечения информации из текстов», номер гос. регистрации 01201455353.

Список литературы

1. *Guarino, N., Giaretta, P.* Ontologies and Knowledge Bases // Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing // International Conference on Building and Sharing Very Large-Scale Know: Towards a Terminological Classification — Amsterdam: Ios Press, 1995, pp. 3–28.
2. *Stojanovic, L.* Methods and Tools for Ontology Evolution, PhD thesis, University of Karlsruhe, 2004, August 5. — 249 p.
3. *Лукашевич, Н. В.* Тезаурусы в задачах информационного поиска. М.: Издательство Московского университета, 2011. — 512 с.
4. *Zablith, F., Antoniou, G., d'Aquin, M., Flouris, G., Kondylakis, H., Motta, E., Plexousakis, D., Sabou, M.* Ontology Evolution: A Process Centric Survey // The Knowledge Engineering Review. — Cambridge University Press, 2013. — pp. 1–31.
5. *Haase, P., Stojanovic, L.* Consistent Evolution of OWL Ontologies // The Semantic Web: Research and Applications // Proceedings of the 2-nd European Semantic Web Conference (ESWC). Lecture Notes in Computer Science — Berlin: Springer-Verlag, 2005. Vol. 3532, pp. 182–197.
6. *Scharrenbach, Th., d'Amato, C., Fanizzi, N., Grutter, R., Waldvogel, B., Bernstein, A.* Unsupervised Conflict-Free Ontology Evolution Without Removing Axioms // Proceedings of the 4th InternationalWorkshop on Ontology Dynamics (IWOD). — Shanghai, China, 2010.
7. *Scharrenbach, Th., d'Amato, C., Fanizzi, N., Grutter, R., Waldvogel, B., Bernstein, A.* Default Logics for Plausible Reasoning with Controversial Axioms // Proceedings of the 6th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW). — Shanghai, China, 2010.
8. *Broer, H. W., Dumortier, F., van Strien, S. J., Takens, F.* Structures in Dynamics: Finite Dimensional Deterministic Studies (Studies in Mathematical Physics). North-Holland, Amsterdam, 1991. — 253 p.
9. *Chueshov, I. D.* Introduction to the Theory of Infinite-Dimensional Dissipative Systems. ACTA Scientific Publishing House, Kharkiv, Ukraine, 2002. — 418 p.
10. *Ceccherini-Silberstein, T., Coornaert, M.* Cellular automata and groups. Springer Monographs in Mathematics. Berlin: Springer-Verlag, 2010. — 439 p.
11. *Непейвода, Н. Н.* Алгебраический подход к управлению // Проблемы управления, 2013, №6, с. 2–14.
12. *Кормалев, Д. А., Куршев, Е. П., Сулейманова, Е. А., Трофимов, И. В.* Извлечение информации из текста в системе ИСИДА-Т // Труды 11-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» (RCDL'2009). — Петрозаводск, Россия, 2009.

Evolving Ontologies in the Aspect of Handling Temporal or Changeable Artifacts

A. A. Demidov

Program Systems Institute of RAS
alex@dem.botik.ru

Abstract. We propose an algebraic approach to building ontologies which capable of evolution under the influence of new facts and which have some internal mechanisms of validation. For this purpose we build a formal model of the interactions of objects, and find out the limitations on transactions with objects imposed by this model. Then, in the context of the formal model, we define basic entities of the model of knowledge representation: concepts, samples, properties, and relationships. In this case the formal limitations are induced into the model of knowledge representation in a natural way.

Keywords: ontology, knowledge representation, ontology evolution, formal system.