INSTITUTE FOR NEURAL COMPUTATION

MACHINE PERCEPTION LAB

REPORT ON RESEARCH ACTIVITIES

2001-2006

FACULTY

Javier R. Movellan, Director.

Marian Stewart Bartlett, Co-Director.

Gwen Littlewort, Co-Director.

GRADUATE STUDENTS

Nick Butcko

Paul Ruvolo

Cynthia Taylor

Ian Fasel (Graduated 2006)

Tim Marks (Graduated 2006)

Jonathan Nelson (Graduated 2005)

John Hershey (Graduated 2003) Microsoft Research Corp.

Postdocs

Ian Fasel

Claudia Lainscsek

J. Cooper Roddey (2002-2004)

Visiting Scholars

Fumihide Tanaka, Sony Research Corporation

Masayuki Hoshino, Sony Research Corporation

Kazuki Aisaka, Sony Research Corporation (2004-2005)

Bjorn Braathen, Norweigan Defense Institute (2000-2001)

Kenta Kawamoto, Sony Research Corporation, (2000-2001)

Research Assistants

Kate Kern

Micah Eckhart

Jake Whitehill

Joel Chenu (2002-2004)

Bret Fortenberry (2003-2005) BU graduate program.

Josh Susskind (2002-2004) U. Toronto graduate program.

Evan Smith (2000-2001). CMU graduate program.

SCIENTIFIC PROGRAMMING AND TECH

Luis Palacios

Andrew Salamon

RESEARCH

The goal of the MPLab is to gain insights into how the brain works by developing embodied systems that solve problems similar to those encountered by the brain. We focus on systems that perceive and interact with humans in real time using natural communication channels (e.g., visual, auditory, and tactile information). To this effect we are developing perceptual primitives to detect and track human faces and to recognize facial expressions, and are developing probabilistic models for integrating multiple sensory modalities, and actions. We are also interested in the role of learning in perception: How adaptation to the statistical structure of the sensory input may contribute to sensory coding in both artificial and biological vision systems. This work requires a multidisciplinary approach that combines mathematical modeling, machine learning techniques, computational modeling of brain function, and behavioral experiments. Applications include teaching robots, automatic tutoring systems, and automatic assessment of affective disorders.

Applications of Personal Robots to Learning and Education

Javier Movellan, Cynthia Taylor, Paul Ruvolo, Fumhide Tanaka (Sony Electronics), Kathryn Owen (UCSD Early Childhood Educaion Center)

The goal of this project is to explore the idea of humanoid robots that teach children skills and assess their development, while interacting with them in an affective and human-like manner. To pursue this idea a humanoid robot, RUBI, is being developed, endowed with real-time machine perception primitives developed at our laboratory. These include face detection and tracking, eye and eye-blink detection, expression recognition, speech detection, recognition of affect in speech, and recognition of common communicative words in English. A touch screen facilitates the presentation of audio-visual material to the children. RUBI is being developed in close collaboration with the teachers and the director of UCSD`s Early Childhood Education Center (ECEC), where it is being tested and evaluated. This is the third year of the project. In year one we immersed researchers and two prototype robots, RUBI and QRIO, a humanoid robot from Sony elecronics, in the daily life of Room 1 at the Early Childhood education center at UCSD. We ran a total of 58 daily sessions, 30 minutes each. The sessions were video taped using 2 synchronized cameras and stored for later analysis. Most importantly, we received ample feedback from children, teachers and parents, that helped us identify the problems that need to be solved to make this technology a useful reality in the classroom. In Year 2 our was on the analysis of the data collected during Year 1 and on the development of new hardware and software tools that incorporate the lessons learned in Year 1. Research activities included: (1) Development of perceptual primitives that work reliably in the classroom. (2) Development of a more sophisticated robot prototype (RUBI-02) that targets the goals we identified in Year 1. (3) Analysis of the 58 hours of video sessions using both quantitative and qualitative approaches. (4) Development of a new operating system (RUBIOS) to support the challenges faced by robots that need to interact with children and teachers in a social manner. In Year 3 we are bringing the new technologies back into the school, to evaluate their effectiveness, and to further progress towards the development of systems that autonomously enrich the educational environment and become useful tools for teachers.

This project received the Best Paper Award at the 14th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2005).
http://mplab.ucsd.edu:16080/%7eboom/paper/Tanaka_ROMAN-05.pdf

Figure 1: Children at the Early Childhood Education Center learn shapes with RUBI (left) and dance with QRIO (right).

Real time face detection and facial feature detection

Ian Fasel, Bret Fortenberry, Javier Movellan

We formulated a probabilistic model of image generation and derived optimal inference algorithms for finding objects and object features within this framework. The approach models the image as a collage of patches of arbitrary size, some of which contain the object of interest and some of which are background. The approach requires development of likelihood-ratio models for object versus background generated patches. These models are learned using boosting methods. One advantage of the generative approach proposed here is that it makes explicit the conditions under which it is optimal, thus facilitating understanding of how the proposed algorithms could be modified to be optimal under challenging image conditions. We applied the approach to the problem of finding faces and eyes on arbitrary images. Optimal inference under the proposed model works in real time and is robust to changes in lighting, illumination, and differences in facial structure, including facial expressions and eyeglasses.

The eye detector enables more precise alignment of face images, including correction of in-the-plane rotations. The precision of the current system is on the order of 1/4 of an iris, similar to the precision obtained by human labelers in our previous study. Furthermore, the system can simultaneously track the eyes and blinks of multiple individuals. We made source code for the face detector freely available at http://kolmogorov.sourceforge.net.

Automatic Analysis of Facial Expressions

Gwen Littlewort, Marian Stewart Bartlett, Javier R. Movellan, Mark Frank (Rutgers University and SUNY Buffalo), Claudia Lianscsek, Jake Whitehill

The goal of this project is to develop systems to automatically code spontaneous facial expressions from video. This project stems from a collaboration between Paul Ekman and Terrence Sejnowski which began with an NSF planning workshop in 1992, and continues with Marian Bartlett, Javier Movellan, and Mark Frank. In 2000-2001 the CIA sponsored a feasibility study of the facial expression recognition systems at the Machine Perception Lab, and the system developed by Takeo Kanade and Jeff Cohn's group at CMU. The goal of the project was to explore the potential of computer vision technology for automatic expression recognition using a realistic dataset of spontaneous expressions. The results were evaluated by a team of computer vision experts (Yaser Yacoob, Pietro Perona) and facial expression experts (P Ekman, M Frank). The expert panel produced a report that concluded that our approach is viable and that a fully automated system for measuring facial actions (FACS) is a realistic goal within the next 10 years. The report identified the most important challenges to fully automated, viable facial expression measurement technology as 1) Collection of large video databases of spontaneous behavior, coded by experts in facial expression measurement including FACS. 2) Creation of a community of interdisciplinary teams working towards automatic facial expression measurement. Our research from 2001–2006 has focused on meeting those challenges. In a collaboration between MPLab and Mark Frank, we collected a large database of spontanous facial behavior which is being coded by Mark Frank, and which will be released to the research community. MPLab has advanced our technology for facial expression recognition, and we now have a fully-automated system, described below. Automated facial expression measurement will benefit basic research in behavioral science and psychiatry, man-machine interaction for education, and homeland security.

Fully automated recognition of expressions of basic emotion. We applied machine learning methods to the problem of recognizing facial expressions in images. The system overview is shown in Figure 3. The system is fully automated and runs in near-real time. Automatically detected faces are passed through a bank of Gabor wavelet filters. The filter bank representations are then channeled to a statistical classifier to code the image in terms of a set of expression dimensions. We conducted a comparison of classifiers, including support vector machines (SVM's), Adaboost (Freund & Shapire, 1996), and Linear Discriminant Analysis (Littlewort et al., in press, Bartlett et al., 2003). We also explored feature selection techniques. We found that best performance in terms of both speed and accuracy was obtained in an approach that employs Adaboost for feature selection on the Gabors prior to classification by SVM's. The system obtained the best performance reported so far on the publicly available Cohn Kanade dataset, with 93% correct generalization to new subjects on a 7-way forced choice. The outputs of the classifier change smoothly as a function of expression intensity and provide information about expression dynamics. The expression recognition system, as well as the FACS detectors described below are presently being compiled into the MPLab Comuter Expression Recognition Toolbox (CERT) which will be released to the research community.

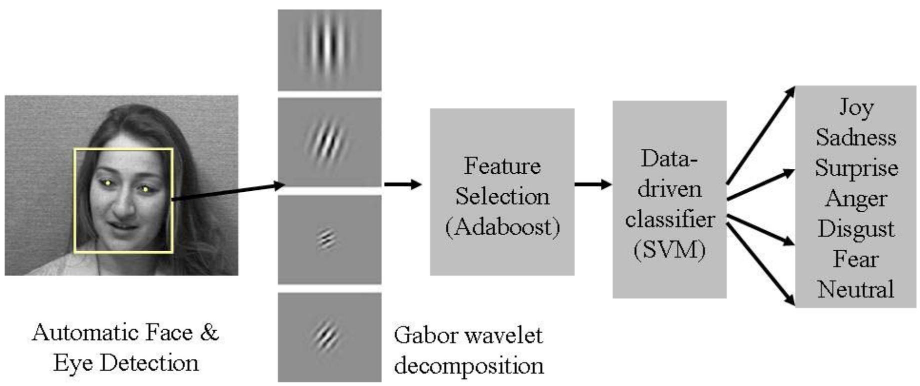

Figure 2. Fully automated facial expression recognition system, using machine learning techniques.

Fully automated facial action coding. The facial action coding system (FACS) (Ekman & Friesen 1978) is the most widely used expression coding system in the behavioral sciences. A human coder decomposes facial expressions in terms of 46 component movements, which roughly correspond to the individual facial muscles. FACS provides a comprehensive language for describing facial expressions and relating them back to what is known about their meaning from the behavioral science literature. Because it is objective, FACS also allows for the discovery of new patterns related to emotional or situational states. FACS was developed for coding by hand, however the time required to manually code facial behavior has been identified as the largest obstacle to research on emotion. MPLab developed a system for fully automatic recognition of facial actions from the Facial Action Coding System, applying the machine learning approach shown in Figure 2. The current system recognizes 20 action units and is 100% automated. The system obtained a mean of 94.5% agreement with human FACS labels for fully automatic recognition of 20 facial actions in posed expression data, and we are presently evaluating it on spontaneous expressions.

Spontaneous expressions in continuous video. Spontaneous facial expressions differ from posed expressions in both which muscles are moved, and in the dynamics of the movement. Advances in the field of automatic facial expression measurement will require development and assessment on spontaneous behavior. We performed preliminary analysis on a task of facial action detection in spontaneous facial expressions. We employ a user independent fully automatic system for real time recognition of facial actions from the Facial Action Coding System (FACS). The system automatically detects frontal faces in the video stream and coded each frame with respect to 20 Action units. The approach applies machine learning methods such as support vector machines and AdaBoost, to texture-based image representations. This work is in collaboration with Mark Frank at Rutgers University. We collected a dataset of 100 subjects participating in a “false opinion” paradigm. See RU-FACS1 dataset in ‘Products.’ Figure 5 shows sample system outputs. The output of the system is a real valued number indicating the distance to the separating hyperplane for each classifier. The output margin for the learned classifiers predicts action unit intensity. Frame-by-frame intensity measurements will enable investigations into facial expression dynamics which were previously intractable by human coding.

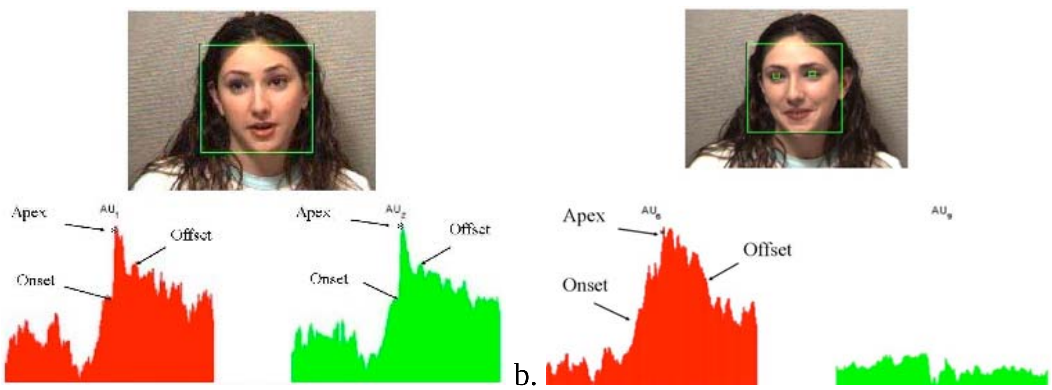

Figure 3. a. Sample system outputs for a 10-second segment containing a brow raise (FACS Code 1+2). System outputs for AU 1(red) and AU 2 (green) are shown. Human codes are overlaid for comparison (onset, apex, and offset). B. Sample system outputs for a sequence containing orbicularis oculi contraction (AU 6). System outputs are shown for AU 6 (red) and AU 9 (green). This sequence contained AU 6 and not AU 9.

Project web site: http://mplab.ucsd.edu/grants/project1/itr.html

Automated measurement of spontaneous facial expressions of genuine and posed pain Gwen Littlewort, Marian Stewart Bartlett, Kang Lee

In this project we apply the Computer Expression Recognition Toolbox (CERT), developed in the Machine Perception Laboratory, to spontaneous facial expressions of pain. In this study, participants submerged their hand in a bath of water for three minutes. Each subject experienced three experimental conditions: baseline, real pain and posed pain. In the real pain condition, the water was 3 degrees Celsius, whereas in the baseline and posed pain conditions the water was 20 degrees Celsius. Applying the CERT system to the pain video data produced a 20 channel output stream, consisting of one real value for each learned AU, for each frame of the video. This data was further analyzed to predict the difference between baseline and pained faces, the difference between expressions of real pain and fake pain, and the difference between faked pain with or without recent experience of real pain. In a preliminary analysis of 5 subjects, the system correctly identified the experimental condition (baseline, posed pain, real pain) for 93 percent of samples in a 3-way forced choice. This is considerably higher than the performance of naive human observers.

Real time video tracking

Josh Susskind, John Hershey, Javier Movellan

Bayesian filtering provides a principled approach for a variety of problems in machine perception and robotics. Current filtering methods work with analog hypothesis spaces and find approximate solutions to the resulting non-linear filtering problem using Monte-Carlo approximations (i.e., particle filters) or linear approximations (e.g., extended Kalman filter). Instead, we proposed digitizing the hypothesis space into a large number, n= 100,000, of discrete hypotheses. Thus the approach becomes equivalent to standard hidden Markov models (HMM) except for the fact that we use a very large number of states. One reason this approach has not been tried in the past is that the standard forward filtering equations for discrete HMMs require order n-squared operations per time step and thus rapidly become prohibitive. In our model, however, the states are arranged in two-dimensional topologies, with location independent dynamics. With this arrangement predictive distributions can be computed via convolutions. In addition, the computation of log-likelihood ratios can also be performed via convolutions. We describe algorithms that solve the filtering equations, performing this convolution for a special class of transition kernels in order n operations per time step. This allows exact solution of filtering problems in real time with hundreds of thousands of discrete hypotheses. We found this number of hypotheses sufficient for object tracking problems. We also propose principled methods to adapt the model parameters in non-stationary environments and to detect and recover from tracking errors.

***G-flow*: Estimating direction of attention and head pose**

Tim Marks, John Hershey, Cooper Roddey, Javier Movellan

Tracking faces as they rotate in 3D becomes a non-rigid tracking problem when we are interested in tracking during speech and changes in facial expression. We developed an algorithm called *G-flow* for 3D tracking of morphable objects, and applied it to the problem of automatic head pose estimation in 3D. *G-flow* is a generative model with an associated stochastic filtering algorithm for simultaneous tracking of 3D position and orientation, non-rigid motion, object texture, and background texture. The model defines a stochastic process that belongs to the class of conditionally Gaussian processes. This allows partitioning the filtering problem into two components: a linear component for texture that is solved using a bank of Kalman filters with time-varying parameters, and a nonlinear component for pose (rigid and non-rigid motion parameters) whose solution depends on the states of the Kalman filters. When applied to the 3D tracking problem, this results in an inference algorithm from which existing optic flow-based tracking algorithms and tracking algorithms based on texture templates emerge as special cases. Flow-based tracking emerges when the pose of the object is certain but its appearance is uncertain. Template-based tracking emerges when the position of the object is uncertain but its texture is relatively certain. In practice, optimal inference under this model integrates optic flow-based and template-based tracking, dynamically weighting their relative importance as new images are presented. Automated head pose tracking will provide information about head movement dynamics that was previously unavailable through hand coding. The non-rigid motion tracking will enhance the image-based MPLab system for facial expression measurement, and the head pose estimates will enable expression measurement with out-of-plane head rotations by enhancing precision of 3D face alignment.

Information maximization in face processing
Marian Bartlett

This work explores principles of unsupervised learning and how they relate to face recognition. Dependency coding and information maximization appear to be central principles in neural coding early in the visual system. I argue that these principles may be relevant to how we think about higher visual processes such as face recognition as well. We first review examples of dependency learning in biological vision, along with principles of optimal information transfer and information maximization. Next, we examine algorithms for face recognition by computer from a perspective of information maximization. The eigenface approach learns first and second-order dependencies among face image pixels, and maximizes information transfer only in the case where the input distributions are Gaussian. Independent component analysis (ICA) learns high-order dependencies in addition to first and second-order relations, and maximizes information transfer for a more general set of input distributions. Face representations based on ICA gave better recognition performance than eigenfaces, supporting the theory that dependency learning is a good strategy for high level visual functions such as face recognition. Finally, we review perceptual studies suggesting that dependency learning is relevant to human face perception as well, and present an information maximization account of perceptual effects such as the atypicality bias, and face adaptation aftereffects.

Characterizing the visuo-spatial properties of American Sign Language
Marian Bartlett, Rain Bosworth, Karen Dobkins (Psychology, UCSD)

The objective of this study is to characterize the visuo-spatial properties of American Sign Language (ASL) in order to enable subsequent investigations of plasticity in the visual system of deaf subjects and native signers. Deaf people have unique exposure to a visual language, ASL, in which the hands form intricate configurations that move through space. Using ASL as an example of a natural visual stimulus that the human visual system is faced with decoding, the current study quantified the power spectra of a set of ASL signs produced by two signers. Fourier transforms were applied to static photographs in order to determine whether signing produces unique spatial frequency and orientation information, compared to other types of images. The power spectra of the signer's arms and hands were compared to those of a "neutral" image of the same person with arms at rest, as well as to the power spectra of faces and natural scenes. In order to assess which spatial frequencies and orientations contained the most *information* about signs, entropy analysis was performed on the power spectra of ASL images, and compared to the entropy of faces and natural scenes. To our knowledge this is the first investigation of entropy of natural scenes and faces, in addition to ASL. We found that while the frequency spectra of the three groups of images were quite similar, consistent with the well reported power law, the entropy of the three sets of images differed. Although natural scenes had more power at low spatial frequencies, the entropy increased with spatial frequency. The entropy of ASL images was band-pass, with peak entropy at about 0.5 cyc/cm (1.3 cyc/deg assuming a 5 ft conversation distance). The entropy of faces was also bandpass, with peak at 1.7 cyc/cm (4.3 cyc/deg).

<u>Developmental Methods for Automatic Discovery of Object Categories</u>

Ian Fasel, Javier Movellan

The goal of this project is to develop systems that autonomously learn to detect and track objects in video images. Current approaches to object detection require the existence of large training datasets for which the object of interest has been manually segmented from the background by human operators. This process is very laborious and time consuming greatly limiting progress in the field.

This work is a computational investigation of the task of locating and recognizing objects in unconstrained images in real-time, and learning to do so with minimal supervision. We take a probabilistic generative modeling approach, which involves formulating analytical models of several real-world vision problems, studying how optimal inference would proceed under such models, developing techniques for learning parameters under these models, and evaluating the performance of the optimal inference algorithms in realistic data. We begin by developing a novel generative model of images under which an image is a collection of sets of pixels, which are generated by different object categories. This provides a novel definition of “object” as a set of pixels that are co-dependent, but conditionally independent of the other sets of pixels in the image. We then develop an algorithm for optimal inference (i.e., detection of objects) and maximum likelihood learning when the segmentation of training images is known. We point out a computational tradeoff between robustness of object detection and precision of localization, and propose context dependent detectors as a way to solve the problem. These techniques are used to develop a state-of-the-art, real-time head, eye, and blink detector. We predict that similar context-dependent detectors may be found in the brain. We develop an algorithm for optimal inference and maximum likelihood learning when the segmentation of training images is unknown. We test this on image datasets labeled with the identity but not the location of objects, and achieve state-of-the-art performance in discovery of object categories. We then test the algorithm in a fully unsupervised context, in which a real-time person detector is learned from just a few minutes of visual information self-labeled through multi-modal contingency detection. This suggests that early face (and other) preferences in humans infants may be evidence for rapid statistical learning rather than innate biases. We develop software for learning robust, real-time object detectors from both labeled and unlabeled examples, including a real-time head, eye, and blink detector available to the public.

Learning about Humans During the First 6 Minutes of Life

Nick Butcko, Ian Fasel. Javier Movellan

There is strong experimental evidence that newborn infants orient towards human faces \(Morton, 1991). While opinions are divided as to whether this preference reflects domain specific knowledge about the appearance of human beings, or general preferences for stimuli that happen to occur in humans, most views agree that the face-preference phenomenon is innate and not learned. Here we explore another hypothesis, *the Rapid Learning Hypothesis* which in the past was rejected as being computationally implausible.

We built a robot with the appearance of a human baby and endowed it with an algorithm that detects contingencies in the auditory domain. Members of our laboratory were then encouraged to sporadically interact with the baby robot. We also endowed the robot with a new machine learning algorithm for discovering visual concepts (Fasel, 2005). The input to the system were the images collected by the baby robot's camera at 30 frames per second. In addition each image was automatically labeled by the auditory contingency detector to indicate whether or not auditory contingencies were present at the time the video frame was captured. Contingency detection provided a noisy signal of the presence of a person in the image. In 70% of the images captured while auditory contingencies were detected, people were visible. In addition, people could appear anywhere in the image plane, sometimes showing their face, sometimes other parts of their body.

In less than 6 minutes of interaction with the world the robot learned to locate people in novel images. In addition, it developed a preference for drawings of human faces over drawings of non-faces, even though it had never been exposed to such schematic face drawings before. During learning, the baby robot was never told whether or not people were present in the images, or whether people were of any particular relevance at all. It simply discovered that to make sense of the images and sounds it received, it was a good idea to use feature detectors that happened to discriminate the presence or absence of people. While in our experiment we used auditory contingency as a training signal for a visual concept learner, other training signals could also have been used. All that is required is for the signal to provide higher than chance information about the presence or absence of people. For example, if a baby or robot is being touched or moved, these could likely serve also as a training signal. The results illustrate that visual preferences of the type typically investigated in human neonates can be acquired very quickly, in a matter of minutes. Previous studies that were thought to provide evidence for innate cognitive modules may actually be evidence for rapid learning mechanisms in a neonate brain exquisitely tuned to detect the statistical structure of the world.

Diffusion Neural Networks

Javier R. Movellan, Paul Miniero, Ruth Williams, Tim Marks

This work explores information processing in neural network architectures with feedback connections, continuous representations and a significant stochastic component. Such networks are stochastic diffusion processes and can be analyzed using techniques from stochastic calculus. As part of this line of research we have shown how such networks can be trained to learn probability distributions of trajectories and to perform optimal inference. We have shown how such architectures could also be used to learn factorial codes using unsupervised learning methods. Finally we have described conditions under which such networks combine information from different modalities in a manner consistent with psychophysical experiments. We have also used the networks to model integration of classical and non-classical receptive field in primary visual cortex.

We have also analyzed the relationship between diffusion networks and the Lucas-Kanade optic flow algorithm. This led us to develop a new filtering algorithm for object tracking which we named G-flow, and is described above. We derived the Hessian matrix (second derivatives) of rotations with respect to the exponential coordinates for rotation. This has enabled more robust application of diffusion networks to 3D nonrigid object tracking.

Information integration: Implications for models of perception and neural coding
Javier Movellan, Jay McClelland (CMU), Thomas Wachtler, Odelia Schwartz, Tom Albright, Terry Sejnowski

Information integration may be studied by analyzing the effect of 2 or more sources (e.g., auditory and visual) on participants' responses. Experiments show that ratios of response probabilities often factorize into components selectively influenced by only 1 source (e.g., 1 component affected by the acoustic source and another 1 affected by the visual source). This is called the Morton-Massaro law (MML). This work identifies conditions in which the law is optimal and notes that it reflects an implicit assumption about the statistics of the environment. Adherence to the MML can be used to assess whether the assumption is being made, and analyses of natural stimuli can be used to determine whether the assumption is reasonable. Feed-forward and interactive models subject to a channel separability constraint are consistent with the law.

The notion of Morton-style factorial coding may help understand information integration and perceptual coding in the brain. We show that by focusing on average responses one may miss the existence of factorial coding mechanisms that become only apparent when analyzing spike count histograms. We show evidence suggesting that the classical/non-classical receptive field organization in the cortex effectively enforces the development of Morton-style factorial codes. This may provide some cues to help understand perceptual coding in the brain and to develop new unsupervised learning algorithms. While methods like ICA (Bell & Sejnowski, 1997) develop independent codes, in Morton-style coding the goal is to make two or more external aspects of the world become independent when conditioning on internal representations.

Explaining eye movements during learning as an active sampling process
Jonathon Nelson, Gary Cottrell (Computer Science, UCSD), Javier Movellan

People are active experimenters, constantly seeking new information relevant to their goals. A reasonable approach to active information gathering is to ask questions and conduct experiments that minimize the expected state of uncertainty, or maximize the expected information gain, given current beliefs (Fedorov, 1972; MacKay, 1992; Oaksford & Chater, 1994). In this paper we present results on an exploratory experiment designed to study people's active information gathering behavior on a concept learning task. The results of the experiment suggest subjects' behavior may be explained well from the point of view of Bayesian information maximization.

Models Of The Development And Dysfunction Of Shared Attention: Testing Parameters Of Social Learning That Might Underlie Autistic Behavior

Javier Movellan, Gedeon Deak (Cogsci, UCSD), Jochen Triesch (Cogsci, UCSD)

The goal of this project is to develop and test neural models of the development of shared attention, and changes in neural learning processes that might impair the acquisition of shared attention in children with autism. Deficits in shared attention are common among children with autism, and might be related to atypical contingency learning that makes it difficult for these children to predict the course of social interactions. We will explore this idea in a multi-method investigation that includes three interrelated studies: 1. Observing the structure and timing of infant-caregiver interactions. 2. Computational models of the development of shared attention. 3. Experimental tests of social contingency learning.

Creating the Next Generation of Intelligent Animated Conversational Agents

Javier Movellan, Ron Cole (CU Boulder)

We propose to develop intelligent animated conversational agents--lifelike computer characters capable of natural face -to-face conversational interaction in specific task domains. Developing these agents requires basic research leading to new and improved language technologies in areas of computer speech recognition, natural language understanding, dialogue modeling, visual recognition of facial movements, expressions and gestures, and generation of natural and expressive auditory and visual speech by three -dimensional animated characters. The research advances and systems developed during the project will be evaluated in the context of language training applications that serve as test beds for research. Learning programs will be developed to enable English - and Spanish -speaking children with speech and reading difficulties to interact with animated agents to learn to speak, understand and read language. The project is high risk, long term and potentially high impact. If successful, the proposed work will produce advances in scientific knowledge in areas of human language technology, human computer interaction and interactive computer -based learning. It will also yield significant benefits to society by offering more effective and assessable multilingual interfaces for learning and information access. Research problems that must be solved to develop effective learning programs with intelligent animated agents include (a) improving capabilities of current multimodal language technologies; (b) integrating these technologies in conversational agents in advanced dialogue systems; and (c) understanding and modeling the auditory and visual behaviors of teachers and students during interactive learning. The proposed project is by necessity both multidisciplinary and multi-institutional.

PRODUCTS

MACHINE PERCEPTION TOOLBOX

This software was developed by personnel in the Machine Perception Lab and is distributed under BSD license. MPT1.0 supplies cross-platform C++ libraries for real-time perception primitives, including face detection, eye detection, blink detection, and color tracking. Soon it will also include expression recognition, predictive color tracking, and tracking based on multisensor fusion. In addition, MPT supplies substantial documentation, including many examples which show how to embed the MPT functions in applications, and which are useful applications in their own right. Finally, Matlab .mex interfaces are provided for calling the library from MatLab. This toolbox is available for free download on our website: http://mplab.ucsd.edu/software/software.html.

COMPUTER EXPRESSION RECOGNITION TOOLBOX (CERT)

The Machine Perception Lab is developing a toolbox for fully automated facial expression recognition which will be made available to the research community. This toolbox includes recognition of expressions of basic emotions as well as comprehensive expression coding in terms of facial actions from the Facial Action Coding System. This toolbox will benefit basic research into emotion, social interaction, and psychiatry by making facial expression measurement more accessible as a behavioral measure, and by providing data on the dynamics of facial behavior at a resolution that was previously unavailable.

SCORE

The Machine Perception lab has developed a toolbox for FACS coding from digital video. This SCORE video coding tool is a modular tool that can be easily modified for other video annotating applications as well. This tool is available for free download on our website http://mplab.ucsd.edu/software/software.html.

Figure 7. Sample subject from the RU-FACS-1 database with SCORE video coding tool.

RUBI and RUBI2 SOCIAL ROBOTS

MPLab developed two social robots as part of the UC Discovery/Sony project. The robots are being used as research tool to understand the organization of behavior in real time, and as a testbed of the perceptual and control primitives developed at our laboratory.

RUBIOS

MPLab developed a software architecture, called RUBIOS, for social robots. The architecture allows parallel operation of multiple control modules organized as stochastic dynamical systems (e.g., diffusion networks). These nodes send temporal utility messages interpreted as the value they are willing to "pay" to other modules for performing their requests. Most importantly these values can be a function of space and time. All modules are endowed with a learning system so as to optimize the long term value accumulated. Adaptive behavior emerges in the system as each of the modules act so as to selfishly achieve their goals. RUBIOS is the software architecture we are already using on RUBI2

RU-FACS-1 DATABASE

This database consists of spontaneous facial expressions from multiple views, with ground truth FACS codes provided by two facial expression experts. This database is being collected by Mark Frank at Rutgers University as part of a collaboration with the Machine Perception Lab at INC to automate FACS. The data collection equipment, environment, and paradigm were designed with advice from machine vision consultant Yaser Yacoob and facial behavior consultant Paul Ekman. The system records synchronized digital video from 4 Point Grey Dragonfly video cameras and writes directly to RAID. 100 subjects were recorded for 2.5 minutes each while participating in a false opinion paradigm in which they were randomly assigned to either lie or tell the truth about their opinion on an issue about which they indicated strong feelings (e.g., Frank & Ekman, 1997; 2004). FACS coding of the database is ongoing. The first 20 subjects are available for release.

GENKI FACE DATABASE

MPLab collected a database of 70,000 images of faces from the Web. The goal was to develop a database for training machine learning based systems from a large variety of camera, lighting, and image conditions. The dataset was labeled for facial expression (smile vs not smile) and position of the eyes. 30,000 images were also labeled for 3D pose parameters. The dataset will be released to the research community in the near future.

IR MARKS DATABASE

MPLab developed a new method for collecting ground-truth face motion data using infrared markings and a rig of visible-light and infrared cameras. We used this new method to collect a new data set of face motion, which we will make available to other researchers.

KOLMOGOROV PROJECT
Dr. Movellan founded the Kolmogorov project that provides a collection of open source software and tutorials on topics related to Machine Learning, Machine Perception and Statistics (http://kolmogorov.sourceforge.net).

PATENTS

1) An optimal controller that finds people via contingency detection. UCSD/Sony patent application. in the USA and Japan.

2) Real Time Expression Recognition. UCSD/Sony patent application

EVENTS

ICDL2004 3rd International Conference on Development and Learning. San Diego, CA, October 20-22, 2004. General Chair: Movellan, J.R. Co-Chairs: Bartlett, M.S. Chiba, A., Deak, G. Jebara, T., Littlewort, G., Triesch, J.

The goal of this international conference, held at the Salk Institute for Biological Studies on October 20-23, 2004, was to bring together leading researchers in neuroscience, machine learning, robotics, and developmental psychology, in order to gain new insights about learning and development in natural organisms and robots. The scope of developmental processes considered was broad, including cognitive, social, emotional, and many other skills exhibited by humans, and other animals. The theme of the conference this year was "Developing Social Brains".

Despite the strides being made in each discipline, further interactions across disciplines will serve to accelerate progress. For example, developmental psychologists can benefit from an understanding of the computational problems underlying the creation of machines that can perceive and interact with humans in the real world. Researchers in robotics can learn from the way biological systems balance learning and innate predispositions during development. Machine learning and machine perception can provide a framework to help neuroscientists understand neural dynamics and neural mechanisms underlying learning and development.

Workshops and Symposia:

- Discovery of Object Categories, Neural Information Procesing Systemts (NIPS) Post-Conference Workshop, Vancouver, December 9, 2005. Organizers: Ian Fasel and Javier Movellan.

- Workshop on Modeling People and Human Interaction, International Conference on Computer Vision (ICCV), Beijing, China, October 2005. Organizers: Remi Ronford, Daniel Gatica-Perez, Pascal Fua, Adrian Hilton. Marian Bartlett.

- NSF Workshop on Perceptive Animated Interfaces and Virtual Humans. San Diego, CA, April 8-9, 2004. Organizers: Cole, R., Movellan, J.R., and Gratch, J.

- DiMI Workshop on Perceptive Social Agents and Robots. San Diego, CA, Jan 8-9, 2003. Organizers: Movellan, J.R., Turk, M., and Cole, R.

- Learning in Vision, Special Session of the 6th International Conference on Knowledge-based and Intelligent Engineering Systems. Crema, Italy, Sept 17, 2002. Organizer: Bartlett, M.S.

- Affective Computing, Neural Information Procesing Systemts (NIPS) Post-Conference Workshop, Breckenridge, CO, December 2, 2000. Organizers: Movellan, J.R., Bartlett, M.S., Cottrell, G., and Picard, R.

OUTREACH

The Machine Perception Lab has also been involved since 2004 in the “Reach for Tomorrow” program which exposes underprivileged high school students to science with the hope of interesting them in attending college. The students have included Native Alaskan high school students from above the arctic circle, as well as students from inner-city Washington DC. The lab puts together a seminar for the students including hands-on demos and a lab tour.

The Machine Perception Lab actively involves students from the Preuss School in robotics projects. The Preuss School is a school for low-income middle and high school students who will be the first generation in their families to attend college, and is housed on the UCSD campus. The students are 60% Hispanic, 13% African American, 21% Asian and 6% White. In 2005-2006 the Robotics club at Preuss built a mobile platform for the RUBI robot. Dr. Movellan and other members of MPLab also guided the Preuss Robotics club in their entry for FIRST Robotics, a national high school computer vision and robotics competition.

RESOURCES

Laboratory: The Machine Perception Laboratory is housed in the new CalIT2 Building at UCSD, recently named Atkinson Hall. The laboratory is approximately 1200 square feet and consists of a large computer lab with access to CalIT2 resources including conference rooms. There is also a robot construction lab housed in Sequoia hall of approximately 400 square feet. I addition, the Machine Perception Lab has an experiment room of approximately 400 square feet, for conducting robotics experiments, and setting up video collection. This room is in the Chemistry Research Building, which is the building next to the robot construction room in Sequia Hall.

Computer: The Machine Perception Lab has an Apple G5 computer cluster consisting of 12 dual G5 processors, 2 ghz each, with a high speed (1GB) Ethernet network switch. The cluster includes fast large network storage in a 2.5 terabyte Xserve RAID with Fibre Channel connections. Each of the 12 computers in the cluster additionally has a 150 GB hard drive. In addition, the lab has 11 workstations, and a Macintosh Xserve with 2 TB of storage capacity used for centralized data storage and analysis. The workstations consist of 1 Mac Workstation, and 9 PC workstations, 5 of which have Athalon 1.1 ghz processors, 100 gb hard drives and 2 gb RAM. All of our systems are wireless networking capable and receive dynamic ip addresses from our local wireless DHCP server. The lab also a tape backup system, and two printers: a 4050 HP laser printer and an HP color Inkjet 2600.

PUBLICATIONS

Machine Perception Lab

MPLab Journal Papers and Dissertations:

Bartlett, M.S. (in press). Information maximization in face processing. *Neurocomputing*. Special Issue on Development and Learning.

Bartlett, M.S., (2001). *Face Image Analysis by Unsupervised Learning*. Foreword by Terrence J. Sejnowski. Boston: Kluwer Academic Publishers.

Bartlett, M.S., Littlewort, G.C., Frank, M.G., Lainscsek,C., Fasel, I., Movellan, J.R. (accepted). Automatic Recognition of Facial Actions in Spontaneous Expressions. *Journal of Multimedia.*

Bartlett, M.S., Movellan, J.R., & Sejnowski, T.J. (2002). Face recognition by Independent component analysis. *IEEE Transactions on Neural Networks* 13(6) p. 1450-64.

Bosworth, R.G., Bartlett, M. S., and Dobkins, K. R. (in press). Image Statistics of American Sign Language: Comparison to Faces and Natural Scenes. *Journal of the Optical Society of America A.*

Cole, R., van Vuure, S., Pellom, B., Hacioglu, K., Ma, J., Movellan, J., Schwartz, S., Wade-Stein, D., Ward, W., and Yang, J. (2003). Perceptive animated interfaces: First steps toward a new paradigm for human computer interaction. *Proceedings of the IEEE, Special Issue on Human Computer Multimodal Interfaces,* 91(9):1391–1405.

Deak, G., Bartlett, M., and Jebara, T. (in press). New trends in integrative cognitive science: Approaches to development and learning. *Neurocomputing*. Special Issue on Development and Learning.

Draper, J., Baek, K., Bartlett, M.S., & Beveridge, J.R. (2003). Recognizing faces with PCA and ICA. *Computer vision and image understanding* 91: Special issue on Face Recognition, p. 115-137.

Fasel, I.R. (2006). Learning Real-Time Object Detectors: Probabilistic Generative Approaches. Doctoral dissertation, Department of Cognitive Science, University of California, San Diego.

Fasel, I., B. Fortenberry, and J. R. Movellan (2005). A generative framework for real-time object detection and classification. *Computer Vision and Image Understanding* 98(1) p. 182-210.

Littlewort, G., Bartlett, M., Fasel, I., Susskind, J., and Movellan, J. (2006). An automatic system for measuring facial expression in video. *Image and Vision Computing* 24(6) p. 615-625.

Marks, T.K. (2006). Facing Uncertainty: 3D Face Tracking and Learning with Generative Models. Doctoral dissertation, Department of Cognitive Science, University of California, San Diego.

Movellan, J.R., Mineiro, P., & Williams, R.J. (2002). A Monte-Carlo EM approach for partially observable diffusion processes: Theory and applications to neural networks. *Neural Computation* 14(7):1507–1544.

Movellan J.R. and J.Nelson (2001) Probabilistic functionalism: A unifying paradigm for the cognitive sciences. *Behavioral and Brain Sciences* 24(4).

Movellan J.R. and J.L. McClelland (2001) The Morton-Massaro law of information integration: Implications for models of perception. *Psychological Review*, (1):113--148, 2001.

Susskind, J.M., Littlewort, G.C., Bartlett, M.S., Movellan, J.R., and Anderson, A.K. (in press). Human and computer recognition of facial expressions of emotion. *Neuropsychologia.*

<u>MPLab Articles in Peer Reviewed Proceedings:</u>

Bartlett, M.S., Littlewort, G., Braathen, B., Sejnowski, T.J., & Movellan, J.R. (2003). A prototype for automatic recognition of spontaneous facial actions. In S. Becker & K. Obermayer, (Eds.) *Advances in Neural Information Processing Systems, Vol 15*. p 1271-1278. MIT Press.

Bartlett, M., G. Littlewort, I. Fasel, and J. Movellan (2003). Real time face detection and expression recognition: Development and application to human-computer interaction. In *CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction*.

Bartlett, M.S., Littlewort, G.C., Lainscsek, C., Fasel, I., Frank, M.G., Movellan, J.R. (2006). Fully automatic facial action recognition in spontaneous behavior. *7th International Conference on Automatic Face and Gesture Recognition*, p. 223-228.

Bartlett, M.S., Littlewort, G., Frank, M.G., Lainscsek, C., Fasel, I., and Movellan, J. (2005). Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior. *IEEE International Conference on Computer Vision and Pattern Recognition*. p. 568-573.

Bartlett, M., G. Littlewort, C. Lainscsek, I. Fasel, and J. Movellan. Machine learning methods for fully automatic recognition of facial expressions and facial actions (2004). In *IEEE International Conference on Systems, Man & Cybernetics*, The Hague, Netherlands, p. 592-597.

Bartlett, M.S., J.R. Movellan, G. Littlewort, B. Braathen, Frank. M. G., and T. J. Sejnowski (2004). Towards automatic recognition of spontaneous facial actions. In P. Ekman, editor, *What the Face Reveals*. Oxford University Press. p. 393-426.

Bartlett, M.S., Movellan, J.R. and Sejnowski, T.J. (2005). Face modeling by information maximization. Invited chapter in *Face Processing: Advanced Modeling and Methods*. R. Chellappa and W. Zhao, Eds., Academic Press.

Bosworth, R. G., Wright, C. E., Bartlett, M. S., Corina, D., and Dobkins, K. R. (2003). Characterization of Visual Properties of Spatial Frequency and Speed in American Sign Language. In A. Baker, B. van den Bogaerde, & O. Crasborn (Eds.) *Cross-Linguistic perspectives in Sign Language Research; Selected papers from TISLR 2000*, Signum Press. *pg. 265-282*.

Braathen, B., Bartlett, M.S., Littlewort, G., and Movellan, J.R. (2001). First Steps Towards automatic recognition of spontaneous facial action units. *Proceedings of the ACM Conference on Perceptual User Interfaces*.

Braathen, B., Bartlett, M.S., Littlewort-Ford, G., and Movellan, J.R. (2002). An approach to automatic recognition of spontaneous facial actions. *Fifth International Conference on automatic face and gesture recognition p. 231-235*.

Braathen, B., Bartlett, M.S. Littlewort-Ford, G. and Movellan, J.R. (2001). 3-D head pose estimation from video by stochastic particle filtering. *Proceedings of the 8th Annual Joint Symposium on Neural Computation*.

Deak, G.O., and I. R. Fasel and J. R. Movellan (2001). The emergence of shared attention: using robots to test developmental theories. *Proceedings of the first international workshop on epigenetic robotics: modeling cognitive development*, p. 95-104.

Fasel, I.R., Bartlett, M.S. & Movellan, J.R. (2002). A comparison of Gabor filter methods for automatic detection of facial landmarks. *5th International Conference on automatic face and gesture recognition p. 345-350.*

Fasel I., and G. O. Deak and J. Triesch and J. R. Movellan (2002). Combining embodied models and empirical research for understanding the development of shared attention. *Proceedings of the International Conference on Development and Learning* (ICDL02).

Fasel, I.R., Smith, E.C., Bartlett, M.S. & Movellan, J.R. (2000) A comparison of Gabor filter methods for automatic detection of facial landmarks. *Proceedings of the 7th Annual Joint Symposium on Neural Computation*.

Fasel, I. and Movellan, R. R. (2001) Meta analysis of neurally inspired face detection algorithms. *Proceedings of the 8th Annual Joint Symposium on Neural Computation*.

Gray, M.S., T.J. Sejnowski, and J.R. Movellan (2001) A comparison of image processing techniques for visual speech recognition applications. In T.Dietterich, editor, *Advances in Neural Information Processing Systems*, number 14. MIT Press, Cambridge, Massachusetts.

Kanda, T., N. Miralles, M. Shiomi, T. Miyashita, I. Fasel, J. R. Movellan, and H. Ishiguro (2004). Face-to-face interactive humanoid robot. *IEEE 2004 International Conference on Robotics and Automation*, 2004.

Littlewort, G., M. Bartlett, I. Fasel, J. Susskind, and J. Movellan. Dynamics of facial expression extracted automatically from video (2004). In *IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Face Processing in Video*, 2004.

Littlewort, G., M. Bartlett, C. J, I. Fasel, T. Kanda, H. Ishiguro, and J. Movellan (2004). Towards social robots: Automatic evaluation of human-robot interaction by face detection and expression classification. In *Advances in neural information processing systems*, volume 16. MIT Press, Cambridge, MA, p. 1563-1570.

Marks, T., Hershey, J., Movellan, J.R., & Roddey, J.C. (2005) Joint tracking of pose, expression, and texture using conditionally Gaussian filters. *Advances in Neural Information Processing Systems* 17. MIT Press.

Marks, T.K., J. Hershey, J. C. Roddey, and J. R. Movellan (2004). 3D tracking of morphable objects using conditionally gaussian nonlinear filters. *Computer Vision and Pattern Recognition*, 2004.

Marks, T.K. and Movellan, J.R. (2001). Diffusion networks, product of experts, and factor analysis. *Proc. 3rd Intl. Conference on Independent Component Analysis and Blind Signal Separation*.

Marks, T.K. and Movellan, J.R. (2006). Invisible face painting: Using Infrared marks for 3D measurements during video collection. *Proceedings of the CVPR workshop on Performance Evaluation of Tracking and Surveillance*.

Movellan, J.R. (2005). An infomax controller for real time detection of contingency. In *Proceedings of the International Conference on Development and Learning (ICDL05)*, Osaka, Japan, 2005.

Movellan, J. R., J. Hershey, T. K. Marks, and C. Roddey. (2004). 3D tracking of morphable objects using conditionally gaussian non-linear filters. *CVPR Workshop on Generative Models*.

Movellan, J. R., J. Hershey, and J. Susskind (2004). Large scale convolutional HMMs for real time video tracking. *Computer Vision and Pattern Recognition*.

Movellan, J.R., P.Mineiro, and R.J. Williams (2001) Partially observable SDE models for image sequence recognition tasks. In T.Dietterich, editor, *Advances in Neural Information Processing Systems*, number 14. MIT Press, Cambridge, Massachusetts.

Movellan, J.R., F. Tanaka, B. Fortenberry, and K. Aisaka. (2005). The RUBI project: Origins, principles and first steps. In *Proceedings of the International Conference on Development and Learning (ICDL05)*, Osaka, Japan, 2005.

Movellan, J. R., T. Wachtler, T. D. Albright, and T. Sejnowski (2003). Morton-style factorial coding of color in primary visual cortex. In *Advances in Neural Information Processing Systems*, number 15. MIT Press, Cambridge, Massachusetts.

Movellan, J.R. and J. S. Watson (2002) The development of gaze following as a

Bayesian systems identification problem. *Proceedings of the International Conference on Development and Learning* (ICDL02).

Nelson, J.D. and J.R. Movellan (2001) Active inference in concept induction. In T.Dietterich, editor, *Advances in Neural Information Processing Systems*, number 14. MIT Press, Cambridge, Massachusetts.

Nelson, J.D. and J. B. Tenenbaum and J. R. Movellan (2001). Active Inference in Concept Learning, *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, p. 692-697.

Smith, E., Bartlett, M.S., and Movellan, J.R. (2001). Computer recognition of facial actions: A study of co-articulation effects. *Proceedings of the 8th Annual Joint Symposium on Neural Computation*.

Tanaka, F., B. Fortenberry, K. Aisaka, and J. R. Movellan (2005). Plans for developing real-time dance interaction between qrio and toddlers in a classroom environment. In *Proceedings of the International Conference on Development and Learning (ICDL05)*, Osaka, Japan, 2005

Tanaka, F., B. Fortenberry, K. Aisaka and J.R. Movellan (2005). Developing Dance Interaction between QRIO and Toddlers in a Classroom Environment: Plans for the First Steps. *Proceedings of the 2005 IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN)*, p.223-228 Nashville, U.S.A., August 2005. (Best Paper Award).

Tanaka, F. and J.R. Movellan (2006). The RUBI Project: Designing Everyday Robots by Immersion. *Fifth International Conference on Development and Learning* (ICDL) Bloomington, U.S.A., June 2006.

Tanaka, F. and J.R. Movellan (2006). Behavior Analysis of Children's Touch on a Small Humanoid Robot: Long-term Observation at a Daily Classroom over Three Months. *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), in press* Hatfield, United Kingdom, September 2006.

Tanaka, F., J. R. Movellan, B. Fortenberry, and K. Aisaka (2006). Daily HRI evaluation at a classroom environment: Reports from dance interaction experiments. In *Proceedings of the 2006 Conference on Human-Robot Interaction (HRI)*, Salt Lake City, 2006.

MPLab Conference Presentations:

Bartlett, M.S. (2003). Unsupervised learning in face recognition. Invited talk, About Faces: A multidisciplinary Approach to the Science of Face Perception. Princeton University, Princeton, NJ, September 19-21.

Bartlett, M.S. (2004). Information maximization in face processing. Poster presentation, Proceedings of the 2nd International Conference on Development and Learning.

Bartlett, M.S., G. Littlewort, C. Lainscsek, M.G. Frank, & J.R. Movellan (2004). Towards Automatic Measurement of Spontaneous Facial Behavior. Oral presentation. International Conference on Research on Emotion. July 6-11 New York, New York.

Bartlett, M.S., Movellan, J.R., Littlewort, G., Susskind, J. (2003). Dynamic face recognition technology. Oral presentation, DARPA Human ID Symposium, Arlington, VA, Sept 29-30.

Chenu, Fortenberry, Movellan (2004). RUBI: A robotic platform for real time social interaction. 11th Joint Symposium on Neural Computation. Poster presentation.

Chenu, Fase,l, Kanda, Ishiguro, Movellan (2003). Automatic evaluation of human-robot interaction. 10th Joint Symposium on Neural Computation. Poster presentation.

Cohn, J., M.S. Bartlett, J. Chenu & Z. Ambador (2004). Digital Scoring Tools. Poster presentation. International Conference on Research on Emotion. July 6-11 New York, New York.

Bret Fortenberry, Joel Chenu, Javier R. Movellan (2004) RUBI: A Robotic Platform for Real-time Social Interaction. Poster presentation, Proceedings of the 2nd International Conference on Development and Learning, pg. 34-37.

Fortenberry, B., Bartlett, M.S., Roddey, J.C., Fasel, I., Marks, T., Littlewort, G., Chenu, J., Movellan, J.R. (2004). Machine perception primitives for real time social interaction. Demonstration, 18th Conference on Neural Information Processing Systems.

Littlewort, G.C. Bartlett, M.S., Fasel, I., Susskind, J.M. and Movellan, J.R. (2003). Towards sociable robots: Implementation and Application of a system for fast, accurate and robust automatic face detection and expression recognition. Joint Symposium on Neural Computation, University of California, Irvine, May 15, 2003.

Littlewort, G., Bartlett, M., Fasel, I., Susskind, J., Movellan, J. (2003). Towards sociable robots: Implementation and application of a system for fast, accurate, and robust automatic face detection and expression classification. 10th Joint Symposium on Neural Computation. Poster presentation.

Littlewort, G., Bartlett, M., Fasel, I., Susskind, J., Movellan, J. (2002). Finding emotions in a crowd: Real time, Fully automatic coding of facial expressions from video. Oral presentation, 9th Joint Symposium on Neural Computation.

Littlewort, G.C., M.S. Bartlett, I. Fasel, J. Chenu, T. Kanda, H. Ishiguro, J.R. Movellan. (2004). Facial Expression in Social Interactions: Automatic Evaluation of Human-Robot Interaction. Poster presentation, Proceedings of the 2nd International Conference on Development and Learning.

Littlewort, G., Bartlett, M.S., and Lee, K. (2006). Faces of Pain: Automated measurement of spontaneous facial expressions of genuine and posed pain. Proceedings of the 13th Joint Symposium on Neural Computation, San Diego, CA.

Littlewort, G. ., Bartlett, M.S. and Movellan, J.R. (2001). Your Eyes Smiling? Detecting genuine smiles with support vector machines and Gabor wavelets. *Proc.8th Annual Joint Symposium on Neural Computation.*

Marks, T., Roddey, C., Hershey, J., & Movellan, J.R. (2003). Determining 3D structure from video images using G-flow. Demonstration, 17th Conference on Neural Information Processing Systems.

Movellan J. R., Hershey J., Marks T., and Roddey C. (2003) GFlow: A generative model for fast tracking using 3D deformable models. In *DARPA Symposium on Human ID*, Washington DC, September 29-30.

Movellan, J.R., Gwen Littlewort, G.C. and Bartlett M.S. (2003). Development, evaluation and application of a system for real time fully automatic face finding and emotion recognition. In *Learning Workshop of the Computational and Biological Learning Society*, Snowbird, Utah, April, 2.

Movellan, J.R., (2003). Building perceptive social robots: A new research paradigm for the cognitive sciences. In *International Symposium on Emergent Mechanisms of Communication*, Awaji Yumebutai International Conference Center, Japan, Feb 28-March 3.

Movellan, J.R., (2004). Finding People by Contingency: An Infomax Controller Approach. Poster presentation, Proceedings of the 2nd International Conference on Development and Learning.

Movellan, J.R., (2002). Real time object detection and classification. In *Institute of Statistical Mathematics*, Tokyo, November 22, 2002.

Movellan, J.R., (2002). Contingency, robotics and development. In *Annual Workshop of the Information Processing Society of Japan, Kansai area*, Osaka, October 31, 2002.

Movellan, J.R., (2002). The emergence of joint attention as a real-time skill aquisition process. In *Institute of Advanced Studies*, Kyoto, October 26, 2002.

Movellan, J.R., and Fasel I (2002). A new non-local algorithm for tracking deformable objects. In *Digital Creatures Laboratory*, Sony Corporation, Tokyo, October 14, 2002.

Movellan, J.R., (2002). Expression recognition and E-learning. In *Workshop for Research and Development of Human Communication Technologies for Conversational Interaction and Learning*, NSF, Merida MX, April 18-21, 2002.

Movellan, J.R., (2002). Automatic FACS coding. In *Workshop on Next Generation Face Recognition,*DARPA, March 4-5, 2002.

Movellan, J.R., (2001). Information integration in the brain. In *Program on Dynamics of Complex Systems in Science and Engineering*, Northwestern University, May 18, 2001.

Nelson, J., Cottrell, G., Movellan, J.R. (2004). Explaining eye movements during learning as an active sampling process. Proceedings *of the 3rd[d] International Conference on Development and Learning*.

Schwartz, O., J. R. Movellan, T. Wachtler, Albright T., and Sejnowski T. J., (2003). Spike count distributions factorability, and contextual effects in area V1. In *The Twelfth Annual Computational Neuroscience Meeting*, Spain, July 6, 2003.

Susskind, J., John Hershey, Javier Movellan (2004). Exact Inference in Robots Using Topographical Uncertainty Maps. Poster presentation, Proceedings of the 2nd International Conference on Development and Learning.

Susskind, J.M., J. Movellan, M.S. Bartlett, G. Littlewort, A.K. Anderson (2005). Comparison of Human and Computer Recognition of Facial Emotion. Proceedings of the of the Cognitive Neuroscience Society. New York, NY. April, 2005.

Vural, E., Bartlett, M.S. and Movellan, J.R. (2006). Machine learning systems for detecting driver drowsiness. Proceedings of Machine Learning, Theory, Applications, Experiences: A workshop for women in machine learning.

MPLAB GRANTS 2000-PRESENT

CURRENT

HOMELAND SECURITY / NAVAL RESEARCH LABORATORY BAA 55-05-03
Title: Automatic Facial Expression Recognition
PI: M. Bartlett, J. Movellan, G. Littlewort
Project Period: 4/1/05-3/31/09

Invited submission

NSF SCIENCE OF LEARNING CENTER
PI: Gary Cottrell
Co-PI's: Movellan, Chiba. Sejnowski
Title: The Temporal Dynamics of Learning
Project Period: 10/01/2006-9/31/2011

NSF CRI
PI: M. Bartlett
Title: Automated Facial Expression Measurement: Toolbox and Database
Project Period: 3/1/05-2/31/07

NSF ADVANCE FELLOWS AWARD
PI: Littlewort
Title: Theory and Application of Machine Perception to Understanding of Human Behavior
Project Period: 6/1/2004 – 5/31/2007

UC DISCOVERY PROGRAM/SONY dig03-10202
PI: J. Movellan
Title: Applications of Personal Robots to Learning and Education
Project Period: 9/1/2006-8/31/2007

NSF ADVANCED LEARNING TECHNOLOGIES
PI: Javier Movellan
Title: Computational Analysis of Real Time Social Interaction: Theory and Application
Project Period: 11/01/2006-10/31/2009

PREVIOUS

UC DISCOVERY PROGRAM/SONY dig03-10181
PIs: J. Movellan and T. Lee
Title: Applications of Personal Robots to Learning and Education
Project Period: 9/7/2005-9/6/2006

UC DISCOVERY PROGRAM/SONY dig03-10158
PIs: J. Movellan and T. Lee
Title: Applications of Personal Robots to Learning and Education
Project Period: 9/8/2004-9/7/2005

NSF IIS-0223052
PI: J. Movellan
Title: Diffusion-Based Approaches to Unsupervised Learning and Inference
Project Period: 8/1/2002-7/31/2004 (extended to 2005)

NSF IIS-0220141
PI: J. Movellan, M. S. Barlett, T. Sejnowski (in collaboration with M. Frank, Rutgers)
Title: Collaborative Research: ITR: Automatic Analysis of Spontaneous Facial Expressions
Project Period: 9/1/2002-8/31/2004 (extended to 2005)

NSF IIS-0329287
PI: J. Movellan, Co-PI's J. Triesch, V. De Sa
Title: ITR: Developmental Methods for Automatic Discovery of Object Categories
Project Period: 8/1/2003-7/31/2005

NAAR
PI: Triesch
Co-PI: Movellan
Title: The MESA project: Modelling the Emergence of Shared Attention
Project Period: 8/1/04 – 7/31/06

CIA 2000-I-058500-000
PI: T. Sejnowski
Co-I: Movellan, Bartlett
Title: RFP 00-I-007: Facial Analysis Coding System Project
Project Period: 7/24/00-9/30/2001

NIH F32 MH12417
PI: M. Bartlett
Title: Computer Vision Analysis of Dynamic Facial Behavior
Project Period: 6/1/99-5/31/02

ONR N00014-02-1-0616
PI: J. Movellan
Co-Investigator: M. Bartlett

Exploring the Next Generation of Dynamic Face Recognition Technology
Project Period: 6/1/02-12/31/03

UC DISCOVERY PROGRAM/SONY dig02-10158
PIs: J. Movellan and T. Lee
Title: Affective Computing for Personal Robots
Project Period: 8/12/03-8/11/2004

UC DIMI 01-10130 / SONY
PI: J. Movellan and T. Lee
Title: Affective Computing for Personal Robots
Project Period: 9/10/02-9/09/03 (extended to 12/31/03)

UC DIMI 00-10084 / SONY
PIS: J. Movellan and T. Lee
Title: Affective Computing for Personal Robots
Project Period: 5/14/2001-5/13/2002

SONY (UCSD 21-5056)
PI: J. Movellan
Title: Recognition of Face, Emotion Expression, Sound and Speech
Project Period: 7/15/2000-1/14/2001

UC DIMI Program Opportunity doa01-28
PI: J. Movellan
Title: Integrating Computer Animated Characters ad Machine Perception: An Exploration of Emerging Research and Technologies
Project Period: 8/6/02-6/30/03

UNIVERSITY OF COLORADO (NSF IIS-0086107-COLE) S0000018582
PI: J. Movellan
Title: ITR: Creating the Next Generation of Intelligent Animated Conversational Agents
Project Period: 9/1/2000-8/31/04

UC DAVIS MIND INSTITUTE
PI: Gedeon Deak
Co-investigator: **Javier Movellan**
Title: Testing Neuro Models Of Normal And Austistic Development Of Shared Attention

Department: Cognitive Science
Grant period: 8/1/02-7/31/03