Facial Expression Invariant Head Pose Normalization using Gaussian Process Regression

Ognjen Rudovic
Comp. Dept
Imperial College
London, UK
o.rudovic@imperial.ac.uk

Ioannis Patras
Elec. Eng. Dept
Queen Mary University
London, UK
i.patras@elec.qmul.ac.uk

Maja Pantic
Comp. Dept
Imperial College, London, UK
EEMCS, Univ. Twente, NL
m.pantic@imperial.ac.uk

Abstract

We present a regression-based scheme for facial-expression-invariant head pose normalization. We address the problem by mapping the locations of 2D facial points (e.g. mouth corners) from non-frontal poses to the frontal pose. This is done in two steps. First, we propose a head pose estimator that maps the input 2D facial point locations into a head-pose space defined by a low dimensional manifold attained by means of multi-class LDA. Then, to learn the mappings between a discrete set of non-frontal head poses and the frontal pose, we propose using a Gaussian Process Regression (GPR) model for each pair of target poses (i.e. a non-frontal and the frontal pose). During testing, the head pose estimator is used to activate the most relevant GPR model which is later applied to project the locations of 2D facial landmarks from an arbitrary pose (that does not have to be one of the training poses) to the frontal pose. In our experiments we show that the proposed scheme (i) performs accurately for continuous head pose in the range from 0° *to* 45° *pan rotation and from* 0° *to* 30° *tilt rotation despite the fact that the training was conducted only on a set of discrete poses, (ii) handles successfully both expressive and expressionless faces (even in cases when some of the expression categories were missing in certain poses during the training), and (iii) outperforms both 3D Point Distribution Model (3D-PDM) and Linear Regression (LR) model that are used as baseline methods for pose normalization. The proposed method is experimentally evaluated on data from the BU−3DFE facial expression database.*

1. Introduction

Rigid motion of the face accounts for a great amount of variance in its appearance in a 2D image array. Simultaneously, the non-rigid deformations of the face (from facial expression variation and identity variation) cause more subtle variations in the 2D image. An individual's identity and/or his or her facial expression, however, are captured by these small variations alone and are not specified by the variance due to the rigid head motion. Thus, it is necessary to compensate or normalize a face for position so that the variance due to this is minimized. Consequently, the small variations in the image due to identity, facial expression, and so on, will become the dominant source of variance in an image and can thus be analyzed for recognition purposes [12].

The main aim of head pose normalization is to generate a 'virtual' view, *i.e*. to normalize an input face image to a predefined pose (*e.g*. frontal), before further analysis is conducted. A standard solution to this problem in computer vision applications is to use a 3-D or 2-D face-shape model. In [3], a 3-D morphable model is used to estimate 3-D facial shape and bring it into frontal view where further analysis of faces is carried out. In [17], a 3D-PDM and normalized SVD decomposition are used to simultaneously recover facial expression and pose parameters. A similar 3D-PDM is employed in [16] to separate the rigid head rotation from non-rigid facial expressions. A rigid face shape model is applied to build person-dependent descriptors that were later used to decompose facial pose and expression simultaneously [10]. Well-known face-shape models in 2-D are 2-D PDM and Active Appearance Model (AAM), the latter of which fits an input face image to a pre-learned face model and consists of separated shape and appearance models [6]. In the work presented in [9], a non-frontal view image is warped onto the 2-D PDM and the target virtual view is synthesized via Thin Plate Splines-based texture mapping. In [7, 1], a multi-view AAM, where the frontal view of the input face can be easily synthesized, has been proposed. Overall, although the head pose normalization can be achieved to some extent by means of the 3-D/2-D face-shape models, a disadvantage of these methods is the use of generative models and/or fitting techniques that can fail to

978-1-4244-7030-3/10/$26.00 ©2010 IEEE

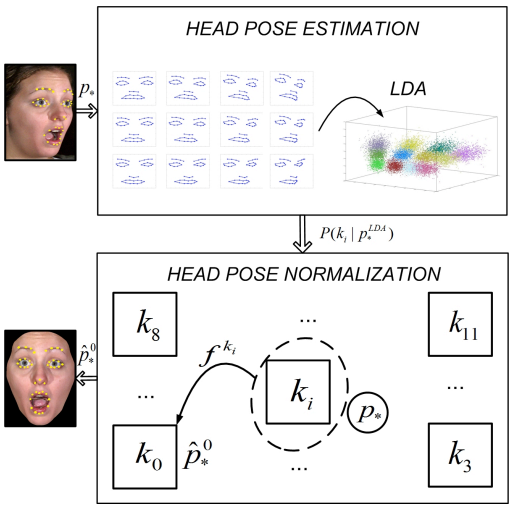


Figure 1. The proposed approach. Twenty points marked with yellow color are used as an input to the pose estimation. In the second block we couple two GP to obtain more accurate predictions in the frontal pose.

converge. In addition, most of these methods are computationally expensive and in need of time-consuming initialization process (e.g. due to manual annotation of more than 60 facial landmark points).

Next to 3-D/2-D model-based methods, another approach to head pose normalization is to apply a 2-D face-shape-free method. In [5], an affine transformation is used to independently map different face regions from a discrete set of non-frontal views to the frontal view. In [4], the Local Linear Regression (LLR) method is proposed. In this method, the whole surface of the face is partitioned into multiple uniform blocks, and the mappings between the corresponding blocks in non-frontal poses and in the frontal pose are learned using linear regression model. Similarly, in [8], the non-frontal facial region is partitioned into multiple facial components where linear transformations are applied to different components for the generation of their frontal counterpart. The underlying assumption of the aforementioned methods is that by partitioning the whole face surface into multiple patches, the mapping of each patch to its counterpart in the frontal view can be modeled using a linear transform [4]. However, none of the aforementioned works has analyzed the problem of virtual pose generation (i.e. head pose normalization) in the presence of facial expressions. Yet, as the rigid head motions and non-rigid facial expressions are non-linearly coupled in 2D [17], the underlying assumption that the mappings between different image patches can be represented as a linear transformation may not hold anymore. Moreover, the above models have been tested only on a set of the poses that were used to learn the mappings, and were not tested for unknown poses, i.e. the poses that were not used for learning the mappings.

In this paper we propose a 2-D face-shape-free method for head pose normalization. In contrast to above-described methods, the proposed method is facial-expression invariant and based on geometric (as opposed to appearance) features. Specifically, we address the problem by mapping the locations of 2-D facial points from non-frontal poses to the frontal pose. The proposed two-step approach is illustrated in Fig. 1. In the first step of the approach, we propose a head pose estimator that maps the input 2D facial point locations into a head-pose space defined by a low dimensional manifold attained by means of multi-class LDA. To deal with the problem of facial landmark localization, any state-of-the-art facial-expression-and-head-pose-invariant facial point detector can be used (*e.g.* see [14]). In this work we have used manual facial landmark annotations instead, which allows us to test the performance of the method alone, without the effect of landmark detection errors. In the second step, we use a Gaussian Process Regression (GPR) [13] model to learn mappings between the 2D locations of landmark points in non-frontal poses to their location in the frontal pose, and we do so for each pair of target poses (*i.e.* non-frontal and frontal). The contributions of the proposed approach can be summarized as follows.

1. We propose a novel facial-expression-invariant head pose normalization approach based on geometric features that can handle expressive faces in the range from $0°$ to $+45°$ pan rotation and from $0°$ to $+30°$ tilt rotation. The proposed approach performs accurately for continuous head pose (*i.e.* for unknown poses in the given range) despite the fact that the training was conducted only on a set of discrete poses. It also can successfully handle the problem of having an incomplete training dataset (e.g. when instances of certain facial expressions are not included in the training dataset for the given discrete pose).

2. We propose modeling the mappings between different head-poses using the state-of-the-art probabilistic non-linear regression model, namely, Gaussian Process Regression model. We experimentally show that the proposed scheme for facial-expression-invariant head pose normalization outperforms both the baseline face-shape methods for pose normalization like 3D PDM and the baseline face-shape-free methods such as Linear Regression (LR) model.

The rest of the paper is organized as follows. In Section 2 we present our facial-expression-invariant head pose normalization approach. In Section 3 we present and discuss experimental results, and in Section 4 we conclude the paper.

2. Facial-expression-invariant Head Pose Normalization

In this section we describe the proposed approach for facial-expression-invariant head pose normalization given the 2D locations $p = [p_1^x .. p_L^x \, p_1^y .. p_L^y]_{1\times d}$ of $L = d/2$ facial landmarks of an facial image observed in an arbitrary pose. The proposed approach consists of two main steps: (i) head pose estimation by using a pose classifier on p, (ii) pose normalization by mapping the positions p of the facial landmarks from a non-frontal pose to the corresponding 2D positions p^0 in the frontal pose. We assume that we have training data for each of P discrete poses and the correspondences between the points for each target pair of poses (non-frontal and frontal pose). In our case, we discretize the head pose space into $P = 12$ poses evenly distributed across the range from $0°$ to $+45°$ pan rotation and from $0°$ to $+30°$ tilt rotation, with an increment of $15°$. We denote with $D^k = \{p_1^k, ..., p_N^k\}$ the data set in pose k containing N vectors of facial landmark locations, and with $D = \{D^0, ..., D^k, ..., D^{P-1}\}$ the whole training data set. In what follows, we first introduce the proposed head pose estimation approach. Then, we describe the base GPR model used to independently learn $P-1$ mapping functions $\{f^1, .. f^k .., f^{P-1}\}$, one per each pair of target poses, $(k, 0)$. Note that while in the training phase landmark vector p is associated to one of P discrete poses, during testing the pose associated with p is unknown and may not belong to the discrete set of poses (here the constraint is that the pose belongs to the aforementioned range of pan and tilt rotation). We use the output of the pose estimator to select the most relevant GPR model, f^k, which then maps the points p from an arbitrary pose to the frontal pose.

2.1. Head Pose Estimation

Various head pose estimation methods based on appearance and/or geometric features are proposed in the literature [11]. In this paper, we propose to estimate the probability that input facial landmark vector p is associated to a head pose belonging to a discretized head-pose space. To obtain a low-dimensional expression-invariant head-pose manifold, we apply multi-class LDA. Prior to learning this low-dimensional manifold, we first normalize all examples from D (i.e. training data containing 2D locations of facial landmarks in P poses) to remove the scale and translation components. To do this end, we transform the vector of facial points p into image coordinates (p_i^x, p_i^y), $i = 1...L$, and compute the gravity center (p_c^x, p_c^y) and the scaling parameter sc as

$$p_c^x = \frac{1}{L}\sum_{i=1}^{L} p_i^x \;\;,\;\; p_c^y = \frac{1}{L}\sum_{i=1}^{L} p_i^y$$
$$sc = \frac{1}{L}\sum_{i=1}^{L} ((p_i^x - p_c^x)^2 + (p_i^y - p_c^y)^2)^{\frac{1}{2}}$$

where the normalized facial landmark positions are

$$p_i^{x_n} = \frac{((p_i^x - p_c^x)^2 + (p_i^y - p_c^y)^2)^{\frac{1}{2}}}{sc}(p_i^x - p_c^x)$$
$$p_i^{y_n} = \frac{((p_i^x - p_c^x)^2 + (p_i^y - p_c^y)^2)^{\frac{1}{2}}}{sc}(p_i^y - p_c^y).$$

The normalized vector $p^n = [p_1^{x_n}, .., p_L^{x_n}, .., p_1^{y_n}, .., p_L^{y_n}]$ is later used as an input to the multi-class LDA. We use LDA since it finds a simple linear transform S that, given a training set of normalized facial landmark points D^n along with the target pose labels, is used to project the high-dimensional input data p to a low dimensional manifold that best represents pose variations while ignoring the other sources of variation such as facial expressions, identity-specific variation across different individuals, etc. The facial landmark vectors projected onto the low-dimensional manifold, $p_{lda} = S \cdot p$, are used to model a head pose k by a Gaussian conditional distribution

$$P(p_{lda}|k) = \frac{1}{(2\pi)^{d/2}\,|\Sigma_k|^{1/2}} \cdot \exp(-\frac{1}{2}\left[p_{lda} - \mu^k\right]^T \Sigma_k^{-1}\left[p_{lda} - \mu^k\right]) \quad (1)$$

where μ^k and Σ^k are the mean and covariance computed from D_k^n. By applying the Bayes' rule to Eq.(1), we obtain the probability of pose k given the low dimensional vector p^{lda} as

$$P(k|p^{lda}) \propto P(p^{lda}|k)P(k) \quad (2)$$

where we assume an uniform prior, $P(k) = 1/P$, for each of P discrete head poses.

2.2. Gaussian Process Regression (GPR) Model

GPR model has gained increased popularity in statistical machine learning as it offers a principled nonparametric Bayesian framework for inference, model fitting and model selection [13]. Formally, for a given set of N input vectors p_i^k in pose k along with the target values p_i^0 in the frontal pose, we use the GPR model to find a smooth function f^k: $\mathbb{R}^d \rightarrow \mathbb{R}^d$ that maps the input facial landmark locations to the corresponding frontal facial landmark locations. Assuming a Gaussian noise ε_i with zero mean and covariance matrix $\sigma_n^2 I$, this is expressed by $p_i^0 = f^k(p_i^k) + \varepsilon_i$. Further, a zero mean Gaussian process prior is placed over the function f^k that is $f^k \sim GP(0, K + \sigma_n^2 I)$. Here, $K(D^k, D^k)$ denotes $N \times N$ covariance matrix obtained by applying the kernel function

$$k(p_i^k, p_j^k) = \sigma_s^2 \exp(-\frac{1}{2}(p_{ig}^k - p_{jg}^k)^T W (p_{ig}^k - p_{jg}^k)) + \sigma_l p_i^k p_j^k + \sigma_b \quad (3)$$

to the pairs (p_i^k, p_j^k), where $i, j = 1..N$. σ_s and $W = diag(w_1, ..., w_d)$ are the parameters of the Radial Basis

Function (RBF) with different length scales for each input dimension (each coordinate of each landmark point), σ_l is the process variance which controls the scale of the output function f^k, and σ_b is the model bias. We use this compound kernel function as it performed best in a pilot study which compared single RBF, MLP, Polynomial and Linear kernel, as well as their compound versions. In addition, the applied kernel has been widely used in literature due to its ability to handle both linear and non-linear data structures [2]. For a new facial landmark vector p_*^k from pose k we obtain the predictive mean $f^k(p_*^k)$ and the corresponding covariance $V^k(p_*^k)$ as

$$f^k(p_*^k) = k_*^T(K + \sigma_n^2 I)^{-1} D^0 \quad (4)$$

$$V^k(p_*^k) = k(p_*^k, p_*^k) - k_*^T(K + \sigma_n^2 I)^{-1} k_* \quad (5)$$

with $k_* = k(D^k, p_*^k)$, where $k(\cdot,\cdot)$ is given in Eq. (3). The kernel parameters $\theta = \{\sigma_s, W, \sigma_l, \sigma_b, \sigma_n\}$ are found by maximizing the log marginal likelihood, $\mathcal{L} = \log p(D^0|\theta)$, using the conjugate gradient algorithm. Because we are solving the multivariate regression problem here, we assume that the output dimensions (each coordinate of each landmark point in p_i^0) are identically distributed. This allows us to have the same covariance function for each output dimension [13].

2.3. Algorithm Summary

We give a summary of our approach to facial-expression-invariant head pose normalization in Alg. 1. In the off-line phase of the algorithm, we first apply multi-class LDA to the training data D to find the transformation matrix S that maps the input data to the low dimensional manifold of head poses and estimates the probability that the input data is in each of the discrete training poses. In the second step of the off-line phase, the training data is registered by defining reference faces, one for each pose k, and by mapping the landmark points in pose k to the corresponding reference face using an affine transform. The registration is carried out using three referential points: the nasal spine point and the inner corners of the eyes (see Fig. 1) that are chosen since they are stable facial points and the contractions of the facial muscles do not affect them. The registered training data are used to train $P-1$ GPR models, one for each pair of the non-frontal and the frontal pose. The learned models are stored for later use in the on-line part of the algorithm. In the on-line phase, an input vector of the facial landmark positions p_* in an arbitrary pose is first subjected to the proposed head pose estimation. The output of this step are the probabilities of p_* being in each of the P training poses. In the second step, we register p_* to the most likely pose k_{max} found in the previous step. Note that the affine transformation used to register points p_* to pose k_{max} helps both scaling of p_* and reducing the error introduced due to the discretization of the pose space (i.e. $\pm 7.5°$). Finally, if the retrieved pose is not the frontal pose, we select the most relevant GPR model, $f^{k_{max}}$, that is further used to make predictions of the facial landmark locations in the frontal view, $\hat{p}_*^0$, given the registered facial landmark vector $p_*^{k_{max}}$.

Algorithm 1 Head Pose Normalization

OFF-LINE: Learning model parameters
1. Apply LDA to compute transformation S, and learn μ_k and σ_k for each head pose $k = 0..P-1$ (Sec.2.1)
2. Form pairs of registered facial landmark locations, (D_{reg}^k, D_{reg}^0), $k = 1..P-1$ (Sec.2.3)
3. Learn GPR models $\{f^1, ..., f^{P-1}\}$ for $P-1$ target pairs of poses (Sec. 2.2)

ON-LINE: Inference of facial landmark locations p_* in an arbitrary pose
1. Apply pose estimation to obtain $P(k|p_*^{lda})$, where $p_*^{lda} = S \cdot p_*$ and $k = 0..P-1$ (Sec. 2.1)
2. Register p_* to pose $k_{max} = \max_k(P(k|p_*^{lda})), p_*^{k_{max}}$
3. Predict locations of facial points in frontal pose

if $(k_{max} == 0)$ **then**
$\hat{p}_*^0 = p_*^{k_{max}}$
else
$\hat{p}_*^0 = f^{k_{max}}(p_*^{k_{max}})$
end if

3. Experiments

The experimental evaluation of the proposed methodology has been carried out using the BU-3D Facial Expression (BU3DFE) database [15]. BU3FEDB contains 3D range data of 7 facial expressions performed by 100 subjects (60% female of various ethnic origin) which are shown in Fig. 2. All facial expressions except Neutral were sampled in four different levels of intensity. We generate 2D multi-view images of facial expressions from the available 3D data by rotating 39 facial landmark points provided by the database creators (see Fig. 4), which were used further as the features in our study. The data in our experiment include images of 50 subjects (54% female) from $0°$ to $45°$ pan rotation, and $0°$ to $30°$ tilt rotation (see Fig. 1), with $5°$ increment, resulting in 1250 images for each of 70 poses. The training data is subsampled from this dataset to include images of expressive faces in 12 poses only ($15°$ increment in pan and tilt angles). These data (referred to as BU-TR dataset in the text below) as well as the rest of the data (referred to as BU-TST dataset and used to test the performance of the proposed method in case of unknown head poses) were partitioned into five parts in a person-independent manner for use in a 5-fold cross validation procedure. The rest of this

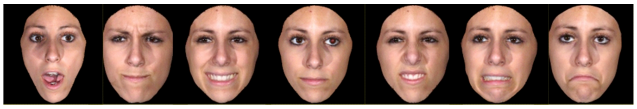


Figure 2. The examples of images from BU3FE dataset representing seven facial expressions in the frontal view: Surprise, Angry, Happy, Neutral, Disgust, Fear and Sad

section is organized as follows. First, we evaluate the accuracy of the proposed head pose estimator. Then, we present the experiments aimed at evaluation of the accuracy of the proposed head pose normalization method. To measure the accuracy of the method, we used the Root-Mean-Squared-Error (RMSE) that is defined as $\sqrt{\frac{1}{d}\|\Delta p\|^2}$, where Δp is the facial deformation vector representing the difference between the predicted positions of the facial landmarks in the frontal pose and the ground truth (the manually annotated landmarks in the frontal pose).

We evaluate the performance of the proposed head pose estimator when trained on BU-TR dataset and tested on a subset of BU-TST dataset containing only the facial data in the discrete poses used for training. A five-fold person-independent cross validation was performed, where data of 40 subjects from BU-TR were used for training, and data of 10 unknown subjects from BU-TST were used for testing the pose estimator. As can be seen from Table 1, an average classification error rate of 6.9% has been attained overall. It is interesting to note that correct pose estimation was easier to obtain for facial data in poses being far from the frontal pose than for facial data in poses being closer to the frontal pose, in which case confusions between neighbouring near-frontal poses were often encountered. This is, however, a natural consequence of the fact that the facial point distribution is rather similar for various near-frontal poses and that it is rather different for various poses being far away from the frontal pose.

	+45	+30	+15	+0
+30	0.9±1.3%	2.3±1.8%	5.1±1.1%	12.3±2.3%
+15	2.3±2.3%	3.9±1.6%	7.6±1.2%	14.8±2.5%
+0	3.9±2.2%	7.5±2.1%	10.2±2.2%	11.9±2.1%

Table 1. Average error rate for head pose estimation across 12 training poses where training/testing of the pose estimator was performed in a person-independent manner using data from BU-TR/BU-TST dataset. The head-pose classification was performed by selecting the most likely pose

To evaluate the proposed head pose normalization approach, first we trained 11 GPR models (as described in Alg. 1) using the data from BU-TR dataset, and then we tested the trained models on the whole BU-TST dataset (including the facial data in unknown poses). We compare the performance of the proposed GPR-based method to that achieved by the 'baseline' methods for pose normalization, namely, the 3D-PDM [17] and the standard LR model [2].

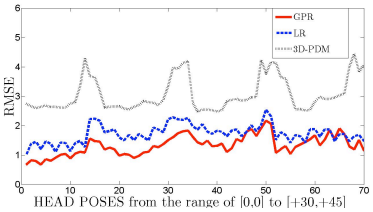
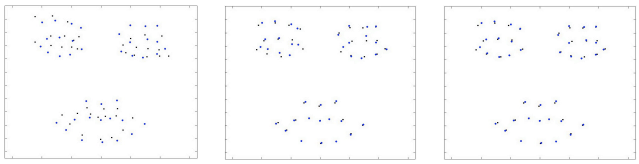


Figure 3. Comparison of head pose normalization methods GPR, LR, 3D-PDM trained on BU-TR (12 head poses) and tested on BU-TST (70 head poses) in a person-independent manner in terms of RMSE



(a) (b) (c)

Figure 4. Prediction of the facial landmarks in the frontal pose for an BU3DFE image of Happy facial expression in pose $(+45^\circ, +30^\circ)$ obtained by using (a) 3D-PDM (b) LR and (d) GPR. The blue $\bigtriangledown$ represent the ground truth and the black $\square$ are the predicted points. As can be seen, the alignment of the predicted and the corresponding ground truth facial landmarks is far from perfect in case of 3D-PDM

The 3D-PDM was trained using the corresponding frontal 3-D data from BU-TR dataset, retaining 15 modes of non-rigid shape variation. LR models were trained in a similar way as the GPR models, i.e. for each pair of target poses. Fig. 3 shows the comparative results in terms of RMSE of the tested head pose normalization methods along with the results obtained when no pose normalization is performed and only the translation component has been removed. As can be seen, both GPR- and LR-based methods outperform the 3D-PDM method for pose normalization. Judging from Fig 4, this is probably due to the fact that the tested 3D-PDM was not able to accurately model the non-rigid facial movements present in facial expression data.

The performance of the aforementioned models in the presence of noise in test data was evaluated on the BU-TST data corrupted by adding four different levels of noise. As can be seen from Table 3, even in the presence of high levels of noise the performance of GPR-based method is comparable to that of 3D-PDM achieved for noise-free data. Although the GPR-based method outperforms the LR-based method (as can be seen from Fig. 3), the performance of GPR- and LR-based can be considered comparable in the aforementioned experiments where the utilized data were balanced (i.e. when the method is trained on data containing examples of all facial expression categories and their intensity levels in all target poses). However, we further tested the ability of the GPR- and LR-based models to handle unseen non-rigid face motion, *i.e.* novel facial expres-

	$\sigma = 0$	$\sigma = 0.5$	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$
3D-ASM	2.9±0.5	2.8±0.6	3.3±0.8	3.6±0.6	3.8±0.6
LR	1.5±0.3	1.7±0.3	1.9±0.2	2.7±0.2	3.7±0.2
GPR	1.3±0.2	1.5±0.3	1.7±0.3	2.5±0.2	3.4±0.1

Table 2. Comparison of head pose normalization methods 3D-PDM, LR and GPR, trained on BU-TR (12 head poses) and tested on BU-TST (70 head poses) corrupted with different levels of Gaussian noise with standard deviation σ, in a person-independent manner, in terms of RMSE

sions. This is important when testing the generalization ability of the learned models. It is also central to the ability of the proposed model to learn the mappings even when examples of some facial expressions are not available in all poses. To evaluate this ability of the LR and GPR models, we did the following. We used the data of 40 subjects from BU-TR dataset for training the LR and GPR models. The used training data contained examples of the apex of six facial expressions (i.e. intensity-level-4 examples except for the facial expression Neutral where only one example per person is available). The testing was performed on intensity levels 2, 3 and 4 of the missing expression category (except Neutral), using data of 10 subjects from BU-TST dataset. As can be seen from Table 4, GPR-based method outperforms LR-based method. This clearly shows that the GPR-based head-pose normalization method is more robust to non-rigid deformations that are not seen in the training set than is the case with the LR-base method. Moreover, GPR-based method generalizes better than LR-based method even when just a small amount of training data is provided.

4. Conclusion

We presented a novel 2D-shape-free regression-based scheme for facial-expression-invariant head pose normalization that is based on 2-D geometric features. Experimental results show that the proposed method outperforms both the baseline 3D-face-shape method, i.e. 3D-PDM, and the baseline 2D-shape-free method, i.e. LR-based method. Moreover, we show experimentally that the proposed GPR-based method performs accurately for continuous head pose (i.e. for unknown poses in the given range of poses) despite the fact that the training was conducted only on a discrete (small) set of poses. Finally, experimental results clearly show that the proposed method can successfully handle the problem of having an incomplete training dataset, e.g., when examples of certain facial expressions were not included in the training dataset for the given pose.

Acknowledgments

The work of Ognjen Rudovic is funded in part by the European Community's 7th Framework Programme [FP7/2007-2013] under grant agreement no. 211486 (SEMAINE). The work of Maja Pantic is funded in part by the European Research Council under the ERC Starting Grant agreement no. ERC-2007-StG-203143 (MAHNOB). The work of Ioannis Patras is partially supported by EPSRC project EP/G033935/1.

	Disgust	Angry	Fear	Happy	Sad	Surprise	Neutral
LR	2.3±1.1	1.8±0.9	1.7±0.7	2.1±1.2	1.9±0.8	2.1±1.2	1.7±0.8
GPR	1.4±0.4	1.2±2.2	1.2±0.4	1.3±0.8	1.1±0.4	1.7±0.6	1.2±0.3

Table 3. RMSE for pose normalization obtained by using LR- and GPR-based methods in the presence of incomplete training dataset

References

[1] A. Asthana, R. Goecke, N. Quadrianto, and T. Gedeon. Learning based automatic face annotation for arbitrary poses and expressions from frontal images only. In *CVPR*, pages 1635–1642, 2009. 1

[2] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 2007. 4, 5

[3] V. Blanz, P. Grother, J. P. Phillips, and T. Vetter. Face recognition based on frontal views generated from non-frontal images. In *CVPR*, pages 454–461, 2005. 1

[4] X. Chai, S. Shan, X. Chen, and W. Gao. Local linear regression (llr) for pose invariant face recognition. In *AFGR*, pages 631–636, 2006. 2

[5] X. Chai, S. Shan, and W. Gao. Pose normalization for robust face recognition based on statistical affine transformation. In *ICSP*, pages 1413–1417, 2003. 2

[6] T. Cootes and C. Taylor. Active shape models - smart snakes. In *BMVC*, pages 266–275, 1992. 1

[7] T. Cootes, G. Wheeler, K. Walker, and C. Taylor. View-based active appearance models. *Image and Vision Computing*, 20(9-10):657 – 664, 2002. 1

[8] S. Du and R. Ward. Component-wise pose normalization for pose-invariant face recognition. In *ICASSP*, pages 873–876, 2009. 2

[9] D. Gonzalez-Jimenez and J. Alba-Castro. Symmetry-aided frontal view synthesis for pose-robust face recognition. In *ICASSP 2007*, pages II–237 –II–240, 2007. 1

[10] S. Kumano, K. Otsuka, J. Yamato, E. Maeda, and Y. Sato. Pose-invariant facial expression recognition using variable-intensity templates. *Int'l J. Computer Vision*, 83(2):178–194, 2009. 1

[11] E. Murphy-Chutorian and M. M. Trivedi. Head pose estimation in computer vision: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(4):607–626, 2009. 3

[12] M. Pantic and L. J. M. Rothkrantz. Automatic analysis of facial expressions: The state of the art. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22:1424–1445, 2000. 1

[13] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005. 2, 3, 4

[14] M. Valstar, B. Martinez, X. Binefa, and M. Pantic. Facial point detection using boosted regression and graph models. In *CVPR*, 2010. 2

[15] J. Wang, L. Yin, X. Wei, and Y. Sun. 3d facial expression recognition based on primitive surface feature distribution. In *CVPR*, pages 1399–1406, 2006. 4

[16] T.-H. Wang and J.-J. J. Lien. Facial expression recognition system based on rigid and non-rigid motion separation and 3d pose estimation. *Pattern Recognition*, 42(5):962 – 977, 2009. 1

[17] Z. Zhu and Q. Ji. Robust real-time face pose and facial expression recovery. In *CVPR*, pages 681–688, 2006. 1, 2, 5