

"How old are you?" : Age Estimation with Tensors of Binary Gaussian Receptive Maps

John A. Ruiz-Hernandez
john-alexander.ruiz-hernandez@inrialpes.fr

James L. Crowley
james.crowley@inrialpes.fr

Augustin Lux
augustin.lux@inrialpes.fr

INRIA Grenoble Rhones-Alpes
Research Center and Laboratoire
d'Informatique de Grenoble (LIG)
655 avenue de l'Europe
38 334 Saint Ismier Cedex, France

Abstract

In this paper we describe experiments with a method to automatically estimate human age from facial images. This system extends recent results with the use of a tensorial representation from Gaussian receptive field responses for face recognition to the problem of estimating age. Among other results, we show that inclusion of fourth order Gaussian receptive fields can improve recognition. We describe an optimal tensorial configuration and compare the use of two different configurations with Multilinear Principal Component Analysis to reduce tensor order and Relevance Vector Machines as regressor. Experimental results are demonstrated for the FG-NET and the MORPH aging datasets.

1 Introduction

Human faces are an important non-verbal source of information for human interaction. In addition to expressions of emotion, the human face communicates gender, ethnic origin and age [19]. Automatic estimation of age from facial images has been used for: Age-Specific Human Computer interaction (ASHCI) [15, 19], security [18, 19], missing individuals retrieval [24, 19], internet access control for minors [9, 15], and surveillance monitoring of alcohol or cigarettes vending machines [7, 8], appearance prediction across aging [22], and targeting of publicity [8].

Despite the number of potential applications, automatic image-based age estimation remains a challenging problem. Compared with other facial variations, aging effects are very dependent on genetics [9], life style, location of residence [10] and weather conditions [24]. Furthermore, males and females age differently, and the apparent effects of aging are often masked by makeup and facial accessories [22]. Accommodating the influence of individual differences to provide a general method for estimating age based on facial images remains an open problem.

Most automatic image-based age estimation systems are composed by combining two components [24]: an image representation and a age estimation process. Five different approaches are commonly used for image representation:

1) anthropometric models [13, 18] have been considered for coarse age estimation and simulation of changes in appearance with age.

2) Active Appearance Models (AAMs) [9] have been widely used as a feature for age estimation. For example, Geng *et al.* [10] have described the AGing pattErn Subspace (AGES) that uses AAMs as face feature to create a sequence of individual aging's images sorted in time order to make a representative subspace.

3) Age manifolds where the underlying face age structure is projected on a low-dimensional feature space. For example, Guo *et al.* [8] describes the use of a Orthogonal Locality Preserving Projection (OLPP) [11] to model facial aging variations in a low dimensional space.

4) Feature-Based models that describe facial aging structure with a set of discriminant texture features. For example, Gabor filters have been described by Guo *et al.* [10] as discriminant features for age estimation.

5) Aging simulation models. Suo *et al.* [22] has described the use of a Markov process using a sparse graph representation trained with a large annotated face dataset.

Image representation is completed by a second component that provides the actual age estimation. Examples include multiclassification [15], regression [8, 10] and hybrid methods [9].

The methods described in this paper employ a set of Binary Gaussian receptive field maps calculated using a scale-invariant half-octave Gaussian pyramid [5]. Image representations estimated from Gaussian derivative operators have been widely used in invariant object recognition [21, 30], face recognition [20, 27], image tracking and scene reconstruction [16, 25, 26]. In the experiments described below, binary Gaussian receptive field maps are used as a robust image representation that encodes facial changes due to the aging process. A tensorial representation is used to encode the shape changes in the aging process and to conserve the 3D structure of the scale space. Multi-linear Principal Component Analysis (MPCA) [17] is used to reduce the dimensionality in the tensorial space. Relevance Vector Machines (RVM) [24] are used as a regression method for estimating age.

The presentation is organized as follows: Binary Gaussian Receptive Field maps are discussed in Section 2. Section 3, presents the tensorial representation for facial aging images while section 4 presents two optimal configurations for automatic-image-based age estimation. Section 5 provides comparative experimental evaluations against other state of the art approaches. Some conclusions are presented in section 6.

2 Binary Gaussian Receptive Maps

2.1 Steerable Gaussian Receptive Fields

In Neuro-science, the receptive field of a sensory neuron is a region of space in which the presence of a stimulus will alter the firing of that neuron. For mammals, receptive fields have been identified for neurons of the auditory system, the somatosensory system, and the visual system. Young *et al.* [5] has reported that receptive fields in the visual cortex can be well modeled using Gaussian derivative operators up to fourth order. From this, Koenderink *et al.* [2] argues that the local visual appearance in an image neighborhood can be represented by a local Taylor series expansion of the neighborhood, computed using local Gaussian derivatives. The coefficients of this series constitute a feature vector that represents image appearance and can be used for indexing, matching and recognition. Following the terminology in [21], we refer to the use of Gaussian derivatives for images analysis as Gaussian receptive Maps.

A Gaussian scale space is computed as:

$$I(x, y, \sigma) = G(x, y, \sigma) * I(x, y) = \frac{1}{2\sigma\pi} e^{-\frac{x^2+y^2}{2\sigma^2}} * I(x, y) \quad (1)$$

$$G_k^m = \frac{\partial I(x, y, \sigma)}{\partial^k x \partial^m y}$$

Where σ is the size of the support in terms of the second moment (or variance), I is the image and $*$ is the convolution operator. From the preceding equation, the steerable filter [1] response for the Gaussian derivative up to fourth order for an arbitrary orientation θ could be defined as follows:

$$G_1^\sigma(\theta) = \cos(\theta)G_1^0 + \sin(\theta)G_0^1$$

$$G_2^\sigma(\theta) = \cos^2(\theta)G_2^0 - 2\sin(\theta)\cos(\theta)G_1^1 + \sin^2(\theta)G_0^2$$

$$G_3^\sigma(\theta) = \cos^3(\theta)G_3^0 - 3\sin(\theta)\cos^2(\theta)G_2^1 + 3\sin^2(\theta)\cos(\theta)G_1^2 - \sin^3(\theta)G_0^3 \quad (2)$$

$$G_4^\sigma(\theta) = \cos^4(\theta)G_4^0 - 4\cos^3(\theta)\sin(\theta)G_3^1 + 6\sin^2(\theta)\cos^2(\theta)G_2^2 - 4\cos(\theta)\sin^3(\theta)G_1^3 + \sin^4(\theta)G_0^4$$

The use of Gaussian derivative operators to model receptive fields is motivated by their capacity to describe image neighborhoods and also by their representation of image appearance in specific orientations and frequencies [1]. This representation is well suited for object recognition and can be efficiently computed at half-octave scales using a linear complexity $O(N)$ algorithm for computing a Gaussian pyramid. More details concerning linear complexity calculation of the half octave pyramid may be found in [1].

2.2 Generating Binary Gaussian Receptive Maps

This section presents the Binary Gaussian Receptive Maps. Binary Gaussian Receptive Maps provide a robust representation that encodes changes in facial appearance due to aging. To compute such maps, a Local Binary Pattern (LBP) [2] is applied over each Gaussian Receptive field to assign a label to each pixel of the image by thresholding the 3×3 neighborhood of each pixel with the center pixel value and considering the result as a binary or decimal number. This operator encodes a set of micro-patterns from the neighborhood appearance in Gaussian Receptive Fields. The mathematic expression for this operation is:

$$BG_n^\sigma(\theta) = LBP(G_n^\sigma(\theta)) \quad (3)$$

Where $BG_n^\sigma(\theta)$ is the Binary Gaussian Receptive map, n is the order of the Gaussian derivative, σ is the scale factor, θ is the steering angle and LBP is the Local Binary Pattern. An example of $BG_n^\sigma(\theta)$ calculated at different scale orders and orientations is shown in the figure 1.

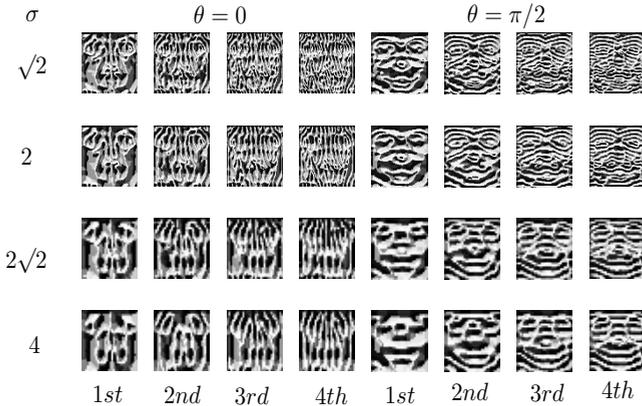


Figure 1: Example of Binary Gaussian Receptive Maps computed at different scales ($\sigma = \sqrt{2}, 2, 2\sqrt{2}, 4$), orders ($n = 1, 2, 3, 4$) and orientations ($\theta = 0, \pi$)

3 Tensorial Representation of $BG_n^\sigma(\theta)$

In addition to information about local appearance, as captured by the Binary Gaussian Receptive maps, age estimation also requires information about 3D structure of feature space that encodes shape information. To encode such information, we propose a tensorial representation for the Binary Gaussian Receptive maps. To compute this representation, first we divide each $BG_n^\sigma(\theta)$ of each image into non-overlapping rectangular sub-regions with a specific size. A set of histograms is then computed for each sub-region and finally each histogram is organized in four different 3-D tensors, where each tensor corresponds to a specific derivative order of the Binary Gaussian Receptive maps. The characteristic equation of $\mathcal{T}(BG_n^\sigma(\theta))$ is shown as follows:

$$\mathcal{T}(\mathbf{BG}^\sigma(\theta))_{\mathbf{n}} \in \mathbb{R}^{N_\theta \times N_{\text{pos}} \times N_{\text{bins}}} \quad \mathbf{n} = 1, 2, 3, 4 \quad (4)$$

Where n and N_θ are the order and orientation angles for the gaussian derivatives respectively, N_{pos} is the number of non-overlapping positions in the map and N_{bins} is the number of bins used in the construction of each local histogram.

Recognition in a high-dimensional tensor representation can suffer from the well known "curse of dimensionality". This representation may be projected to a more compact uncorrelated vectorial representation, by performing Multilinear Principal Component Analysis (MPCA) [17] over each tensor. Mathematically we describe this operation over each $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_{\mathbf{n}}$ as follows:

$$\mathbf{y}_{\mathbf{n}} = \text{MPCA}(\mathcal{T}(\mathbf{BG}^\sigma(\theta))_{\mathbf{n}}) \quad \mathbf{y}_{\mathbf{n}} \in \mathbb{R}^{m \leq M} \quad \mathbf{n} = 1, 2, 3, 4 \quad (5)$$

Where $\mathbf{y}_{\mathbf{n}}$ is the vectorial representation of the n 'th $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_{\mathbf{n}}$, m is the number of retained entries ranked in order of descending variance and M is the number of tensor samples used to calculate the orthonormal matrix in the MPCA algorithm.

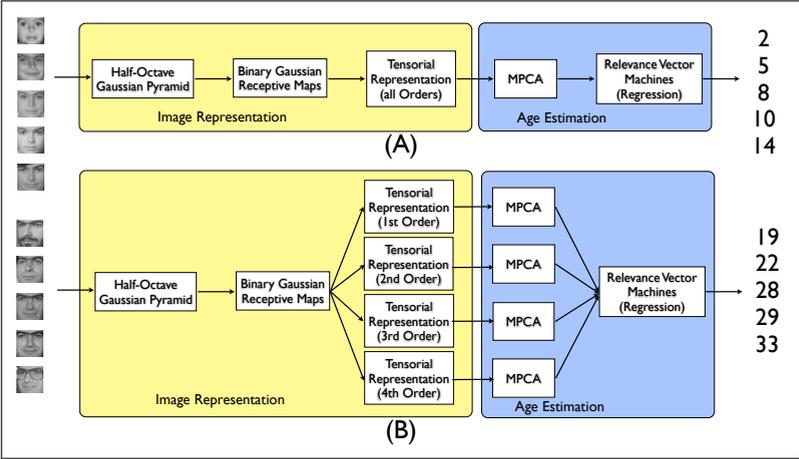


Figure 2: Proposed Configurations for Age Estimation. (a) The tensors are fused into one before applying MPCA (b) MPCA is applied at each tensor and then the resulting vectors are fused into one

4 Optimal Configurations for Age Estimation

We have compared two possible algorithms for automatic age estimation using $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_n$. The structures for these two algorithms are shown in the figure 2. These algorithms are inspired in [20, 23]. The first algorithm concatenate the tensors $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_1$, $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_2$, $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_3$ and $\mathcal{T}(\mathbf{BG}^\sigma(\theta))_4$ to form a 4-D tensor \mathcal{T}_T as shown in the next equation:

$$\begin{aligned} \mathcal{T}_T &= [\mathcal{T}(\mathbf{BG}^\sigma(\theta))_n] \quad \mathbf{n} = 1, 2, 3, 4 \\ \mathcal{T}_T &\in \mathbb{R}^{N_n \times N_\theta \times N_{\text{pos}} \times N_{\text{bins}}} \end{aligned} \quad (6)$$

MPCA is then applied to this tensor \mathcal{T}_T to obtain a vector $Y_T \in \mathbb{R}^{m \leq M}$ with a dimension similar to the equation 5.

$$Y_T = \text{MPCA}(\mathcal{T}_T) \quad Y_T \in \mathbb{R}^m \quad (7)$$

The second algorithm computes MPCA over each tensor separately as shown in equation 5 to obtain the vectors y_1, y_2, y_3 and y_4 , these vectors are concatenated to form a single vector $Y_F \in \mathbb{R}^{4 \times m}$, where $m \leq M$.

$$Y_F = [y_n] \quad Y_F \in \mathbb{R}^{4 \times m} \quad n = 1, 2, 3, 4 \quad (8)$$

The relative advantage of each of the two algorithms are discussed in the next section.

Finally, we use Relevance Vector Machines (RVM) [24] rather than Support Vector Machines (SVM) as a regression algorithm. RVM is preferred because it does not required setting the free parameters required by SVM, and does not require cross-validation based post optimizations.

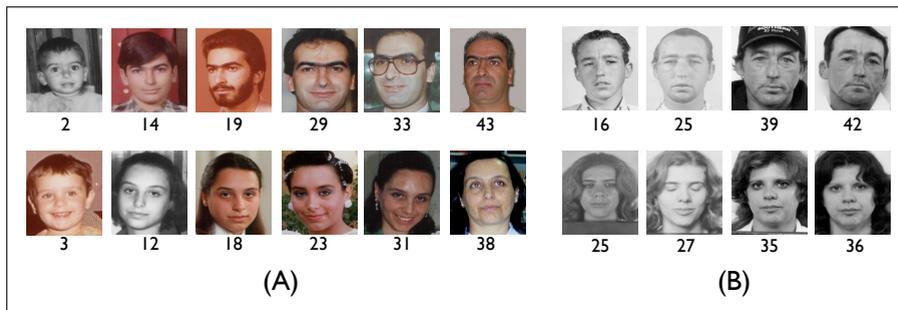


Figure 3: Sample images of different individuals from: (A) the FG-NET dataset and (B) the MORPH dataset

5 Experimental Evaluation

We have performed several experiments to compare different approaches for estimating age from facial images. Two publicly available databases have been used in our experiments: The FG-NET [1] database and the MORPH [2] database. Some examples of images from these datasets are shown in figure 3.

The FG-NET Aging Database contains 1,002 face images of 82 subjects from multiple races with age ranges from 0 to 69 years. Each image in the database has 68 labelled facial landmarks characterizing shape features not used in our approach. Leave-One-Person-Out (LOPO) mode is used for testing our approaches in the FG-NET database. In this mode, the images of one person are used as the test set and those of the others are used as the training set. After 82 folds each subject has been used as a the test set once and the final results are calculated based in the result of each fold. The MORPH Aging Database contains 1,724 face images of 515 subjects. In our experiment with this database, we use the same testing protocol used by Geng *et al.* [2]. The images on these dataset are only used to test the algorithms trained on the FG-NET database. In addition, because all subjects in the FG-NET database are caucasian descent, only the 433 images of caucasian descent in the MORPH database are used as the test set.

In all of our experiments, the images were cropped using manually located eye positions and normalized in size to 64×64 pixels. The Binary Gaussian Receptive Maps are calculated in a half-octave Gaussian pyramid with four levels ($\sigma = \sqrt{2}, 2, 2\sqrt{2}$ and 4). A border of 4 pixels in each pyramid level is left untested for faces to avoid problems related with image borders.

To calculate histograms, we used a sub-region size of 8×8 pixels, removing two bins corresponding to values of 0 and 255 per each histogram. The remaining bins are grouped to form a 127 bin histogram. We used the publicly available *Matlab*[®] implementation of the MPCA and RVM algorithms provided by [1]¹ and [2]² respectively. After applying MPCA, each final vector is normalized to unit standard deviation. For the RVM algorithm we used a Gaussian kernel $k(x, y) = e^{-\|x-y\|^2/q}$ with a scale parameter q determined using a tuning dataset chosen randomly from the training dataset (the value of q is reported in each experiment and is the only one fixed by tuning). The performance of age estimation

¹<http://www.dsp.utoronto.ca/haiping/index.php>

²<http://www.vectoranomaly.com/downloads/downloads.htm>

TensorConf	Order	Orientations	
		6	8
Y_T (equation 7, $m = 1000$) + RVM	$n = 1, 2, 3$	5.25 ($q = 28.72$)	5.16 ($q = 30.82$)
	$n = 1, 2, 3, 4$	5.25 ($q = 29.58$)	5.23 ($q = 30.82$)
	$n = 3, 4$	5.53 ($q = 29.58$)	5.48 ($q = 30.82$)
Y_F (equation 8, $m = 1000$) + RVM	$n = 1, 2, 3$	5.23 ($q = 50.00$)	5.18 ($q = 50.00$)
	$n = 1, 2, 3, 4$	5.17 ($q = 66.33$)	4.96 ($q = 66.33$)
	$n = 3, 4$	5.49 ($q = 38.73$)	5.46 ($q = 38.73$)

Table 1: MAEs on the FG-NET database for different tensorial configurations

is measured by the mean absolute error (MAE) and the cumulative score (CS). The MAE is defined as the average of the absolute errors between the estimated ages and the ground truth ages $MAE = \sum_{k=1}^N |Age_k - \hat{Age}_k|/N$, where Age_k is the ground truth age for the test image k , \hat{Age}_k is the estimated age, and N is the total number of test images. MAE is only an indicator of average performance for age estimators, it does not provide enough information of how accurate the estimators might be. The accuracy can be estimated by the cumulative score (CS) that is defined as $CS(j) = N_{e \leq j}/N \times 100\%$, where $N_{e \leq j}$ is the number of test images on which the estimator makes an absolute error non-higher than j years.

5.1 Comparing Y_T and Y_F performances

Our first experiment investigates the performance of each type of configuration for the age estimation problem. We compared the configurations mentioned in the section 4 on the FG-NET database and we report the results in table 1. Different orientations and derivative orders were tested to obtain the best configuration. We used orientations between 0 and π and derivative orders up to four. Experimental results show that the best performance can be achieved with 8 orientations ($0, \pi/7, 2\pi/7, 3\pi/7, 4\pi/7, 5\pi/7, 6\pi/7$ and π) and four derivative order, organized in a Y_F configuration (equation 8), the second best result was achieved with 6 orientations ($0, \pi/5, 2\pi/5, 3\pi/5, 4\pi/5$ and π) and three derivative orders in a Y_T configuration. The best results were highlighted in the table 1 and used in the following experiments.

5.2 Experimental Results in the FG-NET database

We compared the two best tensorial configurations of table 1 with the most relevant results of the state-of-the-art in age estimation. In table 2, we report the results for seven different age groups between 0 and 69 years. In this table we observed that our method outperforms other methods for age groups 40 - 59 and 20-29 years. For other age groups, our method is competitive and sometimes superior with competing approaches in age estimation. The experiments the robustness of our method to textural changes that occur in the periods from adulthood to old age.

In the FG-NET database, the MAEs of our method are 5.16 for the Y_T configuration and 4.96 for the Y_F configuration. Comparisons with alternative approaches are reported in table 3. Our results are comparable with the latest state-of-the-art methods in automatic age estimation in the FG-NET database.

Comparisons of cumulative scores (CS) on the FG-NET database are shown in Figure 4. We can observe that despite the MAE results for our methods, our approaches outperforms

Range	# img.	$Y_T + RVM$	$Y_F + RVM$	BIF [10]	RUN [12]	QM [13]	MLP [14]
0-9	371	3.14	3.19	2.99	2.51	6.26	11.63
10-19	339	4.05	3.90	3.39	3.76	5.85	3.33
20-29	144	4.72	4.29	4.30	6.38	7.10	8.81
30-39	70	10.08	9.17	8.24	12.51	11.56	18.46
40-49	46	14.15	13.76	14.98	20.09	14.80	27.98
50-59	15	22.06	20.06	20.49	28.07	24.27	49.13
60-69	8	33.12	32.25	31.62	42.50	37.38	49.13
Total	1002	5.16	4.96	4.77	5.78	7.57	10.39

Table 2: MAE (years) at different age groups on FG-NET.

Method	QM [13]	MLPs [14]	RUN [12]	BM [15]	LARR [8]	PFA [9]	BIF [10]	$Y_T + RVM$	$Y_F + RVM$
MAE	6.55	6.98	5.78	5.33	5.07	4.97	4.77	5.16	4.96

Table 3: MAE (years) comparisons on FG-NET.

state-of-the-art methods in low age error levels ($Error\ level \leq 4$), with almost 5% of improvement in accuracy $CS_{\leq 4} = 73\%$, in addition for high error levels our method has an $CS_{\leq 10} = 88\%$ similar to BIF [10] ($CS_{\leq 10} = 89\%$) that as far as we know the best result in the FG-NET dataset.

5.3 Experimental Results in the MORPH(test) Database

More experiments were conducted on the MORPH(test) aging database. From the results in the table 4, the MAEs results of our method are 6.19 and 6.76 for Y_F and Y_T respectively, those results outperforms the AGES method [7] in almost two years of difference and other methods like SVM and WAS with a difference of almost three years.

The CS curves on the MORPH database are shown in Figure 5. Our method outperforms other methods in error levels for all of the age groups with an $CS_{\leq 10} = 84\%$ against the AGES lda method [7] with a $CS_{\leq 10} = 78\%$.

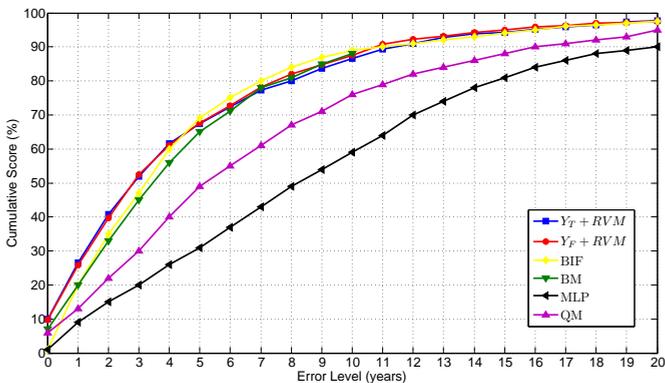


Figure 4: Cumulative scores on the FG-NET database

Method	WAS [□]	SVM [□]	AGES [□]	AGES _{lda} [□]	Y_T +RVM	Y_F +RVM
MAE	9.32	9.23	8.83	8.07	6.77	6.19

Table 4: MAE (years) comparisons on MORPH (Test Set).

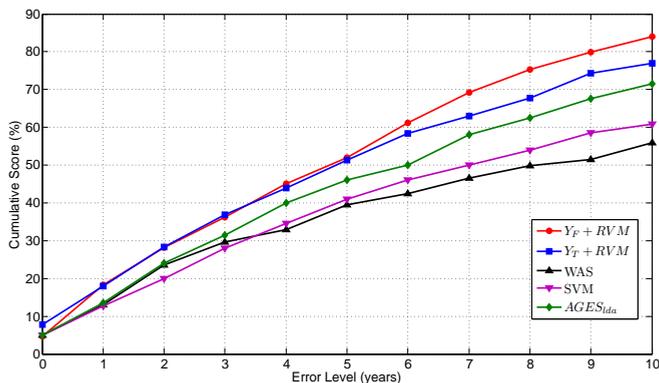


Figure 5: Cumulative scores on the MORPH (test) database

6 Conclusion

In this paper, we have introduced a new approach for estimating age from facial images. Our method uses a simple set of Binary Gaussian Receptive Maps calculated with a Half-Octave Gaussian pyramid. We have proposed a tensorial representation that conserves the natural structure of the information contained in the binary maps. Two algorithms for age estimation have been proposed and customized to solve this problem. Finally the Relevance Vector Machines (RVM) has been investigated for age estimation based in the tensorial representation.

References

- [1] The fg-net aging database. <http://www.fgnet.rsunit.com/>. accessed: April 2010.
- [2] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE PAMI*, 28(12):2037–2041, 2006.
- [3] D. Cai, X. He, J. Han, and H-J. Zhang. Orthogonal laplacianfaces for face recognition. *IEEE TIP*, 15(11):3608–3614, 2006.
- [4] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor. Active appearance models. *ECCV*, pages 484–498, 1998.
- [5] J.L. Crowley and O. Riff. Fast computation of scale normalised gaussian receptive fields. In *Proc. Scale Space Methods in Computer Vision*, pages 584–598, 2003.
- [6] W. T. Freeman and E. H. Adelson Y. The design and use of steerable filters. *IEEE PAMI*, 13:891–906, 1991.

- [7] X. Geng, Zhi-Hua Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE TPAMI.*, 29(12):2234–2240, 2007.
- [8] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE TIP*, 17(7):1178–1188, 2008.
- [9] G. Guo, Y. Fu, Charles R. Dyer, and Thomas S. Huang. A probabilistic fusion approach to human age prediction. *IEEE CVPR Workshop*, 2008.
- [10] G. Guo, G. Mu, Y. Fu, and Thomas S. Huang. Human age estimation using bio-inspired features. *IEEE CVPR*, pages 112–119, 2009.
- [11] Karl Ricanek Jr. and Tamirat Tesafaye. Morph: A longitudinal image database of normal adult age-progression. *IEEE FG*, 2006.
- [12] J.J Koenderink and A.J van Doorn. Representation of local geometry in the visual system. *Biol. Cybern.*, 55(6):367–375, 1987. ISSN 0340-1200.
- [13] Y.H. Kwon and N. da Vitoria Lobo. Age classification from facial images. *Comput. Vis. Image Underst.*, 74(1):1–21, 1999.
- [14] A. Lanitis, C. J. Taylor, and T. F. Cootes. Toward automatic simulation of aging effects on face images. *IEEE TPAMI.*, 24(4):442–455, 2002.
- [15] A. Lanitis, C. Draganova, and C. Christodoulou. Comparing different classifiers for automatic age estimation. *IEEE TSMCB*, 34(1):621–628, 2004.
- [16] David G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60: 91–110, 2004.
- [17] Haiping Lu, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos. MPCA: Multilinear principal component analysis of tensor objects. *IEEE TNN*, 19(1):18–39, 2008.
- [18] N. Ramanathan and R. Chellappa. Face verification across age progression. *IEEE TIP*, 15(11):3349–3361, 2006.
- [19] N. Ramanathan, R. Chellappa, and S. Biswas. Computational methods for modeling facial aging: A survey. *J. Vis. Lang. Comput.*, 20(3):131–144, 2009.
- [20] John A. Ruiz-Hernandez, James L. Crowley, Antoine Meler, and Augustin Lux. Face recognition using tensors of census transform histograms from gaussian features maps. *BMVC.*, September 2009.
- [21] B. Schiele and J.L. Crowley. Recognition without correspondence using multidimensional receptive field histograms. *IJCV*, 36:31–50, 2000.
- [22] Jinli Suo, Song-Chun Zhu, Shiguang Shan, and Xilin Chen. A compositional and dynamic model for face aging. *IEEE TPAMI.*, 32(3):385–401, 2010.
- [23] X. Tan and B. Triggs. Fusing gabor and lbp feature sets for kernel-based face recognition. In *Analysis and Modelling of Faces and Gestures*, pages 235–249, oct 2007.

-
- [24] Michael E. Tipping. Sparse bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.
- [25] Engin Tola, Vincent Lepetit, and Pascal Fua. Daisy: An efficient dense descriptor applied to wide baseline stereo. *IEEE PAMI*, 99(1), 2009.
- [26] Simon A. J. Winder and Matthew Brown. Learning local image descriptors. In *IEEE CVPR*, June 2007.
- [27] J. Wright and Gang Hua. Implicit elastic matching with random projections for pose-variant face recognition. In *IEEE CVPR*, June 2009.
- [28] Shuicheng Yan, Huan Wang, Xiaoou Tang, and Thomas S. Huang. Learning auto-structured regressor from uncertain nonnegative labels. *IEEE ICCV*, 2007.
- [29] Shuicheng Yan, Huan Wang, Xiaoou Tang, Jianzhuang Liu, and Thomas S. Huang. Regression from uncertain labels and its applications to soft biometrics. *IEEE TIFS*, 3(4):698–708, 2008.
- [30] J.J. Yokono and T. Poggio. Oriented filters for object recognition: an empirical study. In *IEEE FG*, 2004.
- [31] Richard A. Young, Ronald M. Lesperance, and W. Weston Meyer. The gaussian derivative model for spatial-temporal vision: I. cortical model. *Spatial Vision*, 2001:3–4, 2001.