Facial age estimation and gender classification using Multi Level Local Phase Quantization

Salah Eddine Bekhouche*, Abdelkrim Ouafi*, Azeddine Benlamoudi†, Abdelmalik Taleb-Ahmed‡, Abdenour Hadid§
*Laboratory of LESIA, University of Biskra, Algeria. salah@bekhouche.com, ou_karim@yahoo.fr
†Laboratory of LAGE, University of Ouargla, Algeria. be.azzeddine@gmail.com
‡LAMIH, UMR CNRS 8201 UVHC, University of Valenciennes, France. taleb@univ-valenciennes.fr
§Center for Machine Vision Research, University of Oulu, Finland.

***Abstract*—Facial demographic classification is an attractive topic in computer vision.Attributes such as age and gender can be used in many real life application such as face recognition and internet safety for minors. In this paper, we present a novel approach for age estimation and gender classification under uncontrolled conditions following the standard protocols for fair comparaison. Our proposed approach is based on Multi Level Local Phase Quantization (ML-LPQ) features which are extracted from normalized face images. Two different Support Vector Machines (SVM) models are used to predict the age group and the gender of a person. The experimental results on the benchmark Image of Groups dataset showed the superiority of our approach compared to the state-of-the-art.**

Keywords*—Age estimation, Gender classification, Local Phase Quantization, Support Vector Machines

I. Introduction

Age estimation and gender classification belong to soft biometrics that provides ancillary information of an individual's identity information. Age estimation is defined as the task of determining the age of a person based on biometric features. It can be considered as either a multi-class problem (age classification) or a regression problem (age estimation) [1]. Gender classification is a binary classification problem (male vs. female).

There are several studies on automatic demographic classification (age, gender and ethnicity) from facial images. Recently, several applications have emerged to make the demographic classification usefull. These applications include access control, re-identification in surveillance videos, law enforcement, integrity of face images in social media, intelligent advertising, and human-computer interaction, to cite few [2].

Most of age estimation and gender classification studies have utilized face images captured in controlled conditions. Many variations may appear in the image of a person's face. These variations may affect the ability of the computer vision system to estimate the age or recognize the gender. They are due to many factors which can be divided into: human factors (race, facial expression etc.) and capture process factors (pose, illumination, quality etc.).

Gallagher et al. [3] studied contextual features for capturing the structure of people images. Instead of treating each face independently, they extracted features which cover the structure of the group from persons' faces. Firstly, they used the social context features, then they tried to use appearance features. Finally, they combined context and appearance features together achieving the accuracy of 42.9% and 74.1% for age and gender respectively.

Shan [4] investigated age estimation and gender classification by treating each face independently. He focused on appearance features exactly on Local Binary Patterns (LBP) and Gabor features as face representation, then he adopted Adaboost to learn the discriminative local features. The best performance in his experiments is with the boosted LBP based on SVM classifier with an accuarcy of 50.3% for age and 74.9% for gender.

Li et al. [5] focused on facial age estimation based on ordinal discriminative feature learning. They tried to remove redundant information from both the locality information and ordinal information by minimizing non linear correlation and rank correlation, using different feature selection algorithms (Laplacian Score, LAR, Fisher Score, Rank Boost and PLO). The Piecewise Linear Orthonormal (PLO) gave the best results for the age estimation with an accuracy of 48.5%.

Ylioinas et al. [6] studied automatic age classication using LBP variants. They proposed a method based on a combination of LBP variants encoding the structure of elongated facial micro-patterns and their strength. Their experimental results gave an accuracy of 51.7% for age classification.

Fu et al. [7] wrote a survey about age synthesis and estimation. They divided the age estimation systems into two concatenated modules: age image representation and age estimation techniques. There are five main models in age image representation, they are: Anthropometric Models, Active Appearance Models, Aging Pattern Subspace, Age Manifold and Appearance Models. And the age estimation techniques are: Classification, Regression and Hybrid techniques. They mentioned the majority of the aging databases including the Images of Groups Dataset. They summarize and compare different age estimation methods on different databases.

Ng et al. [8] wrote a survey about human gender recognition from face, gait and body. They mentioned some face preprocessing methods that may be applied in gender classification systems like normalization for contrast and brightness. They also categorized feature extraction methods for face gender classification into geometric-based and appearance-based methods. As in any survey, a comparaison of different methods and databases was summarized.

In our work, we propose a new approach to estimate the age and recognize the gender in uncontrolled conditions based on facial images. We uses faces detected directly from the images without alignment, also it extracts the facial features using the multi level local phase quantization (ML-LPQ) which are powerful means of texture description. These texture features are then classified by multi-class Support Vector Machine (SVM) via one-against-one strategy. We conduct experiments on the Images of Groups Dataset. We also compare multi block and multi level local phase quantization results.

The rest of the paper is organized as follows: in Section II we explain our approach. The experiments and the results are given in Section III. In section IV we give the conclusion and the future plans.

II. PROPOSED APPROACH

The local phase quantization (LPQ) method is based on the blur invariance property of the Fourier phase spectrum [9]. It uses the local phase information extracted using the 2-D DFT or, more precisely, a short-term Fourier transform (STFT) computed over a rectangular $M-by-M$ neighborhood at each pixel position x of the image $f(x)$ defined by:

$$F(u,x) = \sum_{y \in N_x} f(x-y) e^{-j2\pi u^T y} = w_u^T f_x \qquad (1)$$

where w_u is the basis vector of the 2-D Discrete Fourier Transforms (DFT) at frequency u, and f_x is another vector containing all M^2 image samples from N_x [9].

In [10], we divided the face ROI (region of interest) into 3×4 sub-blocks and applied Local Binary Patterns (LBP) features on each sub-block. This method is called multi block local binary patterns (MB-LBP). Inspired by MB-LBP we propose MB-LPQ (see figure 1). The MB-LPQ method provides the most important information, i.e it gives better results and these results differentiates according to the division of the face ROI and experiments proved this.

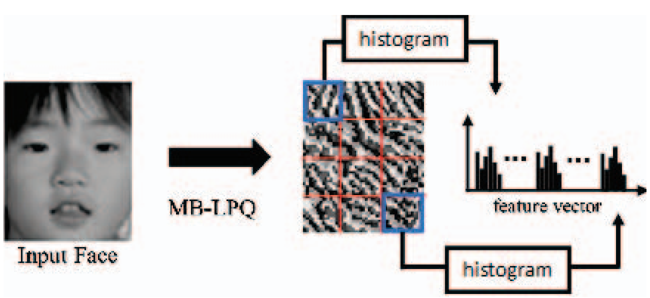

Fig. 1: Example of MB-LPQ features extraction with 4×3 sub-blocks.

Based on the LPQ method, instead of using single MB-LPQ division we use Multi Level Local Phase Quantization (ML-LPQ). The main idea of ML-LPQ is to extract features from different MB-LPQ divisions and then combine them. In other words, extracting features from the whole image, then dividing the image into 2^2 sub-blocks and extracting the features from each sub-blocks and so on until we reach the intended level. The final result of ML-LPQ is $1^2+2^2+....+n^2$ histograms. We combine these histograms to get the feature vector. Figure 2 explains our approach.

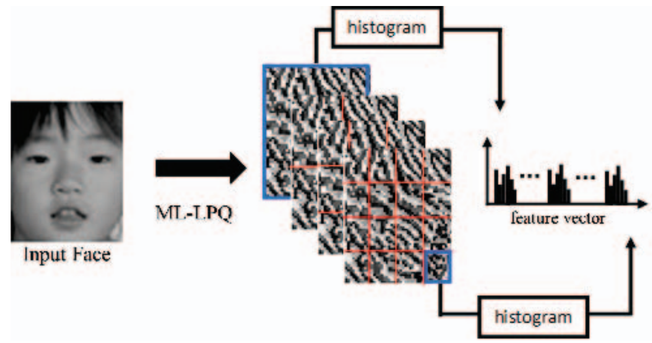

Fig. 2: Example of ML-LPQ features extraction with 4 level.

To classify each subject into his/her age group and gender, we use multi-class Support Vector Machine (SVM) via one-against-one strategy with a non-linear RBF kernel. Multiclass SVMs are generally considered as unreliable for arbitrary data, since in many cases no single mathematical function exists to separate all classes of data from one another. The one-against-one strategy, also known as one-versus-one method, an SVM is constructed for every pair of classes by training it to discriminate the two classes. Thus, for a problem with c classes, $c(c-1)/2$ SVMs are trained to distinguish the samples of one class from the samples of another class. The maxwins strategy is commonly used to determine the class [11].

III. EXPERIMENTS

To evaluate the performance of the proposed approach, we use the Images of Groups database[1] (see Fig. 3) which contains 5080 images with 28231 faces labeled with age and gender. There are seven age categories as follows:0-2, 3-7, 8-12, 13-19, 20-36, 37-65, and 66+ years, roughly corresponding to different life stages. In some images people are sitting, laying, or standing on elevated surfaces. People often have dark glasses, face occlusions, or unusual facial expressions [3].

A. Experimental Settings

As we are not studying face detection we use the faces that detected by [3], they manually add missed faces, but 86% of the faces are automatically found. Following the BeFIT[2] recommendations, we trained the classifiers using a random selection of 3500 faces, having equal number of samples per each age category (500 per class). Tests was performed on an independent uniformly distributed set of 1050 face samples (150 per class). For the gender classification we used 23218 face samples as a train and 1881 as a test as in [3]. The table below shows the distribution of the 25099 faces used in the experiments.

B. Experimental Results

The performance of the age estimation is measured by the accuracy of an exact match (AEM) and the accuracy of

[1] http://chenlab.ece.cornell.edu/people/Andy/ImagesOfGroups.html
[2] http://fipa.cs.kit.edu/425.php

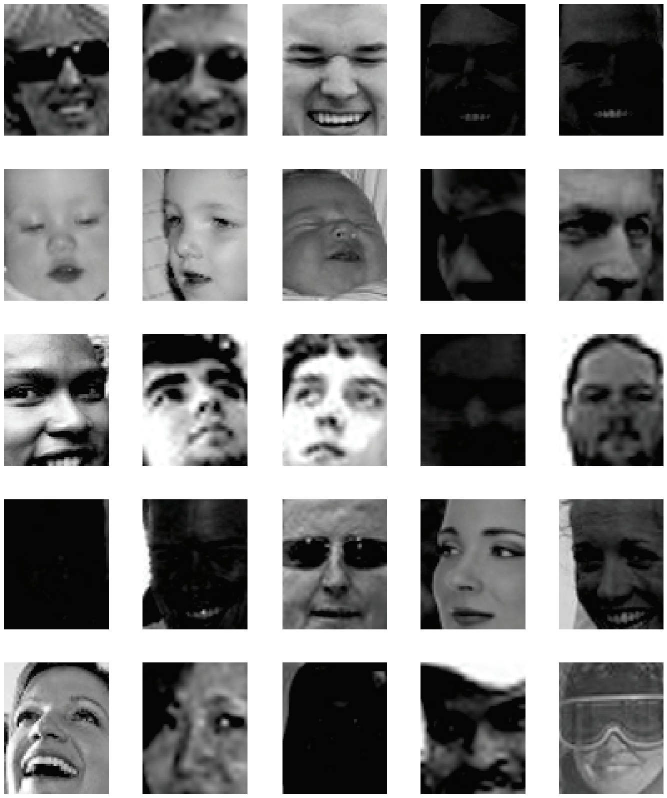

Fig. 3: Example of detected faces from Image of Groups database.

	0-2	3-7	8-12	13-19	20-36	37-65	+66
Male	382 44	664 56	386 45	704 61	6452 440	2887 241	501 56
Female	337 44	625 55	365 44	641 60	5988 439	2799 241	487 55
Total	719 88	1289 111	751 89	1345 121	12440 879	5686 482	988 111

TABLE I: The age/gender distribution of the faces used for the gender classification

allowing an error of one age category (AEO) (e.g. a 0-2 year old predicted as 3-7 year old). They are defined as follows [5]:

$$AEM = \frac{\hat{N}_m}{N'} \times 100\% \quad (2)$$

$$AEO = \frac{\hat{N}_o}{N'} \times 100\% \quad (3)$$

where $\hat{N}_m$ is the number of an exact match with N' test images,and $\hat{N}_o$ is the number of correct prediction when allows an error of one age category. The measure of the gender classification is evaluated by:

$$Acc = \frac{\hat{N}}{N'} \times 100\% \quad (4)$$

where $\hat{N}$ is the number of true classificaion.

To obtain the best results, we consider the different LPQ parameters, LPQ with 5×5 local window size gives the best results. We also compare different MB-LPQ uniform divisions, the results are shown in Table II.

MB-LPQ divisions	AEM	AEO	ACC
2×2	39.42%	70.05%	64.36%
3×3	45.32%	80.96%	71.93%
4×4	**50.03%**	**84.27%**	**75.16%**
5×5	47.69%	81.73%	72.46%

TABLE II: Comparaison between the diffrent MB-LPQ uniform divisions

Finally, we apply ML-LPQ and change the level until the results began to fall. Table III shows the changes in the results between different levels. All these experiments are evaluated by multi-class SVM based on one-class SVM.

ML-LPQ level	AEM	AEO	ACC
1	29.14%	54.76%	63.10%
2	39.71%	69.33%	69.91%
3	51.14%	83.52%	75.06%
4	**56%**	**88.8%**	**79.1%**
5	53.52%	85.81%	76.50%

TABLE III: Comparaison between different levels of ML-LPQ

We have obtained AEM = **56%** and AEO = **88.19%**. The confusion matrix of the age estimation is shown in Table IV. From the table, the classification of infants (0-2 years old) and over-ages (66+ years old) have the highest accuarcy compared to the other age groups. However for other groups we remark that it is difficult to determine the exact group due to several reasons. The most important is the converged ages like 3-7 and 8-12 groups which represent the childhood.

	0-2	3-7	8-12	13-19	20-36	35-65	66+
0-2	**116**	27	3	2	1	1	0
3-7	19	**87**	33	4	3	3	1
8-12	5	38	**66**	23	9	5	4
13-19	3	7	28	**73**	27	9	3
20-36	1	5	14	31	**72**	23	4
35-65	1	4	5	8	38	**67**	27
66+	1	3	2	3	4	30	**107**

TABLE IV: Confusion matrix of age estimation

The confusion matrix of the gender classification is shown in Table V. From this table, we see that our approach recognizes males easily compared to females. For this classification, we have obtained an accuracy of **79.1%** (cf. Table VI).

We conducted experiments on the Images of Groups dataset [3] for age estimation and gender classification. Table VI shows the comparaison against the state-of-the-art. Its confirm the superiority of our approach in the case of gender classification and age classification with exact match.

	male	female
male	802	141
female	251	687

TABLE V: Confusion matrix of gender classification

Approach	Age		Gender
	AEM	AEO	ACC
Context [3]	32.9%	64.4%	66.9%
LaplacianScore [5]	35.5%	74.5%	N/A
Appearance [3]	38.4%	71.3%	69.6%
FisherScore [5]	42.8%	83.7%	N/A
Appearance+Context [3]	42.9%	78.1%	74.1%
Gabor + Adaboost [4]	43.7%	80.7%	70.2%
LAR [5]	44.8%	84.9%	N/A
RankBoost [5]	44.8%	84.5%	N/A
LBP + Adaboost [4]	44.9%	83.0%	71.0%
boosted Gabor + SVM [4]	48.4%	84.4%	73.3%
PLO [5]	48.5%	88.0%	N/A
boosted LBP + SVM [4]	50.3%	87.1%	74.9%
CLBP_M + SVM [6]	51.7%	88.7	N/A
Ours	**56%**	**88.8%**	**79.1%**

TABLE VI: Age estimation and gender classification results and comparison to state-of-the-art

IV. Conclusion

In this paper, we described a novel approach for demographic (age group and gender) classification based on Multi Level Local Phase Quantization features and One-class SVM with non-linear kernel (RBF) classifier. The experimental results showed that our approach provides a better performance than previous approaches. As a future work, we will study the effect of age on gender recognition and the other human factors (race, facial expression, ..etc.) on age estimation. We will also consider feature selection to boost the performance.

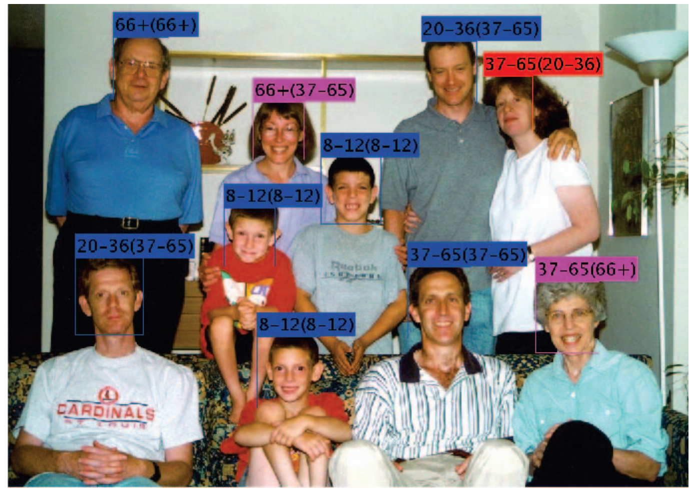

Fig. 4: Examples of age estimation (ground truth in parentheses) and gender classification (male: blue, female: magenta and wrong gender: red).

References

[1] P. Grd, "Introduction to human age estimation using face images," *Research Papers Faculty of Materials Science and Technology Slovak University of Technology*, vol. 21, no. Special Issue, pp. 24–30, 2013.

[2] H. Han and A. K. Jain, "Age, gender and race estimation from unconstrained face images," Department of Computer Science, Michigan State University, East Lansing, Michigan, Tech. Rep. MSU-CSE-14-5, July 2014.

[3] A. Gallagher and T. Chen, "Understanding images of groups of people," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, June 2009, pp. 256–263.

[4] C. Shan, "Learning local features for age estimation on real-life faces," in *Proceedings of the 1st ACM International Workshop on Multimodal Pervasive Video Analysis*, ser. MPVA '10. New York, NY, USA: ACM, 2010, pp. 23–28.

[5] C. Li, Q. Liu, J. Liu, and H. Lu, "Learning ordinal discriminative features for age estimation," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, June 2012, pp. 2570–2577.

[6] J. Ylioinas, A. Hadid, and M. Pietikainen, "Age classification in unconstrained conditions using lbp variants," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, Nov 2012, pp. 1257–1260.

[7] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 11, pp. 1955–1976, 2010.

[8] C. Ng, Y. Tay, and B.-M. Goi, "Recognizing human gender in computer vision: A survey," in *PRICAI 2012: Trends in Artificial Intelligence*, ser. Lecture Notes in Computer Science, P. Anthony, M. Ishizuka, and D. Lukose, Eds. Springer Berlin Heidelberg, 2012, vol. 7458, pp. 335–346.

[9] V. Ojansivu and J. Heikkilä, "Blur insensitive texture classification using local phase quantization," in *Image and signal processing*. Springer, 2008, pp. 236–243.

[10] S. Bekhouche, A. Ouafi, A. Taleb-Ahmed, and A. Hadid, "Facial age estimation using bsif and lbp," in *Proceeding of the first International Conference on Electrical Engineering ICEEB14*, in press.

[11] J. Milgram, M. Cheriet, R. Sabourin *et al.*, "one against one or one against all: Which one is better for handwriting recognition with svms?" in *Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.