Age Estimation from Face Images: Challenging Problem for Audience Measurement Systems

Vladimir Khryashchev, Alexander Ganin, Olga Stepanova, Anton Lebedev
Yaroslavl State University
Russia
vhr@yandex.ru, angnn@mail.ru, dcslab@uniyar.ac.ru, lebedevdes@gmail.com

Abstract—**The real-time audience measurement system consists of five consecutive stages: face detection, face tracking, gender recognition, age classification and in-cloud data statistics analysis. The challenging part of such system is age estimation algorithm on the basis of machine learning methods. The face aging process is determined by different factors: genetic, lifestyle, expression and environment. That is why same age people can have quite different rates of facial aging. We propose a novel algorithm consisting of two stages: adaptive feature extraction based on local binary patterns and support vector machine classification. Experimental results on the FG-NET, MORPH and our own database are presented. Human perception ability in age estimation is studied using crowdsourcing which allows a comparison of the ability of machines and humans.**

I. INTRODUCTION

Automatic video data analysis is a very important problem in modern society. A lot of different algorithms, using popular techniques as principal component analysis, histogram analysis, artificial neural networks, Bayesian classification, adaptive boosting learning, and many others, have been proposed in the field of computer vision and object recognition over recent years [1, 2].

A wide variety of information that can be extracted from a face image, such as identity, age, gender, ethnicity, and scars, marks and tattoos. The identification characteristic of face images has been well explored in real-world applications, such as passports and driver licenses control [1-3]. Despite the broad interest for face recognition algorithms, there is only a limited number of research on how to estimate and use the demographic information (age, gender, ethnicity) contained in face images [4].

In this paper we focus on age estimation, whose objective is to determine the specific age or age group of a subject based on preliminary detected face region. Among its possible applications one should note electronic customer relationship management (such systems assume the usage of interactive electronic tools for automatic collection of age information of potential consumers in order to provide individual advertising and services to clients of various age groups), security control and surveillance monitoring (for example, an age estimation system can warn or stop underage drinkers from entering bars or wine shops, prevent minors from purchasing tobacco products from vending machines, etc.), biometrics (when age estimation is used as a part that provides ancillary information of the users' identity information, and thus decreases the whole system identification error rate). Besides, age estimation can be applied in the field of entertainment, for example, to sort images into several age groups, or to build an age-specific human-computer interaction system, etc. [1].

In order to organize a completely automatic system, classification algorithms are utilized in the combination with a face detection algorithm, which selects candidates for further analysis. In papers [3, 4] we propose a system which extracts all the possible information about depicted people from the input video stream, aggregates and analyses it in order to measure different statistical parameters (fig. 1). The quality of face detection step is critical to the final result of the whole system, as inaccuracies at face position determination can lead to wrong decisions at the stage of recognition [5, 6]. To solve the task of face detection AdaBoost classifier, described in paper [7], is utilized. Detected fragments are preprocessed to align their luminance characteristics and to transform them to uniform scale. On the next stage detected and preprocessed image fragments are passed to the input of gender recognition classifier which makes a decision on their belonging to one of two classes («Male», «Female») [4, 13]. Same fragments are also analyzed by the age estimation algorithm. The proposed gender and age classifiers are based on non-linear Support Vector Machines (SVM) classifier with Radial Basis Function (RBF) kernel [37]. To extract information from image fragment and to move to a lower dimension feature space Local Binary Patterns (LBP) features are utilized [38-41]. To estimate the period of a person's stay in the range of camera's visibility, face tracking [8-11] algorithm is used. It is based on Lucas-Kanade optical flow calculation procedure [12].

The challenging part of such system is age estimation algorithm on the basis of machine learning methods. A number of studies in the biological, psychological, and cognitive science areas, have reported on how the human

ISSN 2305-7254

brain perceives, represents, and remembers faces. In particular, various aspects of human age estimation have been studied in the field of psychology [15]. Psychological studies often have the goal of examining the efforts of a subject's age, gender, and race on the accuracy of age estimation that the subject provides [16]. However, the accuracy of age estimation by human subjects on a large scale has not been reported for most databases used in automatic age estimation research [17]. A variety of research in the area of age classification has been done over last few years [14, 18-34].

The main contribution of this paper are as follows. Human perception ability in age estimation is studied using crowdsourcing expert opinion which allows a comparison of the ability of machines and humans. We propose a novel algorithm consisting of two stages: adaptive feature extraction based on local binary patterns and support vector machine classification. Experimental results on the FG-NET, MORPH [35, 36] and our own database are presented.

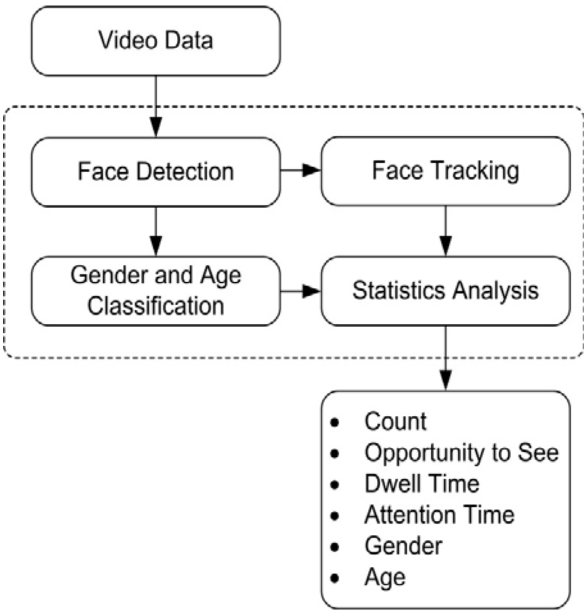

Fig. 1. Real-time audience measurement system based on computer vision and machine learning algorihms

II. FACE AGING DATABASES

There are conventional databases that are widely used in a field of age estimation from facial images. Each image in such data sets contains information about biological age of a person on this image. The most commonly used database is MORPH [35]. It contains over 55 thousand facial images of more than 13000 different persons: men and women of various races, nationalities and ages. Database MORPH is quite big and suitable for both training and testing.

Images of people of Caucasian race (about 17000) have been selected from MORPH database. Faces were automatically allocated using boosting algorithm. Further, false positives samples were removed. Moreover, large number of fragments of unsatisfactory quality was manually discarded. Such fragments contained images with significant defects of lightning, repeated images of the same person of the same age, faces with unnatural facial expression (closed eyes, grimacing, etc.) and faces of persons whose age is visually significantly different from the biological.

For selecting of images a group of experts was involved. It resulted in creating of an image set from MORPH database. This set contains 6513 fragments which are most suitable for carrying out experiments on the automatic age estimation. Statistical distribution of faces depending on the age is represented on Fig. 2.

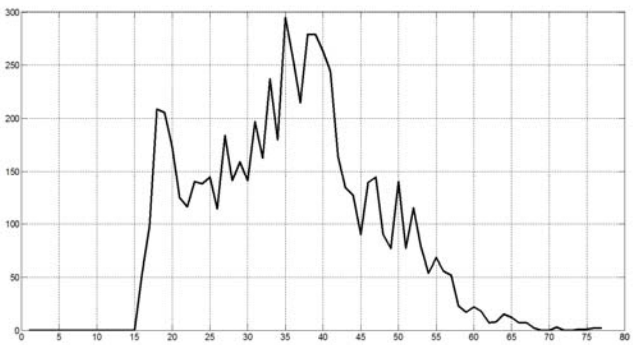

Fig. 2. Statistical distribution of faces depending on the age for the set of images that was selected from MORPH database

It should be noted that the distribution of the ages in the final database of selected fragments is nonuniform. There are no images of people whose age is less than 16 years, and for people who are older than 57 years the number of available fragments on the class is inadequate for training or testing. The most of the images falls on the age range from 30 to 40 years. Thus, the base MORPH has necessary number of images (approximately 100 per class) only in the range of 16 to 55 years. For range extension it is required to develop own database. Particular attention should be given to images of people younger than 16 years and older than 57 years.

Another conventional database is FG-NET [36]. It contains 1002 images of 82 persons (about 12 images of different ages for each person). This database is not enough big for training, despite of it FG-NET is widely used as testing database in literature.

For the research purpose faces were automatically allocated. Further, false positives were removed manually. The resulting set contained 841 person (level of selection for boosting algorithm was 84%). The distribution of faces depending on the age for FG-NET is represented on Fig. 3.

The distribution is nonuniform: most of images falls on the age younger than 20 years. For the rest age ranges number of images decreases. Such distribution may distort the statistics of the test algorithm and make an analysis of its work unreliable. Because of this there is the important task of developing a faces test material having a uniform distribution of the ages. Due to these challenges the task of developing our own database with uniform distribution is being of current interest.

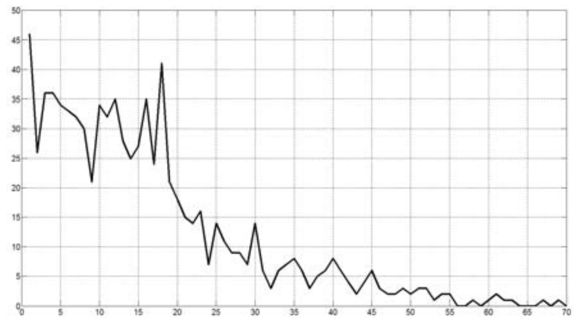

Fig. 3. Statistical distribution of faces depending on the age for the set of images that was selected from FG-NET database

The main problem of creating the specialized database for age recognition task is a lack of information about the real biological age of people on images. However, it is possible to use a visually perceived age as a reference. For this purpose a group of experts is involving to determine approximate age of each person in database in crowdsourcing experiment. Database that is labeled in a such way is suitable for both training and testing of algorithms. Such experiment was carried out for new database (10500 images) that was previously used for training and testing of gender classifier. Each image was evaluated by five experts. The final value of age was obtained by averaging taking into account possible blunders which were not taken into account. The statistical distribution for the resulting database is shown on Fig. 4.

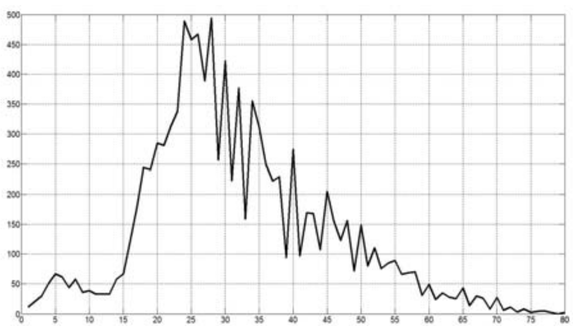

Fig. 4. Statistical distribution of faces depending on the age for own database

The distribution is nonuniform, the most of the images falls on the age range from 16 to 40 years. The developed database is more suitable for training and testing of algorithms because it has more wide age range than MORPH database. However, developed database contains insufficient number of images of young and older people. This challenge leads to creating another database with a controlled number of images for each age.

Control Russian Faces Database (RUS-FD) that was selected from free sources of information (social network Vkontakte with labeled age for person avatar picture) contains 150 images of real-life low resolution (60x60 pixels on each face) Russian-people faces for each age (from 6 to 60 years). The biological age of people on images was known in advance. The accuracy of this information was verified by the expert group.

Thus, four different databases was prepared for training and testing of algorithms of age estimation. Their main parameters are given in Table. I.

TABLE I. PARAMETERS OF TESTING AND CONTROL FACE AGING DATABASES

Database	**Number of images**	**Age range**	**Distribution by age**	**Labeling of age**	**Purpose**
Image set from MORPH	6513	16-75	nonuniform	real biological age	training and testing
Image set from FG-NET	841	1-70	nonuniform	real biological age	testing
Own database	10500	1-80	nonuniform	subjectively perceived age	training and testing
Control database (RUS-FD)	8100	6-60	uniform (150 images per each age)	real biological age	training and testing

III. AGE ESTIMATION ALGORITHM

The proposed age estimation algorithm realizes multiclass classification approach (Fig. 5) where for each age (from 1 to N) a binary classifier is constructed deciding whether a person on input image looks older than the given age or not. Input fragments are preprocessed to align their luminance characteristics and to transform them to uniform scale. Preprocessing includes color space transformation and scaling, both similar to that of gender recognition algorithm.

Additionally image normalization was performed by histogram equalization procedure. Transformation to LBP feature space and SVM training procedure are used for binary classifier construction. To predict direct age binary classifier outputs are statistically analyzed and the most probable age becomes the algorithm output.

To test age estimation algorithms performance standard metrics were calculated:

- Mean Absolute Error (MAE) – mean absolute difference between estimated and real ages.
- Cumulative Score (CS) – the probability that estimated age lies within an interval dx from real age.
- Probability Density Function of age estimation error.

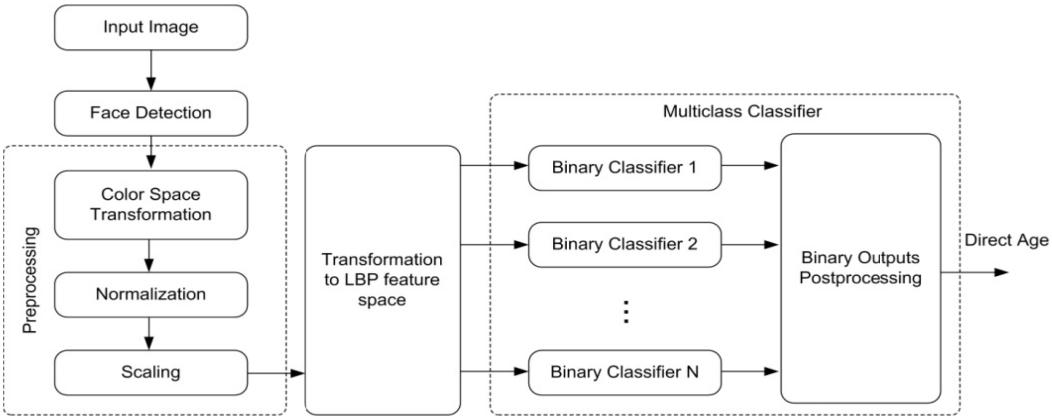

Fig. 5. LBP-SVM age estimation algorithm block diagram

To estimate the proposed algorithm in real-life situation testing firstly performed on FG-NET database. Age on FG-NET database was marked manually by a group of experts to compare subjective estimation with the algorithm performance. The corresponding dependences for LBP-SVM algorithm simulation are presented on Fig. 6 and Fig. 7.

The proposed algorithm shows results comparable to the subjective evaluation in a range of ages from 20 to 35 years. The average absolute error in this range is about 6 years old. Accuracy of LBP-SVM algorithm decreases on senior ages because of MAE grows. In this range (45-60 years), the proposed algorithm yields an expert evaluation approximately 10-15 years in terms of average error.

Cumulative score shows that around 40% of estimations have less than 5 years deviation from true age and 70% - less than 10 years deviation. Subjective evaluation curve on Fig. 7 give us the possible limit for future age estimation algorithm improvement.

Analysis of the error probability density function shows that the proposed algorithm has close to symmetric error distribution. Objective results are not inclined to overestimate the true age, which is typical for the evaluation of experts.

MAE and CS comparison for LBP-SVM algorithm on different test databases is presented on Fig. 8 and Fig.9.

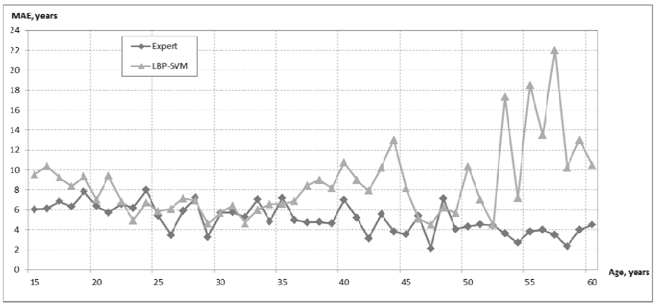

Fig. 6. MAE on FG-NET database for LBP-SVM algorithm

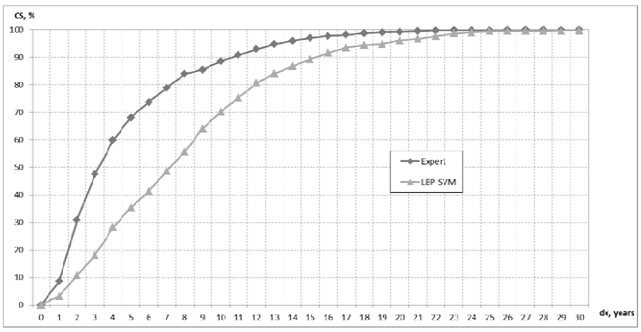

Fig. 7. CS on FG-NET database for LBP-SVM algorithm

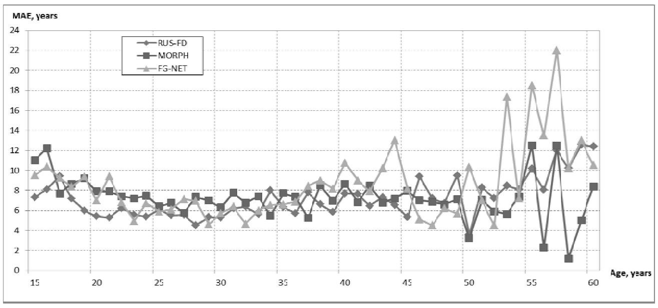

Fig. 8. MAE comparison on different databases for LBP-SVM algorithm

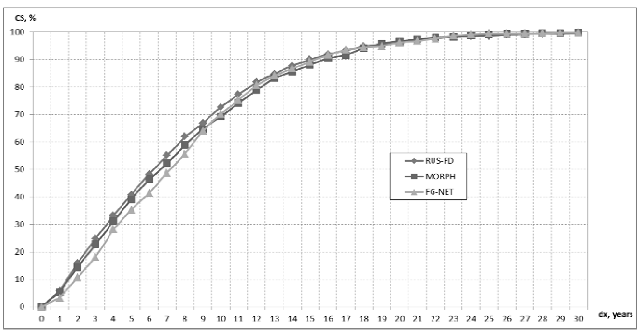

Fig. 9. CS comparison on different databases for LBP-SVM algorithm

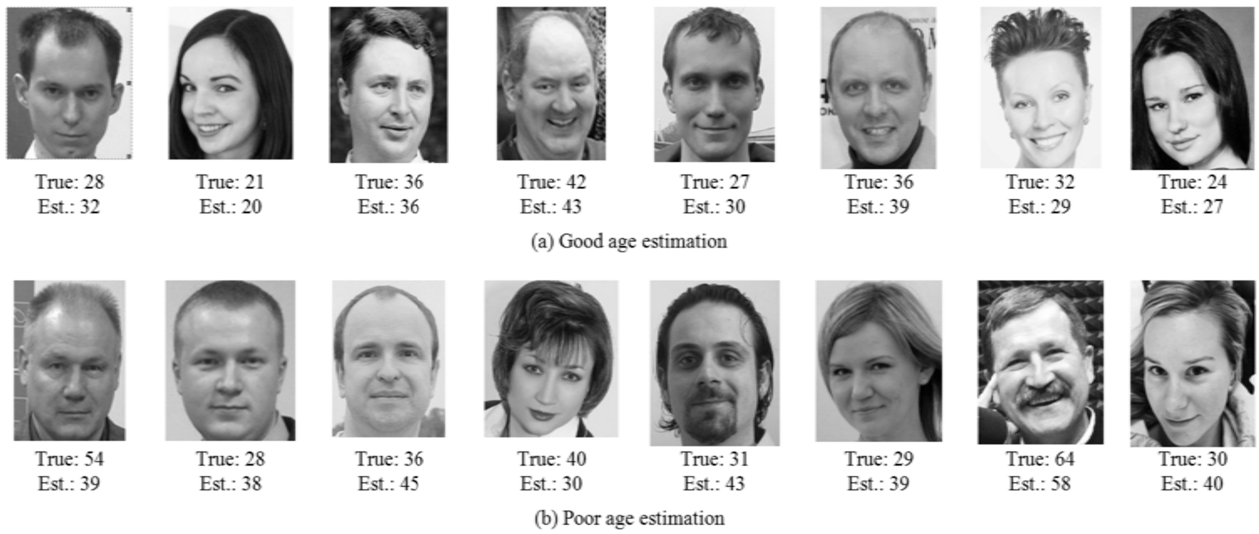

Fig. 10. Examples of age estimation using the proposed algorithm. (a) Good age estimation, (b) Poor age estimation

Total MAE score of LBP-SVM algorithm (learning on real-life dataset) on RUS-FD database is 6.94, MORTH database – 7.29, FG-NET database – 7.47. Subjective estimation MAE is 4.2 indicating that the proposed algorithm still needs much improvement to show results comparable to a human. It's important to notice that if we learn and test the LBP-SVM algorithm on MORTH database we have received total MAE score - 4,86, because the simplicity of MORTH database (the same light condition for all the faces, more than 1 face image for person). The possible ways to improve the accuracy of age classifier are feature set expansion (utilization of a combination of different feature transforms), cost-sensitive SVM learning procedure utilization, pre-processing and post-processing steps.

Examples of face images where LBP-SVM algorithm have good and poor age estimation are shown on Fig. 10.

IV. Algorithms comparison

We summarize published methods and results for age estimation from different face databases in Table II. As you can see from this review the most popular features type are – biologically inspired features (BIF) and it's modification, Gabor and local binary patterns and it's modification. Except MORTH and FG-NET databases in some papers algorithms also test on PCOS (Pinellas County Sheriff's Office), PAL, LFW+ (extended version of Labeled Faces in the Wild) and some private databases Top level age estimation algorithms can reach total MAE score between 4 and 5. CS in the same column reflects the percentage of correct age estimations within 5-year absolute error. In some papers researchers prefer to calculate MAE and CS separately for male and female sub-databases. Our age estimation algorithm provides world-quality results for MORTH database, but focused on real-life audience measurement application in which faces can be looks more or less similar to RUS-FD private database. In this case we can reach total MAE score less than 7.

V. Conclusion

This paper addresses the problem of automatic age estimation from real-life face images acquired. We have proposed an age estimation algorithm consisting of two stages: adaptive feature extraction based on local binary patterns and support vector machine classification. Experimental results on standard FG-NET, MORPH and our private RUS-FD face aging databases are presented. Human perception ability in age estimation is studied using crowdsourcing experiment which allows a comparison of the ability of machines and humans. Experimental results show that age estimation from unconstrained face images remains a very difficult problem. The age estimation algorithm described in this paper has integrated in audience measurement system which can collect and process the videodata in real time.

References

[1] S.Z. Li and A.K. Jain, *Handbook of face recognition, 2nd edition*. Springer-Verlag, 2011.

[2] R. Szeliski, *Computer Vision: Algorithms and Applications*. Springer, 2010.

[3] V. Khryashchev, A. Ganin, M. Golubev and L. Shmaglit, "Audience analysis system on the basis of face detection, tracking and classification techniques", *in Proceedings of the International MultiConference of Engineers and Computer Scientists,* Hong Kong, 2013, pp. 446-450.

[4] V. Pavlov, V. Khryashchev, E. Pavlov and L. Shmaglit, "Application for video analysis based on machine learning and computer vision algorithms", *in Proceedings of the 14th conference of open innovation association FRUCT*, Helsinki, 2013, pp. 90–100.

TABLE II. A COMPARISON OF PROPOSED METHOD WITH ANOTHER AGE ESTIMATION APPROACH

Publication	Features extraction	Using face aging database (#subjects, #images)	Performance measure and accuracy
Fu and Huang [14]	Holistic appearance	Private YGA (1600, 8000)	MAE: 5~6 CS Female:55% CS Male: 50%
Thukral et al. [19]	Landmark based hierarchical approach	FG-NET	MAE: 6.2
Han et al. [17]	Component and holistic biologically inspired features (BIF)	FG-NET MORTH II PCSO (1802, 10036)	FG-NET/MORTH II/PCOS MAE: 4.6 / 4.2 /5.1 CS: 74,8% / 72.4% / 64%
Geng et al. [24]	Holistic appearance, principal component analysis (PCA)	FG-NET MORTH	FG-NET/MORTH MAE: 6.8 / 8.8 CS: 65% / 46%
Suo et al. [25]	Holistic and local topology, 2D shape, color, and gradient	FG-NET Private (NA, 8000)	FG-NET/Private MAE: 6.0 / 4.7 CS: 55% / 66%
Guo et al. [26]	Holistic BIF	FG-NET Private YGA (1600, 8000)	MAE: 4.8 / F: 3.9, M: 3.5 CS: 47% / F: 75%, M: 80% MAE: 8.6
Choi et al. [27]	Holistic appearance, Gabor, LBP	FG-NET PAL (NA, 430) Private BERC (NA, 390)	FG-NET/PAL/BERC MAE: 4.7 / 4.3 /4.7 CS: 73% / 70% / 65%
Guo and Wang [28]	Holistic BIF, partial least squares (PLS)	PAL (590, 844) FACES (171, 1026)	PAL/FACES MAE: 6.1 / 8.1
Chao et al.[29]	Label-sensitive relevant component analysis	FG-NET	MAE: 4.4
Nguyen et al. [30]	LBP, multilevel LBP, Gabor	PAL (NA, 430)	MAE: 6.53
Han et al. [31]	BIF	Images of Groups LFW+ FG-NET	Images of Groups / LWF+ Age Group: 68.1%, 66.7% FG-NET MAE: 4.5
Shan [32]	LBP, Gabor	Images of Groups	Age Group: 55.9%
Ylioinas et al. [33]	LBP	Images of Groups	Age Group: 51.7%
Alnajar et al. [34]	Orientation histogram of local gradients	Images of Groups	Age Group: 56.5%
Proposed	LBP	Private RUS-FD, MORTH FG-NET	RUS-FD/ MORTH MAE: 6.94 / 4.86 CS: 42% / 56%

[5] D. Kriegman, M.H. Yang and N. Ahuja, “Detecting faces in images: a survey”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, No 1, 2002, pp. 34-58.

[6] E Hjelmas, “Face detection: a survey”, *Computer vision and image understanding,* vol. 83, No 3, 2001, pp. 236-274.

[7] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features”, *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2001, pp. 511-518.

[8] A. Yilmaz, O. Javed, and M. Shah, “Object tracking: a survey”, *ACM Comput. Surv*, vol. 38, 2006, article 13, 45 pages.

[9] D. Comaniciu, V. Ramesh, and P. Andmeer, “Kernel-based object tracking”, *IEEE Trans. Patt. Analys. Mach. Intell.* vol. 25, 2003, pp. 564-575.

[10] J. Shi and C. Tomasi, “Good features to track”, *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 1994, pp. 593-600.

[11] H. Tao, H. Sawhney, and R. Kumar, “Object tracking with bayesian estimation of dynamic layer representations”, *IEEE Trans. Patt. Analys. Mach. Intell.* vol. 24, 2002, pp. 75-89.

[12] B. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision”, *in Proceedings of Imaging Understanding Workshop*, 1981, pp. 121-130.

[13] E. Makinen and R. Raisamo, “An experimental comparison of gender classification methods”, *Pattern Recognition Letters*, vol. 29, No 10, 2008, pp. 1544-1556.

[14] Y. Fu and T.S. Huang, “Age synthesis and estimation via faces: a Survey”, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, No 11, 2010, pp. 1955-1976.

[15] M. Rhodes, “Age estimation of faces: a review”, *Applied Cognitive Psychology,* 23(1), 1995, pp. 38-59.

[16] M.C. Voelkle, U.L.N.C. Ebner and M. Riediger, “Let me guess how old you are: effects of age, gender, and facial expression on perception of age”, *Psychology and Aging*, 27(2), 2012, pp. 265-277.

[17] H. Han, C. Otto and A. Jain. “Age estimation from face images: Human vs machine performance”, *in Proceedings of*

the 6th IARP international conference on Biometric, Madrid, Spain, 2013.

[18] E.C. Sung, J.L. Youn, J.L. Sung, R.P., Kang and K. Jaihie, “A comparative study of local feature extraction for age estimation”, *IEEE Int. Conf. on Control Automation Robotics & Vision (ICARCV),* 2010, pp. 1280-1284.

[19] P. Thukral, K. Mitra and R. Chellappa, “A hierarchical approach for human age estimation”, *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 1529-1532.

[20] G. Guodong and M. Guowang, “Human age estimation: what is the influence across race and gender”, *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010, pp. 71-78.

[21] L. Zhen, F. Yun and T.S. Huang, “A robust framework for multiview age estimation”, *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010, pp. 9-16.

[22] G. Guodong and W. Xiaolong, “A study on human age estimation under facial expression changes”, *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 2547-2553.

[23] L.W. Hee, W. Jian-Gang, Y. Wei-Yun, L.C., Xing and P.T. Yap, “Effects of facial alignment for age estimation”, *IEEE Int. Conf. on Control Automation Robotics & Vision (ICARCV)*, 2010, pp. 644-647.

[24] X. Geng, Z.-H. Zhou and K. Smith-Miles, “Automatic age estimation based on facial aging patterns”, *IEEE Trans. PAMI,* 29(12), 2007, pp. 2234-2240.

[25] J. Suo, S.-C. Zhu, S. Shan and X. Chen, “A compositional and dynamic model for face aging”, *IEEE Trans. PAMI,* 32(3), 2010, pp. 385-401.

[26] G. Guo, G. Mu, Y. Fu and T.S. Huang, “Human age estimation using bio-inspired features”, *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2009, pp. 112-119.

[27] S. E. Choi, Y.J. Lee, S.J. Lee, K.R. Park and J. Kim, “Age estimation using hierarchical classifier based on global and local facial features”, *Pattern Recognition*, 44(6), 2011, pp. 1262-1281.

[28] G. Guo, and X. Wang, “A study on human age estimation under facial expression changes”, *in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2012, pp. 2547-2553.

[29] W.-L. Chao, J.-Z. Liu and J.-J. Ding, “Facial age estimation based on label-sensitive learning and age-oriented regression”, *Pattern Recognition*, 46(3), 2013, pp. 628-641.

[30] D.T. Nguyen, S.R. Cho, K.Y. Shin, J.W. Bang and K.R Park, “Comparison study of human age estimation with or without preclassification of gender and facial expression”, *The Scientific World Journal* , article ID 905269, 2014, 15 pages.

[31] H. Han and A. Jain, “Age, gender and race estimation from unconstrained face images”, *MSU Technical Report (2014): MSU-CSE-14-5.*

[32] 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,* 2010, pp. 23-28.

[33] J. Ylioinas, A. Hamid and M. Pietikäinen, “Age classification in unconstrained conditions using LBP variants”, *in Proceedings of 21st International Conference of Pattern Recognition (ICPR),* 2012, pp. 1257-1260.

[34] F. Alnajar, C. Shan, T. Gevers and T. Geusebroek, “Learning-based encoding with soft assignment for age estimation under unconstrained imaging conditions”, *Image Vision Comput.*, 30(12), 2012, pp. 946-953.

[35] K. Ricanek and T. Tesafaye, “MORPH: a longitudinal image database of normal adult age-progression”, *IEEE 7th Int. Conf. on Automatic Face and Gesture Recognition*, 2006, pp. 341-345.

[36] The FG-NET Aging Database, Web: http://fipa.cs.kit.edu.

[37] E. Alpaydin, *Introduction to Machine Learning*. The MIT Press, 2010.

[38] M. Pietikäinen, and T. Ojala, “Texture analysis in industrial applications”, *in Image Technology - Advances in Image Processing, Multimedia and Machine Vision, (Ed. J.L.C. Sanz),* Springer-Verlag, 1996, pp. 337-359.

[39] T. Mäenpää, “The local binary pattern approach to texture analysis - extensions and applications”, Dissertation, University of Oulu, 2003, 78 p + App.

[40] M. Pietikäinen, “Image analysis with local binary patterns”, *in Image Analysis, SCIA 2005 Proceedings, Lecture Notes in Computer Science 3540,* 2005, Springer, pp. 115-118.

[41] Da Bangyou, Sang Nong, “Local binary pattern based face recognition by estimation of facial distinctive information distribution”, *Optical Engineering,* vol. 48, Issue 11, 2009, pp. 117203-117203-7.