International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-6, Issue-5, May 2020 Pages 07-11

An Effective Age Estimation Method using Scattering Transform

Omaima FathElrahman Osman

Abstract— **With the increasing demand of automatic recognition and surveillance systems, research on human faces such as face recognition, face detection, facial expression recognition and gender classification, have attracted significant attention in both computer vision and pattern recognition. Compared to these faces related research, face age estimation is one of the popular research topics. It is an important technique that used in many real-world applications, such as security control and monitoring, soft biometrics and Electronic Customer Relationship Management (ECRM). In this paper a new method for age estimation was proposed. Where an efficient descriptor, scattering transform, which scatters the Gabor coefficients and pooled with Gaussian smoothing in multiple layers, is evaluated for facial feature extraction. Then support vector regression was used for age prediction. The experimental results show that scattering transform-based descriptor yielded a superior result on benchmark datasets FERET and PAL, with MAE 2.12 and 2.22 respectively.**

Index Terms— **Age estimation, scattering transform, texture feature and support vector regression.**

I. Introduction

Facial age estimation can be defined as: Label a face image automatically with the exact age (year) or the age group (year range) of the individual face. By extracting important features from faces of known ages, the age can be estimated for an individual face by solving the inverse problem using the same feature extraction technique [1].

Automatic age estimation from face image is one of the popular research topics that used in many real-world applications such as security control and monitoring, demographic analysis, Electronic Customer Relationship Management (ECRM), forensic art, social psychology, and entertainment [2].

In **Access Control and Surveillance Monitoring**, an accurate age estimation system can prevent children from entering to unauthorized places or websites, when a monitoring camera is used with the system, and prevent child and adult seniors from using danger games in a theme park. To control access in a cigarette vending machine, a Japanese company developed vending machine that check a user age by counting wrinkles and skin sags to prevent anyone under the legal age (20 years old) from buying a cigarette. Age estimation system also can be used in health care systems for customized services such as robotic nurse and intelligent intensive care unit [2]. Age estimation as soft biometrics can be used to improve the recognition accuracies in basic biometric system [3].

The effective marketing strategy which called **Electronic Customer Relationship Management (ECRM)** can benefit from age estimation systems by targeting specific customers in same age group for specific advertisements. It can provide important information for marketing study such as the number of young and adult customers who have visited a mall along with the preferences of each age group [4]. However, it is difficult to access this information while retaining privacy, but using automatic age estimation system can help to perform this task readily without violating the human privacy as the face image of a customer can be captured to identify age and delete the images instantly [2].

Recently, age estimation system is introduced as an application in some smart phones for **Entertainment** purposes such as iPhone Age Detector. When many photos and videos are captured, age information can organize the images to make them easy to access and find the related images when needed (Automatic album management). This also can be used for age-based image and video retrieval systems (information retrieval), where users could have the ability to restore their images by determining a wanted age-group in e-photo albums [4].

Age estimation also can be used with age synthesis in **Forensic Art** to find lost people. When the photos of missing children or any other family members are outdated, the system can predict the age and use age synthesis to produce updated face image [5].

Although there are several techniques that are used in a literature to solve age estimation problem, it is yet largely unsolved and remained a challenging issue, because it is affected mainly by human aging process, including different changes in human face. In addition to these normal changes, there are many external factors such as environmental influence, solar radiation, lifestyle, disease, drug use, and psychological stress that can have effect on face aging process and make it uncontrollable and personalized process [6].

Therefore, an efficient age estimation method is still needed. In this paper a new method was proposed to estimate the age from local face features, where Scattering Transform (ST) was used as a local descriptor.

ST include multiscale and multi-direction co-occurrence information. It is computed with a cascade of wavelet decompositions and complex modulus. This scattering representation is locally translation invariant and linearizes deformations [7].

Omaima FathElrahman Osman, Department of Computer Science, College of Computer Science and Information Technology, Sudan University of Science and Technology, Khartoum, Sudan

This paper is structured as follows: Section II provides a brief overview of previous works on age estimation; Section III outlines the proposed method; Section IV presents the results and discussion; finally, conclusions are summarized in Section V.

II. Related works

This section contains a review on face age estimation techniques. Most of these techniques composed of two phases: features extraction and age learning.

A. Features Extraction

The performance of age estimation system is affected mainly by features extracted from face images. Facial features can be classified as local, global, and hybrid.

Local features include the geometry of facial components including age spots, depth, wrinkles, and freckles. The first publication on face age estimation by Kwon and Lobo [8] using local features. They outlined a computational theory for age classification in which they used a combination of distance ratios between facial parts and wrinkle analysis to classify people into children, young adults, and senior adults. To differentiate baby faces from the two older groups, a set of distance ratios was computed between the face parts (mouth, nose, forehead, chin, and eyes). Snakelets analysis are used to find wrinkles to differentiate young adults from senior adults. They used only 47 face images in their experiment [8]. However, the anthropometric model proposed in [8] consider the geometry information only rather than texture information to be appropriate for identifying young age group. It is not useful for the identification of the adults because most changes pertaining to them appear as texture.

Ylioinas et al. [9] proposed an age estimation model using Kernel Local Binary Pattern (KLBP) to represent facial features. Firstly, the face was aligned using similarity transformation and the face pose was corrected using a flipping operation to evaluate the pose correction method. For feature extraction the method compared between a histogram and a kernel density estimate for LBP. Finally, a Support Vector Regression (SVR) was applied for age estimation. Ylioinas et al. [58] found that the kernel method was more powerful than the histogram using pose corrected images, which achieved 5.09 and 5.23 respectively for MAE. Using the original images, the MAE is 5.20 for kernel and 5.39 for histogram, which means that the pose correction enhanced the age estimation performance [9].

In 2013 Gunay et al. [10] also extracted local features to estimate the age using Radon transform. After the pre-processing step which included cropping, scaling and resize images to 88 × 88, The Radon transform was applied for an image matrix to compute the projection in specified direction. The result from this projection is that a new image which is a sum of the intensities of the pixels in each direction, includes more geometric information than the original image. However, the Radon projections produce a high dimension feature vector, so that the PCA was applied to reduce the dimension. Finally, multiple linear regression (MLR) was applied for age learning. Local radon feature achieved a better result on FG-NET (6.18) than MORPH (6.65) and FERET (6.98) [10], which means that the radon feature work better on images with lower resolution.

El-Dib and Onsi [11] investigated the ability of wrinkled appearance in eyes, the face parts excluding the forehead and the whole face to estimate the age. They found that the wrinkled appearance in eyes has a significant feature related to the age, when it was compared with the other parts. In 2015, Ng et al. [12] investigated the influence of wrinkles on specific age estimation. They localized and presented the wrinkles in a significative way to estimate the age. They proposed a multiscale aging patterns (MAP) method, which extract the features directly from local patches. MAP was generated by applying a multi-scale filter to different face parts and wrinkles, and thus transformed into meaningful aging patterns. Wrinkles appear in a different size. To present this feature, MAP extends a method that used dynamic scale filter called Hybrid Hessian Filter (HHF), which is a wrinkle detection method proposed by Ng et al. [65]. Based on multi-level analysis of Hessian eigenvalues, the local behavior of the face image is emphasized and wrinkles were identified subsequently. Then a multi-scale aging patterns were constructed from extracted wrinkles. Finally, supervised learning algorithm was applied to learn and predict the age. This work demonstrated the ability of wrinkles to estimate the age on high resolution images, where the MAE is 4.87 on FERET [12].

On the other hands, global features contain personal features such as gender, identity, ethnicity, and expression. Subspace learning techniques have been used extensively to extract global features. These include Locality Sensitive Discriminant Analysis (LSDA), Marginal Fisher Analysis (MFA) [13], PCA, Neighbourhood Preserving Projections (NPP), Locality Preserving Projections (LPP), and Orthogonal LPP (OLPP) [14]. However, Active Appearance Model (AAM) [15] is a most widely used method to extract global features for age estimation and to provide information about the shape and appearance of a face.

AAM is a statistical appearance models proposed by Cootes (2001) [15]. It combines shape and texture variation. Building a face model require face images marked with points to determine the main features. AAM model applied Procrustes analysis to align the sets of points (represented as a vector) and build a statistical shape model. It then warps images by the mean shape, and normalized texture vector by applying a linear transformation. To build a texture model eigen-analysis was applied. Finally, a combined appearance model was generated by learning the correlations between shape and texture. Although AAM considers shape and texture features, it produces a high dimensional feature vector; and this may contain an outlier in the age labelling space that may not directly reflect the identical age labels. Some valuable parts of the data (such as wrinkles) that represent discriminatory features might have been lost during dimensionality reduction process [2].

In 2016 Bukar et al. [16] proposed supervised Appearance Model (sAM) for age and gender estimation that improves the AAM by using the partial least-squares (PLS) as the core of the model rather than PCA. They claimed that them proposed model (sAM) effectively represents the face features than AAM, whereas the PLS preserves worthy parts of the data that represent discriminatory features. PLS creates latent features via a linear combination of the predictor variables and response (class labels). Because sAM considers both shape and texture features, it was used as feature extractor in age estimation model, whereas ordinary least-square (OLS) and quadratic function (QF) regressions were used for age learning with QF showing better results than OLS. The sAM model represents the faces in a convenient way to solve both age and gender classification problems, but it does not outperform the methods that used AAM.

In various face related applications, the local and global features were combined as hybrid features to offer superior performance. Choi et al. [17] introduced hybrid features representation and hierarchical classifier to increase the age estimation accuracy. For local features extraction, a set of region-specific Gabor filters were used to extract wrinkle features. LBP was used for skin features, and AAM was used for global features extraction. SVM and SVR were used for age classification and age regression. However, the combination of extracted features using these three methods (Gabor, LBP and AAM) produced a high dimension feature vector which can increase the computation time.

Gunay and Nabiyev [18], [19] conducted several experiments to estimate the age using different local and global methods to produce hybrid-based age estimation algorithms. In 2015 they proposed an age estimation method called "Global and Local feAture-based Age estiMation (GLAAM)" that used AAM to extract global features and 2D-DCT for local features. PCA was applied to reduce the dimensionality after concatenating the local and global features. Finally, the age was estimated with multiple linear regression [18]. In 2016 Gunay and Nabiyev [19] repeated the experiment using different methods for local features. They applied Gabor filters to extract wrinkles and LBP for skin features, and applied PCA to reduce the dimensionality. Based on three types of features that were extracted (global, wrinkles and skin), They modelled three aging functions separately with multiple linear regression, then they got the final decision by performing a decision level fusion to combine the results. They applied Gabor filter and LBP for local features to produce better result than 2D-DCT, where the MAE is 5.39 years for first experiment (AAM and 2D-DCT) and 4.87 years for second one (AAM, Gabor and LBP) [19].

B. *Age Learning Techniques*

Once the features are extracted from the face images, learning technique is needed to classify or predict the age. In general, age learning techniques can be classified into three groups: age groups classification, regression or single-level age estimation and hierarchical age estimation. Classification is an approach used to classify the age into multiple age groups such as babies, young, middle-aged adults, or old adults [1].

Classification methods based on conventional machine learning algorithms such as SVM, Linear Discriminant Analysis (LDA), Artificial Neural Networks (ANN), and nearest neighbour classifier were applied publicly for age group classification. Ng et al. [1] developed ANN with three layers: input, hidden and output to classify the age into two groups and applied their proposed method to investigate the effectiveness of wrinkles in age classification, which called Local Wrinkle-based Extractor (LOWEX) that used Canny edge detection to detect the wrinkles from different face regions. Their proposed method with ANN achieved 80% on FG-NET.

Regression approaches consider age label as a set of sequential numbers (e.g., 10, 11, 12,…). This sequence fits the ordinal nature of age label. This is a very popular age learning technique it includes many methods that can be used for regression such as Support Vector Regression (SVR), Quadratic Regression, Multiple Linear Regression, and Multilayer Perceptron (MLP). Ng et al. [20] applied SVR with their proposed feature descriptor Hybrid Aging Patterns to estimate the age. They achieved a MAE of 3.02 on FERET.

Hierarchical age estimation combines the classification and regression methods to predict age. It provides more accurate result and simplifies the computational load, where the data was firstly classified into age groups, then the age was predicted from these pre-defined groups. Choi et al. [17] applied age group classification followed by age regression to predict the age using SVM and SVR respectively. The hybrid features were extracted using AAM, LBP and Gabor. By using this combination of hybrid and hierarchical approach to predict the age, the MAE was 4.6 on FG-NET.

More recently, convolutional networks and deep learning approach have been successfully applied to several tasks related to facial analysis, including age estimation. Deep learning was introduced by Wang et al. [21] for a first time to solve age estimation problem. Their CNN architecture consists of three convolution layers, two pooling layers and full connection layer. Wang et al. [21] yield good results on FGNET and MORPH. However, this CNN was used to represent the features then traditional linear SVR was applied for age prediction. However, Due to the lack of face aging datasets, deep learning was used as features extractor (off the shelf CNN).

III. Proposed Methodology

The proposed method aims to estimate the age from local features in three steps: pre-processing step, features extraction using Scattering Transform (ST), and age prediction using SVR.

A. *Face Detection*

As in many face researches, the face was detected firstly as a pre-processing step. In this paper on-line Face++ detector was used for face detection. A total of 88 landmarks were detected using Face++ (as shown in Figure 1 (a)). Once the face was detected, a linear transformation is determined between each image and template (mean shape image) through the Procrustes analysis. In this work, the local transformation is defined as the projection of a shape S to a new shape 'S' using three landmark points: the center landmarks of eyes and mouth. Then, a warped image is generated by an affine geometric transformation [12] as shown in Figure 1 (b). the final images that used in experiment is shown in Figure 1 (c).

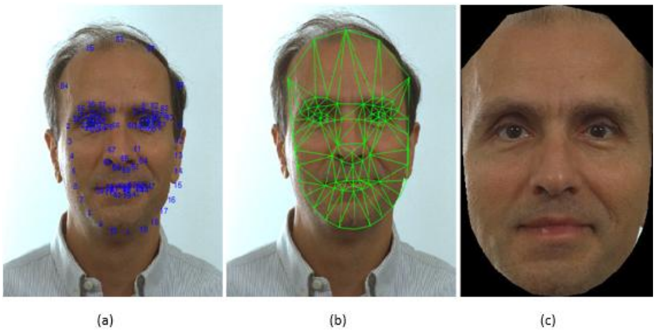

Figure 1: Face detection: (a) Original image obtained from FERET dataset; (b) Facial landmark detection using Face++; and (c) Warped image.

B. *Feature Extraction using Scattering Transform*

A scattering transform is a local descriptor introduced by Stephane Mallat [22]. It is computed with a cascade of wavelet decompositions and complex modulus. Scattering operators provide much richer descriptors of complex structures such as corners, junctions and multi-scale texture variations. These coefficients are locally translation invariant and they linearize small deformations. They are computed with a convolution network which cascades contractive wavelet transforms and modulus operators, as shown in Figure 2

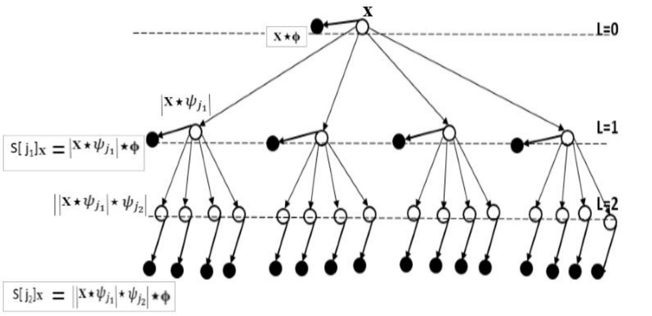

Figure 2: Scattering Transform Architecture.

The wavelet is very localized at high frequency (sensitive to translation). To build a wavelet coefficient that is an invariant to translation and stable to deformations, compute the average of the partial derivative amplitudes of "f" in the "K" directions ϒ ϵ Γ with a low-pass filter φj:

$$|f \star \psi_{j,\gamma}| \star \phi_J(x) \tag{1}$$

Which forms the first layer coefficients of scattering transform that shown in Figure 2. However, due to the averaging, this wavelet might lose some information that is essential to discriminate textures. Scattering operators recover part of the information that lost in averaging by convolutions with wavelets the complex phase is removed by a complex modulus then averaged, that is clearly shown in Figure 2 second layer, where there is only one operator that repeated iteratively. The output in this network at all layers (the output at each layer is demonstrated as a black node).

IV. RESULTS AND DISCUSSION

The performance of the proposed methods is tested on the FERET dataset [23] and PAL [24]. FERET is a comprehensive database with 2366 images of 994 subjects that presents multiple problems related to face recognition such as illumination variations, pose variations, facial expressions, etc. In addition, it consists of a few hundred age-separated face images of subjects with the age difference of 18 months or more and the age range is between 10 and 70 [23].

The Productive Aging Lab Face database (PAL) consist of 1,142 face images for 575 subjects in age range from 18 to 93 year, with neutral expression except some smiling expression. PAL dataset was developed to be more representative of age groups across the lifespan, with a special emphasis on recruiting older adults, which allow researchers interested in using facial stimuli access to a wider age range of adult faces than other datasets [24].

The performance metrics that used to evaluate the proposed method is Mean Absolute Error (MAE).

The MAE is defined as the average of the absolute errors between estimated age and the ground truth.

$$MAE = \frac{1}{X}\sum_{n=1}^{X} |a'_n - a_n| \tag{2}$$

TABLE 1: Comparison of MAE results on the FERET database.

Method	MAE
MAP [12]	4.87
Radon Transform [10]	6.89
FAM+BIF [25]	3.28
FAM-KLBP [25]	3.29
FAM+MWP [20]	3.02
Proposed method using ST	**2.12**

Aging is a complicated process, different subjects at same age may have a different feature. To represent a facial image as a feature vector that is insensitive to translation and small displacement but reactive to large deformation, scattering transform was used.

The scattering transform of an input signal "x" is defined as the set of all paths that "x" might take from layer to layer. In this sense, the architecture of a scattering network closely resembles a convolutional deep network as shown in Figure 2. Scattering coefficients provide new local descriptors, carrying co-occurrence information at different scales and orientations (4 scales and 8 orientations in this experiment). Despite the remarkable successes of deep convolution networks, the properties and optimal configurations of these networks are complex. ST network architecture is optimized to retain important information while avoiding useless computations. Table 1 and 2 shows that ST outperforms the state-of-the-art methods on FERET and PAL datasets.

TABLE 2: Comparison of MAE results on the PAL database

Method	MAE
AAM+LBP+Gabor [17]	4.32
FAM+MWP [20]	6.50
Proposed method using ST	**2.22**

V. CONCLUSION

Face age estimation is one of the popular research topics. It is an important technique that used in many real-world applications. In this paper a new and effective method was proposed for age estimation using scattering transform to extract the discriminative aging features. The experimental results showed that the proposed method yielded a superior result on FERET and PAL with MAE 2.12 and 2.22 years respectively.

REFERENCES

[1] C.-C. Ng, M. H. Yap, N. Costen, and B. Li, "An investigation on local wrinkle-based extractor of age estimation," in Computer Vision Theory and Applications (VISAPP), 2014 International Conference on, vol. 1. IEEE, 2014, pp. 675–681.
[2] 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.
[3] A. Dantcheva, P. Elia, and A. Ross, "What else does your biometric data reveal? a survey on soft biometrics," IEEE Transactions on Information Forensics and Security, vol. 11, no. 3, pp. 441–467, 2016.
[4] G. Guo, "Human age estimation and sex classification," in Video Analytics for Business Intelligence. Springer, 2012, pp. 101–131.
[5] K. Scherbaum, M. Sunkel, H.-P. Seidel, and V. Blanz, "Prediction of individual non-linear aging trajectories of faces," in Computer Graphics Forum, vol. 26, no. 3. Wiley Online Library, 2007, pp. 285–294.
[6] M. Albert, K. Ricanek, and E. Patterson, "A review of the literature on the aging adult skull and face: Implications for forensic science research and applications," Forensic Science International, vol. 172, no. 1, pp. 1–9, 2007.
[7] K.-Y. Chang and C.-S. Chen, "A learning framework for age rank estimation based on face images with scattering transform," IEEE Transactions on Image Processing, vol. 24, no. 3, pp. 785–798, 2015.
[8] Y. H. Kwon and N. da Vitoria Lobo, "Age classification from facial images," Computer Vision and Image Understanding, vol. 74, no. 1, pp. 1–21, 1999.
[9] J. Ylioinas, A. Hadid, X. Hong, and M. Pietikainen, "Age estimation using local binary pattern kernel density estimate," in International Conference on Image Analysis and Processing. Springer, 2013, pp. 141–150.
[10] A. G¨unay and V. V. Nabiyev, "Age estimation based on local radon features of facial images," in Computer and Information Sciences III. Springer, 2013, pp. 183–190.
[11] M. Y. El Dib and H. M. Onsi, "Human age estimation framework using different facial parts," Egyptian Informatics Journal, vol. 12, no. 1, pp. 53–59, 2011.
[12] C.-C. Ng, M. H. Yap, N. Costen, and B. Li, "Will wrinkle estimate the face age?" in Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. IEEE, 2015, pp. 2418–2423.
[13] G. Guo, G. Mu, Y. Fu, C. R. Dyer, and T. S. Huang, "A study on automatic age estimation using a large database," in ICCV, 2009, pp. 1986–1991.
[14] Y. Fu, Y. Xu, and T. S. Huang, "Estimating human age by manifold analysis of face pictures and regression on aging features," in Multimedia and Expo, 2007 IEEE International Conference on. IEEE, 2007, pp. 1383–1386.
[15] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," IEEE Transactions on Pattern Analysis & Machine Intelligence, no. 6, pp. 681–685, 2001.
[16] A. M. Bukar, H. Ugail, and D. Connah, "Automatic age and gender classification using supervised appearance model," Journal of Electronic Imaging, vol. 25, no. 6, pp. 061 605–061 605, 2016.
[17] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim, "Age estimation using a hierarchical classifier based on global and local facial features," Pattern Recognition, vol. 44, no. 6, pp. 1262–1281, 2011.
[18] G¨unay and V. V. Nabiyev, "Age estimation based on aam and 2d-dct features of facial images," International Journal of Computer Science and Applications, vol. 6, no. 2, 2015.
[19] Gunay and V. V. Nabiyev, "Age estimation based on hybrid features of facial images," in Information Sciences and Systems 2015. Springer, 2016, pp. 295–304.
[20] C.-C. Ng, M. H. Yap, Y.-T. Cheng, and G.-S. Hsu, "Hybrid ageing patterns for face age estimation," Image and Vision Computing, 2018, 69, pp.92-102
[21] X. Wang, R. Guo, and C. Kambhamettu, "Deeply-learned feature for age estimation," in 2015 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2015, pp. 534–541.
[22] J. Bruna, and S. Mallat "Classification with scattering operators." In CVPR 2011, pp. 1561-1566. IEEE, 2011.
[23] P. J. Phillips, P. J. Rauss, S. Z. Der et al., FERET (face recognition technology) recognition algorithm development and test results. Army Research Laboratory Adelphi, MD, 1996.
[24] M. Minear and D. C. Park, "A lifespan database of adult facial stimuli," Behavior Research Methods, Instruments, & Computers, vol. 36, no. 4, pp. 630–633, 2004.
[25] C.-C. Ng, Y.T Cheng, G.S Hsu, and M.H Yap. "Multi-layer age regression for face age estimation." In 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), pp. 294-297. IEEE, 2017.

Omaima FathElrahman Osman received her BSc degree in Computer and Information System from Sudan University of Science and Technology (SUST), Sudan in 2009. She received the MSc degree in Computer Science from Khartoum University in 2012. She is now a PhD student at SUST. Her research interests include face analysis and pattern recognition.