

An Averaging Technique for Improving Human Age Estimation Algorithms

Alireza Keshavarz Choobeh⁺

Young Researchers Club, Buinzahra Branch, Islamic Azad University, Buinzahra, Iran

Abstract. Accurate age estimation is one of the most important issues in human communication. It is essential part of human-computer interaction. In this paper, an averaging technique is proposed which is able to further improve several individual age estimation algorithms. Both mathematical and experimental proofs show, if the individual age estimators are diverse in error, then to improve the results, we can make the ensemble age estimator using the best selected individual age estimators.

Keywords. Age estimation, Averaging technique, Face images, Machine vision.

1. Introduction

Human face as a source of various important information plays significant role during communication between peoples. One of this information is the age. For having better interactions between peoples, it is necessary to estimate the age of peoples correctly. A survey of age estimation algorithms can be found in [1]. Age estimation has a large body of applications. Some applications that pass through the author's mind include the following.

- Preventing people to access on computer programs, internet websites and electronic devices that are harmful for specific ages.
- Indexing of face images based on the individual's ages for age-based retrieval of face images from videos.
- Summarizing videos focused on specific ages.
- Improving the accuracy and speed of face recognition systems.

In this paper has been shown that if the estimated age of the best individual age estimators that are diverse in error are averaged, the results can be improved. The reminder of the paper is organized as follows. Section 2 gives an overview of the system. Section 3 gives mathematical proof for efficiency of the proposed averaging technique. Sections 4 and 5 briefly reviews the Gabor filter and the modified census transform for generating appropriate texture features. The experimental results are presented in Section 6, and finally, Section 7 provides concluding remarks.

2. System Overview

The system diagram is shown in Fig. 1. The system inputs are facial shape and gray-level intensities of shape-free face images. The facial shape are represented

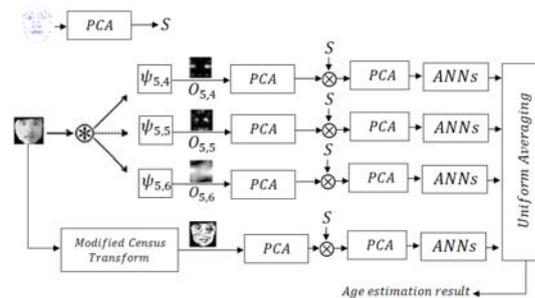


Fig. 1. Block diagram of the ensemble age estimator.

⁺ Corresponding author.
 E-mail address: Keshavarz_c@yahoo.com

by the coordinates of a number of landmarks placed at predefined positions of the face images. Before using the facial shapes, they are aligned by Procrustes analysis to minimize shape variation due to scaling, translation, and rotation [2], [3]. Then each image is warped to the mean shape using a piece-wise affine (linear) triangle based warping algorithm [4] so that shape variation within the training set is eliminated and obtaining shape-free faces. The Modified Census Transform and the best 2-D Gabor filters are used for extracting facial texture features from internal facial region. The shape and texture are both vectorized for image representation. Principal component analysis (PCA) [5], Artificial Neural Networks (ANNs) [6], and uniform averaging make the other parts of the system.

3. An Averaging Technique and Mathematical Proof for Its Efficiency

The averaging operation is to take a linearly weighted summation of the individual estimator outputs. The output of the combination is:

$$f_{ens(x_p)} = \sum_{i=1}^L w_i f_i(x_p) \quad (1)$$

where L is the number of estimators, f_i is the output of the i th estimator that is referred to as the individual estimator, w_i is the corresponding non-negative real-valued combination weight, x_p is a testing datapoint, and f_{ens} is a convex combination of the individual estimators that is referred to as the ensemble estimator. The weights can be non-uniform but have the constraint that they sum to one: $\sum_{i=1}^L w_i = 1$; and also the uniform combination in (2), where all weights is equal at $w_i = \frac{1}{L}$.

$$\bar{f} = \frac{1}{L} \sum_{i=1}^L f_i \quad (2)$$

If the ensemble is a uniformly weighted convex combination, we have the Bias-Variance-Covariance decomposition [7], [8]. So we have the mean squared error of the ensemble estimator \bar{f} as (3).

$$E\{(\bar{f} - d)^2\} = \overline{bias}^2 + \frac{1}{L} \overline{var} + (1 - \frac{1}{L}) \overline{covar} \quad (3)$$

where d is the target value of an arbitrary testing datapoint, $\overline{bias} = \frac{1}{L^2} \sum_{i=1}^L \sum_{p=1}^L (E\{f_i(x_p)\} - d(x_p))$, $\overline{var} = \frac{1}{L^2} \sum_{i=1}^L \sum_{p=1}^L E\{(f_i(x_p) - E\{f_i(x_p)\})^2\}$, $\overline{covar} = \frac{1}{L^3(L-1)} \sum_{i,j=1, i \neq j}^L \sum_{p=1}^L E\{f_i(x_p) - E\{f_i(x_p)\}\} (f_j(x_p) - E\{f_j(x_p)\})$.

The generalization error of an individual estimator can be broken down into two components: bias and variance. These two usually work in opposition to each other. Attempts to reduce the bias component will cause an increase in variance and vice versa [8]. Techniques in the machine learning literature are often evaluated on how well they can optimize the trade-off between these two components. The bias can be characterized as a measure of how close your estimator is to its target. The variance is a measure of how stable the solution is. For example, in a slightly different training data, an estimator with high variance will tend to produce wildly varying performance. The covariance term is a measure of error correlation between individual estimators. We can see that the error of an ensemble of estimators depends critically on the amount of covariance term. As (3) shows, we would ideally like to decrease the covariance, without causing any increases in the \overline{bias}^2 or \overline{var} terms. To achieve this goal, an averaging technique which is described in the following is proposed.

To improve the age estimation results by averaging technique in comparing to each individual age estimator, it is necessary that first, different representations of face images are provided to capture the error diversity between individual age estimators. Second, the best individual age estimators participate in averaging. The first decreases error covariance of ensemble estimator and the second decreases the error bias and variance of each individual age estimator. So based on the (3), the total error of ensemble estimator that comprises of these three terms decreases. From this point of view, different textural features from face images are evaluated. To generate textural features of face images, the total 64 Gabor filters and the Modified Census transform are used. The Gabor filter bank is selected based on minimum linear correlation between filters to reduce the redundancy of generated textural features. Selecting the Gabor filters based on the minimum linear correlation between filters is a factor to diversify the age estimation error between individual age estimators. Using each textural feature, a model of face appearance is built. The parameters of each model with Artificial Neural Networks (ANNs) classifier individually build the age estimators that have different errors in the age estimation results. Among the Gabor filters, some of them that have minimum age estimation errors are

selected. The uniform average of the age estimation results of the individual age estimator that built using best Gabor filters and Modified Census transform make the age estimation system that improves the age estimation performance compared to all the component individual age estimators.

4. Gabor Filter-Based Feature Extraction

The 2D Gabor filters take the form of a complex plane wave modulated by a Gaussian envelope function [9]:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{-(\|k_{\mu,v}\|^2 \|z\|^2 / 2\sigma^2)} \left[e^{ik_{\mu,v}z} - e^{-\frac{\sigma^2}{2}} \right] \quad (4)$$

where $k_{\mu,v} = \frac{k_{max}}{fv} e^{i\pi\mu/N}$, $z = (x, y)$, μ and v define the orientation and scale of the Gabor filters, k_{max} is the maximum frequency, and f is the spacing factor between kernels in the frequency domain. $\sigma = 2\pi$, $k_{max} = \pi/2$ and $f = \sqrt{2}$.

Gabor features of a facial image are extracted based on its Gabor representations, which are obtained by convolving the facial image with Gabor filter. Let $I(x,y)$ be an image, the convolution of image $I(x,y)$ and a Gabor filter is denoted as follows:

$$O_{\mu,v}(z) = I(z) \otimes \psi_{\mu,v}(z) \quad (5)$$

where $O_{\mu,v}(z)$ is called as the Gabor representation of facial image $I(z)$ at orientation μ and scale v . In this paper, for extracting discriminating information of different orientations and scales as much as possible, Wang et al scheme was adopted [8]. In this scheme, the orientation and scale set is $U = \{\mu: 0 \leq \mu \leq 7\}$ and $V = \{v: -2 \leq v \leq 5\}$ respectively. Note that the magnitude of Gabor filter response was used as the textural features.

5. The Modified Census Transform

The modified census transform [10] is a non-parametric local transform which is defined as an ordered set of comparison of pixel intensities in a local neighborhood. Let $N(x)$ and $\bar{I}(x)$ are a local spatial neighborhood of the pixel at x including it and the intensity mean on this neighborhood respectively. The modified census transform at x is defined using (6).

$$\Gamma(x) = \bigotimes_{y \in N} \xi(\bar{I}(x), I(y)) \quad (6)$$

here \otimes is the concatenation operation and $\xi(I(x), I(x'))$ is a comparison function as shown using (7).

$$\xi(I(x), I(x')) = \begin{cases} 1 & I(x) < I(x') \\ 0 & \text{Otherwise} \end{cases} \quad (7)$$

6. Experimental Results

To study of the automatic age estimation system, I used the FG-NET Aging Database [11]. The age estimation performance are evaluated by the Mean Absolute Error (MAE) [1].

Sixty-eight landmark points and 18000 pixels from internal region of shape-free faces are used. In total, 130 textural features were applied to each of the ANNs. The ANNs architecture is as follows: 130 input units, 5 hidden units, which have been experimentally determined, and one output unit, representing the corresponding age of each face. All neurons have linear threshold functions.

Among the individual age estimators based on the Gabor filters, the best 2D Gabor filters ($(\mu, v) = \{(5,4), (5,5), (5,6)\}$) achieves the mean absolute error of 5.82, 5.72, and 5.85 years respectively [12]. The MAE of the individual age estimator based on the Modified Census Transform is 5.49 years [13]. The ensemble age estimator improves the age estimation MAE compared to the individual age estimators up to 4.85 years.

7. Summaries

An ensemble age estimation algorithm which can improve individual age estimation algorithms is proposed. In the first step, several individual age estimators which are diverse in error are created. Different representation of face images using 2D Gabor filters and the Modified Census Transform are used to create diversity in error between the component individual age estimators. Among them, the four individual age estimators that have the best performance are selected. The four individual age estimators use a same classification model on different feature representations of same facial images. Averaging of the age

estimation results of these best age estimators gives us an ensemble age estimator that has the better performance compared to all the individual estimators. Mathematical proofs guarantees further improvement of the ensemble age estimation algorithm when the best individual age estimators are diverse in error.

8. References

- [1] Y. Fu; G. Guo; T.S. Huang. Age Synthesis and Estimation via Faces: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2010, pp: 1955-1976.
- [2] C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society B*, 1991, PP: 285-339.
- [3] J. C. Gower. Generalized Procrustes analysis. *Psychometrika* 40, 1975, PP: 33-51.
- [4] A. Goshtasby. Piecewise linear mapping functions for image registration. *Pattern Recognition*, 1986, PP:459-466.
- [5] I.T. Jolliffe. *Principal Component Analysis*, second ed. Springer-Verlag, 2002.
- [6] M. Riedmiller; H. Braun. A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. *Proc. IEEE Int'l Conf. Neural Networks (ICNN '93)*,1993,PP:568-591.
- [7] N. Ueda; R. Nakano. Generalization error of ensemble estimators. *Int. Conf. on Neural Networks, ICNN96*, 1996, PP: 90-95.
- [8] G. Brown. Diversity in Neural Network Ensembles. Ph.D. Thesis, University of Birmingham, 2003.9.
- [9] Wang L; Li Y; Wang C. 2D Gaborface representation method for face recognition with ensemble and multichannel model. *Image and Vision Computing*, 2008, 26, PP: 820-828.
- [10] B. Froba; A. Ernst. Face detection with the modified census transform. *Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, 2004.5, PP: 91-96.
- [11] The FG-NET Aging Database. [http://webmail.cycollege.ac.cy/~alanitis /tmp/](http://webmail.cycollege.ac.cy/~alanitis/tmp/), 2007.
- [12] A.K. Choobeh. Human Age Estimation Using 2D Gabor Filter. *International Conference on Computer Technology and Science (ICCTS)*. Forthcoming 2011.7.
- [13] A.K. Choobeh; M.J.P. Jalali. Automatic Human Age Estimation Based on Neural Networks and the Modified Face Model. *3rd International Conference on Software Technology and Engineering (ICSTE)*. 2011.7, PP: 111-115.