

The Third Information Systems International Conference

Human Facial Age Classification Using Active Shape Model, Geometrical Feature, and Support Vector Machine on Early Growth Stage

Yosi Kristian^{a,b}, Endang Setyati^{a,b}, Yuliana Melita Pranoto^a, Arya Tandy Hermawan^a, Michael Putra Wijaya^{a*}

^aDepartment of Information Technology Sekolah Tinggi Teknik Surabaya, Surabaya 60284, Indonesia

^bDepartment of Electrical Engineering Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

Abstract

Human face is an integral part for delivering an amount of nonverbal information to facilitate communication. One of the important is human age, but accurately extracting human age from his face is not easy. Facial aging process is different on early growth (child to adult) and adult aging (adult to senior). On early growth and development of the face, from birth to adulthood, the greatest change is the shape change. But on adult aging the most significant change is on the skin (textural change), the shape change still continue but much less dramatic compared to early growth. In our research we develop a two stage facial age classification system to classify human age into age classes. This paper is the first part of our two stage facial age classification. In this paper we develop facial age classification system for the early growth stage, we classify human age into six classes which are 0-2, 3-4, 5-6, 7-10, 11-16, and 17+. Because the early growth have grater impact on shape change, we only use geometrical feature to classify the human face. We use Haar-like cascade to perform the face detection and then we use Active Shape Model to extract the geometrical features on human face. For the classifier we use Support Vector Machine(SVM) classifier with Radial Basis Function (RBF) kernel. This system is developed with OpenCV library and C++ language. The results of classification based on geometrical features achieved an accuracy rate of 71.25% using FG-NET dataset.

© 2015 Published by ISICO and/or peer-review under responsibility of Institut Teknologi Sepuluh Nopember

Keywords: Facial Ace Classification, Active Shape Model, Support Vector Machine, Geometrical Feature

* Corresponding author.

E-mail address: yosik81@gmail.com.

1. Introduction

Modern systems are expected to have an ability to accurately recognize and interpret human face. An example of system that can be used in the real world is age estimation from human face. According to [1] facial age estimation is a process to label a face image automatically with the exact age (year) or the age group (year range) of the individual face. By using facial age estimation system, computer can predict the age of a person through the person face. It can be used for a computer to perform a specific reaction of people with a particular age group. For example, if the age estimation is applied to the vending machine, the vending machine can sell goods such as cigarettes or liquor that only allow adults to buy [2]. For on-line advertising firm, they can choose which commercial to display based on the age of the viewer [3]. For adult content provider, they can enforce an age filtering that restrict minor watching their content, and still many more use full example for age estimation system.

Regardless of its great use, estimating human age based on facial feature is still a hard task. Each human age differently beside based on gene, external factor like their lifestyle, illness, environment, weather, and location also contribute to the aging process [3]. How to extract general discriminative aging features while reducing the negative influence of individual differences still remains an open problem.

On [1] they distinguish aging process into two stage, early growth and adult aging. During the early growth and development of the face, from birth to adulthood, the greatest change is the craniofacial growth (shape change). This stage is between born until 17 years old. With the growth of the cranium, the forehead slopes back, shrinks, and releases spaces on the surface of the cranium, while the facial features, such as eyes, nose, ears, and mouth, expand their areas and tend to cover these interstitial spaces. Cheeks extend to larger areas and the chin becomes more protrusive. The facial skin relatively does not change too much compared with the craniofacial growth. Skin color may also change a little bit. During adult aging, from adulthood to old age, the most perceptible change becomes skin aging (texture change). The shape change still continues, but less dramatically, mostly due to typical patterns in skin and tissue.

There are several research trying to estimate age from human facial feature, for example [4] which is focused on non linear characteristic of human aging pattern, [5] focusing on label distribution, and [6] which examine textural and geometrical feature to conduct an age estimation. Age estimation can be considered a regression problem if we tried to predict the exact age of a person like in [3], but it can also be considered a classification problem if we use ranges to estimate person age like in [7].

In [8] they use a hierarchical age classification system that we are also trying to accomplish. We use different geometrical feature, different textural feature, different age group and different classification technique.

2. Our Age Estimation System

Based on two stage of aging process, we design our system to conduct two stage of age estimation. Then we choose classification because aging process are very different between person so classify on a range is much more reasonable than specific estimation. On Early growth stage we use geometrical feature to estimate age, and for adult aging we use textural feature. Our process flow can be seen on Fig. 1. We start with an image, then detect the face region using Viola Jones Haar Like Cascade [9], we continue by fitting a previously trained Active shape Model [10]. We only use geometrical feature to perform the first stage age classification because the early growth aging process mainly concentrated on shape change. In this paper we only discuss on the first stage of this integrated system. We divide our first age class into 6 group: 0-2, 3-4, 5-6, 7-10, 11-16, and 17+ this group is simple modification of Kilinc age group on [11]. In the next 4 sub chapter we will discuss the 4 sub process for our first stage age classification.

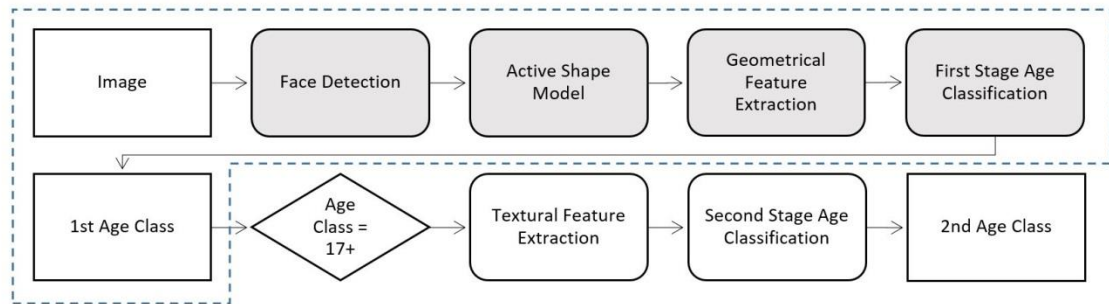


Fig. 1. Infant Face Detection

2.1. Face Detection

Viola Jones haar-like cascade from [12] and [9] is one of the well known methods for object detection. This method is very popular and have high speed and high accuracy. The name comes from intuitive similarity to haar wavelets and it is commonly used in the real-time face detector. The word "cascade" in the classifier name means that the resultant classifier consists of several simpler classifiers (stages) that are applied subsequently to a region of interest until at some stage the candidate is rejected or all the stages are passed. [13]

The term "haar-like" origins from the calculation of Viola and Jones which similar with haar wavelet transform method. Basically, haar-like cascade working with changing a region near an image into pixel area based on its classifier then calculate the intensity difference of each pixel area. The resulting difference is used to categorize each area for an object detection. [12]

A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. Basically, haar-like features working with changing a region near an image into pixel area based on its classifier. The next step will be to calculate the intensity difference of each pixel area. The resulting difference is used to categorize each area for an object detection. Example of Haar-like feature implementation for face detection can be seen on Fig. 2.



Fig. 2. Face Detection

2.2. Dataset

In this research we use FG-NET dataset, a public dataset contain of 1002 personal photographs of 82 subjects with age ranged between 0 to 69. This dataset contain of mostly frontal face but some are quite tilted. Few example of images in FG-NET dataset can be seen on fig 3.



Fig. 3. FG-NET Dataset Example

2.3. Facial Active Shape Model

Active Shape Model (ASM) [10][14] is a method where the model is iteratively changed to match that model into a model in the image. This method is using a flexible model which acquired from a number of training data sample. Given a guessed position in a picture, ASM iteratively will be matched with the image. By choosing a set of shape parameter b for Point Distribution Model, the shape of the model can be defined in a coordinate frame which centered in the object. Instance X in the images model can be constructed by defining position, orientation, and scale. The ASM starts the search for landmarks from the mean shape aligned to the position and size of the face determined by previous face detector. It then repeats the following two steps until convergence:

1. Suggest a tentative shape by adjusting the locations of shape points by template matching of the image texture around each point.
2. Conform the tentative shape to a global shape model. The individual template matches are unreliable and the shape model pools the results of the weak template matchers to form a stronger overall classifier.

On our previous research [15] we also use Facial ASM to extract geometrical feature. In this research we trained our ASM using FG-NET dataset with 68 landmark point. The result for frontal face picture can be seen on the first 4 image on Fig 4 shows good point fitting. But some tilted head pose resulting in missed ASM fitting, can be seen on the last 2 on Fig. 4.

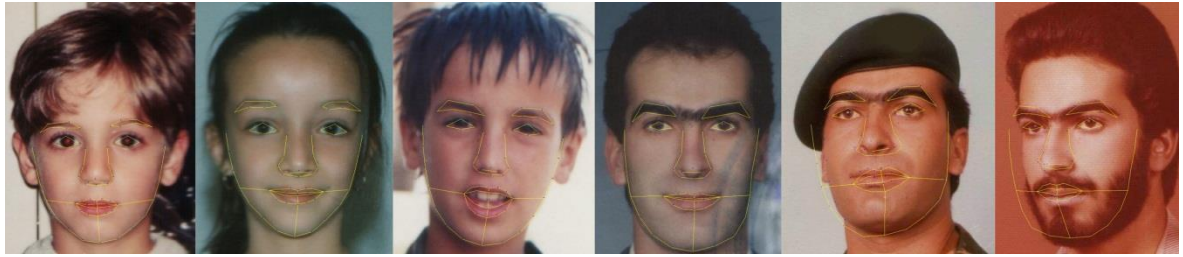


Fig 4. Succesful Facial Active Shape Model

2.4. Geometrical Feature Extraction

Extracting geometrical feature from ASM landmark points is quite straight forward. We extracted 49 features from ASM landmark points. These 49 feature consist of:

1. Five features from distance between face borders on the left and right, see Fig. 5 top row first column.
2. Fifteen features from distance between tip of the nose to all face border landmarks, see Fig. 5 top row second column.
3. Five features from distance between center of mouth to chin landmarks, see Fig. 5 top row third column.
4. Five features from distance between tip of the nose to all face border landmarks, see Fig. 5 top row third column.
5. Six features from distance between tip of the nose to all nose landmarks, see Fig. 5 top row last column.
6. Six features from distance between eyes to lip center landmark, see Fig. 5 bottom row first column.
7. Six features from distance between eyes to tip of the nose landmark, see Fig. 5 bottom row second column.
8. Three features from distance between tip of the nose to upper lip landmark, see Fig. 5 bottom row third column.
9. Three features from distance between eyes landmark, see Fig. 5 bottom row last column.

The distances are then normalized by dividing them with the face width to form the geometrical features for the face. These features are then trained to a machine learning, to form the classifier for our facial age estimation.



Fig. 5. Geometrical Feature Extraction

Table 1. Various Kernel SVM Accuracy

Kernel	Accuracy
Linear	59.38%
Sigmoid	60.27%
Radial Basis	71.25%

2.5. SVM Classifier

To classify the features discovered previously, we use Support Vector Machine (SVM) with Sequential Minimal Optimization [16]. We choose SVM because of its large margin characteristic, with it we can achieve better classifier for our problem. The other strong characteristic of SVM is application of kernel trick. By using kernel we change the dimension of our feature to distance measurement between data and landmark points using a specific function. This function is what we called kernel. SVM works really well with kernel trick, but not all classifier can achieve same advantage as SVM in using kernel. In this research we divide our dataset into training set (700 data) and testing set (302 data). We also tried 3 kernel variation: linear, sigmoid and radial basis.

3. Experimental Result

The result of 3 kernel variation to classify FG-NET dataset can be seen on Table 1. The result clearly shows that radial basis kernel perform much better than linear and sigmoid kernel on this age classification system. Table 2 shows confusion matrix for SVM with radial basis kernel. Example of age correctly estimated are shown on top row Fig. 6 and the incorrectly classified are shown on bottom row Fig. 6. The miss classified is mainly caused by failed ASM to fit correctly to the face, and the miss fitting is caused by non frontal or tilted head pose, also by poor lighting and blurry image of the dataset. Facial hair also contribute in degrading our accuracy.

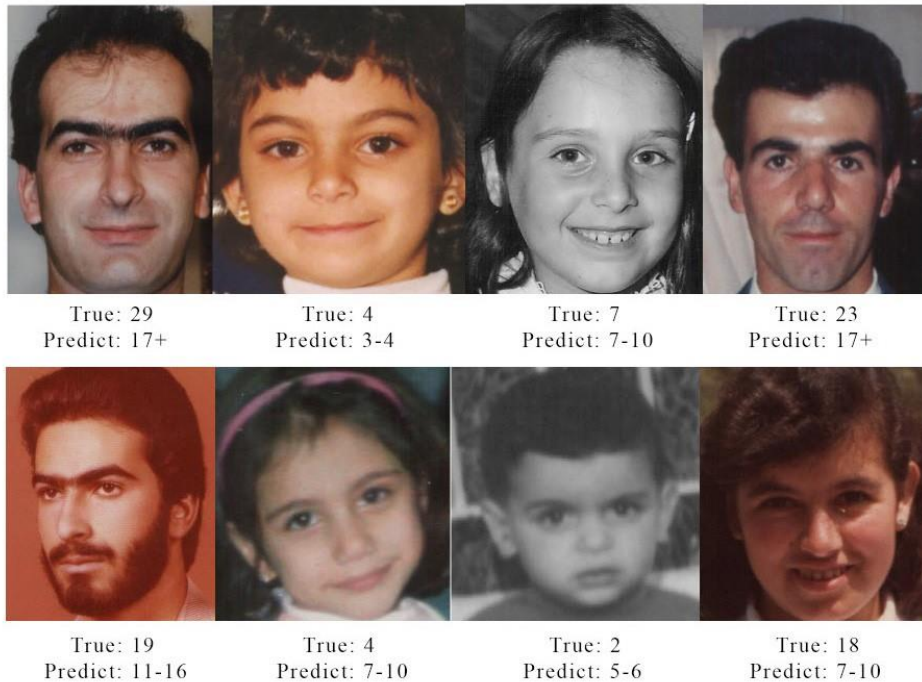


Fig. 6. Correctly Classified Age (Top Row) and Incorrectly Classified Age (Bottom Row)

Table 2. My Caption

Age Prediction	Real Age Class					
	0-2	3-4	5-6	7-10	11-16	17+
0-2	82	14	6	6	3	1
3-4	10	58	16	13	5	2
5-6	1	6	13	4	2	3
7-10	8	9	17	81	26	11
11-16	0	3	3	18	65	22
17+	2	8	8	23	38	345

4. Conclusion and Future Work

The combination of geometrical feature with SVM could perform well to estimate human age. The most significant factor that affecting our classifier accuracy is tilted head pose, blurry image, facial hair, and inconsistent lighting. In FG-NET dataset there are about 20% pictures that are having tilted head pose, blurry or low quality picture, and also inconsistent lighting.

Currently we are working on the second stage classification engine which is to classify the adult aging stage. Because the adult aging stage impact mostly on skin textures we are using textural feature on specific area such as below the eyes and sides of the mouth. After we finished the second stage classifier

we will combine these two classifiers into a hierarchical classifier that can predict age class 0-2, 3-4, 5-6, 7-10, 11-16, 17-28, 29-39, 40-54, 55+.

Acknowledgements

This research is funded by Direktorat Penelitian dan Pengabdian kepada Masyarakat, Ministry of Research, Tech-nology and Higher Education through Hibah Bersaing program of 2015.

References

- [1] Y. Fu, G. Guo, T. S. Huang, Age synthesis and estimation via faces: A survey, *Pattern Analysis and Machine Intelligence*, IEEE Transactions on 32 (11) (2010) 1955–1976.
- [2] X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, H. Dai, Learning from facial aging patterns for automatic age estimation, in: *Proceedings of the 14th annual ACM international conference on Multimedia*, ACM, 2006, pp. 307–316.
- [3] G. Guo, Y. Fu, C. R. Dyer, T. S. Huang, Image-based human age estimation by manifold learning and locally adjusted robust regression, *Image Processing*, IEEE Transactions on 17 (7) (2008) 1178–1188.
- [4] X. Geng, K. Smith-Miles, Z.-H. Zhou, Facial age estimation by nonlinear aging pattern subspace, in: *Proceedings of the 16th ACM international conference on Multimedia*, ACM, 2008, pp. 721–724.
- [5] X. Geng, C. Yin, Z.-H. Zhou, Facial age estimation by learning from label distributions, *Pattern Analysis and Machine Intelligence*, IEEE Transactions on 35 (10) (2013) 2401–2412.
- [6] M. Kilinc, Y.S. Akgul, Human age estimation via geometric and textural features., in: *VISAPP* (1), 2012, pp. 531–538.
- [7] N. Hewahi, A. Olwan, N. Tubeel, S. El-Asar, Z. Abu-Sultan, Age estimation based on neural networks using face features, *Journal of Emerging Trends in Computing and Information Sciences* 1 (2) (2010) 61–67.
- [8] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, J. Kim, Age estimation using a hierarchical classifier based on global and local facial features, *Pattern Recognition* 44 (6) (2011) 1262–1281.
- [9] P. Viola, M. J. Jones, Robust real-time face detection, *International journal of computer vision* 57 (2) (2004) 137–154.
- [10] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Active shape models-their training and application, *Computer vision and image understanding* 61 (1) (1995) 38–59.
- [11] M. Kilinc, Y. S. Akgul, Automatic human age estimation using overlapped age groups, in: *Computer Vision, Imaging and Computer Graphics. Theory and Application*, Springer, 2013, pp. 313–325.
- [12] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Vol. 1, IEEE, 2001, pp. I–511.
- [13] E. Setyati, Y. Kristian, Y. M. Pranoto, D. Alexandre, Augmented reality 3d eyeglasses frame simulator using active shape model and real time face tracking, in: *ICT and Knowledge Engineering (ICT&KE)*, 2013 11th International Conference on, IEEE, 2013, pp. 1–5.
- [14] T. F. Cootes, C. J. Taylor, Active shape models-smart snakes, in: *BMVC92*, Springer, 1992, pp. 266–275.
- [15] Y. Kristian, M. Hariadi, M. H. Purnomo, Ideal modified adachi chaotic neural networks and active shape model for infant facial cry detection on still image, in: *Neural Networks (IJCNN)*, 2014 International Joint Conference on, IEEE, 2014, pp. 2783–2787.
- [16] J. Platt, et al., Fast training of support vector machines using sequential minimal optimization, *Advances in kernel methodssupport vector learning* 3.