An Image Mining System for Gender Classification & Age Prediction Based on Facial Features

Dhanashree M. Shirkey, S.R.Gupta

Abstract:- The face recognition system with large sets of training sets for personal identification normally attains good accuracy. In the project, we proposed algorithm for Feature Extraction based Face Recognition, Gender and Age Classification with only small training sets and it yields good results even with one image per person. This process involves three stages: Pre-processing, Feature Extraction and Classification. The geometric features of facial images like eyes, nose, mouth etc. are located by using Feature extraction algorithm and face recognition is performed. Based on the texture and shape information, gender and age classification is done by comparing histogram of the query image and the histogram of the images in dataset respectively. By using the proposed work , ratio of 100% for face matching, 90% for gender classification ,and 85% for age classification can be achieved .

Keywords:
Face Detection, Skin Color Segmentation, Face Features extraction, Features recognization,Fuzzy rules,Histogram,Image mining

1. *Introduction*

Human's faces reveal various information including gender, age and ethnicity. They provide important cues for many applications, such as biometric authentication and intelligent human-computer interface. In this paper, we present a new method that can identify humans' genders from their face images. In the past, many researches devote to finding good image features for gender recognition. Among them, Adaboost [1] [2] [3] [4] is a good tool for feature selection. There are many gender-recognition algorithms constructed based on AdaBoost. Wu and Ai [5] proposed a boosting method based on the look-up-table (LUT) weak classifiers. In their work, they train three different detectors (gender, ethnicity, and age) to acquire demographic information by using the two-class Real AdaBoost algorithm [2].

Manuscript Received on May 2013.

Ms. Dhanashree M. Shirkey, M.E(Scholar), Department of CSE, PRMITR, Badnera, India

Prof.S.R.Gupta, Asstt.Professor, Department of CSE,PRMITR, Badnera, India

Shakhnarovich and Viola [6] used Viola and Jones' cascaded AdaBoost method [1] to train a face detector, which is a linear combination of the weak classifiers selected from rectangle features. They then selected a set of weak classifiers from the face detector for gender and ethnicity recognition. Instead of rectangle features, Shummet and Henry [7] designed a new set of weak classifiers which use the relationship between two pixels' intensities as features, and show that the recognition rate can be further improved. Instead of AdaBoost, a high. However, the computational loads of these two approaches are high. They are thus not suitable for real-time applications. Among the above, the Shakhnarovich and Viola's method [6] is probably the most efficient one in terms of the computational cost for real applications. It is because that some of the rectangle features used for face detection are re-used for gender recognition. Hence, in this method, we do not need to recomputed the features once the face has been detected, and so the entire computational time (including both face detection and gender recognition) can be reduced. In addition, it is also well known that the evaluation of rectangle features can be considerably speeded up by using the integral-image technique [1]. However, the recognition rate of [6] still needs a considerable improvement. In this paper, we develop a fast gender recognition algorithm based on rectangle features too. Rectangle features can be used to describe sub-regions of a human face, and hence pixel-wise data can be transformed into component wise data. We cascade several rectangle features into a feature vector. The features are then served as a descriptor for faces to identify the gender. We did a comparative study by employing features in various kinds of classifiers, such as the nearest-neighbor (NN), principle component analysis (PCA), and nonlinear SVMs, and suggest an effective detector for Gender recognition. Unlike the method of Shakhnarovich and Viola [6], we do not restrict ourselves to select the rectangle features only from those used to construct the face detector. In other words, we allow the rectangle features to be selected arbitrarily in a large feature pool. In this way, the rectangle features selection can be more discriminative, and hence our approach is more accurate for gender recognition. Although in our approach, the rectangle features used for gender recognition and face detection may not be the same, they still share the same integral image. Hence, we can still re-use the integral image, originally built for face detection, to efficiently evaluate the rectangle features for gender recognition. The total computational time (including both face detection and gender recognition) can thus be saved as well

2. *Literature Survey*

Identifying the age and gender information of a face is a challenging task and has gained significant attention recently. Metze et al. (2007) compared four approaches for age and gender recognition from telephone speech; these included a parallel phoneme recognizer system to compare the Viterbi decoding scores for each category-specific phoneme recognizer.The discrete cosine transform is applied to the cepstral coefficients and the cepstral trajectories corresponding to lower (3-14 Hz) modulation frequencies provide best discrimination. Prosodic features (pitch, energy, formants, vocal tract length warping factor, speaking rate, etc) and their functional can also be added to the cepstral features at the frame or utterance level to enhance the performance (Spiegl et al., 2009; Meinedo and Trancoso, 2010; Gajˇsek et al., 2010; Eyben et al., 2009; Wolters et al., 2009). In addition to the prosodic features, novel lexical level features like word-class frequencies have also been proposed for age recognition purpose (Wolters et al., 2009). In the fuzzy SVM modeling method proposed by Nguyen et al. (2010), a fuzzy membership is assigned as a weight to each training data point to increase the robustness against noise and outliers. Furthermore, techniques from speaker verification and language identification applications such as GMM-SVM mean super vector systems (Bocklet et al., 2008), nuisance attribute projection (NAP) (Dobry et al., 2009), anchor models (Dobry et al., 2009; Kockmann et al., 2010) and Maximum-Mutual-Information (MMI) training (Kockmann et al., 2010) have been successfully applied to speaker age and gender identification tasks to enhance the performance of acoustic level modeling. In Dobry et al.(2009), anchor modeling utilizes a back end SVM to model the distribution of similarity scores between training data and all the anchor speaker models.

3. *Proposed Algorithm*

3.1 Gender Recognization

1. Input an Image
2. Detect skin area in Input Image.
3. Detect Features like eyes and mouth in skin region.
4. If Features detected then go to step 5 else step 1.
5. Crop Face.
6. Load Database Male Females features.
7. Locate Features in a face area.
8. Count Male & female Features.
9. Filter Counted features into strong & weak features.
10. Form Fuzzy Conclusion from features & Display gender result.

3.2 Age Prediction

3.2.1 Training

1. Select an Input Image.
2. Detect skin area in Input Image.
3. Detect Features like eyes and mouth in skin region.
4. If Features detected then go to step 5 else step
5. Crop Face.
6. Save Face into Database with its age.
7. Repeat step 1 to 6 for 100 images(Training Images)

3.2.2 Testing

1. Select an Input Image.
2. Detect skin area in Input Image.
3. Detect Features like eyes and mouth in skin region.
4. If Features detected then go to step 5 else step
5. Crop Face.
6. Load faces Images from training directory & Match with input face image using histogram matched approach.
7. Retrieve Match image age from database.
8. Display Result.
9. Stop

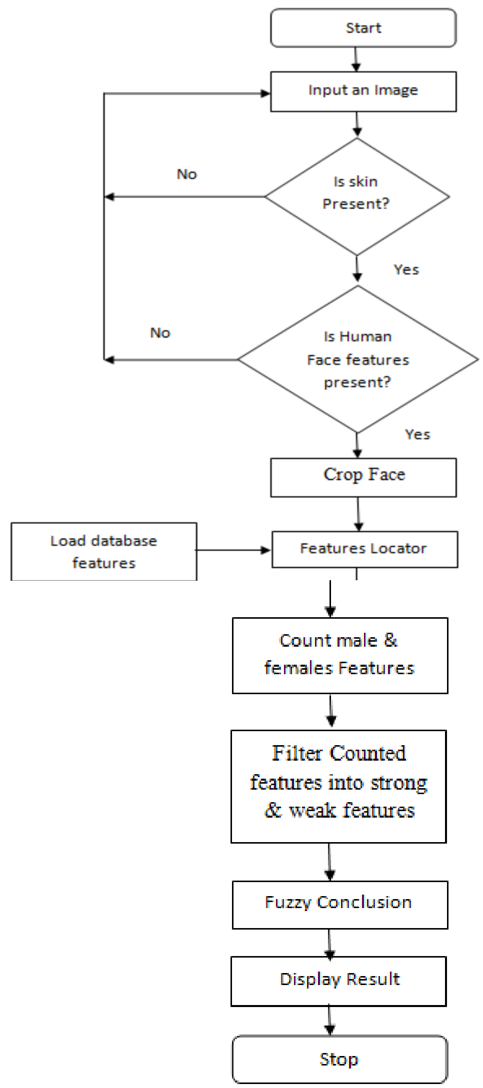

Figure 3.1 Data Flow diagram of Gender Recognization

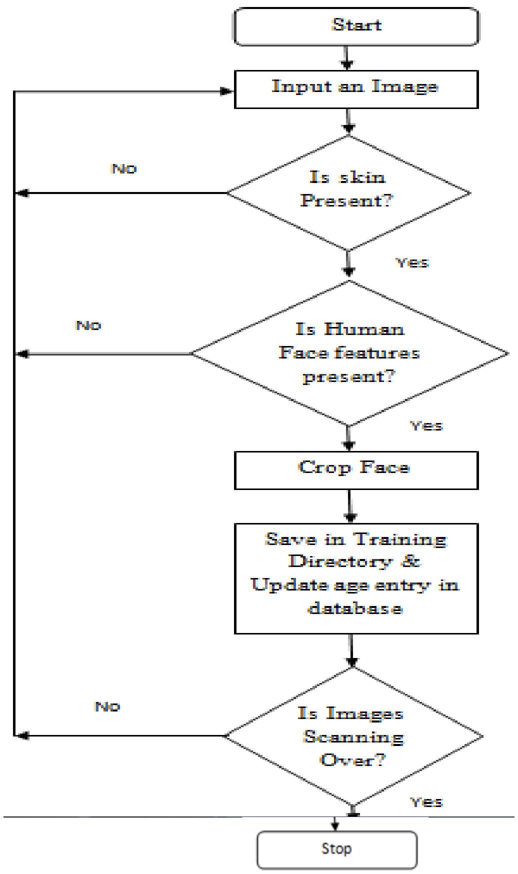

Figure 3.2 Data Flow diagram of Gender age Prediction(Training)

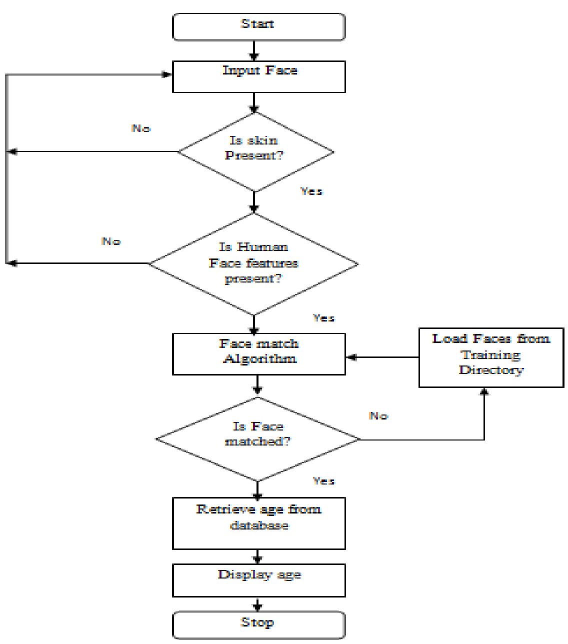

Figure 3.3 Data Flow diagram of Gender age Prediction(Testing)

3.3 Face Detection

Skin color plays a vital role in differentiating human and non-human faces. From the study it is observe that skin color pixels have a decimal value in the range of 120 to 140. In this project, we used a trial and error method to locate skin color and non skin color pixels. But many of the times, system fails to detect whether an image contains human face or not (i.e. for those images where there is a skin color background).an image is segmented into skin color and non-skin color pixels with the equations

$140 \le |Pxy| \le 120$ ex. 3.1

$where\ Pxy = pixel\ at\ position\ xy$

The skin pixels values are set to 1(i.e. #FFFF) and non skin pixels are set to 0(i.e. 0000). The pixels are collected and set as per equation

If

$\lim_{i\to 1} n\ (\int_1^3 120 \le |Pxy| \le 180) = 1$--------eq3. 2

Else

$\lim_{i\to 1} n\ (\int_1^3 180 \le |Pxy| \le 120) = 0$-------eq 3.3

$where\ n\ = total\ number\ of\ pixels\ of\ input\ image$

Literature review point out that, FACS system technique is based on face features extractions like eye, nose, mouth, etc. In this project, we minimize the number of features (i.e. only eyes and mouth) but given the more weight age for fuzzy rules formations from these extracted features. Face extractions consist of following steps

- Let W and H are the width and height of skin and non-pixel image as shown in fig 3.1.1
- Read the pixel at position (0,H/2) which is a middle of i.e. left side of image.
- Travers a distance D_1 = W/6 in horizontal direction to get the start boundary pixel of skin region.
- Travers a distance D_2= H/6 from a pixel position (W/6, H/2) in upward directions. Same may do in downward direction and locate the points X_1, X_2.
- Travers a distance D_3=W/3 from the point X_1 and locate the point X_3. Same do from the point x_2 and locate the point X_4.
- Crop the square image as shown.

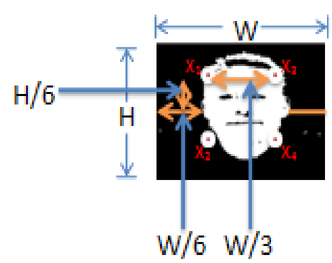

Figure 3.4 Detected Face Area

Figure 3.5 Face Features Extraction

Human face is made up of eyes; nose, mouth and chine etc. there are differences in shape, size, and structure of these organs. So the faces are differs in thousands way. One of the common methods for face expression recognition is to extract the shape of eyes and mouth and then distinguish the faces by the distance and scale of these organs. The face feature extractions consist of following steps

- Let W and H are width and height of an image shown in Fig 3.2.3
- Mark pixel P_i (W/2, H/2) as centre of image.
- Travers a distance H/8 from the pixel P_i towards upward and mark a point K_1.
- Travers a distance W/3 from the point K_1 towards leftward and mark a point K_2.
- Travers a distance H/10 towards downward from the point K_2 and mark a point K_3.
- Travers a distance W/4 from the point K_3 towards right and mark the point K_4.
- Travers a distance H/10 from the point K_4 toward up and mark the point K_5.
- Same steps are repeated for extracting the right eye and mark the point N_2, N_3, N_4, and N5.
- Travers a distance H/8 from the point P_i towards downward and mark the point M_1.
- Travers a distance W/6 towards left and right from the point M_1 and marks the point M_2 and M_3.
- Start with the point M_2 traverse a distance H/10 towards downward and mark the point M_4.
- Travers a distance W/6 from the point M_4 towards right and mark the point M_5. Same may do from point M_5 and mark the point M_6.
- Travers the distance H/10 from M_6 towards up that meets to the point M_3.
- See the below image.

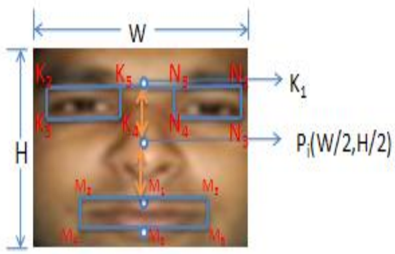

Figure 3.6 Face crop

- Dist $|P_i - K_1| = H/8$
- Dist $|K_1 - K_2|$ = Dist $|M_1 - M_2|$ = Dist $|M_1 - M_3|$ = Dist $|M_4 - M_5|$ = Dist $|M_5 - M_6|$ = W/3
- Dist $|K_2 - K_3|$ = Dist $|K_4 - K_5|$ = Dist $|N_2 - N_3|$ = Dist $|N_4 - N_5|$ = Dist $|M_2 - M_4|$ = Dist $|M_1 - M_5|$ = Dist $|M_3 - M_6|$ = H/10
- Dist $|K_3 - K_4|$ = Dist $|K_5 - K_2|$ = Dist $|N_3 - N_4|$ = Dist $|N_5 - N_2|$ = W/4

3.4 Features Location

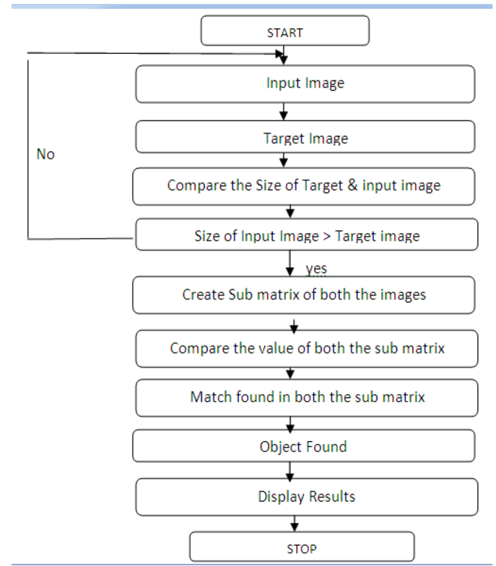

Figure 3.7 Data Flow diagram for locating features of training database into input image face region

In this project we create a system that is useful to find the object that is present inside the image. In this we first take a image as an input for which we have to find the object. After selecting the input image we have to focus on the target image. After getting both input image and target image we compare the size of both the image. If the size of input image is greater than target image then we proceed with our system otherwise we are going to focus on the input image. If the above criteria get satisfied then we create the sub matrixes of both the images for example we create 3 X 3 matrixes. After creating the sub matrixes we compare both the sub matrixes of both the image. This Matrix can be created with the help of pixels present in both the images. If matching found then we concludes that the given object is found in the other image. The object is shown as an output by creating red boxes on that object. The output object image is displayed and the co ordinates of that object that is the height and width from top and bottom in the form of coordinates is shown with the help of im tool.

3.5 Fuzzy Rules & Conclusion

Sr no	*Male Features*	*Female Features*	*Fuzzy Conclusion*
1	Beard		Male
2	Eyebrows		Male
3	-	Bindi	Female
4	-	Neckless	Female
5	-	Haar	Female
6	-	Ear ring	Female
7	Beard	Bindi	Male
8	Eyebrows	Bindi	Female
9	Beard	Bindi,Neckless	ambiguous
10	Bottleneck	Bindi,Neckless	female

Table 3.1 Fuzzy Rules

We forms too many fuzzy rules that gives us appropriate result.

4. Experiment Results

4.1 Project GUI

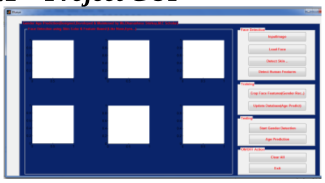

Figure 4.1 Main Form

Figure 4.2 Face Detection

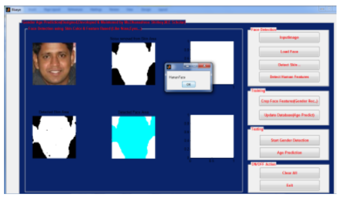

Figure 4.3 Skin Detection

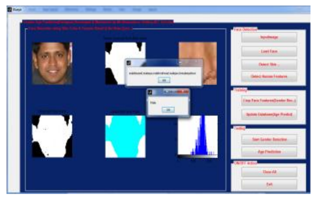

Figure4.4 Features & Gender Detected

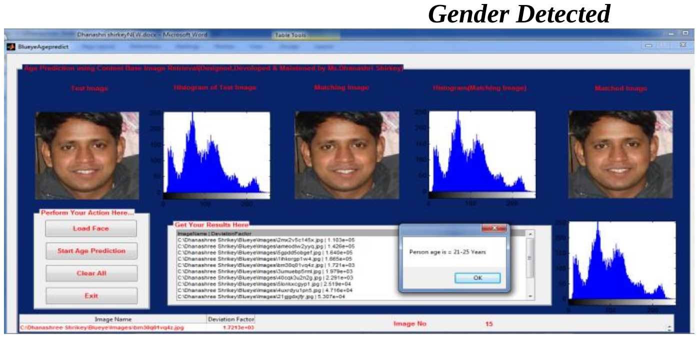

Figure 4.5 Age Prediction

Sr No	*Images tested*	*Gender Predicted*	*% of Gender prediction*	*Age Predicted*	*% of Age Prediction*
1	5	5	100	4	80
2	10	7	70	6	60
3	15	12	80	10	66.66
4	20	16	80	14	70

- Male & Female Features for gender recognization =50M/50F
- Training directory faces for age prediction=100

Table 4.1 Proposed Methodology result

5. Conclusion

This paper presented an approach to age and gender classification from still image. We find that the proposed methodology system performs on average comparably to Human visualizes. The results of a user study [7] shows that use patterns of Proposed systems for senior citizens differ significantly from those of adults or young faces. We believe the overall acceptance of proposed systems can be increased significantly by providing tailored versions of such systems, which adapt characteristics such as the degree of automation in a caller pre-selection scenario, order of presentation of options. In these scenarios, age and gender classification is not used to limit access (e.g. as in protection of minors), but to increase user satisfaction by providing individualized services even in the absence of knowledge about the face identity.

References

1. Edward D. Mysak, “Pitch duration characteristics of older males,” Journal of Speech and Hearing Research, vol. 2, pp.46–54, 1959.
2. Sue E. Linville, Vocal Aging, Singular Publishing Group, SanDiego, CA; USA, 2001.
3. Christian M¨uller, Frank Wittig, and J¨org Baus, “Exploiting speech for recognizing elderly users to respond to their special needs,” in Proc. Euro speech 2003, Geneva; Switzerland, Sept.2003, ISCA.
4. Nobuaki Minematsu, Mariko Sekiguchi, and Keikichi Hirose,“Automatic estimation of one’s age with his/ her speech based upon acoustic modeling techniques of speakers,” in Proc.ICASSP 2002, Orlando, FL; USA, May 2002, IEEE.
5. Izhak Shafran, Michael Riley, and Mehryar Mohri, “Voice signatures,”in Proc. ASRU 2003, U.S. Virgin Islands, Dec. 2003, IEEE.
6. Susanne Sch¨otz, “Automatic prediction of speaker age using CART,” Term paper for course in Forensic Phonetics,G¨oteborg University
7. European Language Resources Association (ELRA), “http://www.speechdat.org/,” http://www.elra.info/.
8. Jitendra Ajmera, “Effect of age and gender on LP smoothed spectral envelope,” in Proc. Speaker Odyssey. 2006, IEEE.
9. Loredana Cerrato, Mauro Falcone, and Andrea Paoloni, “Subjective age estimation of telephonic voices,” Speech Communication,vol. 31, no. 2–3, pp. 107–102, 2000.
10. The FG-NET Aging Database. http://www.fgnet.rsunit.com/.
11. Y. Fu, Y. Xu, and T. S. Huang. Estimating human age by manifold analysis of face pictures and regression on aging features. Proceedings of the IEEE Multimedia and Expo, pages 1383–1386, 2007.
12. G. Guo, G. Mu, Y. Fu, C. Dyer, and T. Huang. A study on automatic age estimation using a large database. International Conference on Computer Vision in Kyoto (ICCV 2009), pages 1986–1991, 2009.
13. A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for no orthogonal problems. Techno metrics, 12(3):55–67, 1970.
14. T. Kanamori, S. Hido, and M. Sugiyama. A least-squares approach to direct importance estimation. Journal of Machine Learning Research, 10:1391–1445, 2009.
15. K. J. Ricanek and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. Proceedings of the IEEE 7th International Conference on Automatic Face and Gesture Recognition (FGR 2006), pages 341–345, 2006.
16. B. Schoolkopf and A. J. Smola. Learning with Kernels, MIT Press, Cambridge, MA, USA, 2002.
17. H. Shimodaira. Improving predictive inference under covariate shift by weighting the log-likelihood function. Journal of Statistical Planning and Inference, 90(2):227–244, 2000.
18. M. Stone. Cross-valedictory choice and assessment of statistical predictions, Journal of the Royal Statistical Society, Series B, 36:111-147, 1974.