

An Integrated Approach to Classify Gender and Ethnicity

Md Azher Uddin
Computer Science & Engineering
Kyung Hee University
South Korea
azher006@yahoo.com

Shayhan Ameen Chowdhury
Computer Science & Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
shayhan@yahoo.com

Abstract—Faces express many social indications, including gender, ethnicity, age, expression and identity, most of them have drawn thriving attention from various research communities, for instance neuroscience, computer science and psychology. In this paper, we propose a new approach to classify gender and ethnicity by merging both texture and shape features extracted from face images. Gabor filter is used to extract the texture features and histogram of oriented gradients (HOG) is used to extract the shape features from face images. In order to achieve higher performance we combined both texture and shape features. After combining, the size of feature vector obtained is in a high dimension, to decrease the dimensionality Kernel PCA has been implemented. Finally, to classify the gender and ethnicity we used Support Vector Machine. The experimental result shows the effectiveness of proposed framework.

Keywords—Gender Recognition; Ethnicity Recognition; Gabor filter; Histogram of oriented gradients; Kernel PCA; Support Vector Machine.

I. INTRODUCTION

In the area of computer vision, facial image investigation is one of the important research issue. The faces of human beings provide various demographic properties for example identity, gender, ethnicity and age, significantly contributing in this area. From the different demographic properties, gender and ethnicity remains unchanged throughout whole life and significantly supports face recognition systems. Moreover, it also has a high degree of interest to be applied for development in various areas including access control, image retrieval and surveillance etc.

This paper proposes a framework to classify gender and ethnicity. The framework composed of three processing steps: 1) feature extraction by utilizing histogram of oriented gradients (HOG) and Gabor filter, 2) dimensionality reduction using Kernel PCA and 3) classification using Support Vector Machine (SVM). Our method has several benefits compared to other feature extraction method as Gabor filter is invariant to translation, scaling, rotation and illumination problem. We use FERET dataset [1] to test our system. In FERET dataset, the ethnicity classification is divided into three class: Caucasian, African, and Asian oriental with male and female gender for each ethnicity.

The remaining paper is structured as follows. The following section introduces a review of related works that have been applied. In section III, we discuss the System

overview. Section IV shows results evaluation of our system. Lastly, we draw conclusion in section V.

II. RELATED WORK

Several systems has been already developed to classify gender and ethnicity. This section presents some previous related works to classify gender and ethnicity which is significant to our framework.

In [2], authors used Gabor wavelet features and SVM classifier to classify ethnicity with an average accuracy of 96%, 93% and 94% for Asian, European, and African respectively. They used HOIP dataset for their experiments. LBPH (Local Binary Pattern histogram) features and Adaboost are used in [3], to classify gender, ethnicity and age. In [4], a method is proposed by using Oriented Gradient Maps (OGMs) to focus local geometry with texture information of whole faces and feature selection is done using Adaboost. The proposed method achieved 98.3% accuracy on FRGC v2.0 dataset. Local Binary Patterns and SVM classifier used by authors in [5] and they also used FRGC face dataset. In [6], experimental comparison has been performed for gender classification.

III. SYSTEM OVERVIEW

Our approach is decomposed into three steps: feature extraction, dimensionality reduction and classification. In feature extraction phase, we extract both shape and texture information by histogram of oriented gradients (HOG) and Gabor filter respectively. Then we combined both feature vector into one. The block diagram in figure 1 illustrates the framework to classify gender and ethnicity.

A. Pre-processing

Pre-processing is a process which is done on the input image before feature extraction phase. Preprocessing phase contains image enhancement, image conversion, image resize etc. To improve the picture contrast image enhancement is used so that picture can become clear. Image conversion is used to convert the RGB image to gray scale image. The RGB model contains of the three colors: red, blue and green whereas the subsequent gray level image are composed of black and white pixels. The range of resultant gray level image will be between 0 to 255 values. Image resize is used to resize image into a fixed size. All images are scale normalized to 100 × 120 pixels.

B. Feature extraction

Choosing important features is a crucial phase in any recognition or classification system. In our work, we extracted both texture and shape features. Histogram of oriented gradients (HOG) and Gabor filter are used to extract features from the gray scale image.

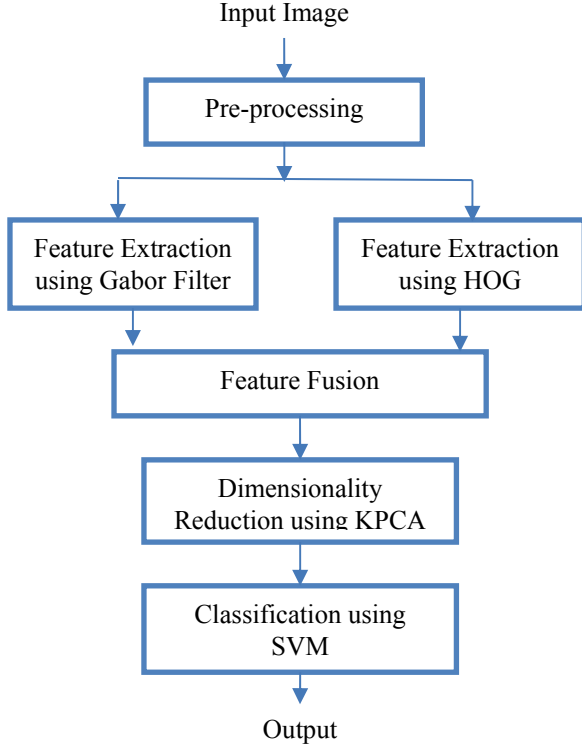


Fig. 1. Proposed framework

1) Gabor Filter

Gabor filter is a robust feature vector. This technique extracts texture information. Gabor filtering is effectively implemented in several computer vision fields including texture analysis, image smoothing, face recognition, fingerprint recognition etc. The Gabor filter is a band-pass filter, where a harmonic function multiplied by a Gaussian function is used to define the impulse response of the Gabor filter. Hence, a two-dimensional Gabor filter establishes a complex sinusoidal plane of certain frequency and orientation modulated by a Gaussian envelope [8]. Gabor filter is invariance to translation, scaling, non-uniform lighting condition and rotation. In spatial domain, a 2D Gabor filter [9] is defined as:

$$G(x, y) = \frac{f^2}{\pi\gamma\eta} \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp(j2\pi f x' + \phi) \quad (1)$$

$$x' = x \cos\theta + y \sin\theta \quad (2)$$

$$y' = -x \sin\theta + y \cos\theta \quad (3)$$

Where, θ denotes the orientation of Gabor function, σ is the standard deviation, f is the frequency of the sinusoidal factor, γ is the spatial aspect ratio and ϕ is the phase offset.

Our framework implements twenty-four Gabor filters in eight orientation angles and three scales. Then based on (5×5) downsampling factor, downsampling is done. The dimension of the feature vector obtained by Gabor filter is $(100 \times 120 \times 3 \times 8) / (5 \times 5)$ which is 11520. However, even after downsampling the dimension of feature vector achieved is still huge.

2) Histogram of oriented gradients (HOG)

The Histogram of Oriented Gradients (HOG) by Dalal et al. [10] is a powerful descriptor for human detection and tracking. HOG features has various application in image processing. HOG use intensity gradient distribution to describe the local feature of an object and its shape. They are calculated by the following steps:

- i. For every pixel $I(x, y)$, the horizontal and vertical gradient are calculated according to the following equation:

$$G_x = I(x + 1, y) - I(x - 1, y)$$

$$G_y = I(x, y + 1) - I(x, y - 1) \quad (4)$$

- ii. Then, the gradient magnitude and orientation are found by:

$$|G| = \sqrt{G_x^2 + G_y^2} \quad (5)$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (6)$$

- iii. Lastly, the histogram of gradient orientations is evaluated to get the HOG features.

In our work, face images are divided into numerous blocks and for every block a local histogram of the gradient orientations is calculated. Then, HOG features are obtained by combining all local histograms. In this research, the HOG over 100×120 images is calculated where each block contain 2×2 cells with size 8×8 and 9 histogram bins. The dimension of HOG feature vector is $(11 \times 14 \times 4 \times 9) = 5544$.

In the feature fusion phase, the both feature vector for Gabor filter and HOG are combined into one feature vector simply by concatenation. This feature fusion yields a feature vector for the face image with size 17064.

C. Dimensionality reduction

Since the size of feature vector is so large, it not only requires more computational time sometime it may also reduce the accuracy of classification. Therefore, dimensionality reduction technique has been applied. Although there are different dimensionality reduction method are available but we used Kernel PCA (KPCA) as it offers a better understanding about what nonlinear features are needed to extract. KPCA is a nonlinear dimensionality technique and it is an extended type of Principle component analysis (PCA). PCA is a linear feature extraction process frequently applied in various pattern recognition system, such as [11]. Kernel based tactics has been applied in numerous systems such as in [12-13]. The features from a low dimensional space is group into a more high dimensional space by KPCA. Due to this, the features can be linearly separable. Kernel PCA [14] calculates

the principal eigenvectors of a matrix, as an alternative of those of the covariance matrix. The re-establishment of conventional PCA in kernel space is simple, due to a kernel matrix is identical to the product of the features in the high-dimensional space that is built by utilizing kernel function. The utilization of PCA in kernel space gives Kernel PCA the estate to handle nonlinear mappings. The kernel matrix K calculated by Kernel PCA can be defined as,

$$k_{ij} = \kappa(x_i, x_j) \quad (7)$$

Where κ is a kernel function. Then, the following alteration of the entries are employed to centralize the kernel matrix K

$$k_{ij} = k_{ij} - \frac{1}{n} \sum_l k_{il} - \frac{1}{n} \sum_l k_{jl} + \frac{1}{n^2} \sum_{lm} k_{lm} \quad (8)$$

The centering process coincides to deducting the mean of the features in conventional PCA. It guaranteed that the features in the high-dimensional space are zero-mean. Then, the principal d eigenvectors v_i are calculated. The eigenvectors of the covariance matrix α_i are scaled forms of the eigenvectors of the kernel matrix v_i

$$\alpha_i = \frac{1}{\sqrt{\lambda_i}} v_i \quad (9)$$

The feature is projected onto the eigenvectors of the covariance matrix to obtain the low-dimensional data demonstration. The result of the projection is specified by,

$$Y = \{\sum_j \alpha_1 \kappa(x_j, x), \sum_j \alpha_2 \kappa(x_j, x), \dots, \sum_j \alpha_d \kappa(x_j, x)\} \quad (10)$$

Where κ is the kernel function. As Kernel PCA is a kernel-based method, the mapping done by Kernel PCA very much relies on the selection of the kernel function κ . Feasible choices for the kernel function consist of polynomial kernel, Gaussian kernel and linear kernel. Gaussian kernel function has been used in our research work. The size of the feature vector has been decreased to 320 from 17064 using nonlinear dimensionality reduction approach Kernel PCA.

D. Classification

The training and testing procedures are finished by utilizing Support Vector Machine (SVM) classifier to classify gender and ethnicity. SVM [15] is a classification tool that uses machine learning idea.

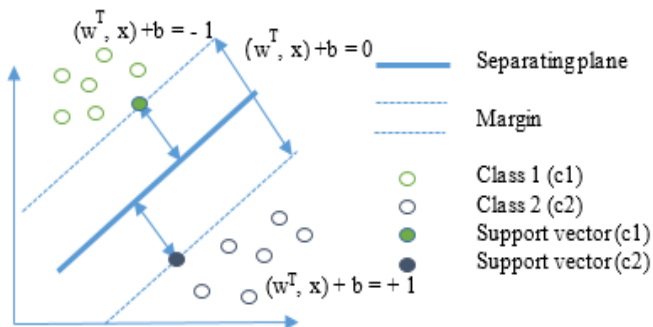


Fig. 2. SVM classifier using linear separating hyperplanes

We are given a set of n data points $\{x_i, y_i\}$; where, $i = 1, \dots, n$, $y_i \in \{-1, +1\}$ and $x_i \in \mathbb{R}^n$. SVM method goals at finding a classifier of form,

$$y(x) = \text{sign}(\sum_{i=1}^n \alpha_i y_i k(x, x_i) + b) \quad (11)$$

Where α_i are positive real constants and b is a real constant; $k(x, x_i)$ is the kernel function which can be defined by,

$$k(x, x_i) = \varphi(x)^T \varphi(x_i) \quad (12)$$

Where φ is a nonlinear mapping. In our work, we used the Gaussian radial basis function that can be defined by,

$$k(x, z) = e^{-\gamma \|x - z\|^2} \quad (13)$$

In high dimensional space, the data can be divided by a linear hyperplane, depending to the following equations,

$$\begin{cases} w^T \cdot x_i + b \geq +1, & \text{if } y_i = +1 \\ w^T \cdot x_i + b \leq -1, & \text{if } y_i = -1 \end{cases} \quad (14)$$

In this research, we develop one SVM classifiers to classify gender and ethnicity. The classifier employs binary tree structure and applies one against rest for sub-classification. The classifier contains 5 sub-classifiers. Figure 3 shows the SVM classifier. Each level contains a leaf node and a sub-classifier except the final level that contains two leaf nodes. Every leaf node denotes positive output ($y_i = +1$), on the other hand each sub-classifier is a negative output ($y_i = -1$).

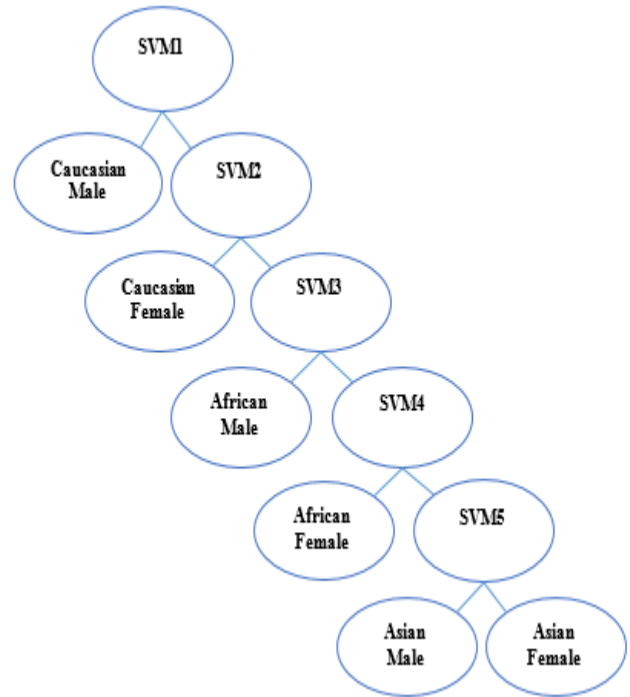


Fig. 3. Binary tree based SVM classifier

IV. RESULT EVALUATION

In this segment, we perform experiments to illustrate the efficiency of the framework for gender and ethnicity classification. We have used FERET dataset for training and

testing our system. Dataset contains 403 male individuals and 403 female individuals with a total of 2400 images. In our work, we have employed 1010 images for training and 1390 images for testing the efficiency of the framework. Fig. 4 shows some sample images. Before the training processing we have done few preprocessing on the original dataset, for instance cropping, rotating and scaling. And the Table I shows the result for training set. Here we have compared the result for Gabor filter features and HOG features. And also we have showed the result after combining both features.



Fig. 4. Sample images for gender and ethnicity classification

TABLE I. Classification Result for Different Groups

| Condition | Number of Images used for testing | Gabor Filter + KPCA | HOG + KPCA | Integrated Approach + KPCA |
|------------------|-----------------------------------|---------------------|------------|----------------------------|
| Caucasian Male | 240 | 93.6 % | 91.1 % | 97.3 % |
| Caucasian Female | 230 | 94.2 % | 87.6 % | 95.9 % |
| African Male | 240 | 91.5 % | 89.3 % | 93.6 % |
| African Female | 230 | 88.7 % | 84.8 % | 96.8 % |
| Asian Male | 230 | 93.2 % | 82.6 % | 93.4 % |
| Asian Female | 220 | 89.6 % | 90.1 % | 95.6 % |

V. CONCLUSION

In our work, we have used an integrated approach to classify gender and ethnicity that combines Gabor filter features and Histogram of Oriented Gradients (HOG) features of a face image. The experimental result shows that integrated approach works better than the single feature method. To extract the texture features Gabor filter features are used and to extract shape features HOG features are used. Since

dimensionality reduction technique is used in our system, so it requires less computation time.

REFERENCES

- [1] P. J. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, 2000, pp. 1090-1104.
- [2] S. Hosoi, E. Takikawa, and M. Kawade, "Ethnicity estimation with facial images," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, 2004, pp. 195-200.
- [3] Z. Yang and H. Ai, "Demographic classification with local binary patterns," in *Proc. International Conference on Biometrics*, 2007, pp. 464-473.
- [4] H. X. Ding, D. Huang, Y. H. Wang, and L. M. Chen, "Facial ethnicity classification based on boosted local texture and shape descriptions," in *Proc. IEEE Int. Conf. Workshops Autom. Face Gesture Recognition*, 2013, pp. 1-6.
- [5] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard, "Soft biometric classification using periocular region features," in *Proc. IEEE 4th Int. Conf. Biometrics: Theory Appl. Syst.*, 2010, pp. 1-7.
- [6] E. M'akinen and R. Raisamo, "An experimental comparison of gender classification methods," *Pattern Recognition Letters.*, vol. 29, pp. 1544-1556, July 2008.
- [7] MdAzher Uddin, Joolekha Bibi Joolee and Shayhan Ameen Chowdhury, "Bangladeshi Vehicle Digital License Plate Recognition for Metropolitan Cities Using Support Vector Machine", in *Proc. International Conference on Advanced Information and Communication Technology*, 2016.
- [8] R. Movellan, "Tutorial on Gabor Filters," 2008. Available: <http://mplab.ucsd.edu/tutorials/gabor.pdf> [Accessed: December 12th 2014].
- [9] M. Haghighat, S. Zonouz and M. Abdel-Mottaleb, "Identification using encrypted biometrics", *Computer analysis of images and patterns (CAIP)*, Springer, pp. 440-448, 2013.
- [10] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, in: *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 886 - 893, 2005.
- [11] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on*, vol. 12, pp. 103- 108, 1990.
- [12] MiladMoghaddam, ManoochehrNahvi and Reza Hassanzadeh Pak, "Static Persian Sign Language Recognition Using Kernel-Based Feature Extraction", 7th Iranian conference on Machine Vision and Image Processing (MVIP), pp. 1-5, 2011.
- [13] A. Kocsor and L. Toth, "Kernel-based feature extraction with a speech technology application," *Signal Processing*, *IEEE Transactions on*, vol. 52, pp. 2250-2263, 2004.
- [14] L.J.P. van der Maaten, E.O. Postma, and H.J. van den Herik, "Dimensionality Reduction: A Comparative Review. Tilburg University Technical Report", TiCC-TR 2009-005, 2009.
- [15] Christopher J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, Vol. 2, pp.121-167, 1998.