Improved gender classification using Discrete Wavelet Transform and hybrid Support Vector Machine

Jakub Walczak, Adam Wojciechowski
Institute of Information Technology, Łódź University of Technology
ul. Wólczańska 215, 90-924 Łódź, Poland
http://it.p.lodz.pl

Abstract. Gender recognition, across different races and regardless of age, is becoming an increasingly important technology in the domains of marketing, human-computer interaction and security. Most state-of-the-art systems consider either highly constrained conditions or relatively large databases. In either case, often not enough attention is paid to cross-racial age-invariant applications. This paper proposes a method of hybrid classification, which performs well even with a small training set. The design of the classifier enables the construction of reliable decision boundaries insensitive to an aging model as well as to race variation. For a training set consisting of one hundred images, the proposed method reached an accuracy level of 90%, whereas the best method known from the literature, tested under the restrictions imposed on the database, achieved only 78% accuracy.

Key words: computer vision, gender recognition, DWT, SVM.

1. Introduction

The ability to correctly determine gender is important for both inter-human communication and advanced human-computer interfaces, which tend to simulate real life interaction. Knowing the gender balance of an audience is a valuable information in the field of marketing, allowing message content to be adjusted so that advertisements can reach their targets more effectively. Algorithms for gender recognition need to be reliable under difficult conditions, since they are frequently employed in places where good facial exposition is rare. In this paper, we present an improved method of gender classification – robust against facial expressions, race and age. An additional advantage is that in the classifications based on a small training data set the method outperforms the current state-of-the-art solutions.

In our research, we used a randomly selected set of 200 subjects (100 samples for training and 100 for testing), from the FERET database [1, 2] including a range of races, ages and facial expressions (see fig. 1). During the investigation, a 100×100 pixel box was used to limit the visible area of each face.

While most of the solutions in the literature ([3, 4, 5]) provide high classification accuracy due to the large training datasets used, our method focuses on a combined

Fig. 1. Exemplary faces used in the project

classification model applicable even to small training sets (like in our case, 100 training samples).

2. Related work

An approach to gender recognition developed by Nazir *et al.* in [3] used the Discrete Cosine Transform to extract image traits. The collected features were then classified using the k-Nearest Neighbors (kNN) algorithm. This method attained accuracy of 99.3% with the Stanford University medical student frontal facial images database (SUMS) of 400 images, 200×200 pixels in size. The drawback of these studies is that the training set comprised good quality and well-lit subjects of approximately the same age. The application of the kNN classifier also does not seem to be a good solution for larger datasets, since all objects must be stored, making the method demanding in terms of memory.

Another method has been suggested by Alrashed *et al.* [5] who used eye region images processed either with the Discrete Cosine Transform or the Discrete Wavelet Transform. The features obtained were classified using a Support Vector Machine with a Radial Basis Function kernel. The accuracy of the systems reached 99.62% for DCT and 99.49% for DWT. Tests were performed on two databases: *Faces94* [6] and *AR database* [7, 8]. The data set comprised 691 images of males and 692 of females. Subjects with sunglasses were excluded, making classification easier. In addition, the first dataset contained images captured under similar lighting conditions.

Singh *et al.* in [4] proposed a method for feature extraction based on Local Binary Patterns and the Histogram of Oriented Gradients [9]. A Support Vector Machine was employed for classification. The accuracy level reached 89.43% for LBP and 95.56% for HOG for *Indian face database* [10]. Although this study took into account variations in lighting, it considered solely the subjects of South Asian ethnicity – focusing on a narrow range of human appearances.

It is worth noting that each of the aforementioned methods relied on the immensity of information provided by large datasets rather than providing a well-generalized gender model based on classification of carefully chosen features. In this paper, we present

Fig. 2. Example of ambiguous face according to our method

a novel method of gender recognition which combines multiple classifiers to boost precision in even small training sets [11].

3. Method

The method, developed by inference from empirical results, is composed of three main stages: preprocessing, features retrieval and classification [12]. In the first stage, an input image face is prepared for further processing by means of noise reduction including Gaussian filtering. The image is then subjected to feature extraction, performed using Discrete Wavelet Transform [5], since DWT and DCT achieved the best precision, and DWT was found to be better at preserving high and low frequency information [13]. As Aguado *et al.* observed, filtering either high or low frequencies deteriorates the precision of gender recognition [14]. The feature vector calculated with DWT is finally passed to our hybrid classifier.

The superiority of the Support Vector Machine over other standard classifiers was confirmed by experimentation. This dominance was also achieved by the application of kernel transformation. It was noticed that the gender model, in the context of DWT coefficients, was highly non-linear and transformations of higher orders worked better. Two particular kernels were determined as paramount transformations: the Quadratic Kernal Function (QKF) and Radial Basis Function (RBF). Classification results for both SVMs indicated that, in most cases, they provided the same answer (*woman* or *man*). In the majority of the other cases, QKF SVM was superior over RBS, since for true decisions it revealed higher classification probability (expressed in (3) and (4) futher in the text), enforcing the right decision of the classification ensemble.

The final classifier was composed of two separate Support Vector Machines, one with a Quadratic Kernel Function $K_q(\beta, \delta) = (\beta^T \delta)^2$ and the other with a Radial Basis similarity Function $K_r(\beta, \delta) = exp(\gamma||\beta - \delta||)$. For both SVMs, the success rate was calculated according to the following formulas, respectively.

$$s_q = \sum_{i=1}^{m} a_i y_i K_q(\beta_i, \beta) + b \, , \tag{1}$$

$$s_r = \sum_{i=1}^{m} a_i y_i K_r(\beta_i, \beta) + b \, , \tag{2}$$

where β_i is the feature vector of the i-th training example, y_i is a class label taking values -1 for class *woman* and 1 for class *man*, m is the number of training observations, a_i is a coefficient associated with an observation (zero if observation is not a support vector), b is a scalar, β is a feature vector of the test object and $K(\beta_i, \beta)$ is a kernel function.

Posterior probabilities $P(s_q)$ and $P(s_r)$ are then measured for two Support Vector Machines. The probability is a function of the success rate s. For the Quadratic kernel, the probability function is presented in (3), while that for the Radial Basis Function kernel it is shown in (4).

$$P(s_q) = \frac{1}{1 + e^{A_1 s_q + B_1}} \, , \tag{3}$$

$$P(s_r) = \frac{1}{1 + e^{A_2 s_r + B_2}} \, , \tag{4}$$

where s_q and s_r are success rates for the Quadratic kernel and RBF kernel respectively. A and B are function parameters.

The functions (3) and (4) provide a qualitative measure of a subject's membership to the class with label **1** – *man*.

The value of $P(s_q)$, denoted in short as x, is the probability of being classified as a *man* by the SVM with a Quadratic kernel. The probability of being classified as a *woman* by the same classifier is $1 - P(s_q)$ (denoted as x̂) according to the probability theorem. In the second equation, the value of $P(s_r)$ (y) indicates the probability of being a *man* according to the SVM with RBF kernel, and the probability of being a women for this classifier is $1 - P(s_r)$ denoted as ŷ.

Based on the complementarity of the two SVMs, one can assume that the arithmetical average of posterior probability for each class gives a more reliable probability than each classifier separately. Hence, having calculated the posterior probability for the two SVMs, the mean is calculated according to standard formulas

$$z = (x + y)/2 \, , \tag{5}$$

$$\hat{z} = (\hat{x} + \hat{y})/2 \, . \tag{6}$$

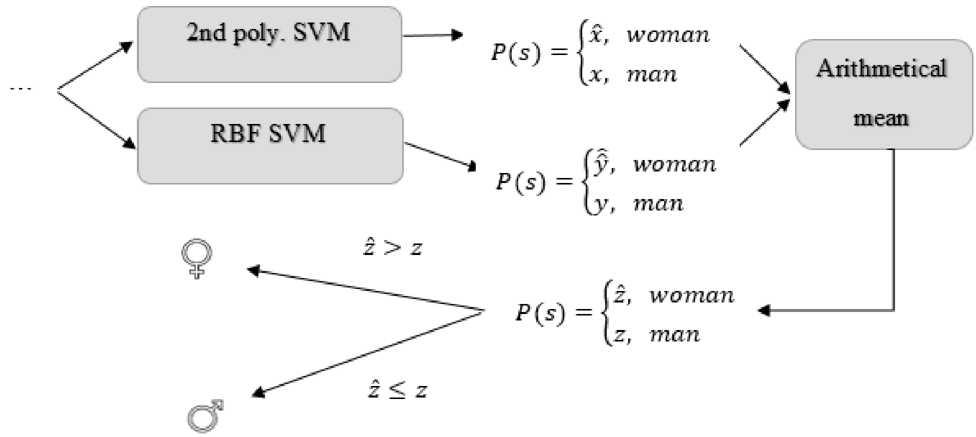

Fig. 3. Proposed classification method scheme

The final decision is then made depending on the values determined for z and $\hat{z}$. If the average posterior probability z is greater or equal to $\hat{z}$ then the examined object is classified as a *man*, otherwise it is classified as a *woman*. The classification method is visually presented in Figure 3.

4. Test and Results

The tests consisted of two parts: validation of systems described in the literature [3, 4, 5] and validation of the proposed hybrid method. Validation was carried out with tests performed on half of the previously described set (100 subjects out of 200) from the FERET database, randomly selected. This set contains people of various races (Caucasian, Black and Asian), of various ages (from round 20 up to around 60), under varying lighting conditions (uniform light, spot light from different directions) and with different facial expressions (neutral, smiling, grinning, sad, disgusted). The images were taken from gray-scaled thumbnail photographs of the *FERET database*. This database was chosen due to the versatility that it provides with regard to age, race, facial expressions and lighting. In spite of the fact that it is not a new database, it is still valid since the general perception of gender does not change over time. To obtain a qualitative comparison, an accuracy factor was used

$$accuracy = \frac{number\ of\ correct\ classifications}{capacity\ of\ test\ set}\ . \tag{7}$$

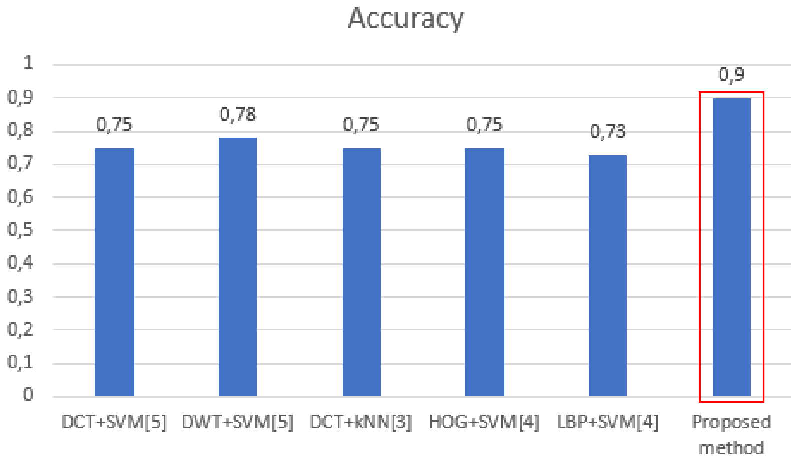

Fig. 4. Comparison of the pure state-of-art algorithms with the proposed system

Below, Figure 4 shows the accuracy of algorithms in the literature for the proposed image set, in comparison to the proposed system. The first two methods presented in the chart were proposed in [5]. They used either Discrete Cosine Transform or Discrete Wavelet Transform together with the SVM classification method. The variant with DWT achieved better results (78%) than that with DCT (74%). The results were worse than those reported by the authors, which may have been due to the significantly smaller database used in our study. The third bar depicts the accuracy achieved by the algorithm presented in [3]. It uses Discrete Cosine Transform applying the kNN classifier. It gained precision of 75%. Again, the lower accuracy with respect to that reported by the authors may have been caused by differences in databases. Nazir *et al.* drew on a larger database of students, hence the ages of the subjects were also quite similar. The next two methods suggested in [4] used a Histogram of Oriented Gradients and Local Binary Patterns as feature descriptors. Both methods were tested with the Support Vector Machine Classifier. The variant with HOG obtained a slightly better accuracy (75%) than that with LBP (73%). The authors trained and tested their system on an Indian database, rejecting cross-racial comparisons. The database used in our study contains subjects of different races, which has significant implications for the accuracy of the method.

As shown in Figure 4, the system from the literature with the highest accuracy (78%) was the Discrete Wavelet Transform with Support Vector Machine. The algorithms suggested by Nazir *et al.* [3], Alrashed *et al.* [5] and Singh *et al.* [4] achieved at most 78% accuracy with pictures from the *FERET database*. However, the accuracy of the proposed method reached 90%. This means that with the proposed system 10 more people in every hundred were classified correctly.

In this way, the superiority of SVM in terms of gender classification [15] was clearly confirmed.

The execution time was measured to assess the feasibility of real-time applications. For tests performed with images of 100×100 px, the accuracy of the method was preserved and the time required to perform the classification was 0.2 s. This enables complete classification of five faces per second.

5. Conclusion

The hybrid classification method presented here, consisting of two Support Vector Machines, one with Quadratic and one with Radial Basis kernels, deals very well with gender recognition, even when the subjects are of different races and ages, have different facial expressions, or are taken under different lighting conditions. The results compare favorably to those for algorithms presented in the literature, in [3, 4] and [5], which performed considerably worse with the small training data set.

Further research should focus on making the proposed method insensitive to rotation and scale and on making the whole system applicable for real use without any manual pre-processing. Finally, low-dimensional image representation pointed out by Amine *et al.* in [16] might be used to improve the accuracy and efficiency.

References

[1] P. J. Phillips, J. Huang H. Wechsler and, and P. Rauss. The FERET database and evaluation procedure for face recognition. *Image and Vision Computing*, 13:259–306, 1998. `doi:10.1016/S0262-8856(97)00070-X`.

[2] P. Phillips, H. Mon, S. A. Rizvi, and P. J. Rauss. The FERET evaluation methodology for face recognition algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22:1090–1104, 2000. `doi:10.1109/34.879790`.

[3] M. Nazir, M. Ishtiaq, A. Batool, M. A. Jaffar, and A. M. Mirza. Feature selection for efficient gender classification. In *Recent Advances in Neural Networks, Fuzzy Systems & Evolutionary Computing. Proc. 11th WSEAS Int. Conf. on Neural Networks and 11th WSEAS Int. Conf. on Evolutionary Computing and 11th WSEAS Int. Conf. on Fuzzy Systems*, NN'10/EC'10/FS'10, pages 70–75, Iasi, Romania, 2010. dl.acm.org/citation.cfm?id=1863431.1863444.

[4] V. Singh, V. Shokeen, and B. Singh. Comparison of feature extraction algorithms for gender classification. *International Journal of Engineering Research and Technology*, 2(5):1313–1318, May 2013. `www.ijert.org`.

[5] H. A. Alrashed and M. A. Berbar. Facial gender recognition using eyes images. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(6):2441–2445, 2013. `www.ijarcce.com`.

[6] L. Spacek. Collection of facial images: Faces94. The Ohio State University. `http://cswww.essex.ac.uk/mv/allfaces/faces94.html` [Online; accessed 09 Sep 2016].

[7] A. M. Martinez. AR Face Database. The Ohio State University. `http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html` [Online; accessed 09 Sep 2016].

[8] A. M. Martinez and R. Benavente. The AR face database. CVC Technical Report #24, June 1998.

[9] A. Wojciechowski. Potential field based camera collisions detection in a static 3D environment. *Machine Graphics & Vision*, 15(3/4):665, 2006.

[10] R. Sharma and M. S. Patterh. Indian face age database: A database for face recognition with age variation. *International Journal of Computer Applications (0975 -8887)*, 126(5):21–27, 2015.

[11] V. N. Pawar and S. N. Talbar. Hybrid machine learning approach for object recognition: Fusion of features and decisions. *Machine Graphics and Vision*, 19(4):411–428, 2010.

[12] R. Staniucha and A. Wojciechowski. Mouth features extraction for emotion classification. In *Proc. 2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 1685–1692. IEEE, Sept 2016. `ieeexplore.ieee.org/document/7733480/`.

[13] A. Katharotiya, S. Patel, and M. Goyani. Comparative analysis between DCT & DWT techniques of image compression. *Journal of Information Engineering and Applications*, 1(2):9–17, 2011. `http://iiste.org/Journals/index.php/JIEA`.

[14] L. Aguado, I. Serrano-Pedraza, S. Rodríguez, and F. J. Román. Effects of spatial frequency content on classification of face gender and expression. *The Spanish Journal of Psychology*, 13(2):525–537, 2010. `doi:10.1017/S1138741600002225`.

[15] S.A. Khan, M. Katameneni, and P. M. Latha. A comparative analysis of gender classification techniques. *Middle-East Journal of Scientific Research*, 20(1):1–13, 2014. `www.idosi.org/mejsr/mejsr.htm`.

[16] A. Amine, S. Ghouzali, M. Rziza, and D. Aboutajdine. An improved method for face recognition based on SVM in frequency domain. *Machine Graphics and Vision*, 18(2):187–199, 2009.