Combined Face Recognition Classifiers based DCT and LBP Feature Sets

Sanaa Ghouzali, Mohamed El Aroussi, Aouatif Amine, Mohammed Rziza, Driss Aboutajdine

GSCM-LRIT, Faculty of Sciences
Mohammed V University
Rabat
Morocco
sg363@cornell.edu

ABSTRACT. Most existing face recognition approaches have limited performance in uncontrolled environments. Effective face recognition requires several different kinds of feature sets to be taken into account which can integrate heterogeneous and complementary information of the input face images. In this paper, we propose to fuse two commonly used face recognition algorithms based on Discrete Cosine Transform (DCT) and Local Binary Patterns (LBP) feature extraction. The classification scheme is ensured using the SVM one-against-one strategy. We have investigated the impact of the information fusion on the recognition rate. Several combination strategies are compared at the score/decision level. Experiments conducted on Yale and AT&T face databases show that the proposed classifier combination approaches outperform individual classifiers.

RÉSUMÉ. Les performances de la plupart des approches existantes pour la reconnaissance de visage sont limites dans des environnements incontrls. L'efficacit de reconnaissance des visages ncessite plusieurs diffrentes sortes de caractristiques prendre en compte pouvant intgrer des informations htrognes et complmentaires de l'image du visage en entre. Dans ce papier, nous proposons de fusionner deux algorithmes de reconnaissance de visages couramment utiliss, la transform en Cosinus discret (DCT) et les LBP. Le systme de classification est assur en utilisant les SVM avec la stratgie un-contre-un. Nous avons tudi l'impact de la fusion d'information sur le taux de reconnaissance. Plusieurs stratgies de combinaison sont compares au niveau score/dcision. Des expriences entames sur les bases de visages Yale et AT&T montrent que la combinaison de classifieurs augmentent la performance des classifieurs individuels.

KEYWORDS : Face recognition, DCT, LBP, SVM, Feature Extraction, and Classifier Fusion.

MOTS-CLÉS : Reconnaissance de Visages, DCT, LBP, SVM, Extraction de caractristiques, Fusion de classifieurs.

1. INTRODUCTION

Among face recognition algorithms, the most popular are appearance-based approaches [21, 3, 4]. These approaches exhibit good performance and robustness against noise in controlled environments but still do not perform well in many real-world situations, where the query test face appearance is significantly different from the training face data, due to variations in pose, lighting and expression. It is often the case that no single feature descriptor is rich enough to capture all of the classification information available in the pattern image. Thus, one of the key challenges for improving face recognition performance is finding and combining efficient and discriminative information about face patterns that are resistant to all kind of variations [23]. Worth noting, by observing the errors misclassified by the different approaches, one can observe that a certain classifier is better suited for the recognition of a certain patterns than another one and therefore, some recognition errors committed by the best approach can be well resolved by the inferior methods. These observations motivated the relatively recent interest in combining different classifiers which integrates various information sources or different type of feature sets [20].

Information fusion can be considered at the feature level or at the classifier level. The feature level fusion is believed to provide better recognition results than classifier level fusion since the features contain richer information about the input data than the matching score or the output decision of a classifier/matcher. However, selecting appropriate and complementary component features is crucial for good performance. In [20], Tan et al. combined two local appearance descriptors, Gabor wavelets [22] and Local Binary Patterns (LBP) [15]. Both feature sets are high dimensional so the authors used PCA to reduce the dimensionality prior to normalization and integration. The Kernel Discriminative Common Vector [7] method is then applied to the combined feature vector. In [14], PCA, ICA and LDA are used together to provide the component subspaces for classifier combination. Each test sample is separately projected into these three subspaces and the resulting distance matrices are then fused to make the final decision using two combination strategies, either the sum rule or an RBF network [5]. The experiment results conducted on several challenging databases showed that both combination approaches outperform individual classifiers.

Motivated in part by the work presented in [20], we propose in this paper a face recognition approach using two complementary feature extraction algorithms, Discrete Cosine Transform (DCT) [18] and Local Binary Patterns (LBP) [15]. The Support Vector Machine (SVM) with one-against-one strategy [6] is used for classification. The reasons underlying the choice of using these feature sets and Support Vector Machines are the following: from one hand, DCT and LBP coefficients have been chosen for their complimentary in the sense that LBP captures small appearance details of facial appearance and texture in the spatial domain while DCT encodes facial texture and edge information in the frequency domain. On the other hand, even if a considerable dimensionality reduction is obtained by these feature extraction techniques with respect to considering the whole image, the resulting space is still large. Standard classifiers could be affected by the so called curse of dimensionality problem; SVMs, instead, are well suited to work in very high dimensional spaces. Each feature set is classified separately using SVM to obtain the individual scores which will be then transformed into posteriori probabilities prior to combination at the score level fusion. Several combination strategies have been investigated including Sum rule, Product rule, Max rule, Min rule and Majority vote rule. Experiment results conducted on Yale and AT&T face databases show that combining DCT-based SVM and LBP-based SVM classifiers at the decision level gives better performance than individual classifiers.

The remainder of this paper is organized as follows: section 2 gives an overview of the use of the DCT and LBP as means of feature extraction algorithms for face representation. In section 3 a brief description of the face recognition based SVM is given. The proposed classifier combination scheme is presented in section 4; experimental results of the proposed technique along with comparative analysis are discussed in section 5. Finally, in section 6 we draw conclusions and give avenues for future work.

2. Feature Extraction Methods

2.1. Discrete Cosine Transform

High information redundancy and correlation in face images result in inefficiencies when such images are used directly for recognition. Discrete Cosine Transform (DCT) is a predominant tool first introduced by Ahmed et al. [1]. Since then, it was widely used as a feature extraction and compression in various applications on signal and image processing and analysis due to its fine properties, i.e., de-correlation, energy compaction, separability, symmetry and orthogonality [18]. In pattern recognition techniques to make the model computationally efficient, transform orthogonality is as important as the class separation in applications like face recognition. Unlike Gabor elementary functions used in [20], which are a set of overlapping functions and not mutually orthogonal, the DCT basis functions are orthogonal. In addition to its de-correlation characteristics, this property renders some reduction in the pre-computation complexity. Furthermore, DCTs are used to reduce image information redundancy because only a subset of the transform coefficients are necessary to preserve the most important facial features.

The local information of a candidate face can be obtained by using block-based DCT as follows: A face image is divided into blocks of 8 by 8 pixels size. Each block is then represented by its DCT coefficients. From the obtained DCT coefficients only a small, *generic* feature set is retained in each block. Ekenel et al. [9] have proved that the highest information necessary to achieve high classification accuracy is contained in the first low frequency DCT coefficients via zigzag scanning.

2.2. Local Binary Pattern

The Local Binary Pattern operator was first introduced by Ojala et al [15] who showed the high discriminative power of this operator for texture classification. An extension to the original operator was made in [16] and called uniform patterns. The idea behind the LBP uniform is to detect characteristic (local) textures in image, like spots, line ends, edges and corners. Through its recent extensions, the LBP operator has been made into a really powerful measure of image texture, showing excellent results in terms of accuracy and computational complexity in many empirical studies. Moreover, LBP's are resistant to lighting effects in the sense that they are invariant to monotonic gray-level transformations, and they have been shown to have high discriminative power for texture classification [15].

For face recognition, LBP method was firstly introduced by T. Ahonen et al [2] consisting on dividing the face into a regular grid of cells and histogramming the uniform

LBP's within each cell. Finally, the cell-level histograms are concatenated to produce a global descriptor vector. Therefore the LBP method is applied on images (of faces) to extract features which can be used to get a measure for the similarity between these images. This model contains information on three different levels: (1) LBP code labels for the local histograms (pixel level), (2) local histograms (region level) and (3) a concatenated histogram which builds a global description of the face image (image level).

3. SVM classifier for face recognition

Recently, the Support Vector Machine learning (SVM) has been gaining popularity in the field of pattern classification due to its promising empirical performance, moderate computation complexity and its strong mathematical foundation. More details about SVM can be found in [6]. SVM are binary classifiers and different approaches like "one-against-all" and "one-against-one" are built to extend SVM to the multi-class classification case for face recognition [6]. For a K-class classification task, the common method is to use "one-against-all" [19] principle to construct K binary SVMs. Each SVM distinguishes one class from all other classes. The final output is the class that corresponds to the SVM with the highest output value. Another major method is the "one-against-one" method [13]. This method consists in building up all possible $K(K-1)/2$ binary SVMs representing all possible pairs out of K classes, each of which is used to discriminate two of the K classes only. When a testing pattern is to be classified, it is presented to all the SVMs, each providing a partial answer that concerns the two involved classes. Different schemes are used to combine the results of binary SVMs. In the classification stage, a majority voting strategy is often used: each binary classification is considered to be a voting where votes can be cast for all data points, in the end point is designated to be in a class with maximum number of votes.

4. Classifier Combination Scheme

A combined system can be based on one or a combination of the following fusion scenarios: in the first scenario, all the classifiers use the same representation of the input pattern whereas in the second scenario, each classifier uses its own representation of the input pattern. In other words, the features extracted from the pattern are unique to each classifier. In our work we focus on classifier combination in the second scenario. The input data is processed with different feature extraction algorithms in order to create templates with different information content. To make a stronger final classifier, several classifiers based on distinct features are combined. Kittler et al. [12] have demonstrated that combining the scores of several classifiers can lead to better recognition results. Some of the rules used to combine the classifiers at the score level are [12]: Sum rule, Product rule, Max rule and Min rule.

In order to employ these schemes, the matching scores of the $K(K-1)/2$ binary SVM classifiers are converted into posteriori probabilities. A probability associated with a classifier is often very useful and it provides confidence about the classification result. Platt [17] introduced the sigmoid function as the probability model to fit $P(y=1|f)$ directly. The parametric model is shown in Eq. 1.

$$P(y=1|f) = \frac{1}{1+\exp(Af+B)} \qquad [1]$$

where A and B are scalar values, which are fit with maximum likelihood estimation. f is the decision function of the binary SVM.

The proposed classifier combination scheme is a two-step fusion which consists first on a fusion of the probabilities at the output of the individual binary SVM classifiers using several fusion rules like the Sum rule, Product rule, Max rule and Min rule. The combined results are then fused again to generate a single scalar score, which is then used to make the final decision via the Majority vote rule. An illustration of this combination scheme is presented in the Figure 1.

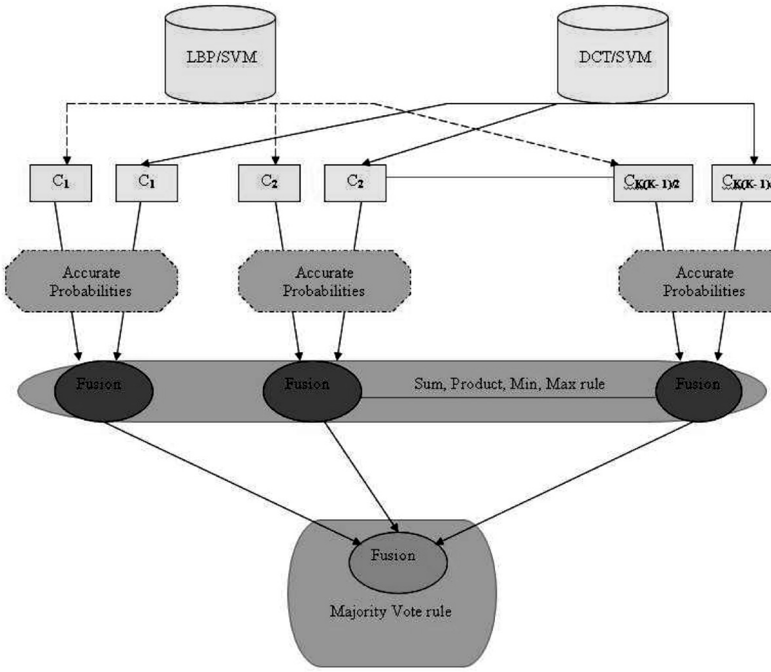

Figure 1. *Diagram of the Score Level Fusion Scheme.*

5. Experiment results

In this section, the DCT and LBP features are extracted from the a face database to construct different feature sets of face information. Multiple classifier combination schemes have been studied under different fusion techniques. To assess the robustness of our method against different facial expressions, pose and lighting conditions, we have choose Yale [4] and AT&T [10] face databases. The Yale face database contains 165 grayscale images of 15 individuals that include variation in both facial expression and lighting. The AT&T database contains 400 images of 40 subjects that include variation in facial expression and pose. The entire face database is divided into two parts. Six images of each subject are used to construct the training data and the remaining ones are used for testing. The face images with illumination from sides and with glasses are put in the test set on purpose in order to harden the testing conditions. All the faces are then scaled to

	DCT-SVM	LBP-SVM	Sum Rule	Product Rule	Max Rule	Min Rule
Recognition Rate	90,21%	88,08%	93,62%	92,77%	92,34%	91,91%

the size 104×104 pixels, aligned according to the eye positions. Figure 2 depicts some sample images from the database.

Figure 2. *Face samples from the YALE and AT&T face databases.*

In our study, we have choose to retain the lowest 3 DCT coefficients in each 8×8 pixels block of the face image; the remaining coefficients form a one dimensional feature vector of $(3 \times 13 \times 13 = 507)$ size. To extract the LBP feature set of face images, we have choose to use the uniform LBP's within the image as presented in [20][2] using an operator of 3 by 3 local neighborhood around each pixel, thresholding the pixels of the neighborhood at the value of the central pixel and using the resulting binary-valued image patch as a local image descriptor of 35×35 pixels grid cell. Finally, the local descriptors are histogrammed to produce a global descriptor vector of $((P \times (P-1)+3) \times 3 \times 3 = 531)$ size with $P = 8$ (pixel's neighbors).

Once the DCT and LBP feature sets are extracted, each feature set is presented to all the SVM classifiers. The SVM with a 2nd degree polynomial kernel has been found in our simulations to outperform linear and RBF kernel functions. In the present work, the library LIBSVM [8] was used. This library implements the SVM with one-against-one voting terminology to handle more than two classes. In this pair-wise classification, we need to train $k(k-1)/2$ SVMs representing all possible pairs out of k classes. The i^{th} individual binary SVM classifier provides a partial score that determines whether the input vector is "class m" or "class n". These partial scores are first separately transformed into a well calibrated probabilities prior to combination. Several fusion rules are investigated including the Sum, Product, Min and Max rule. The combined results are then fused again to generate a single scalar score, which is then used to make the final decision via the Majority vote rule. Since there are more than two classes, the combined decision is correct when a majority of the decisions are correct, but wrong when a majority of the decisions are wrong and they agree. A rejection is considered neither correct nor wrong, so it is equivalent to a neutral position or an abstention. The results of the classifier combination obtained using the different rules are summarized in table 1.

The results for the score level fusion, as shown in table 1, demonstrate that it is possible to achieve good fusion performance for a specific database by carefully choosing the fusion technique. We observe that a combined classifier employing the Sum of prob-

lin Rule
)1,91%

abilities method provides the best classification results and appears to produce the most reliable decisions. The Sum rule can be viewed to be computing the average probability for each binary classifier over all the classifier outputs. Another advantage of the Sum rule is that it is completely data independent as it requires no tuning set to effectively fuse matching scores. We have also observed at the second-step fusion that Majority Vote rule tends to be sensitive to the performance of the worst of its component classifiers. Moreover, the Majority Vote rule is less prone to ambiguity and may not be a good strategy in the case that two classes have identical votes.

6. CONCLUSIONS AND FUTURE WORKS

In this paper we have presented a combined appearance-based face recognition approach, which uses two different representations of the face image. The underlying algorithm utilizes the block-based DCT and the uniform LBP for local representations of the face image. Indeed, these two feature sets capture different and complementary information. We investigated the impact of information fusion at the score level. In this scheme, score-level fusion benefited from using a mixture of different feature types. We conducted extensive experiments on the Yale and AT&T face databases using various classifier combination schemes such as sum, max, product and majority vote rule. It's apparent that the combined classifier outperforms the accuracy of either alone when the best fusion rule is selected. The sum classifier combination rule has proved to be not only a very simple and intuitive technique of improving the reliability of decision making based on different classifier scores but also remarkably robust. As a future work, we want to try different classifier combination such as support vector machine, radial basis function networks and belief functions.

7. ACKNOWLEDGMENTS

This work is supported in part by the Maroc Telecom IAM project 'Access Control', No. 105 909 05/PI.

8. References

[1] N. Ahmed, T. Natarajan and K.R. Rao,Discrete Cosine Transform, *IEEE Transactions Computers*, vol. 23, 1974, pp 90-94.

[2] T. Ahonen, A. Hadid and M. Pietikainen, Face recognition with local binary patterns, *In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004*. LNCS, vol. 3021, 2004, pp 469-481. Springer, Heidelberg

[3] M.S. Bartlett, H.M. Lades and T.J. Sejnowski, Independent component representations for face recognition, *in Proceedings of the SPIE*, vol. 3299, 1998, pp 528-539.

[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19(7), Jul. 1997, pp 711-720.

[5] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, UK; 1995.

[6] J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, vol. 2(2), 1998, pp 121-167.

[7] H. Cevikalp, M. Neamtu, M. Wilkes, Discriminative common vector method with kernels, *IEEE TNN*, vol. 17(6), 2006, 1550-1565.

[8] C.-C. Chang and C.-J. Lin: LIBSVM : a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm

[9] H. K. Ekenel and R. Stiefelhagen, Analysis of Local Appearance-Based Face Recognition: Effects of Feature Selection and Feature Normalization, *IEEE Computer Society CVPRW'06*, 2006, 34-40.

[10] F. Samaria and A. Harter, Parameterisation of a stochastic model for human face identification, *in Proc. 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota FL, Dec. 1994.

[11] A. Jain, K. Nandakumar and A. Ross, Score normalization in multimodal biometric systems, *Pattern Recognition*, vol. 38(12), 2005, 2270-2285.

[12] J. Kittler, M. Hatef, R.P. Duin and J. Matas, On combining classifiers, *IEEE TPAMI*, vol. 20(3), 1998, pp 226-239.

[13] U. Kreβel, Pairwise classification and support vector machines. *In Advances in Kernel Methods: Support Vector Learnings, MIT Press*, Cambridge, MA, 1999, pp 255-268.

[14] X. Lu and Y. Wang and A. K. Jain, Combining Classifiers For Face Recognition, *ICME 2003. Multimedia and Expo.*, 2003.

[15] T. Ojala, M. Pietikainen and D. Harwood, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition*, vol. 29, 1996.

[16] T. Ojala, M. Pietikainen and T. Maenpaa, Multiresolution gray-scale and rotation invarianat texture classification with local binary patterns. *IEEE TPAMI*, vol. 24(7), 2002, pp 971-987.

[17] J. Platt,. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. *In Advances in Large Margin Classiers*, 2000, pp 61-74.

[18] K. Rao and P. Yip, *Discrete Cosine Transform Algorithms, Advantages*, Applications Academic, New York, NY; 1990.

[19] B. Schölkopf and A. J. Smola, *Learning with Kernels*, MIT Press, 2002.

[20] X.Y. Tan and B. Triggs, Fusing Gabor and LBP Feature Sets for Kernel-Based Face Recognition, *AMFG07*, 2007, pp 235-249.

[21] M. Turk and A. Pentland, Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, vol. 3(1), Mar. 1991, pp 71-86.

[22] L. Wiskott, J.-M. Fellous, N. Kruger and C. von der Malsburg, Face recognition by elastic bunch graph matching, *IEEE TPAMI*, vol. 19(7), 1997, pp 775-779.

[23] W. Zhao, R. Chellappa, P.J. Phillips and A. Rosenfeld, Face recognition: A literature survey, *ACM Computing Survey*, vol. 34(4), 2003, pp 399-485.