

Dissimilarity Representations Based on Multi-Block LBP for Face Detection

Yoanna Martínez-Díaz, Heydi Méndez-Vázquez, Yenisel Plasencia-Calaña,
and Edel B. García-Reyes

Advanced Technologies Application Center. 7th Avenue # 21812 % 218 and 222,
Siboney, Playa, P.C. 12200, Havana, Cuba
{ymartinez,hmendez,yplasencia,egarcia}@cenatav.co.cu

Abstract. Face representation is one of the open problems in face detection. The recently proposed Multi-Block Local Binary Patterns (MB-LBP) representation has shown good results for this purpose. Although dissimilarity representation has proved to be effective in a variety of pattern recognition problems, to the best of our knowledge, it has never been used for face detection. In this paper, we propose new dissimilarity representations based on MB-LBP features for this task. Different experiments conducted on a public database, showed that the proposed representations are more discriminative than the original MB-LBP representation when classifying faces. Using the dissimilarity representations, a good classification accuracy is achieved even when less training data is available.

Keywords: face detection, dissimilarity representation, MB-LBP.

1 Introduction

Face detection is the first step in a face recognition system. The more precise this step, the more accurate any further processing [1]. Existing methods to detect faces in images can be grouped into four categories: knowledge-based methods, feature invariant approaches, template matching methods and appearance-based methods [2]. In general, appearance-based methods have shown the best performance, being the most studied and used [1]. However, how to represent and classify faces, are two key factors which are still unsolved problems [1].

Different kinds of features have been proposed for face detection. They are based on physical properties of faces, such as color intensity, texture, vertices, shape and facial features. Among them, the Haar-like features have been one of the most widespread and effectively used [3]. They encode the differences in average intensities of adjacent rectangular regions and can be efficiently computed through the integral image technique. However, the complete (or exhaustive) set of Haar-like features is very large, containing a lot of redundant information. Moreover, these features are not robust to extreme lighting conditions [1].

Object detectors based on the Local Binary Patterns (LBP) descriptor have become more popular, achieving highly competitive results [4]. This can be

attributed to the tolerance of this operator against monotonic illumination variations and its computational simplicity [5]. Among different extensions of the original LBP operator, the Multi-Block LBP (MB-LBP) representation [6] has shown to be more robust to noise and stable for face detection [4]. The MB-LBP compares average intensities of adjacent rectangular blocks in a neighborhood instead of single pixels intensities. Compared to original LBP and Haar-like features, it can capture more information about image structures and shows a better performance [6]. The MB-LBP also allows generating an exhaustive set of features with different scales and locations [6]. This set is much smaller than the Haar-like one, however it is still large and contains redundant information.

When exhaustive sets of features are used, boosting-based methods are usually applied to select the most discriminative features and to build a strong classifier [3,6]. The main drawbacks of these methods are that they need a large set of training samples and the learning process is time consuming [1,7]. To overcome some of these limitations, other learning strategies can be used for appearance-based face detection [1,2,7].

A powerful alternative to those representations based on features are the dissimilarity representations (DR) [8]. They are based on comparisons among objects and can be obtained on top of feature representations. In this way, an object is described by means of its dissimilarities with a set of prototypes [8]. If a suitable dissimilarity measure is defined for the problem at hand, classifiers in a dissimilarity space (DS) can provide very good results [8,9].

In this paper we propose new dissimilarity representations based on MB-LBP features for the problem of face detection, which show to be more discriminative than the original MB-LBP features space. The rest of this paper is organized as follows. In Section 2 we review the MB-LBP representation and the methods used for feature selection and classification. In Section 3, the dissimilarity representations based on MB-LBP features are presented. Experimental results and discussion are drawn in Section 4. Finally, section 5 concludes this paper.

2 Multi-Block Local Binary Patterns Representation

Multi-Block Local Binary Patterns (MB-LBP) features were inspired by the traditional Haar-like rectangular features [6]. In this case, the differences among rectangular regions are encoded by the LBP operator [5]. MB-LBP compares the average intensity of the central rectangle, r_c , with the average intensities of its adjacent rectangular blocks $\{r_1, \dots, r_8\}$ in a 3×3 neighborhood [6]. Afterwards, it encodes the responses in a binary number, which can be further converted to a decimal number:

$$\text{MB-LBP} = \sum_{i=1}^8 S(r_i - r_c)2^i \quad (1)$$

where,

$$S(X) = \begin{cases} 0 & \text{if } X \leq 0 \\ 1 & \text{if } X > 0 \end{cases} \quad (2)$$

Since MB-LBP compares regions instead of pixels, they can capture larger scale structures than original LBP. Compared with Haar-like features, MB-LBP can encode more information, describing more complex image structures [6]. Similar to Haar-like features, the average intensities of rectangular blocks can be computed at constant time by using the integral image [6]. Given a sub-window size, the number of rectangles at various scales, locations and aspect ratios of MB-LBP features (exhaustive set) is about 5% of the total number of Haar-like features. Even when the exhaustive set of MB-LBP features is much smaller than the Haar-like features, it also contains a mass of redundant information [6]. Hence, a feature selection method should improve the classification accuracy and make the training process less expensive.

2.1 Feature Selection and Classification

Despite the successful advances on face detection [1,2], obtaining discriminative features and classifiers to achieve a good compromise between computational cost and classification accuracy, remains as a challenging issue. In previous works, AdaBoost algorithm [10] has been used both to select the most discriminative MB-LBP features and to construct a strong classifier [6]. In the boosting mechanisms used for face detection, each weak classifier is usually trained with a single feature and needs to evaluate thousands of features to minimize an error function. This process is very time consuming and usually requires a large set of training samples for a better performance [7]. Furthermore, if the data is noisy, the classifier can be trapped by local minima.

Other feature selection methods and learning approaches have been used to avoid the above limitations in the context of face detection, such as the Sequential Forward Selection algorithm [7] and the Support Vector Machines (SVM) [1]. As an alternative, the Minimal-Redundancy-Maximal-Relevance (MRMR) criterion, which has been effectively used for different face analysis problems [11], can be also used for the purpose of face detection.

Besides the mentioned strategies, there is another possibility which has not been exploited in the context of face detection. The so-called dissimilarity based classification has shown a good performance for high dimensional data such as images and for two class problems [8,9], therefore it could be useful for this task.

3 Dissimilarity Representations Based on MB-LBP

The DS was proposed in the context of dissimilarity-based classification. It was postulated as an Euclidean vector space, implying that classifiers proposed for feature spaces can be used there as well [8]. Under this framework, the objects are represented in terms of their dissimilarities to a set of so-called prototypes. Given a training set T , the prototypes $R = \{r_1, r_2, \dots, r_k\}$ can be a subset from T , or T itself. The dissimilarity representation of an object x is then defined as:

$$D(x, R) = [d(x, r_1) \ d(x, r_2) \ \dots \ d(x, r_k)], \quad (3)$$

where $d(x, r_i)$ is the dissimilarity value between x and each prototype $r_i, i = 1 \dots k$. Finally, a dissimilarity matrix $D(T, R)$ of all objects against all prototypes is obtained.

More compact representations can be achieved using prototype selection methods [9]. They find the prototypes that better represent the data in the DS. In this way, a DS of lower dimensionality can be constructed, leading to lower computational and storage costs. Besides, the selection of a suitable dissimilarity measure is a key point in DR [8]. The aim of this work is to introduce the use of DR for face detection, so we have built two new DS based on MB-LBP features using Euclidean and L1 distances as dissimilarity measures.

4 Experimental Results and Discussion

In this section, two experiments are conducted in order to evaluate the DR based on MB-LBP features, for face representation and classification in the process of face detection. For this purpose we have used a dataset of 12788 gray-scales images of 24×24 pixels, divided into 4916 faces and 7872 non-faces [12]. Different configurations and learning schemes are evaluated, and three terms are used to measure the performance of the methods: False Positive rate (FP), False Negative rate (FN) and Error Rate (ER). FP measures the non-faces incorrectly classified as faces; FN indicates how many face images were incorrectly classified as non-faces; and ER is the ratio between the number of misclassified images (faces and non-faces) to the total number of analyzed images.

4.1 Classification in DS Based on the Exhaustive Set of MB-LBP

The first experiment was developed to compare our two DR based on the exhaustive set of MB-LBP features with the original MB-LBP features space. AdaBoost classifier [10] is then used on the different representations. Since AdaBoost selects a subset of features (prototypes in the DS) for classification, we decided to use all training data as prototypes. A subset composed by 6000 images (2000 faces + 4000 non-faces) of the dataset was used. It was randomly split into 2000 images for training and 4000 for testing. To reduce the variance of the estimation, this procedure was repeated 10 times and the comparison is then based on the averaged classification accuracy from all partitions. For each one of the 10 generated partitions, an AdaBoost classifier with 50 weak learners was trained and tested. From now on we will refer to this as Protocol 1.

Table 1 shows the performance of the AdaBoost classifier using the proposed representations and the original MB-LBP features space (FS), in terms of the average FN, FP and ER for the 10 partitions. In all cases, standard deviations are lower than 0.01. It can be seen from the table that the lowest errors are obtained for the two proposed DS, outperforming remarkably the original MB-LBP representation and achieving more than 99% of correct classification. The Euclidean distance on the MB-LBP features is slightly better than the L1 distance, but in general both measures performed very well. We can benefit from

Table 1. Performance of AdaBoost algorithm using MB-LBP features in DS and FS

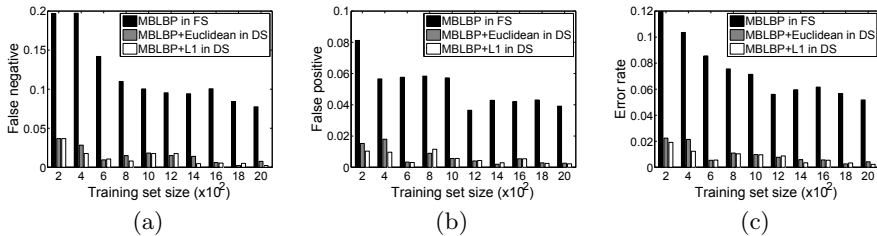
	Average FN	Average FP	Average ER
MB-LBP in FS	0.0908	0.0379	0.0554
MB-LBP + Euclidean in DS	0.0037	0.0028	0.0031
MB-LBP + L1 in DS	0.0045	0.0034	0.0037

the use of DS constructed on top of MB-LBP features with this two common distances, since a more discriminative representation is obtained.

Moreover, the DR allow reducing the dimensionality of the training data. For the given sub-window size of 24×24 , the dimension of the original MB-LBP exhaustive set is 8644. Despite the fact that we did not use prototype selection, in this case, the dimensionality of the DS obtained for training is only 2000. In order to analyze the behavior of the proposed representations when less training data is used, different training set sizes are evaluated. Figure 1 shows the obtained classification errors for the different representations. As it can be appreciated, the DS based on Euclidean and L1 distances always outperform the original representation of MB-LBP features, using AdaBoost classifier. We can see that, when the dissimilarities among objects are taken into account, a good classification accuracy can be obtained using even small training set sizes.

4.2 Classification in DS Based on a Subset of MB-LBP

In the second experiment our aim is to corroborate the robustness of the proposed representations using a different configuration. In this case, the MRMR criterion is used to select the most discriminative MB-LBP features and the two DS are obtained based on the reduced set of features. The basic idea of MRMR is to use information theory, finding the features that are maximally relevant on the target class but minimally redundant [11]. It allows one to select good features at very low cost, so the computation of the dissimilarity matrix become less expensive. Moreover, it removes the redundant information which might have influence on the DR. The two obtained DS are then compared with the reduced original MB-LBP features space, using AdaBoost and SVM classifiers. In this experiment the SVM classifier was used with a linear kernel, with the aim at making the computational cost less expensive.

**Fig. 1.** Performance of MB-LBP features in DS and FS using AdaBoost algorithm with different training set sizes, in terms of (a) FN, (b) FP and (c) ER

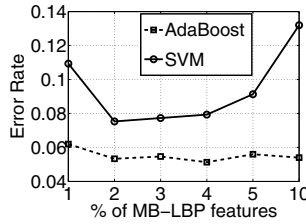


Fig. 2. Performance of Adaboost and SVM classifier with different dimensions of MB-LBP features selected by MRMR criterion

In order to choose an optimal number of features selected by the MRMR criterion, we have tested various percentages of the total number of features (8644), using the two classifiers in the original space. For this purpose, we have used a different subset of 3000 images, randomly divided into two equal parts for training and testing. The obtained error curves are plotted on Fig. 2. We decided to keep only 2% of the total number of features. It can be seen that, for this value a low error is achieved by the two classifiers.

We have constructed the DR based on the reduced set of MB-LBP features using both AdaBoost and SVM classifiers. The results obtained for AdaBoost are shown on Table 2. For this experiment we used the Protocol 1 defined on subsection 4.1. It can be seen from Table 2 that also using the reduced set of features, the DR remarkably outperform the original MB-LBP representation. Notice that although these results do not improve the ones presented on Table 1, the computation of the dissimilarity matrices is less expensive in this case.

We have decided to use a prototype selection method before applying the SVM classifier, with the aim to obtain a more compact DR, making the learning process less expensive. In order to choose a suitable method and an appropriate number of prototypes for the classification in the DS, we compared three methods with different numbers of prototypes. Based on results of previous works [9] we have tested: RandomC, KCentres and forward selection optimizing the leave-one-out nearest neighbor error in the training set (FS+NN error). For this experiment, a different subset of the dataset is used, composed by 500 face images and 1500 non-face images. The data was randomly divided into equal parts for training and testing, and the procedure was repeated 10 times. The classification accuracy of the SVM classifier in the obtained spaces, using both Euclidean and L1 distances, are compared in Fig. 3. In general, the KCentres method shows the

Table 2. Performance of AdaBoost algorithm using reduced set of MB-LBP features in DS and FS

	Average FN	Average FP	Average ER
MB-LBP in FS	0.1078	0.0523	0.0706
MB-LBP + Euclidean in DS	0.0050	0.0025	0.0033
MB-LBP + L1 in DS	0.0054	0.0035	0.0041

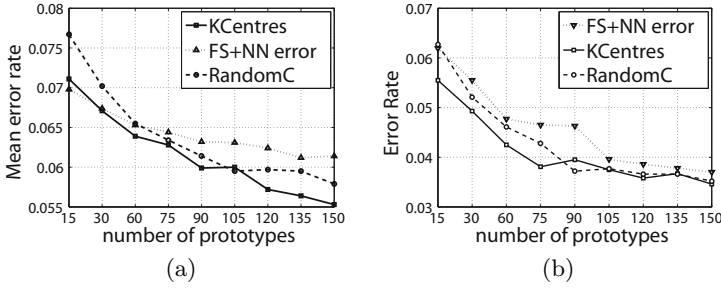


Fig. 3. Average error rates of the SVM classifier in DS using (a) Euclidean distance and (b) L1 distance with three prototypes selection methods

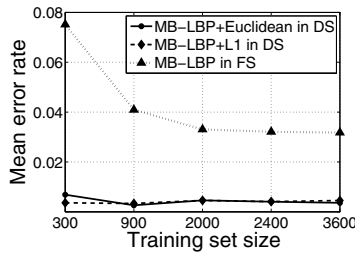


Fig. 4. Performance of SVM classifier with different training set sizes

lowest error rates. We decided to use this method selecting 60 prototypes, since this number of prototypes represents a good compromise between classification accuracy and computational cost.

The behavior of the SVM classifier on the DS obtained by using KCentres with 60 prototypes based on the reduced set of MB-LBP features, is compared with the classification in the reduced original MB-LBP feature space. The same images used on Protocol 1 are used in this experiment, but in this case different training set sizes are tested. For each one of the tested training sizes, the experiment was repeated 10 times, randomly splitting the data for training and testing. The average error rates are shown on Fig. 4. It can be seen that the classification accuracy of the SVM classifier on DS is significantly better for all tested training sizes. Moreover, it is shown that in the original MB-LBP feature space, the size of the training data has a direct influence on the classification accuracy, while the results on the DS exhibit a more stable behavior.

5 Conclusion and Future Work

In this paper we propose new dissimilarity representations based on MB-LBP features for face detection, using Euclidean and L1 distances. The proposed representations have shown to be more discriminative than the original MB-LBP

features space. Moreover, they allow us to obtain a significantly better classification accuracy even when using a few amount of training data. Despite the fact that the redundant information in the exhaustive set of MB-LBP features does not have significant influence on the proposed representations, the use of the reduced set of features makes the computation of the dissimilarity matrices less expensive. This paper can be considered a good starting point for further developments on the use of dissimilarity representations for face detection. As future work we pretend to explore different DR based on other features using suitable dissimilarity measures.

References

1. Zhang, C., Zhang, Z.: A survey of recent advances in face detection. Technical Report MSR-TR-2010-66
2. Yang, M., Kriegman, D., Ahuja, N.: Detecting faces in images: A survey. *IEEE Transactions on PAMI* 24(1), 399–458 (2002)
3. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *IEEE Computer Society Conference on CVPR*, vol. 1, pp. 511–518 (2001)
4. Trefný, J., Matas, J.: Extended set of local binary patterns for rapid object detection. In: *Proceedings of the Computer Vision Winter Workshop 2010* (2010)
5. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition* 29(1), 51–59 (1996)
6. Zhang, L., Chu, R., Xiang, S., Liao, S., Li, S.: Face detection based on multi-block lbp representation. In: *Proceedings of IAPR/IEEE ICB* (2007)
7. Louis, W., Plataniotis, K.N.: Co-occurrence of local binary patterns features for frontal face detection in surveillance applications. In: *EURASIP* (2011)
8. Pekalska, E.: *The Dissimilarity Representation For Pattern Recognition. Foundations and Applications*. World Scientific (2005)
9. Pekalska, E., Duin, R.P., Paclík, P.: Prototype selection for dissimilarity-based classifiers. *Pattern Recognition* 39(2), 189–208 (2006)
10. Freund, Y., Schapire, R.E.: A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting. In: Vitányi, P.M.B. (ed.) *EuroCOLT 1995*. LNCS, vol. 904, pp. 23–37. Springer, Heidelberg (1995)
11. Akbaş, E., Yarman-Vural, F.T.: Design of a Feature Set for Face Recognition Problem. In: Levi, A., Savaş, E., Yenigün, H., Balcısoy, S., Saygın, Y. (eds.) *ISCIS 2006*. LNCS, vol. 4263, pp. 239–247. Springer, Heidelberg (2006)
12. Carbonetto, P.S.: Robust object detection using boosted learning. Technical report, Department of Computer Science, University of British Columbia, Vancouver, Canada (April 2002)