SVM classifier based face detection system using BDIP and BVLC moments

Conference Paper · August 2011

DOI: 10.1109/ATC.2011.6027481

5 authors, including:

Tien Dzung Nguyen
Hanoi University of Science and Technology

Quy Thanh Tran
University of Economics Ho Chi Minh City

Trang Quỳnh
The National Institute of Cardiovascular Diseases

Thang Manh Hoang
Hanoi University of Science and Technology, Hanoi, Vietnam

SVM classifier based face detection system using BDIP and BVLC moments

Tien Dzung Nguyen, Quy Tran Thanh, Thang Man Duc, Trang Nguyen Quynh, Thang Manh Hoang

School of Electronics and Telecommunications, Hanoi University or Science and Technology, Vietnam

Email: ntdzung@ieee.org

***Abstract-* In this paper, a support vector machine (SVM) classifier has been used to detect a face in an authentication application. A face candidate is first allocated from the input frame and then normalized to 200x200 pixels images. The textureness of candidates is then measured by the combination of BDIP and BVLC moments and classified into face and non-face ones by a SVM classifier which is known as efficient classification tool. In SVM learning, a DB of 2500 faces and 2500 non-faces has been created under different light conditions and face expressions. The experiments showed that the effectiveness of the used features for SVM based classification issue in the face-detection system.**

Keywords - BDIP, BVLC, SVM, texture features.

I. INTRODUCTION

With the ubiquity of new information technology and media, more effective and friendly methods for human computer interaction are being developed which do not rely on traditional devices such as keyboards, mice, and displays. Furthermore, the ever decreasing price/performance ratio of computing coupled with recent decreases in video image acquisition cost imply that computer vision systems can be deployed in desktop and embedded systems[1,2,3].

Face detection from a single image is a challenging task because of variability in scale, location, orientation, and pose. Facial expression, occlusion, and light conditions also change the overall appearance of the faces. Face-detection techniques typically can be classified in two categories: model-based and feature-based methods. The first one assumes that a face can be represented as a whole unit. Several statistical learning mechanisms are explored to characterize face patterns, such as neural network, probabilistic distribution, support vector machines, principal components analysis, naïve Bayesian classifier, boosting algorithms, etc. The second method treats a face as a collection of components. Important facial features such as eyes, nose and mouth should be first extracted, and the face can be detected by using their geometrical features and relationships [4, 5, 6, 7, 8].

However, face detection also faced certain difficulties such as:

- The posture of the face in an image can vary.
- The presence of the details is not characteristics of the human face as beard, glasses.
- The different facial expressions on the faces as happy, sad, surprised.
- Conditions of the image, especially the brightness and image quality, camera quality, different size of human face, etc.

The above difficulties have led many methods to solve determine who will face the inevitable flaws and do not achieve certain accuracy.

In our proposed method, from the set of values BDIP and BVLC are calculated on the input image, we layered to form the feature moments then turn into the SVM classifier to detect face or non-face. We have tested our method on several image sets with many different conditions and rates of face detection accuracy over 95% face always.

The paper is organized in three sections. After introduction in section I, section II deals with the proposed method, where the concepts of SVM classifier and BDIP and BVLC moments are discussed to apply into classification and then detection of a face. Finally, the last section discusses about the experimental results and performance evaluation of the proposed method followed by conclusion and further research trend.

II. PROPOSED FACE DETECTION SYSTEM

A block diagram of the proposed face detection system is illustrated in Fig.1. In a captured frame from the camera, Haar features [9] are utilized first to extract all face candidates which next are forwarded to the SVM based detection module. In the pre-processing step, the face candidate images will be normalized to the size of 200x200 pixels and then enhanced by contrast stretching and histogram equalization [10]. In BDIP BVLC moments classification, BDIP and BVLC values of the enhanced images are computed and classified into four classes, and then the first and second moments of BDIP and BVLC values are evaluated, which are used to form a feature vector for solving the classification problem. A SVM classifier finally tests and classifies the input images into face or non-face classes based on the computed features.

A. Conventional features

In this section, we briefly introduce the concepts of BDIP and BVLC moments formed in a feature vector.

a) BDIP

In human vision, edges and valleys [11] in an image are very important features, especially valleys are fundamental in

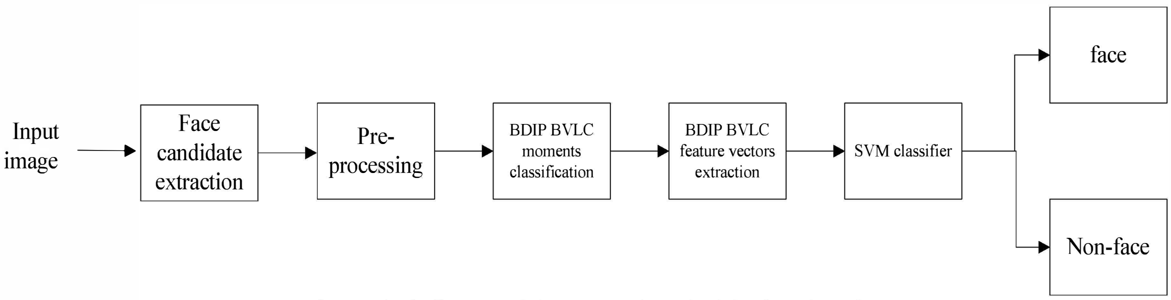


Fig. 1. Block diagram of the proposed method for face detection system

the vision perception of an object shape [12, 13]. BDIP (block difference of inverse probability) [11] is the texture feature which measures the variation in intensities of an image block. It effectively extracts edges and valleys. BDIP of a block of size WxW is defined as:

$$\beta^k(l) = \frac{\frac{1}{W^2}\sum_{(i,j)\in B_l^k}\left[\max\limits_{(i,j)\in B_l^k} I(i,j) - I(i,j)\right]}{\max\limits_{(i,j)\in B_l^k} I(i,j)} \quad (1)$$

where I(i,j) denotes the intensity of a pixel (i,j) in the block B, l(u,v) is the location of block in image and k is the maximum distance of pairs of pixels in block. That means W=k+1. The larger the variations of intensities, the higher the value of BDIP [13].

b) BVLC

BVLC (block variation of local correlation coefficients) [11] represents the variation of block-based local correlation coefficients (LCCs) according to four orientations (-90^0, 0^0, -45^0, 45^0). It is known to measure texture smoothness well. The LCC in each orientation is given by:

$$\rho^k(l) = \frac{\frac{1}{W^2}\sum_{(i,j)\in B_l^k} I(i,j)I(i+\Delta_i(k), j+\Delta_i(k)) - \mu_l\mu_{l+\Delta(k)}}{\sigma_l\sigma_{l+\Delta(k)}} \quad (2)$$

Where l(u,v) denotes the location of the block in an image, μ_l, δ_l the local mean and standard deviation of the block, respectively. Here $\Delta(k) = (\Delta_i(k), \Delta_j(k))$ stands for four orientations in which the block is shifted, respectively. Thus, $\mu_{l+\Delta(k)}$ and $\delta_{l+\Delta(k)}$ represent the mean and standard deviation of the block shifted by Δ(k). And BVLC is defined as:

$$\gamma^k(l) = \max_{\Delta(k)\in O_4}[\rho^k(l)] - \min_{\Delta(k)\in O_4}[\rho^k(l)] \quad (3)$$

$$O_4 = \{(0,k),(k,0),(k,k),(k,-k)\}$$

The larger the degree of roughness there is in a block, the higher the value of BVLC[14]

a) BDIP and BVLC features extraction

An input captured frame is divided into non-overlapping blocks of fixed sizes and then BDIP and BVLC values are computed in each block. These values are classified in four classes in such a way described in Fig.3.

(a) (b) (c)

Fig. 2. Original image and result images of BDIP and BVLC operators. (a) Original image, (b) BDIP operator, (c) BVLC opeator.

At first, the BDIP values in all blocks are divided into two subclasses with threshold is the mean of BDIP values. Then, each subclass is further divided in two subclasses using the mean of all BDIP values in each group as the threshold. BVLC values is also classified in the same way. The first and second moments of the BDIP and BVLC values for each class are then extracted to form an input feature vector for SVM classifier. The feature vector $x^k$for an input image is defined as :

$$x^k = [\mu_n^k(D), \sigma_n^k(D), \mu_n^k(V), \sigma_n^k(V)] \quad (4)$$

Where $\mu_n^k(D), \sigma_n^k(D), \mu_n^k(V), \sigma_n^k(V)$ denote the first and second moments of BDIP and BVLC for the n class for each k, respectively.

The dimension of the feature vector is one of the most important factors in detectionprocess, which determines the amount of storage space for the vector, the detection accuracy and classification time in SVM.

B. SVM classifier

a) Theory of SVM

Let $\{(x_i, y_y), i = 1,2,\dots,N\}$ be a training set input S , $x_i \in R^n$ and labeled with $y_i \in \{-1,1\}$. SVM arlgorithm would learn a linear separating hyperplane to separate samples into two classes. It is equivalent to find hyperplane H: w.x+b=0 and two hyperplanes parallel to H and with equal distances to it: H_1: y=w.x+b=1 and H_2: y=w.x+b=-1 with the condition that there are no data point between H_1 and H_2 and the distance between H_1 and H_2 is maximized. This problem can be resovled if the constraint $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1,\ i=1,\dots,N$ (1) is taken into accout when optimizing the cost function $\mathbf{w}^T.\mathbf{w}/2$.

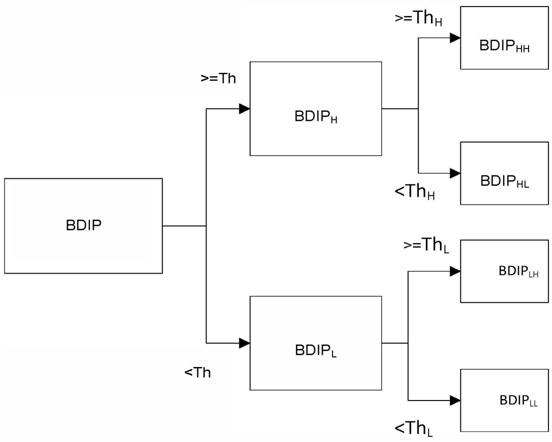


Fig.3. BDIP classification

Introducing Lagrangian multiplexer $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\dots,\alpha_N)\geq 0$ we have the following Lagrangian :

$$\Im(w,b,\alpha)=\frac{1}{2}w.w-\sum_{i=1}^{N}\alpha_i y_i(w.x_i+b)+\sum_{i=1}^{N}\alpha_i \quad (5)$$

Whereas data with linear separability may be analyzed with a hyperplane, linearly non-separable data, which are usually encountered in practical classification problems[16, 17], are analyzed with an appropriate nonlinear operator $\boldsymbol{\varphi}(\cdot)$ for mapping the input feature vector into a higher dimensional feature.

Omitting the details of mathematic derivation we finally can classify a new sample x with a nonlinear classifierwhere K is called a kernel function[18].

$$\mathrm{f(x)} = \mathrm{sgn}\left(\sum_{\mathrm{i=1}}^{\mathrm{N}}\alpha_\mathrm{i}\mathrm{y}_\mathrm{i}\mathrm{K}(\mathrm{x}_\mathrm{i},\mathrm{x})+\mathrm{b}\right) \quad (6)$$

In this paper, we use a nonlinear SVM classifier with Gaussian RBF kernel.

b) Performance indicators

The Receiver Operating Characteristic (ROC) analysisis used to evaluate SVM classifier performance. The constructing ROC curve is based on two statistic factors which are the sensitivity and the specificity, and the accuracyof SVM is then computed [15].

The best possible classifier would yield when the ROC curve tends to the upper left corner representing 100% sensitivity and 100% specificity.

Another parameter is used to estimate SVM performance is AUC. The SVM classifier is called ideal with 100% accuracy when the AUC of its ROC approaches 1 and when AUC equals 0.5, SVM is random classifier. AUC is given by:

$$AUC=\frac{\sum_{i=1}^{n^+}\sum_{j=1}^{n^-}1_{\delta_{ij}}}{n^+n^-} \quad with\ \delta_{ij}=f(x_i^+)-f(x_i^-) \quad (7)$$

Where f(.) is denoted as decision function of classifier, x^+,x^- respectively denote the positive and negative samples and n^+, n^- are respectively the number of positive and negative examples and the$1_{\epsilon>0}$ is defined as 1 if the predicate ϵ is holds and 0 otherwise.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Preliminaries for experiment

The first set consisting of 2500 face image and 2500 non-face images has been prepared for training phase.For the testing phase, the second set of other 250 face images and 250 non-face images has been used to record the performance.

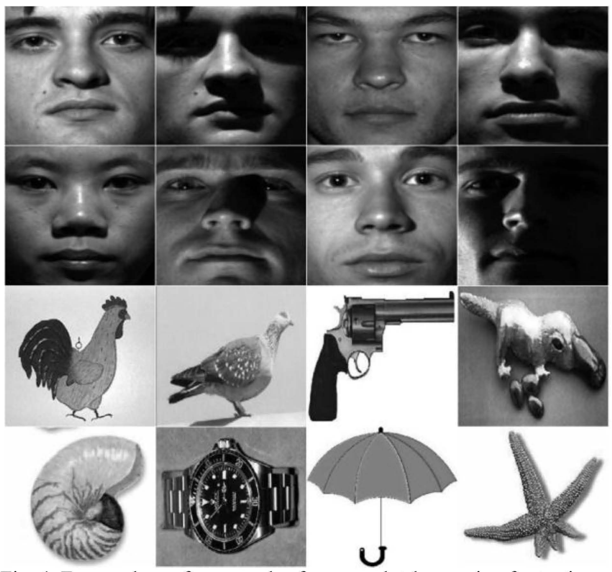


Fig. 4. Face and non-face samples from our database using for testing.

The SVM classifier performance is reflected by ROC curve which is constructed by alternating threshold in each test[19]. In the calculation of BDIP and BVLC values, the block size is experimentally chosen as 2x2. The number of classes of block classification in the computation of BDIP and BVLC moments is determined as 4.

Fig.5. represents ROC curve of the trained SVM classifier when according to the use of BDIP moments, BVLC moments and combination of BDIP and BVLC moments as a feature vector. The feature vector dimension is chosen as 8. BDIP and BVLC vector consist both of the first and second BDIP and BVLC, respectively. And the combination vector consist of the first moment. C=100, $\sigma=2$. The best value of AUC is 0.96 when using the vector with combination of BDIP and BVLC moments.

Fig. 6 shows some samples from the testing set which are incorrectly detected. The reason for this error probably comes from the different variations in the brightness of the face samples or the complex textures of the non-face samples.

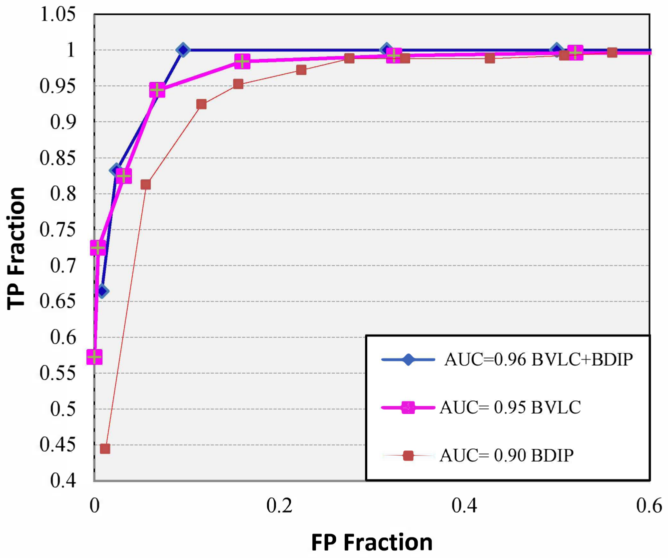
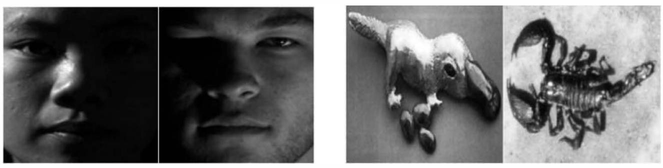


Fig.5. Comparison of ROC curves of the SVM classifiers using various features.



(a) (b)

Fig.6. Some incorrect samples test. (a). False Negative images, (b). False Positive images.

The SVM performance is demonstrated in Table 1, where the trade-off for gettingthe highest accuray rate (95.60%)can be achieved by utilization ofthe combination of BDIP and BVLC moments. The results proved that the proposed approach is effective and auspicious forface detection basedapplication.In future work, the solution to overcome the mentioned error will be addressed.

TABLE I. TESTING RESULTS

Moments	**True Postive**	**True Negative**	**False Negative**	**False Postive**	**Accuracy**
BDIP	196	192	54	58	77.60%
BVLC	236	233	14	17	93.80%
BDIP + BVLC	250	228	0	22	95.60%

IV. REFERENCE

[1] A.Pentland, “Looking at People”, *IEEE Trans.Pattern Analysis and Machine Intelligence*, vol.22, no.1, pp.107-119, Jan.2000.

[2] A.Pentland, “Perceptual Intelligence”, *Comm.ACM*, vol.43, no.3, pp.35-44, 2000.

[3] A.Pentland, T.Choudhury, “Face Recognition for Smart Environments”, *IEEE Computer*, pp.50-55, 2000.

[4] Kyongpil Min, Junchul Chun, Goorack Prak, “A Nonparametric Skin Color Model for Face Detection from Color Images”, *PDCAT 2004, LNCS 3320*, pp. 116-120, Springer-Verlag Berlin Heidelberg, 2004.

[5] Selin Baskan, M. Mete Bulut, Volkan Atalay, “Projection based method for segmentation of human face and its evaluation”, *Pattern Recognition Letters 23*, pp. 1623-1629, Elsevier Science, 2002.

[6] Gines Garcia Mateos, Alberto Ruiz, and Perdo E. Lopez-de-Teruel, “Face Detection Using Integral Projection Models”, *SSPR&SPR 2002, LNCS 2396*, pp. 644-653, Springer-Verlag Berlin Heidelberg, 2002.

[7] V.Jain and E. Learned-Miller, “A benchmark for face detection in unconstrained settings”, *Technical report,* University os Massachusetts, Amherst, 2010.

[8] M.Enzweiler, D.M.Gavrila, “Monocular pedestrian detection: Survey and experiments”, *IEEE Trans. On PAMI,* 31(12): 2179-2195, 2009.

[9] Cuong Nguyen Khac, Ju H.Park, Ho-Youl Jung, “Face detection using Variance based Haar-like feature and SVM”, *World Academy of Science,* Engineering and Technology 60, 2009.

[10] J.C.Terrillon, M.N.Shirazi, D.M.Reynolds, M.Sadek, Y.Sheng, S.Akamatsu,K.Yamamoto, “Invariant Face Detection in Color Image Using Orthogonal Fourier-Mellin Moments and Support Vector Machines”, *ICAPR 2001, LNCS 2013*, pp. 83-92, Springer-Verlag Berlin Heidelberg, 2001.

[11] Chun Y. D., Seo S. Y., and Kim N. C., “Image retrieval using BDIP and BVLC moments,” *IEEE Trans. On CSVT*, Vol. 13, No. 9, Sep. 2003.

[12] D. E. Pearson and J. A. Robinson, “Visual communication at very low data rates,” *Proc. IEEE*, vol.73, pp. 795-812, Apr.1985.

[13] Y. J. Ryoo and N. C. Kim, “Valley operator extracting sketch feature: DIP,” *Electron. Lett.*, vol. 248, pp.461-463, Apr. 1988.

[14] T. D. Nguyen, S. H. Kim and N. C. Kim, “Surface extraction using SVM-based texture classification for 3D fetal ultrasound image,” *Proc. 1st Int. ICCE 2006, Hanoi, Vietnam*, vol. 1, pp. 285-290, Oct.2006.

[15] C. Metz and H. Cronman. “Statistical significance test for bi-normal ROC curves,” *Journal of Mathematical Psychology*, vol. 22,pp. 218-243, 1980 .

[16] J.C.Platt, “Fast training of SVMs using sequential minimal optimization,” *Advance in Kernel Method – Support Vector Learning*, B. Scholkopf, C. Burges and A. Smola, eds. MIT Press, Cambridge, Mass., 1998.

[17] I. El-Naqa, Y. Yang, M. N. Wernick, N. P. Galatsanos, and R. M. Nishikawa, “A support vector machine approach for detection of microcalcifications,” *IEEE Trans. On Medical Imaging*, vol.21, no.12, pp. 1552-1563, Dec. 2002.

[18] Klaus-Robert Müller, Sebastian Mika, Gunnar Rätsch, Koji Tsuda, and Bernhard Schölkopf, “An Introduction to Kernel-Based Learning Algorithms”, *IEEE Transactions on neural networks*, vol.12, no.2, March 2001.

[19] http://www.anaesthetist.com/mnm/stats/roc/Findex.htm, last access April 4th, 2011.