

Joint Integral Histogram based Adaboost for Face Detection System

Ameni Yengui Jammoussi
ICOS Research Unit
University of Sfax, Sfax Engineering School,
BP W, 3038 Sfax, Tunisia

Dorra Sellami Masmoudi
ICOS Research Unit
University of Sfax, Sfax Engineering School,
BP W, 3038 Sfax, Tunisia

ABSTRACT

Face detection is a crucial step in many vision applications. Since the Viola and Jones face detector, many feature extraction approaches based Adaboost are proposed.

This paper presents a novel approach to extract effective features for face detection system. Both LBP and Three Patch LBP (TPLBP) with joint integral histogram are used to extract features. The joint integral histogram was firstly proposed for stereo matching application. Its effectiveness has motivated us to apply it harnessing its advantages. The evaluation of the novel features based Adaboost was done using the CMU-MIT frontal face data set. Experimental results show that its performance is noteworthy especially for the earlier stages. In fact, with few numbers of the new features we can achieve the max detection (1) and reduced false positive rate (0.28).

Keywords

Face detection, Adaboost, TPLBP, LBP, Haar feature, Joint Integral histogram image.

1. INTRODUCTION

Face detection provides interesting challenges due to the requirements for pattern classification, learning and recognition techniques. The face detection problem can be defined as the process of finding faces in an arbitrary image under variations in illumination conditions, background, pose, scale, expression, etc...

It can be treated as a two-class (face/non-face) classification problem. Earlier researches are characterized by the use of simple features such as color, texture. But, these methods break down because of the complexity of image background in real situations.

In the last past years, we have witnessed a considerable progress on face detection by emerging new approaches.

The recent proposed effective methods are learning based ones. They include support vector machine, neuronal network and AdaBoost training algorithms. Viola and Jones algorithm introduced in 2001 [1] falls within face detection based Adaboost approach. Since this work, the state-of-the-art face detection algorithms are the Adaboost based training method and its improvement.

The Adaboost can be seen as a combination of cascade of weak classifiers in order to obtain a strong classifier. These weak classifiers are derived from simple Haar features. To improve the efficiency of this work, many related works are proposed

according to either alternative boosting algorithms or alternative feature classifiers.

Due to the limitation of the Haar-like features in lighting variation conditions, many researchers have been proposed to counter these issues. In the context of face recognition, the Local Binary Patterns introduced by Ojala et al [2] which is a texture descriptor feature shows robustness to illumination variations. Thus, various face detection systems have been proposed using LBP (Hadid et al), Multi-block Local Binary Patterns [3], Locally Assembled Binary features [4] and the HLBP feature which combines the advantages of both haar and LBP [5].

In this paper, we combine different features for improving face detection. The main goal behind this association is to take advantages of their complementary benefits.

Accordingly, we join the discriminating of the Haar-like features and the invariance of the LBP ones.

The remaining of this paper is organized as follows. Section 2 describes an overview of Viola and Jones algorithm. In section 3, the general feature extraction process is firstly presented and secondly the LBP based methods are explained. Section 4 briefly reviews Adaboost based face detection algorithm and section 5 presents experiments and results and shows the performance of our face detector.

Finally, section 6 summarizes our work and draws some conclusions.

2. OVERVIEW OF VIOLA AND JONES FACE DETECTOR

The face detection problem received considerable attention among researches owing to the impressive performance of Viola and Jones face detector. In fact, the boosted-based detector proposed by Viola and Jones [1] is the most important algorithm in real time face detection.

In their work, a frontal face detection system is developed which achieved very good results with respect to accuracy and speed. The success of their work is achieved due to three main following contributions. The first contribution is the cascade structure in order to improve computational efficiency and reduce the number of false positives. The second contribution leads to build the boosted classifier by selecting a few number S of features from a huge amount of computed features. This feature selection process is carried out by using a learning Adaboost algorithm (Freund and Shapire 1995). The Adaboost algorithm, as mentioned above, is able to construct a strong classifier by a linear combination of weak classifiers. These strong classifiers are combined in a cascade structure in order to

improve computational efficiency and reduce the number of false positives.

The last contribution of Viola and Jones framework is a simple rectangular Haar-like feature which can be extracted in one pass time using a new image representation called integral image. Within this representation, the computation of Haar-like features is easier, time constant and scale independent.

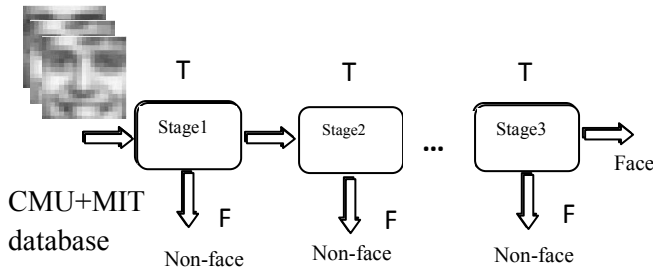


Fig 1: The cascade structure

The number of features in different steps of the cascade structure increases and this makes the strong classifiers more and more discriminating, and even though they are more expensive than the strong classifiers in the early step. Experiments showed that since the first and second layers a large number of negative sub-windows are rejected. These two layers consist only of a small number of features. In the last layers, the classification of face and non-face becomes a hard task. This can be explained by the fact that only face and face-like rare structures reach these steps requiring further processing.

The Viola and Jones technique is considered as a breakthrough in the face detection. But, of course it is not the end of face detection researches. Following this face detector, many researchers present their work mainly addressing two issues: How to develop more effective features to represent the face pattern and how to classify accurately samples using the chosen representation of the face?

From the view of extraction, Haar-like features present some limitations due to its sensitivity to lighting variations.

Thus, there is a need to improve the efficiency of face detection by improving the robustness of the extracted features. In our work, we overcome Haar-like features limitations by introducing LBP based methods.

3. FEATURE EXTRACTION PROCESS

In the Adaboost algorithm, the basic concept is to select most important features. The fundamental features to the whole approach are Haar-like ones which use the integral image for fast evaluation. Nonetheless, Haar-like features and its variants are not the only visual descriptors that are efficient in face detection. Other features like LBP and its variants are a good candidate for the feature extraction process characterized by the robustness to the illumination and the invariance to the rotation of such operator. In this work, the proposed approach consists on creating the image based operator (LBP and its variants) and evaluating the feature template using integral graphs.

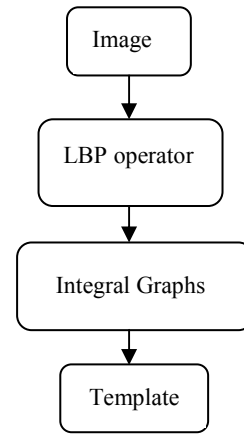


Fig 2: A generic graph of the feature extraction process

3.1 LBP and its variants

3.1.1 Local Binary Patterns (LBP)

LBP was first proposed as a gray level invariant texture primitive extractor. LBP operator describes each feature by its relative gray level to its neighboring pixels. If the gray level of the neighboring pixels is higher or equal, the value is set to one, otherwise is set to zero.

$$LBP_{P,R} = \sum_{k=0}^{P-1} s(g_k - gc)2^k \quad (1)$$

The operator $LBP_{P,R}$ refers to a neighborhood size of P equally spaced pixels on a circle of radius R that form a circularly symmetric neighbor set. The LBP image obtained is characterized by the robustness to the illumination variations that is shown in figure (3). The original image is presented in the left of the top row with other images with different illumination conditions. The corresponding LBP images (bottom row) are identical based on the human vision.

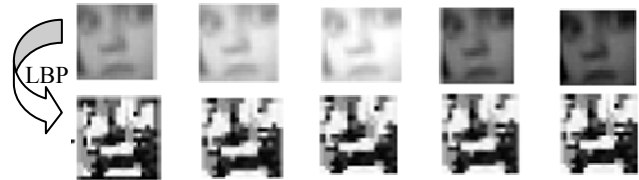


Fig 3: LBP robustness to illumination variation

As mentioned above, the basic LBP operator considers a small neighborhood of pixels and thresholds its values by the central pixel one. Another operator called Three Patch local binary pattern TPLBP [6] compares a center patch with nearby patches instead of single pixel value intensities. The effectiveness of such operator has been demonstrated in the texture segmentation system by its ability to detect local texture properties. In fact, TPLBP has been shown to be highly invariant to local appearance and image style. With the TPLBP, we can extract additional types of local texture information different from the one extracted from the LBP. Thus, as demonstrated in the literature, using patch based descriptors with conjunction of pixel based one can improve the accuracy of the system.

The effectiveness of using the TPLBP alone or in concert with the LBP has motivated us to try it in face detection and to make benefit of its advantages.

3.1.2 Three Patch LBP

For each pixel in the image, a $w \times w$ patch centered on it is considered. And S additional patches distributed uniformly in a ring of radius r around it. The patches along the circle are compared to the central patch [6] (figure 5).

$$TPLBP_{r,S,\omega,\alpha}(P) = \sum_{i=1}^S f(d(C_i, C_p) - d(C_{i+\alpha \text{ mod } S}, C_p))2^i \quad (2)$$

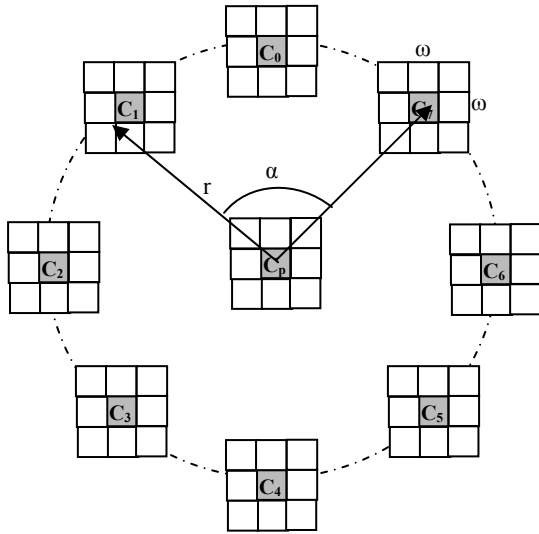


Fig 4: The Three Patch LBP with $\alpha = 2$ and $S = 8$

The TPLBP operator and its variants are robust to variation in illumination and rotation. In the literature, the TPLBP have shown good results in face recognition [6]. Thus, we have used it to evaluate its effectiveness in a face detection system.

3.2 Integral graphs

The first integral graph proposed is the integral image is explained above. The integral image was proposed for fast evaluation of the Haar-like features. For the same purpose, the integral histogram was created through which histogram of any rectangle region in an image can be computed via four references. Recently, a new more generalized technique called joint integral histogram which combines the integral image and integral histogram was proposed in stereo matching application. The effectiveness of the joint integral histogram shown in stereo matching application has motivated us to try it in face detection and to make use of its advantages.

3.2.1 Integral image

The integral image at location x, y is the sum of pixels above and at the left of x, y :

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (3)$$

Where ii is the integral image and i is the original image.

Using the integral image, any rectangular sum can be computed in four array references. In fact, the area of a rectangle R can be calculated as follows using the location of the integral image:

$$R = L_4 - L_3 - L_2 + L_1 \quad (4)$$

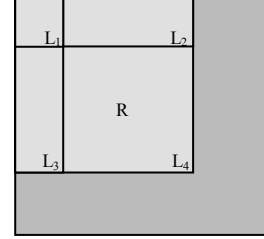


Fig 5: Integral Image

3.2.2 Integral Histogram Image

Like the integral image, the integral histogram image enables us to calculate quickly the features [9]. The concept of the integral histogram image is inspired by the integral image one. So, any rectangle region can be computed via array index operations. The concept of the integral histogram can be explained by a calculation of the integral image to each bin. So, given a $p \times q$ gray image the integral histogram image IHI obtained is a three dimensional array as $IHI[p + 1][q + 1][d]$, where d is the number of histogram bins. The algorithm of the integral histogram is described as follows.

- For a $p \times q$ gray image, create a three-dimensional array $\Upsilon [p + 1][q + 1][d]$ initialized with 0.
- For each bin, calculate the cumulative histogram image. Repeat for $i = 1, \dots, p$: Repeat for $j = 1, \dots, q$:

$$\alpha[k] \leftarrow \alpha[k] + \delta(i, j), k = 0, 1, \dots, d - 1 \quad (5)$$

where $\delta(i, j) = 1$ if the intensity of pixel at location (i, j) belongs to k^{th} bin of histogram; otherwise $\delta(i, j) = 0$.

Once the IHI is computed, the histogram $hr[k]$ of any rectangle region can be determined by 4 array references:

$$hr[k] = IHI[i + \omega][j + h][k] - IHI[i + \omega][j][k] - IHI[i][j + h][k] + IHI[i][j][k],$$

$$k = 0, 1, \dots, d - 1 \quad (6)$$

3.2.3 Joint Integral Histogram

The joint integral histogram is inspired by both the integral images and the integral histograms. The equation of the joint integral is constructed as follows [10]:

$$JIH(x, y, bi) = \sum_{x' \leq x, y' \leq y} \delta(x', y') f(x', y') \quad (7)$$

where $\delta(x', y') = 1$ if $g(x', y') = bi$

As the equation shows, the JIH is composed of a combination of an integral image and an integral histogram.

The contribution of (x, y) to the joint integral histogram is jointly determined by two functions $f(\cdot)$ and $g(\cdot)$. The function $g(x, y)$ determines which bin to increase and $f(x, y)$ determines the value

to increase at that bin. In a joint integral histogram, instead of remembering bin occurrences, the value at each bin indicates an integral defined by two signals. In stereo matching application, the weight of filter could be a function different from the signal to be filtered. To demonstrate the effectiveness of this representation to our work, we conduct two experiments:

(1) the $f(.)$ function is chosen to represent the LBP (or TPLBP) image and the $g(.)$ function represents the integral image of pixels values.

(2) the $f(.)$ function represents the integral image of pixels values and $g(.)$ represents the LBP (or TPLBP) image. We notice the first experiment JIH (1) and the second JIH (2).

3.3 Templates

In this subsection, we introduce the Haar-like features which owe their name to their intuitive similarity with Haar wavelets and were used in the early real-time face detectors. The simple Haar features are reminiscent of Haar basis functions which have been used by Papageorgious et al at 1998. More specifically, they use three kinds of features. A two rectangle feature is the difference between the sum of the pixels within two rectangular regions. Selected regions have the same size and shape and are horizontally or vertically adjacent. A three rectangle feature computes the sum within two outside rectangles subtracted from the sum in a center rectangle.

In this paper, we further exploit the LBP potentiality in extracting orientation and gray level values independent features. After computing the LBP image, we will apply the Haar-like features or the templates based rectangle masks used in [9] (figure8).

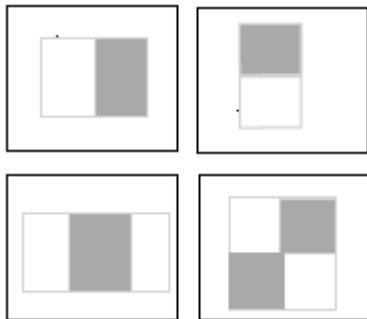


Fig 6: Rectangle features

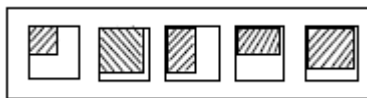


Fig 7: Some examples of the template set

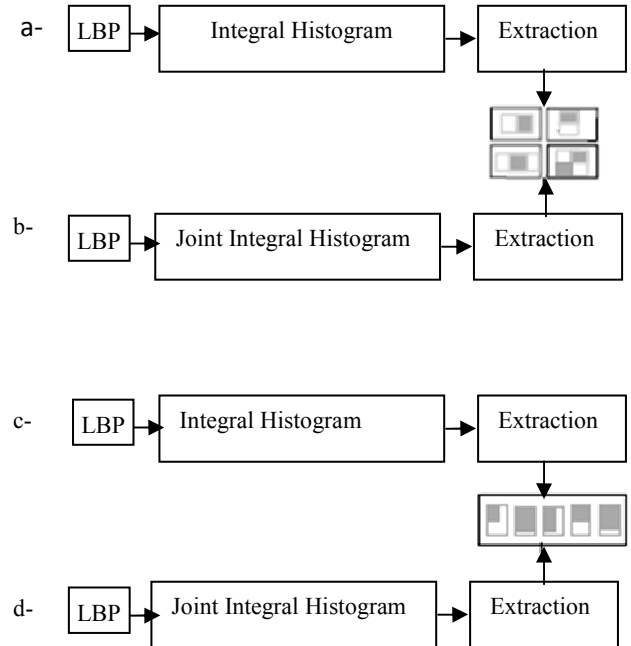


Fig 8: The four possible combinations: a-LBP followed by Integral Histogram with application of Haar-like features b-LBP followed by Joint Integral Histogram with application of Haar-like features c-LBP followed by Integral Histogram with application of rectangle masks d-LBP followed by Joint Integral Histogram with application of rectangle masks

3.4 The proposed framework

Under the learning process, after the application of the transformation (LBP or TPLBP) to the image, we applied the template to extract specific region information using the integral graphs. Based on the integral graphs, we can classify the feature extraction process in two groups: features extracted with histograms (integral or joint) and features extracted with the Haar features or templates based masks. The evaluation of the features extracted was achieved by a comparison with the Haar like-features of the Viola and Jones framework. After the feature extraction process, the best features which separate face and non-face samples are chosen by Adaboost learning algorithm.

4. ADABOOST TRAINING ALGORITHM

The Boosting, as the name implies, is the fact of improving the performance of any learning algorithm, and so reducing the error of weak classifiers. A weak classifier is a classifier with the accuracy a little bit better than 50%. Boosting algorithm is able to construct a strong classifier by a linear combination of weak classifiers chosen from a huge amount of set. Each hypothesis in the training algorithm is constructed using a single feature. The algorithm is described as follows.

- Given example images $(x_1, y_1), \dots, (x_n, y_n)$ where $y_i = 0, 1$ for negative and positive examples respectively.
- Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$, for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives examples.

- For $t = 1 \dots T$:
 - Normalize the weights

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,i}}$$
 so that w_t is a probability distribution.
 - For each feature, j , train a classifier h_j is restricted to using a single feature. The error is evaluated with respect to w_t ,

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$
 - Choose the classifier, h_t , with the lowest error ϵ_t .
 - Update the weights:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-\epsilon_i}$$
 where $\epsilon_i = 0$ if example x_i is classified correctly, $\epsilon_i = 1$ otherwise, and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$

- The final strong classifier is:

$$h(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{Otherwise} \end{cases}$$

5. EXPERIMENTS AND RESULTS

This section is organized as follows. The performance measurements are done on the CMU-MIT data set. First, the evaluation of the performance was measured on a single stage. Then, it was done on cascades of stages.

The CMU-MIT data set consists of 130 images with 507 labeled frontal faces. The system was trained using 5804 mirrored faces and 9050 non faces. For the validation set, we have used 480 faces and 10000 non faces. The faces were cropped to images of size 19*19 pixels. The number of non-faces is higher than the number of faces in order to represent the disparity of existing patterns on real images. In fact, in real images there are much more non-face pattern than faces.

5.1 Performance on single stage

The Haar-like features become a good reference as image features in the context of face detection. So, we compare the performance of our proposed feature to the Haar-like features.

In this work, we conduct two experiments to evaluate our proposed method. (1) Comparing the joint integral histogram to the integral histogram. (2) Comparing the TPLBP and LBP operators as discriminating features. In the first step, we compare the performance of the integral histogram to the joint integral histogram for the two first features selected by the Adaboost algorithm. Comparing the two experiments according to the Joint Integral Histogram, the first one is more efficient than the second. Comparing the JIH to th IH, experimental results show that the joint integral histogram yields a performance improvement compared to the integral histogram (table 2).

Table1. Comparison of the IH and JIH

	LBP+IH	TPLBP+IH	LBP+JIH(1)+Haar
TPR	0.57529	0.68436	0.8072

Table2. Comparison of the JIH(1) and JIH(2)

	LBP+JIH(1)+Haar	LBP+JIH(1)+Haar
TPR	0.8072	0.32753

We trained a single stage based on LBP (or TPLBP) with the joint integral histogram. Figure 9 shows a comparison based true positive rate between LBP and TPLBP.

The obtained results show that the LBP operator is more efficient for fewer numbers of features. The TPLBP operator gets more performance when the number of features is more than eight.

To evaluate the performance of the proposed feature, the true positive rate of the LBP feature based joint integral histogram is compared to the conventional Haar-like features (table3).

Table3. Experimental results on MIT+CMU data set

TPR	LBP+JIH+Haar	TPLBP+JIH+Haar	Haar
2	0.8072	0.44848	0.55508
4	0.80014	0.62543	0.85096
6	0.75052	0.67936	0.87388
10	0.69573	0.75603	0.9613
20	0.75913	0.79394	0.97159

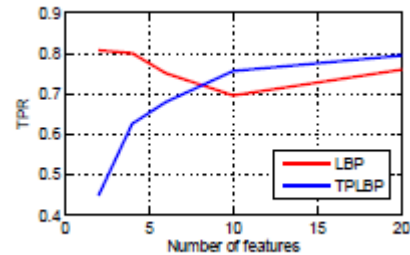


Fig 9: The true positive rate for LBP and TPLBP for different number of features using joint integral histogram with Haar-like features

The performance of the new feature for a single strong classifier can be witnessed only for a few numbers of features. In the next step, we will evaluate the performance of our proposed features on cascade structure.

5.2 Performance on cascades of stages

In this subsection, we compare the performance of the LBP based JIH feature with Haar-like features. Based on Adaboost learning algorithm, three boosting classifiers are trained. We fixed the false positive rate to 0.3 and the true positive rate to 0.99 with hand-tuned stage goals in the four first cascade stages according to the Viola and Jones work [7]. The number of features selected at each stage, detection rate and false positive rate are presented in table 4. Experimental results show with few numbers of the new features (12 features), we can achieve the maximum detection rate and very reduced false positive rate. The first stage is composed by 11 features, then in the following stages, a single feature is sufficient. These very good results are

obtained under a very hard conditions (0.99 for detection rate and 0.3 for false positive rate) comparing to the results of the Haar-like features (0.9 for detection rate and 0.5 for false positive rate). The training results obtained for the Haar like-features are not similar to the ones cited in Viola and Jones features. This can be explained by the fact that the training process is not held under the same training conditions (the training database). Accordingly, experiments show that database hardly affects the performance of the system.

Table 4. Experimental results (JIH+Haar) on MIT+CMU data set

	Features	detection rate (d)	false positive rate (f)
Stage1	11	0.99153	0.4585
Stage2	1	1	0.28378
Stage2	1	1	0.192

As shown in table4 and table5, features based JIH are more and more effective than Haar-like features. However, the computational time of these features is higher. Furthermore, features based JIH increases the computational complexity especially at test time.

Table 5. Experimental results (Haar) on MIT+CMU data Set

	Features	detection rate (d)	false positive rate (f)
Stage1	3	0.91737	0.4299
Stage2	24	0.90069	0.48578
Stage2	28	0.9	0.49733

6. CONCLUSION

In this work, we have proposed a new feature extraction approach which combines an existing image descriptor based Local Binary Patterns with conjunction of the Joint Integral Histogram. Experiments have shown that our proposed features outperform the conventional Haar-like features. In fact, the proposed JIH not only reduce the number of features in each stage but also yields a performance improvement in the detection rate and false positive rate. On the one hand, the max detection rate is obtained by training only 12 features. On the other hand, a single feature is made efficient for training the second and the third stages with achievement of a reduced false positive rate (0.192). Experimental results show that the selected features by Adaboost are very powerful to extract the discriminative information from the image.

7. REFERENCES

- [1] Viola, P. and Jones, J. 2004. Robust Real-time Face Detection, International Journal of Computer Vision. Netherlands.
- [2] Ojala, T., Pietikainen, M. and Maenpaa, T. 2002 Multi-resolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns.
- [3] Zhang, L., Chu, R., Xiang, S. and Li, S.Z. 2007. Face Detection based on Multi-block LBP representation.
- [4] Yan, S., Shan, S., Chen, X. and Gao, W. 2008. Locally Assembled Binary (LAB) Feature with Feature-centric Cascade For Fast Accurate Face Detection. China.
- [5] Roy, A. and Marcel, S. 2009. Haar Local Binary Pattern Feature for Fast Illumination Invariant Face Detection. Lausanne, Switzerland
- [6] Wolf, L., Hassner, T. and Taigman, Y. 2008. Descriptor based Methods in the Wild. Israel.
- [7] Wang, H., Li, P. and Zhang, T. 2007. Histogram Feature based Fisher linear discriminant for face detection.
- [8] Zhang, K., Lafruit, G., Lauwereins, R. and Gool, L.V. 2010 Joint Integral Histogram and its Application in Stereo Matching. International Conference on Image Processing. Hong Kong.
- [9] Trefny, J. and Matas, J. 2010. Extended Set of Local Binary Patterns for Object Detection, Winter Workshop. Czech Republic.
- [10] Anusha Bamini, A.M and Kavitha, T. 2010. Dominant Local Binary Pattern based Face Feature Selection and Detection. International Journal of Engineering and Technology.
- [11] Paisitkriangkrai, S. Shen, C. and Zhang, J. 2010. Face detection with Effective Feature Extraction. CoRR.
- [12] Chang-yeon, J. 2008. Face Detection using LBP features. Final Project Report.
- [13] Zhang, C. and Zhang, Z. 2010. A Survey of Recent Advances in Face Detection. Technical report.
- [14] S. Yan, S. Shan, X. Chen and W. Gao, "Locally Assembled Binary (LAB) Feature with Feature-centric Cascade for Fast and Accurate Face Detection", China, 2008.