Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 3, March 2014, pg.1193 – 1203

SURVEY ARTICLE

Multi-View Face Detection: A Comprehensive Survey

Shivkaran Ravidas[1], M. A. Ansari[2](Member IEEE), Jagdish Kukreja[3]

[1]Dr. Z.H. Institute of Technology and Management, Agra, India
[2]International Islamic University Madinah, KSA (On EOL from GBU, Greater Noida, India)
[3]G.L. Bajaj Institute of Technology and management, Greater Noida, India

mailmekaran@gmail.com, ma.ansari@ieee.org

Abstract: Locating multi-view faces in images with a complex background remains a challenging problem. In recent years many techniques have been applied to this problem, such as view-based learning, neural networks and AdaBoost. Most of the techniques from statistical analysis of face and non-face samples have good performance in detection of frontal faces but locating multi-view faces detection still remains challenging. The task of this paper is to present a comprehensive and critical survey of multi-view face detection. One of the major challenges encountered by face detection lies in the difficulty of handling arbitrary poses variations. As shown in Fig.1, in real-world images, faces have significant variations in orientation, pose, facial expression, lighting conditions, etc. This paper describes various methods and algorithms to detect faces with non-upright (rotated) and non-frontal (profile) faces and survey the progress toward a system which can detect faces regardless of pose reliably. As the number of proposed technique increased, the survey and evaluation become important.

Keywords: Multi view, Face detection, AdaBoost, non-frontal, rotation invariant

I. INTRODUCTION

Face detection can be defined as the process of extracting faces from scenes. So, the system positively identifies a certain image region as a face. An important aspect for automated facial recognition systems is the detection and segmentation of faces in images. According to Yang et al. [1,2] face detection can be described as the process of locating regions of the input image where faces are present. To detect faces accurately, faces need to be located and registered first to facilitate further processing. It is evident that face detection plays an important and critical role for the success of any face processing systems. The face detection problem is challenging as it needs to account for all possible appearance variation caused by change in illumination, facial features, occlusions, etc. In spite of all these difficulties, tremendous progress has been made in the last decade and many systems have shown impressive real-time performance.

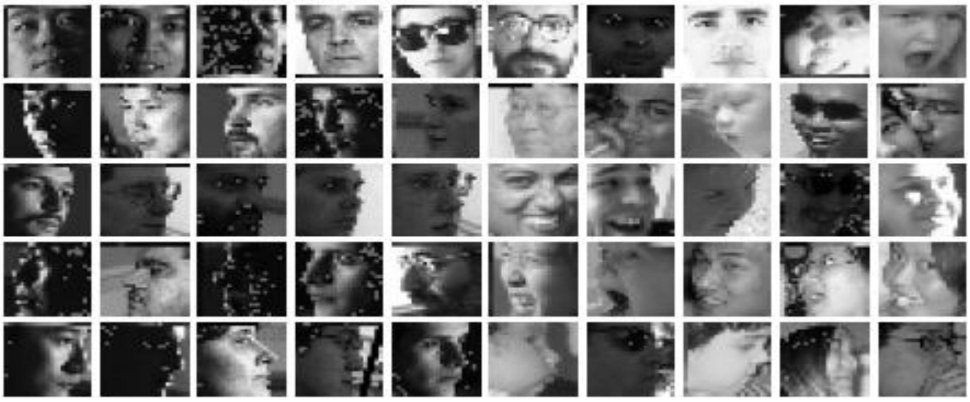

Figure1. Example of face images. Note the huge variations in pose, facial expression, lighting condition etc. [5]

Statistics show that most of the faces in photos and videos are non-frontal [3]. How to detect multi-view faces in photos and videos still remains a difficult problem due to the much more complicated variation within the multi-view face classes. Face detection proposed by Viola and Jones is the first approach for real-time face detection [4]. This approach utilizes the AdaBoost algorithm [5], which identifies a sequence of rectangle features that indicate the presence of a face. There are a number of techniques that can successfully detect frontal upright faces in a wide variety of images [7, 12, 3, 6]. Frontal face detection has been studied for decades. Viola and Jones [12] developed a fast frontal face detection system. In their work, a cascade of boosting classifiers is built on an over-complete set of Haar like features that integrates the feature selection and classifier design in the same framework. Sung Figure1. Example of face images. Note the huge variations in pose, facial expression, lighting condition etc. and Poggio [15] built a classifier based on the difference feature vector which was computed between the local image pattern and the distribution-based model.

The rest of the paper is organized as follows: Section 2 describe how the Viola-Jones framework can be extended to rotated and profile faces. In section 3,vector boosting method will be discussed. Section 4 gives few other prominent multi view face detection methods. Conclusions and future work are given in section 5.

II. VIOLA-JONES FRAMEWORK FOR FACE DETECTOR

There are many proposed approaches for face detection in a wide variety of images. While they can successfully detect frontal upright faces, many natural images include rotated or profile faces that are not reliably detected in the real world.

A. Background

There are only a small number of techniques that address non-frontal or non-upright face detection [8,9,23]. Non-upright face detection, proposed by Rowley et al., [9] is the first reliable approach. They use two neural network classifiers. The first estimates the pose of a face in the image window. The second performs conventional face detection. Faces are detected in three steps: the pose of the face for each image window is first estimated; the estimated pose is then used to de-rotate the image window; the window is then evaluated by the second classifier. Schneiderman et al. [11] develop the algorithm that can reliably detect human faces with out-of-plane rotation. They develop separate detectors that are each specialized to a specific orientation of the face. They have one detector specialized to right profile faces and one that is specialized to frontal faces. To detect left profile faces, they apply the right profile detector to flipped images. Jones and Viola [17] extend this framework [18]. They handle the detection of non-upright faces using a two stage approach. In the first stage the pose of each image window is estimated using a decision tree constructed using features like those described by Viola and Jones [18,17]. In the second stage one of N pose specific detectors is used to classify the window. These methods need a substantial amount of computation. Therefore, this constitutes a bottleneck to the application of face detection in real-time.

B. Viola Jones Face Detector

If one were asked to name a single face detection algorithm that has the most impact in the 2000's, it will most likely be the seminal work by Viola and Jones [17]. The Viola-Jones face detector contains three main ideas that make it possible to build a successful face detector that can run *in real time*: the integral image, classifier learning with AdaBoost, and the attentional cascade structure.

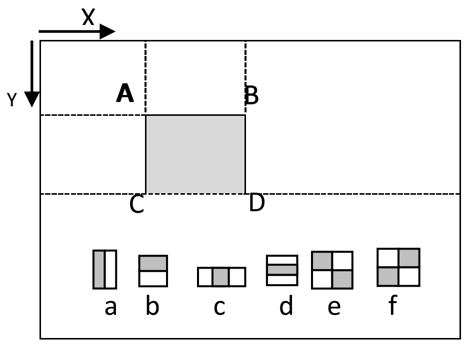

Figure 2. Illustration of the integral image and Haar-like rectangle features (a-f)

C. *The Integral Image*

Integral image, also known as a summed area table, is an algorithm for quickly and efficiently computing the sum of values in a rectangle subset of a grid. The integral image is constructed as follows:

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \qquad (1)$$

where *ii*(*x; y*) is the integral image at pixel location (*x; y*) and *i*(*x0; y0*) is the original image. Using the integral image to compute the sum of any rectangular area is extremely efficient, as shown in Fig. 2. The sum of pixels in rectangle region *ABCD* can be calculated as:

$$\sum_{(x,y)\in ABCD} i(x,y) = \quad ii(D) + ii(A) - ii(B) \quad - \quad ii(C) \qquad (2)$$

The integral image can be used to compute simple Haar like rectangular features as shown in Fig. 2.

D. *AdaBoost Learning*

Boosting is a method of finding a highly accurate hypothesis by combining many “weak” hypotheses, each with moderate accuracy. For an introduction on boosting, we refer the readers to [19] and [20]. The AdaBoost (Adaptive Boosting) algorithm is generally considered as the first step towards more practical boosting algorithms [21, 22]. In [23], presents a generalized version of AdaBoost algorithm, usually referred as *RealBoost*. It has been advocated in various works [24, 25, 26, 27] that Real Boost yields better performance than the original AdaBoost algorithm.

Input

- Training examples $S = \{(xi, zi), i = 1, \ldots N\}$
- T is total number of weak classifier to be trained

Intialize

- Initialize example score $F^o(x_i) = \frac{1}{2}\ln(\frac{N+}{N-})$,

  Where N_+ and N_- are the number of positive and negative examples in the data set.

Adaboost Learning

For $t = 1, \ldots., T$:

1. For each Haar-like feature $h(x)$ in the pool, find the Optimal threshold H and confidence score c_1 and c_2 to minimize the Z score minimum L^t.
2. Select the best feature with the minimum L^t.
3. Update $F^t(x_i)$= $F^{t-1}(x_i)$+$f_t(x_i), i = 1, \ldots., N$
4. Update $W_{+1j,} W_{-1j,}$j=1,2.

Output final classifier F^t(x).

Figure 3. AdaBoost learning pseudo code

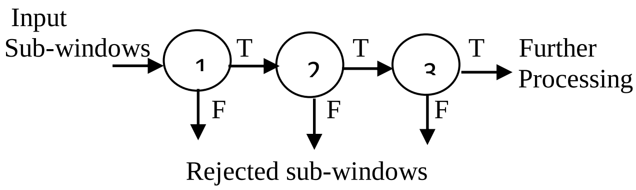

Figure 4. The attenuation cascade.

E. The Attentional Cascade Structure

Attentional cascade is a critical component in the Viola-Jones detector. The key insight is that smaller, and thus more efficient, boosted classifiers can be built which reject most of the negative sub-windows while keeping almost all the positive examples. Consequently, majority of the sub windows will be rejected in early stages of the detector, making the detection process extremely efficient. The overall process of classifying a sub-window thus forms a degenerate decision tree, which was called a “cascade” in [28]. As shown in Fig. 4, the input sub-windows pass a series of nodes during detection. Each node will make a binary decision whether the window will be kept for the next round or rejected immediately. The number of weak classifiers in the nodes usually increases as the number of nodes a sub-window passes.

F. Decision Tree

The pose estimator is thus a multi-class classifier. Furthermore, it needs to be about the same speed as a single face detector to make it advantageous to use over the try-all poses approach. This speed constraint is a stringent one since a pose-specific face detector only evaluates about 8 features per window on average. A decision tree classifier meets these constraints. Unlike a boosted classifier for multi-class problems, it is very straightforward to learn a multi-class decision tree. In addition a decision tree is quite efficient, since only one path from the root to a leaf is evaluated for each example. Even for very large and complex trees the number of features evaluated per window is logarithmic in the total number of nodes in the tree. One criticism of decision trees is that they are somewhat brittle, and sometimes do not generalize well. This is one reason that the face detector uses boosting rather than decision trees. For this application we have found decision trees to be quite reliable.

G. Detectors for each rotation class

The face examples in [17] are the same for each rotation class modulo the appropriate rotation. They only needed to train 3 detectors. One for 0 degrees (which covers -15 degrees to 15 degrees of rotation), one for 30 degrees (which covers 15 degrees to 45 degrees) and one for 60 degrees (which covers 45 degrees to 75 degrees). Because the filters we use can be rotated 90 degrees, any detector can also be rotated 90 degrees. So a frontal face detector trained at 0 degrees of rotation can be rotated to yield a detector for 90 degrees, 180 degrees and 270 degrees. The same trick can be used for the 30 degree and 60 degree detectors to cover the remaining rotation classes. All of the resulting face detectors coincidentally turned out to have 35 layers of classifiers. They all took advantage of diagonal features (although for the frontal, upright detector, the added diagonal features did not improve the accuracy of the detector over previously trained versions). Training was stopped after new cascade layers ceased to significantly improve the false positive rate of the detector without significantly reducing its detection rate. This happened to be after 35 layers in all three cases.

III. VECTOR BOOSTING

HUANG et al [29] propose a novel tree-structured multi-view face detector (MVFD), which adopts the coarse-to-fine strategy to divide the entire face space into smaller and smaller subspaces. It uses “divide and conquer” strategy. For this purpose, a newly extended boosting algorithm named Vector Boosting is developed to train the predictors for the branching nodes of the tree that have multi-components outputs as vectors. MVFD covers a large range of the face space, say, +/-45° rotation in plane (RIP) and +/-90° rotation off plane (ROP). In MVFD, the most straightforward way of extending Viola and Jones’s framework on face detection to train different cascades individually for each view and then use them as a whole. Figure 3 showed a number of detector structures for multi-view face detection. Among these structures, the most straightforward one is Fig. 5(a), the parallel cascade, by Wu et al. [30]. An individual classifier is learned for each view. Given a test window, it is passed to all the classifiers. After a few nodes, one cascade with the highest score will finish the classification and make the decision.

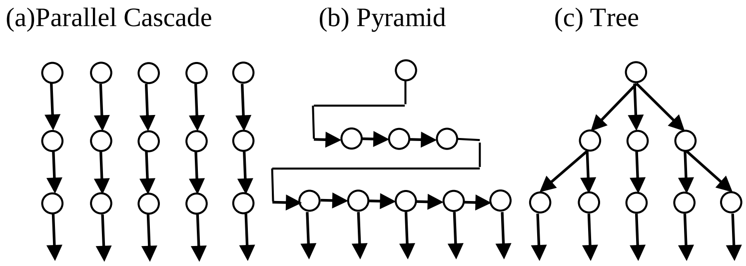

Fig. 5 Different Structures of MVFD

This simple structure could achieve rather good performance, though its running speed is generally slow, and the correlation between faces of different views could have been better exploited. One approach is the pyramid structure [31] that adopts coarse-to-fine strategy to handle pose variance of ROP, Figure 5(b)). Due to the similarities that exist in different poses of faces, the pyramid method treats them as one ensemble positive class so as to improve the efficiency of extracted features. However, the neglect of their intrinsic diversities makes the pyramid method have no discrimination in different poses. As a result, a sample that has passed the parent node has to be sent to all its child nodes (See figure 5(b)), which considerably slows down the decision-making process. On the contrary, another approach, the decision tree method [32], puts emphasis upon the diversities between different poses and the tree works as a pose estimator. With the imperative judgments made by the decision tree, it significantly reduces the time spent on pose estimation. However the results are somewhat unstable that makes its generalization ability not so well.

A. WFS Tree-Structured Detector

There are two main tasks for MVFD problem: one is to distinguish between faces and non-faces; the other is to identify the pose of a face.

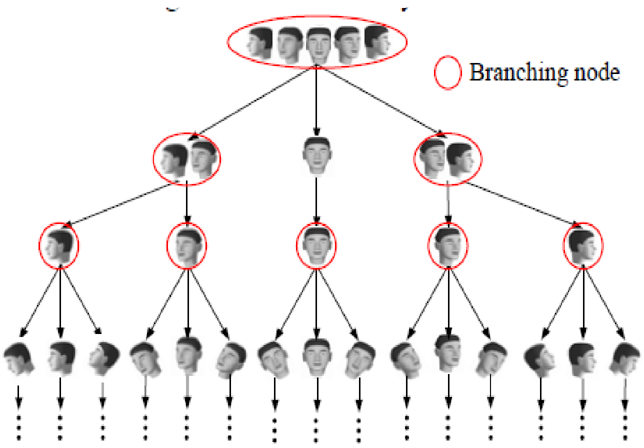

Figure 6. The Structure of the WFS Tree

The first task needs to reject non-faces as quickly as possible, so it is inclined to find the similarities of faces of different poses so as to separate them from non faces, while the latter task focuses on the diversities between different poses. The conflict between the two tasks really leads to the dilemma that either treating all faces as a single class (as in the pyramid method) or different individually separated classes (as in the decision tree method) is unsatisfactory for MVFD problem. To overcome WFS (Width-First-Search) tree structure to balance these two aspects so as to enhance the detector in both accuracy and speed. Pose/Orientation estimation is a non-trivial task, and can have many errors. If a profile face is misclassified as frontal, it may never be detected by the frontal face cascade. Huang et al. [29] and Lin and Liu [33] independently proposed a very similar solution to this issue, which were named vector boosting and multiclass Bhattacharyya boost (MBHBoost), respectively. The idea is to have vector valued output for each weak classifier, which allows an example to be passed into multiple subcategory classifiers during testing and the final results are fused from the vector output. Table I gives most prominent boosting schemes for face detection.

TABLE I

FACE/OBJECT DETECTION SCHEMES TO ADDRESS CHALLENGES INBOOSTING LEARNING.

Challenges in boosting	**Representative work**
General boosting	AdaBoost
	RealBoost
	GentleBoost
	Floatboost
Multiview face detection	Parallel Cascade
	Pyramid structure
	Decieision tree
	MBHBoost

Vector boosting and MBHBoost solved the issue of misclassification in pose estimation during testing. During training, they still used faces manually labeled with pose information to learn the multi view detector. However, for certain object classes such as pedestrians or cars, an agreeable manual pose labeling scheme is often unavailable.

IV. OTHER DETECTION METHODS

In many applications, such as visual surveillance system, human faces in the captured images may not be upright and frontal. In these cases, the detection of faces becomes much more complicated. These non-frontal faces usually contain less information and present more diversity. This fact makes non-frontal detection a lot more sensitive to noise, background, illumination, and facial model. A few methods for non frontal face detection have been proposed in recent years. They could be roughly divided as single-camera systems and multi-camera systems. For single-camera systems, Huang et al.'s2 method provided an important reference. In their system, they proposed a method to construct a rotation invariant multi-view face detector. Their method was composed of a Width-First-Search (WSF) tree detector structure, a Vector Boosting algorithm for learning strong classifiers, a domain-partition-based learning method, sparse features in granular space, and a heuristic search for sparse feature selection. Their system can detect multi-view faces with low computational complexity and high detection accuracy. However, the detection task may fail in some cases, such as low-resolution faces, inter-object occlusions, and incomplete human faces in images. It is also difficult for the method to detect the back side of human heads. Apparently, non-frontal face detection based on a single view of observation would be very difficult. The use of multiple cameras may somewhat relief the difficulties in non-frontal face detection.

Support Vector Machine is a new generation learning system based on recent advances in statistical learning theory. It delivers state-of-the-art performance in real-world applications such as tumor diagnosis [47], text categorization [48] and vehicle classification [49], etc. It is now one of the most important tool for machine learning and data mining. Multiview face detection has also been explored with SVM based classifiers. Li et al. [41] proposed a multi-view face detector similar to the approach in [42, 17]. They first constructed a face pose estimator using support vector regression (SVR), then trained separate face detectors for each face pose. Yan et al. [43] instead executed multiple SVMs first, and then applied an SVR to fuse the results and generate the face pose. This method is slower, but it has lower risk of assigning a face to the wrong pose SVM and causing misclassification. Wang and Ji [44] remarked that in the real world the face poses may vary greatly and many SVMs are needed. They proposed an approach to combine cascade and bagging for multiview face detection. Namely, a cascade of SVMs were first trained through bootstrapping. The remaining positive and negative examples were then randomly partitioned to train a set of SVMs, whose out-puts were then combined through majority voting. Hotta [45] used a single SVM for multiview face detection, and relied on the combination of local and global kernels for better performance.

Sihang ZHOU [46] introduced a novel Multi-Block Local Gradient Patterns (MB-LGP) as the feature set, a four layer face detector, which combined tree structure classifiers with the Support Vector Machine. The MB-LGP is an improved version of Local Gradient Patterns (LGP), it uses the average pixel value in an image block to replace the pixel value in LGP, thus leads to a better representation of large scale structure and greater robustness against noises. The proposed system also incorporated the multi-branch decision trees with support vector machine (SVM) in the classification part to balance the speed and accuracy of the face detector. Multi-Block Local Gradient Patterns were extracted from a 3 × 3 image square block, each sub-block in the square area is in the same size.

Most non-frontal face detector in the literature are based on the view-based method [34] in which several face models are built, each describes faces in a given range of view. Therefore, explicit 3D modeling is avoided. [35] partitioned the views of face into five channels, and developed a multi-view detector by training separate detector networks for each view. [10] studied the trajectories of faces in linear PCA feature spaces as they rotate, and used SVMs for multi-view face detection and pose estimation. The system consists of an array of two face detectors in a view-based framework. Each detector is constructed using statistics of products of histograms computed from examples of the respective view. It has achieved the best detection accuracy in the literature, while it is very slow due to the computation complexity. To address the problem of slow detection speed, Li, *et al*. [13] proposed a coarse-to-fine, simple-to complex pyramid structure, by combining the idea of boosting cascade and view-based methods.

In the last few years there has been a growth of interest in 3D face recognition and detection mainly due to its invariance to pose and lighting (Bowyer et al., 2006), but also to advanced in 3D data acquisition technology. Pamplona Segundo et al [10] proposed 3D face detection for varying poses of images. They describe a real-time 3D face detector based on boosted cascade classifiers that uses a scale-invariant image representation to improve both efficiency and efficacy of the detection process, named orthogonal projection images. In this representation, images are no longer scanned at multiple scales in order to detect faces with different distances in relation to the camera. The proposed detector achieves a high degree of pose invariance by detecting frontal faces not only in the camera viewpoint but also in rotated views of the scene. Koichiro Niinuma et al [16] proposed a fully automatic method for multi-view face detection and recognition system. They first build a 3D model from each frontal target face image, which is used to generate synthetic target face images. The pose of a query face image is also estimated using a multi-view face detector so that the synthetic target face images can be generated to resemble the pose variation of a query face image. Although, approach improves the detection speed significantly, it is still stumped by the following problems: First of all, as the system computation cost is determined by the complexity and false alarm rates of classifiers in the earlier stage, the inefficiency of AdaBoost significantly degrades the overall performance. Secondly, as each boosting classifier works separately, the useful information between adjacent layers are discarded, which hampers the convergence of the training procedure. Thirdly, during the training process, more and more non face samples collected by bootstrap procedures are introduced into the training set; thus it gradually increases the complexity of the classification. In the last stage pattern distribution between face and non-face become so complicated that can hardly be distinguished by Haar-like feature. Finally, view-based method always suffers from the problems of high computation complexity and low detection precision.

Xiangxin Zhu et all [36] presented a single model that simultaneously advances the state-of-the art, for all three detection, pose estimation and landmark localization. They argue that a unified approach may make the problem easier; for example, much work on landmark localization assumes images are pre-filtered by a face detector and so suffers from a near-frontal bias. In their simple approach they utilized encoding elastic deformation and three-dimensional structure and use mixtures of trees with a shared pool of parts .They define a “part” at each facial landmark and use global mixtures to model topological changes due to viewpoint. A part will only be visible in certain mixtures/views by allowing different mixtures to share part templates. This allows to model a large number of views with low complexity. Finally, all parameters of model, including part templates, modes of elastic deformation, and view-based topology, are discriminatively trained in a max-margin framework. Deng Peng et al. [37] proposed a multi view face detection method which combined skin color information and AdaBoost based face detection technique together to improve the detection accuracy and detection speed. First the input image is converted into YCbCr color space, then the image is binarized according to skin threshold and a binary image was obtained in which pixels with value 1 are skin-pixels and pixels with value 0 are non-skin pixels. After that only the skin regions are scanned by multi-AdaBoost detectors which were also trained on skin binarized face images. Before scanned by the multi-AdaBoost detectors, the sub windows are filtered by the ratio of the skin- pixels in the sub windows which can eliminate most of the false face regions.

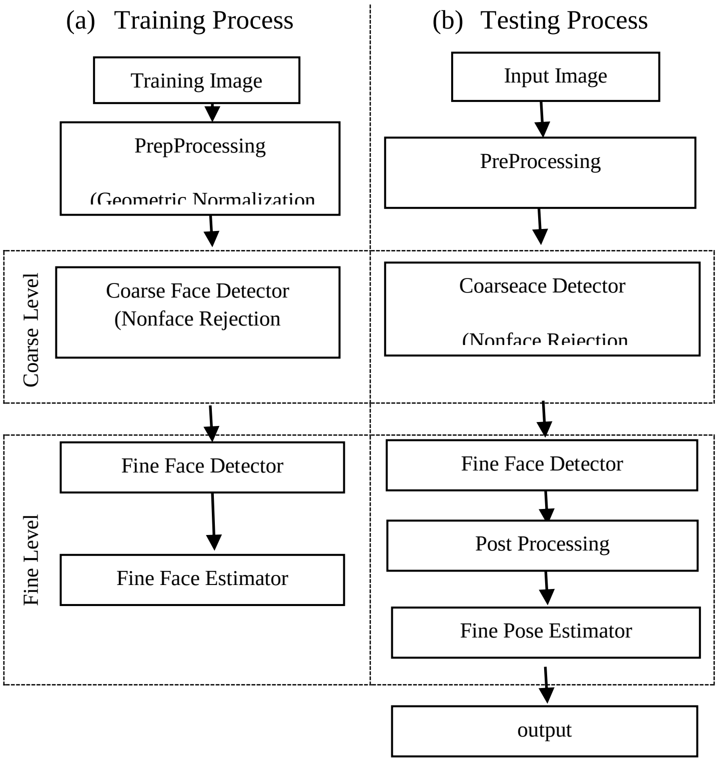

Fig 7. Flowcharts of statistic-based multi-view face detection and pose estimation system: (a) training; and (b) testing process.

Ying Ying et al [38] proposed a novel statistic-based system for automatic multi-view face detection and pose estimation. They constructed a multi-level framework utilizing multiple appearance-based learning methods to build corresponding face detectors and pose estimators, and hierarchically filters human faces. Contributions include the coarse-to-fine structure considering both efficiency and accuracy, different facial features representing low and high-dimensional information and statistic discriminant function regularizing divergent features. Figure 7 presents the flowcharts of training and testing processes of our system. It comprises five modules distributing in coarse and fine level, namely preprocessing, coarse face detector, fine face detector, post processing, and fine pose estimator.

Preprocessing module normalizes image data after variations reduction. Coarse face detector adopts boosted cascade algorithm to eliminate the majority non-face regions efficiently, meanwhile coarsely separates the poses of surviving candidates. In the fine level, fine face detector is applied to further filter out false positive candidates using modified subspace-based algorithm, and with the help of post processing module, fine pose estimator outputs final face results with more accurate location, scale and pose information. According to the two-stage structure, face is detected step by step, meanwhile pose is estimated more and more accurate. Mostly every face recognition systems require faces to be detected and localized in advance. Anvar SM, Yau WY, et al. [39] an approach to simultaneously detect and localize multiple faces having arbitrary views and different scales is proposed. They introduced a face constellation, which enables multi-view face detection and localization. In contrast to other multi-view approaches that require many manually labeled images for training, the proposed face constellation requires only a single reference image of a face containing two manually indicated reference points for initialization. Subsequent training face images from arbitrary views are automatically added to the constellation (registered to the reference image) based on finding the correspondences between distinctive local features. Thus, the key advantage of the proposed scheme is the minimal manual intervention required to train the face constellation. In their approach they combine face detection, localization, and feature selection tasks into a single framework. It is based on establishing a face constellation, which is essentially a reference map indicating how all the other face images in the training set are registered with respect to a reference image. A key feature of this work is the formulation of links to establish a match of a given face with a reference image and to detect and localize faces from arbitrary views at different scales.

Bernd Heisele et al [40] presented a component-based framework for face detection and identification. The face detection and identification modules share the same hierarchical architecture. They both consist of two layers of classifiers, a layer with a set of component classifiers and a layer with a single combination classifier. The component classifiers independently detect/identify facial parts in the image. Their outputs are passed the combination classifier which performs the final detection/identification of the face. They describe an algorithm which automatically learns two separate sets of facial components for the detection and identification tasks. In experiments, the detection and identification systems to standard global approaches is compared. The experimental results show that our component-based approach is superior to global approaches.

The purpose of the training images is to provide enough diversity to the constellation such that it can provide sufficient links that connect various registered faces from different poses, illumination, background clutter, etc.

V. CONCLUSION AND DISCUSSION

Up to now, a lot of algorithms have already been proposed to solve the face detecting problem. The most popular approach is the training-based approach which collects lots of face data to construct a database for training. With the face database, a suitable classifier is learned to detect faces with high detection rate and low false alarm rate. For example, Viola and Jones1 proposed the Adaboosting detection algorithm which is fast, robust and reliable to detecting frontal faces in 2-D images. A single cascade with Haar features has proven to work very well with frontal or near-frontal face detection tasks. However, extending the algorithm to multi-pose/multi-view face detection is not straightforward. Now a days, several algorithms with similar structures have been proposed to improve the accuracy of detection based on AdaBoosting detection algorithm. However, there still exist many difficulties in face detection, one of which is the detection of non-frontal faces. For non-frontal face detection, there appear view-dependent deformation and variation. Hence, these frontal face classifiers usually cannot be directly applied to non-frontal face detection.

REFERENCES

[1] Ming-Hsuan Yang, Member, IEEE, David J. Kriegman, Senior Member, IEEE, and Narendra Ahuja, Fellow, IEEE "Detecting Faces in Images: A Survey", IEEE Transaction on pattern analysis and machine intelligence , Vol 24,No.1 January 2002

[2] Erik Hjelmal and Boon Kee Low," Face Detection: A Survey" Computer Vision and Image Understanding 83, 236–274 (2001)

[3] Kuchinsky A, Pering C, Creech M, Freeze D, Serra B,Gwizdka J. FotoFile: A consumer multimedia organization and retrieval system. In: Proc. ACM SIGCHI, 1999.

[4] P. Viola and M. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, 57(2), 137-154, 2004.

[5] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generaliztion of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*,1997

[5] Cha Zhang and Zhengyou Zhang ," A Survey of Recent Advances in Face Detection", Technical Report , MSR-TR-2010.

[6] Bongjin Jun, Daijin Kim ,"Robust face detection using local gradient patterns and evidence accumulation," pattern recognition, Volume 45, Issue 9, September 2012.

[7] J. Ng and S. Gong. "Performing multi-view face detection and pose estimation using a composite support vector machine across the view sphere". *Proc. IEEE International Workshop on Recognition, Analysis,and Tracking of Faces and Gestures in Real-Time Systems*, pages 14--21, Corfu, Greece, September 1999.

[8] A. Pentland, B. Moghaddam, and T. Starner. "View-based and Modular Eigenspaces of Face Recognition". *Proc. of IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, pp. 84-91, June 1994.Seattle, Washington.

[9] H. A. Rowley, S. Baluja, and T. Kanade. "Neural network-based face detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (1998).

[10] M. Pamplona Segundo, L. Silva, O.R.P. Bellon, S. Sarkar," Orthogonal Projection Images for 3D Face Detection", Accepted Manuscript in *Pattern Recognition Letters, S0167-8655(13)00369-3, 2013*

[11] H. Schneiderman and T. Kanade. "A Statistical Method for 3D Object Detection Applied to Faces and Cars". *Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, 2000

[12] P. Viola and M. Jones. "Robust real time object detection". *IEEE ICCV Workshop on Statistical and Computational Theories of Vision*, Vancouver, Canada, July 13, 2001.

[13] S. Z. Li, *et al*. "Statistical Learning of Multi-View Face Detection". *Proc. of the 7th European Conf .on Computer Vision*. Copenhagen, Denmark. May, 2002.

[14] Ching-chun Huang , Jay Chou , Jia-hau Shiu and Sheng-Jyh Wang ,"Multi-view Face Detection Based on Position Estimation over Multi-camera Surveillance System",

[15] T. Poggio and K. K. Sung. "Example-based learning for view-based human face detection". *Proc. Of the ARPA Image* Understanding Workshop, II: 843-850. 1994.

[16] Koichiro Niinuma, Hu Han, and Anil K. Jain," Automatic Multi-view Face Recognition via 3D Model Based Pose Regularization", IEEE 6th International Conference on Biometrics: Theory, Applications and Systems (BTAS), Sept. 29-Oct. 2, 2013, Washington DC, USA

[17] Jones, M.; Viola, P.,"Fast Multi-view Face Detection", MITSUBISHI ELECTRIC RESEARCH LABORATORIES, TR2003-96 August 2003

[18] P.Viola and M. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, 57(2, 2004.

[19] R. Meir and G. R¨atsch. An introduction to boosting and leveraging. *S. Mendelson and A. J. Smola Ed., Advanced Lectures on Machine Learning, Springer-Verlag Berlin Heidelberg*, pages 118–183, 2003.

[20] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. Technical report, Dept. of Statistics, Stanford University, 1998

[21] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *European Conf. on Computational Learning Theory*, 1994.

[22] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

[23] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *European Conf. on Computational Learning Theory*, 1994.

[24] S. Li, L. Zhu, Z. Zhang, A. Blake, H. Zhang, and H. Shum. Statistical learning of multi-view face detection. In *Proc. Of ECCV*, 2002.

[25] C. Bishop and P. Viola. Learning and vision: Discriminative methods. In *ICCV Course on Learning and Vision*,2003.

[26] H. Schneiderman and T. Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 746-751, 2000.

[27] T. Mita, T. Kaneko, and O. Hori. Joint Haar-like features for face detection. In *Proc. of ICCV*, 2005.

[28] P. Viola and M. Jones. Rapid object detection using a

[29] Chang HUANG, Haizhou AI, Yuan LI and Shihong LAO , "Vector Boosting for Rotation Invariant Multi-View Face Detection", Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on Oct. 2005

[30] B.Wu, H. Ai, C. Huang, and S. Lao. Fast rotation invariant multi-view face detection based on real adaboost. In *Proc.of IEEE Automatic Face and Gesture Recognition*, 2004.

[31] S. Z. Li, L. Zhu, Z. Q. Zhang, et al., "Statistical Learning of Multi-View Face Detection". ECCV 2002.

[32] Jones and Viola, "Fast Multi-view Face etection".MERL-TR2003-96, July 2003.

[33] Y.-Y. Lin and T.-L. Liu. Robust face detection with multiclass boosting. In *Proc. of CVPR*, 2005.

[34] A. Pentland, B. Moghaddam, and T. Starner. "View-based and Modular Eigenspaces of Face Recognition". *Proc. of IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, pp. 84-91, June 1994.Seattle, Washington.

[35] H. A. Rowley, S. Baluja, and T. Kanade. "Neural network-based face detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (1998

[36] Xiangxin Zhu Deva Ramanan,"Face Detection, Pose Estimation, and Landmark Localization in the Wild". 2012

[37] Deng Peng, Pei Mingtao," Multi-view Face Detection Based on AdaBoost and Skin Color", Intelligent Networks and Intelligent Systems, 2008. ICINIS '2008

[38] Ying Ying and Han Wang, "A Statistic-based Approach for Automatic Multi-View Face Detection and Pose Estimation", Robotics, Automation and Mechatronics (RAM), 2011 IEEE Conference.

[39] *Seyed Mohammad Hassan ANVAR,, Wei-Yun YAU, and Eam Khwang TEOH,* "Multi-View Face Detection and Registration Requiring Minimal Manual Intervention"

[40] Bernd Heisele, Thomas Serre and Tomaso POGGIO, "A Component-based Framework for Face Detection and Identification", International Journal of Computer Vision , 2006 Springer Science + Business Media, LLC. Manufactured in the United States. DOI: 10.1007/s11263-006-0006-z

[41] Y. Li, S. Gong, and H. Liddell. Support vector regres-sion and classification based multi-view face detection and recognition. In International Conference on Automatic Face and Gesture Recognition, 2000.

[42] H. A. Rowley, S. Baluja, and T. Kanade. Rotation invari-ant neural network-based face detection. Technical report,School of Coomputer Science, Carnegie Mellow Univ.,CMU-CS-97-201, 1997

[43] J. Yan, S. Li, S. Zhu, and H. Zhang," Ensemble svm re-gression based multi-view face detection system". Technical report, Microsoft Research, MSR-TR-2001-09, 2001

[44] P. Wang and Q. Ji," Multi-view face detection under complex scene based on combined svms" . In Proc. of ICPR, 2004.

[45] K. Hotta. View independent face detection based on com-bination of local and global kernels. In International Con-ference on Computer Vision Systems, 2007.

[46] Sihang ZHOU, Jianping YIN, " Face Detection using Multi-block Local Gradient Patterns and Support Vector Machine", Journal of Computational Information Systems 10: 4 (2014) 1767–1776 Available at http://www.Jofcis.com

[47] S. Wang, J. Wang, H. Chen, et al. SVM-based tumor classification with gene expression data, Advanced Data Mining and Applications. Springer Berlin Heidelberg, 2006: 864-870.

[48] X. Xu, B. Zhang, Q. Zhong. Text categorization using SVMs with Rocchio ensemble for internet information classification, Networking and Mobile Computing. Springer Berlin Heidelberg, 2005: 1022-1031

[49] Y. W. Li, E. Zhu, R. Y. Chen, et al., " A vehicle classification approach based on edge orientation histograms" , Journal of Computational Information Systems 8: 16 (2012) 6979-6989

