REAL TIME FACE DETECTION: SURVEY

Bavitra Krishnan B.E. , Y.R.Packia Dhas M.E.,

***Abstract*-- Face detection on general embedded devices is fundamentally different from the conventional approach on personal computer or consumer digital camera due to the limited computation and power capacity. Numerous applications have been proposed and investigated on embedded smart cameras, such as intelligent surveillance, human computer interface, person authentication, human activity analysis, and so on. Among these applications, which are usually involved with humans, face detection has played a fundamental and first-step role. It also has several applications in areas such as content-based image retrieval, video conferencing, video coding, intelligent human–computer interfaces and crowd surveillance. Numerous techniques have been developed for real time face detection, and the purpose of this paper is to categorize and evaluate the face detection algorithms. Number of techniques for face detection has been proposed during several years. This paper provides a survey of several face detection methods. By surveying various face detection methods, the benefits and limitations are analyzed and parameters like false positive rate, sensitivity, false negative rate, specificity and accuracy are observed.**

Keywords: **content-based image retrieval, intelligent surveillance, human computer interface, person authentication, human activity analysis, video conferencing, real time face detection and crowd surveillance**

I. INTRODUCTION

Face detection is probably the visual jobs which humans can perform effortlessly. Even so, in personal computer vision conditions, this task just isn't easy. A general statement from the problem can be explained as follows: Presented a even now or movie image, discover and localize a great unknown variety (if any) associated with faces. The solution to the issue involves segmentation, removal, and verification of faces and maybe facial characteristics from a great uncontrolled track record. As a visual front end model, a deal with detection system should likewise have the ability to achieve the work regardless associated with illumination, inclination, and video camera distance. This survey aims to offer insight to the contemporary analysis of deal with detection in a very structural way.

Bavitra Krishnan, *Dept. of Electronics and Communication, PET College of Engineering, India*

Y.R.Packiadhas*, Dept. of Electronics and Communication, PET College of Engineering , India*

The Images of face fluctuate as a result of family member camera-face present (frontal, forty-five degree, user profile, benefit down), and several skin capabilities like an eye or perhaps the actual nasal can be in part or perhaps fully occluded. Skin capabilities like beards, mustaches, as well as spectacles might or perhaps will not be provide as well as there may be significant amounts of variability amid these types of factors such as form, color, as well as sizing. The design regarding people can be directly impacted by any person's skin phrase. Encounters could possibly be in part occluded by various other items. In the photograph together with several persons, a number of people might in part occlude various other people. Deal with photographs directly fluctuates regarding distinct rotations about the camera's optical axis. If the photograph can be made, elements like lighting (spectra, source submission as well as intensity) as well as digital camera attributes (sensor result, lenses) affect the appearance of any encounter.

Human being physique can be credited while using sparkling quality of holding a hefty variety of degrees regarding freedom going as far as about 244. Shaping structural in addition to energetic traits for action detection of your mega complex object similar to this is really a challenging chore.

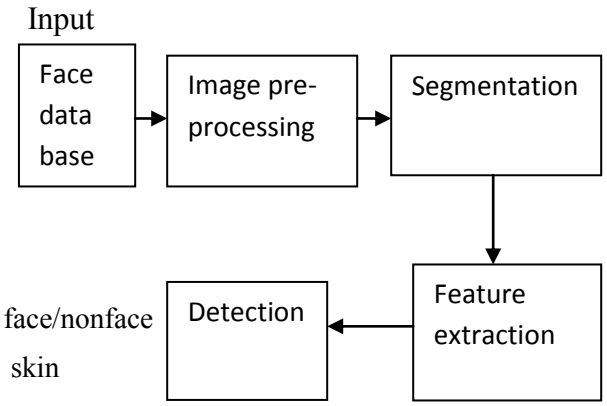

Fig. 1: General Block Diagram for Face detection

All a lot more, evaluating people action can be predominantly difficult because of the complicated, floppy in addition to self occluding character of the expressed people movement. In fact executing on-going activity detection method is usually an overwhelming function when we take into account the hurdles in

ISSN: 2278 – 1323

addition to road prevents at every level of the technique such as setting dysfunction, active light transformations, camera movement's etcetera in the setting subtraction degree, and partial occlusions within the tracking in addition to trait mining levels. The execution of the detection degree mainly will depend on the previous stages as well as on selecting traits for action representation. Large intra-class variations can be caused with the changes with facial look, lighting, in addition to expression. Such variations result in the experience distribution to become highly nonlinear in addition to complex in different space which can be linear for the original photograph space. Modifications in go pose, lights, and skin expression can easily significantly lower the performance of your face discovery system. Approaches with regard to handling varying illumination can be divided directly into four primary categories: (1) extraction of lights invariant characteristics (2) transformation of images with varying illuminations to your canonical representation (3) modeling this illumination modifications (4) utilization of some 3 dimensional face versions whose skin shapes in addition to albedos are generally obtained beforehand. A personal computer based experience detection system created for a real–life program must take into account the pose variant of encounters because encounters often experience large rotations comprehensive. A main challenge with regard to face discovery algorithms is based on the difference faces experience while changing pose.

II. LITERATURE REVIEW

Robert Viola and Michael Jones [1] have described the machine mastering approach pertaining to visual item detection that is capable of processing photos extremely speedily and attaining high recognition rates. His or her work has been distinguished by simply three essential contributions. The 1st was the introduction of any new impression representation referred to as the Integral Image allowing you the features utilized by our detector to get computed rapidly. The next was the learning algorithm which is based on AdaBoost. It select a small amount of critical visible features at a larger established and assure extremely useful classifiers. Your third contribution was a for combining progressively more complex classifiers inside a "cascade" allowing you background regions of the image to get quickly thrown away while shelling out more calculation on promising object-like areas. The cascade can be viewed an item specific focus-of-attention mechanism which not like previous solutions provides record guarantees which discarded areas are not likely to offer the object of interest. Inside the domain of face detection the machine has gave detection rates just like the best previous devices.

Junguk Cho et al. [2] have presented components architecture pertaining to face recognition based system on AdaBoost algorithm using Haar attributes. They get described the hardware style techniques which include image scaling, integral impression generation, pipelined processing along with classifier, and parallel processing multiple classifiers for you to accelerate the processing speed on the face recognition system. Also they have also discussed the optimization on the proposed architecture which is often scalable pertaining to configurable units with variable resources. The proposed system performance may be measured and compared with an similar software rendering. They get showed regarding 35 time's raise of system performance on the equivalent software implementation.

Hsiuao-Ying Chen et al. [3] have proposed the hybrid-boost mastering algorithm pertaining to multi-pose experience detection and facial term recognition. In order to speed-up the detection procedure, the system searches the entire frame for your potential experience regions by employing skin coloration detection and segmentation. Then that scans the epidermis color segments on the image and applies the weak classifiers combined with the strong classifier pertaining to face recognition and term classification. Their proposed system detected human face in different scales, different poses, distinct expressions, partial-occlusion, and defocus. Their key contribution has been the fragile hybrid classifiers selection good Haar-like (local) attributes and Gabor (global) attributes. The multipose experience detection algorithm will also be modified pertaining to facial term recognition. The experimental results get showed how the face recognition system and facial term recognition system have far better performance than the other classifiers.

D. Gobinathan et al. [4] have presented a hybrid means for face recognition in shade images. The well-known HAAR feature-based encounter detector developed by Viola as well as Jones (VJ) that was designed for gray-scale pictures was along with a skin-color filtration system, which supplies complementary facts in shade images. The image was passed by way of a HAAR Feature based encounter detector, which ended up being adjusted in a way that it ended up being operating in a point in its ROC curve which has a low quantity of missed people but a top number involving false detections. Their recommended method features eliminated a number of these false detections. They have likewise used a color pay out algorithm in order to reduce the results of illumination. Their experimental results about the Bao shade face databases have showed that this proposed method was better than the unique VJ algorithm.

Face acceptance was getting much attention in the society involving network multimedia information admittance. Areas such as network safety measures, content indexing as well as retrieval, and movie compression advantages of face acceptance technology because people are the middle of attention in lots of video. Network admittance control by way of face recognition not merely makes cyber criminals virtually difficult to gain

ISSN: 2278 – 1323

access to one's code, but also enhances the user-friendliness within human-computer discussion. Indexing and/or retrieving video data while using appearances involving particular persons will probably be useful for users such as news reporters, political scientists, as well as moviegoers. To the applications involving videophone as well as teleconferencing, the help of face acceptance also comes with a more useful coding system. Shang-Hung Lin [5] has given a great introductory course just for this new facts processing technological know-how. The paper has showed the viewers the generic framework to the face acceptance system, and the variants which are frequently encountered from the face detector.

Images made up of faces were necessary to intelligent vision-based human being computer discussion, and exploration efforts within face control include encounter recognition, encounter tracking, create estimation, as well as expression acceptance. However, many reported methods include assumed that this faces within the image or a graphic sequence are identified as well as localized. To construct fully programmed systems that analyze the knowledge contained within face pictures, robust as well as efficient encounter detection algorithms ended up required. Given 1 image, with regards to face recognition was to recognize all picture regions that contain a face irrespective of its three-dimensional placement, orientation, as well as lighting conditions. Such a difficulty was complicated because people are non-rigid and also have a high level of variability in space, shape, shade, and structure.

The state of the art a hierarchical face detection scheme called as P_FAD is proposed. P-FAD is a three phase adaptive technique, they are: Coarse process, Shift process and Haar detection .The hierarchical detection scheme introduced here is altered to implementing real-time detection scheme with low computation and storage overhead, in which operating units decrease dramatically while the operations on each unit are increasing . To increase the efficiency of the algorithm we use several layers they are Skin detection, Contour point detection, Dynamic group, Region merge & filter and modified V-J detector. The Coarse process is the first process in the P-FAD architecture .Coarse process consist of skin consist of skin detection layer with pixel-level manipulation. Pixel-level manipulation is uses most processing time and resources in a technique. To reduce the time complexity, the basic design principle is uses a high time saving skin detection technique. CbCr subset can eliminate the luminance effect and provide nearly best performance among different color spaces. By using Gaussian mixture models (GMM) we classify a pixel as a skin-pixel or none-skin-pixel. But the time complexity of the pixels is increased due to the judging of the pixels and also the offline training data set are not robust. An adaptive GMM skin color detection algorithm with online learning and simplify judging criteria is introduced to solve these problems. Online training data sets are given and judging time also reduced by using SGM Single Gaussian Model. We should add/remove the SGM to/from GMM If there is no SGM can approximately represent the current skin-color distribution. The skin are detected and processed as a binary image.

The Shift process is the middle tier in the P-FAD architecture. Shift process consist of Contour Points Detection, Dynamic Group and Region Merge & Filter. Here persons face are to be detected and confirmed using these techniques. After we get a binary image from skin detection, much binary image process will follow to form also called region of interest (ROI), such as mask correction (dilation, erosion, and open/close operation), component labeling, region group and so on. Region of interest ROI could be the face candidates. These also increase the computational overhead since pixel-level manipulations are processed here. To handle this problem we use Contour Points Detection, Dynamic Group and Region Merge & Filter.

The contour of a target is the dividing line of foreground and background. When an image is scanned pixel by pixel, the point in the contour is the starting point of an adequately connected component. So we can compare the pixels based on the index and threshold values. Threshold value usually changes on the basic of the image resolution and the image size. The index is formed by the inspiration of “additive increase, multiplicative decrease” (AIMD) method in TCP congestion control. By using AIMD we can calculate very pixels contour point of the binary by the index value derived, we scanned it from left to right and top to bottom. CL and CR are storage array of contour point left and right got from scanning the image left to right.

The contour points are stored by its spatially scanned location in the image, for multiple targets the contour points are to classified and stored in different groups. The target can be mapped if the contour points are closed into a group otherwise it’s a noise. Targets are to be grouped based on this principle dynamically. Here classification is done by dynamic grouping methods during scanning process. If the point is not captured by any existing groups, a new group will be created it is called as dynamic Group creating. Group terminates when the points in a given group have formed a closed area this process is called as Group terminate. If the points in the group have not formed a closed area and the group has not captured new points during a

ISSN: 2278 – 1323

period of time then the group is dead the process is called as Group dead. Region merge & filter is the final process in shift process. Here the candidates face is to be produced and reconstructed. By merging the small adjacent regions of the candidates face like eyebrows, glasses and so on we can construct a candidate's face. By the height-width ratio we can able to filter the non-face regions of the candidate. Haar process is the final stage in the three tier architecture. The above two tier will detect some candidates face in a frame. Here we use a special detector called as V-J detector to finalize the final output of our detection technique. By reducing the sub-windows size from 881484 to several hundred we also reduce the computational overhead of our architecture. Time is also further reduced by the early rejection process and also the cascaded sub-windows.

III. PERFORMANCE ANALYSIS

We need various assessment metric values to be calculated in order to analyze our proposed technique for the pose variation detection. The metric values are found based on True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). The usefulness of our proposed work is analyzed by five metrics namely False Positive Rate (FPR), False Negative Rate (FNR), Sensitivity, Specificity and Accuracy. The demonstration of these assessment metrics are specified in equations that given below.

A.Evaluation Results for face detection based on pose variation

False Positive Rate (FPR)

The percentage of cases where an image was segmented to the shape, but in fact it did not.

$$FPR = \frac{FP}{TN + FP}$$

False Negative Rate (FNR)

The percentage of cases where an image was segmented to the shape, but in fact it did.

$$FNR = \frac{FN}{TP + FN}$$

Sensitivity

The measure of the sensitivity is the proportion of actual positives which are properly detected. It relates to the capacity of test to recognize positive results.

Specificity

The measure of the specificity is the proportion of negatives which are properly detected. It relates to the capacity of test to recognize negative results.

Accuracy

The weighted percentage of pose variation images is correctly classified by the measurement accuracy. It is represented as,

$$Accuracy = \frac{TN + TP}{TP + TN + FN + FP} \times 100$$

For examining the segmentation usefulness our proposed technique is assessed with these above explained assessment metrics False Positive Rate, False Negative Rate, Sensitivity, Specificity and Accuracy.

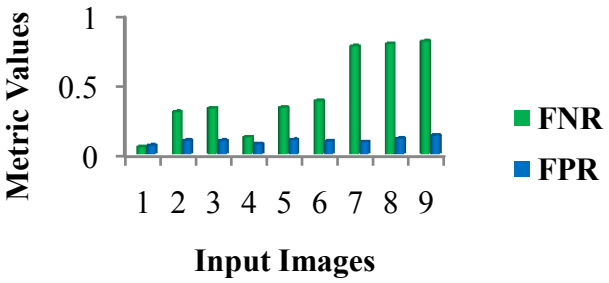

Fig 4: Graph for the comparison of FNR and FPR values for different poses of images

Fig. 4 shows the FNR and FPR values for different poses of images. From the FNR and FPR values it is clear that the values of the straight image are lower than the inclined pose images. Hence it can be shown that the straight pose image attained high accuracy than the inclined images.

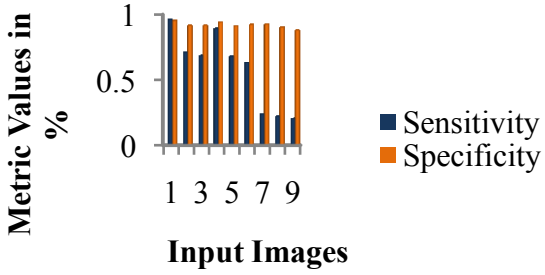

Fig 5: comparison of sensitivity and specificity values

ISSN: 2278 – 1323

Fig. 5 shows the sensitivity and specificity values for different poses of images. From the sensitivity and specificity values it is clear that the values of the straight image are higher than the inclined pose images. Hence it can be shown that the straight pose image attained high accuracy than the inclined images.

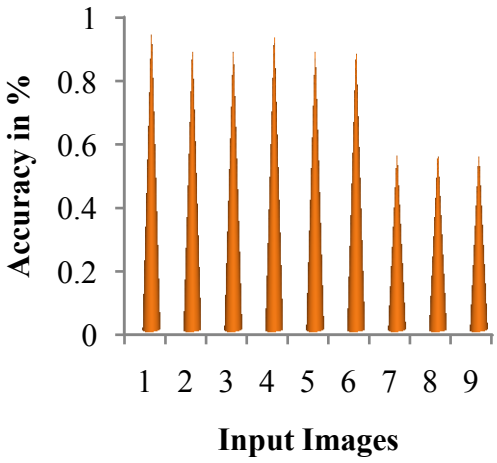

Fig 6: Graph for Accuracy Values of the images

Fig. 6 shows the comparison for the Accuracy Values of the images with different poses. Straight face image has higher accuracy when compared to the inclined images. The Accuracy Values obtained have showed that the straight face image is highly detected than the inclined images.

IV. CONCLUSION

Face detection is a challenging problem in the field of image analysis and computer vision that has received a great deal of attention over the last few years because of its many applications in various domains. Research has been conducted vigorously in this area for the past four decades or so, and though huge progress has been made, encouraging results have been obtained and current face detection systems have reached a certain degree of maturity when operating under constrained conditions; however, they are far from achieving the ideal of being able to perform adequately in all the various situations that are commonly encountered by applications utilizing these techniques in practical life. The ultimate goal of researchers in this area is to enable computers to emulate the human vision system and, as has been aptly pointed out, "Strong and coordinated effort between the computer vision, signal processing, and psychophysics and neurosciences communities is needed" to attain this objective.

REFERRENCES

[1] Paul Viola and Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", In Proceedings of Conference on Computer Vision and Pattern Recognition, 2001.

[2] Junguk Cho, Shahnam Mirzaei, Jason Oberg and Ryan Kastner, "FPGA-Based Face Detection System using Haar Classifiers", In proceedings of ACM, 2009.

[3] Hsiuao-Ying Chen, Chung-Lin Huang and Chih-Ming Fu, "Hybrid-boost learning for multi-pose face detection and facial expression recognition", Elsevier Journal on Pattern Recognition, Vol. 41, pp. 1173 – 1185, 2008.

[4] N.Gobinathan, Abinaya and Geetha. P, "Combining Skin Color based Classifiers and HAAR Feature using VJ Algorithm", International Journal of Research in Engineering & Advanced Technology, Vol. 1, 2013.

[5] Shang-Hung, "An Introduction to Face Recognition Technology", Special Issue on Multimedia and Informing Technologies, Vol. 3, No. 1, 2000.

[6] Ming-Hsuan Yang, David J. Kriegman and Narendra Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, 2002.

[7] R. L. Hsu, A.-M. Mohamed, and A. K. Jain, "Face detection in color images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp.696–706, May 2002.

[8] Q. Wang, P. Zhou, J. Wu, and C. Long, "RaFFD: Resource-aware fast foreground detection in embedded smart cameras," in *Proc. IEEE GLOBECOM Conf.*, Dec. 2012.

[9] C. Zhang and Z. Y. Zhang, "A survey of recent advances in face detection," *Microsoft Res.*, Jun. 2010.

[10] Y. W. Wu and X. Y. Ai, "Face detection in color images using AdaBoost algorithm based on skin color information," in *Proc. 2008 Int.Workshop Knowl. Discov. Data Mining*, 2008, pp. 339–342.

[11] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques", in *Proc. Graphicon-2003*

[12] H. Kruppa, M. A. Bauer, and B. Schiele, "Skin patch detection in realworld images", in *Proc. of the DAGM-Symposium*, 2002, pp. 109-116.

[13] Cahi, D. and Ngan, K. N., "Face Segmentation Using Skin-Color Map in Videophone Applications," *IEEE Transaction on Circuit and Systems for Video Technology*, Vol. 9, pp. 551-564 (1999).

[14] E. Acosta, L. Torres, A. Albiol, and E. J. Delp, "Anautomatic face detection and recognition system for video indexing applications," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol.4. Orlando, Florida, 2002, pp.3644-3647.

[15] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.25, pp.1160-1165, 2003.

[16] Craw, D. Tock, and A. Bennett, "Finding face features," in *Second European Conference on Computer Vision*, 1992, pp.92-96.

[17] Q. Yang and X. Q. Ding, "Discriminant Local Feature Analysis of Facial Images," in *IEEE International Conference on Image Processing*, Vol.2, 2003, pp.863-866.

[18] L. Acasandreni and A. Barriga, "Accelerating Viola-Jones face detection for embedded and SoC environments," in *Proc. 5th ACM/IEEE Int. Conf. Distrib. Smart Cameras*, Aug. 2011, pp. 1–6.

[19] M. Bramberger, J. Brunner, B. Rinner, and H. Schwabach, "Real-time video analysis on an embedded smart camera for traffic surveillance," in *Proc. Real-Time Embed. Technol. Appl. Symp.*, May 2004, pp. 174–178

ISSN: 2278 – 1323

[20] J. Ren, X. Jiang, and J. Yuan, "A complete and fully automated face verification system on mobile devices," *Pattern Recognit.*, vol. 46, no. 1, pp. 45–56, Jan. 2013.
[21] Q.Wang, J.Wu, C. Long, and B. Li, "P-FAD: Real-time face detection scheme on embedded smart cameras," in *Proc. ACM/IEEE Int. Conf. Distrib. Smart Cameras*, Nov. 2012, pp. 1–6.

ISSN: 2278 – 1323