June 2013. Vol. 3, No. 4 ISSN2305-8269

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

EFFICIENT EYES AND MOUTH DETECTION ALGORITHM USING COMBINATION OF VIOLA JONES AND SKIN COLOR PIXEL DETECTION

Ijaz Khan, Hadi Abdullah and Mohd Shamian Bin Zainal

University Tun Hussein Onn Malaysia
Ijaz.uthm@yahoo.com, Hadi.uthm@yahoo.com and shamian@uthm.edu.my

ABSTRACT—

This paper presents improved algorithms for face, eyes and mouth detection in an image. Viola Jones and skin color pixel detection as face detection techniques are widely used. Viola Jones gives accurate face detection but consumes more time whereas skin color pixel detection technique consumes less time but lacks in accuracy. Our design is hybrid of both these techniques which increases accuracy while consuming less time. Viola Jones and other methods can accurately detect faces but in case of facial features detection their accuracy decreases. The focus in this research is mainly on increasing accuracy in detection eyes and mouth while consuming less time. Lastly we have discussed experimentation results comparing standard Viola Jones with our hybrid design for face, eyes and mouth detection.

Index Terms— Viola Jones, Skin color pixels, facial feature, integral images and cascading.

INTRODUCTION

Face and facial features detection has been an active research area in the computer vision field as it has large number of application such as biomedical image analysis [1] human computer interfaces [2],[3] monitoring and surveillance [4], [5] smart rooms [6] intelligent robots [7],[8] image database management system [9] and drivers alertness system [10].

A vast amount of research is performed in the field of face and facial components detection, localization and tracking. Generally, we can classify these approaches into four main groups. (1)Feature invariant approaches: Invariant features, (passive to brightness, and position) are utilized in this approach to detect faces. The relations among face features and the presence of the detected faces are usually described by building up a statistical model. Such face features are Skin Color [11] ,Texture[12] and Facial Features[13]. (2) Template matching method: This technique involves comparison of portions of images against one another. A pattern matching operation is done between the template containing human face and input image. Examples of this method are active Shape Model[14] and shape template[15].(3) Knowledge-based method: in this method the position of the face is localized by finding the invariant feature of the face. Analogy among the features determines the presence of human face in an input image[16]. (4)Appearance-based method: In this method a series of face images are trained that establish a face model for face detection. such as Neural Network [17]and Hidden Markov Model[18].

Face detection proposed by Viola and Jones based on statistic methods is most popular among the face detection approaches. This face detection is a variant of the AdaBoost algorithm [19] which achieves rapid and robust face detection. They proposed a face detection method based on the

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

AdaBoost learning algorithm using Haar features that detected the face successfully with high accuracy. However the accuracy of the method is still not enough when this method is used to detect facial feature.

In our research we have integrated Viola Jones, skin color pixel detection and physical location approximation technique to have a hybrid design which can detect face, mouth and eyes more accurately while consuming less time.

2 Face Detection Techniques

Face detection techniques can be categorized into two major groups that are feature based approaches [20],[21] and image based approaches. Image based approaches use linear subspace method, neural networks [22],[23],[24] and statistical approaches for face detection. Feature based approaches can be subdivided into low level analysis, feature analysis and active shape model.

2.1 Viola Jones technique

Viola-Jones technique [25] is based on exploring the input image by means of sub window capable of detecting features. This window is scaled to detect faces of different sizes in the image. Viola Jones developed a scale invariant detector which runs through the image many times, each time with different size. Being scale invariant, the detector requires same number of calculations regardless of the size of the image.

The system architecture of Viola Jones is based on a cascade of detectors. The first stages consist of simple detectors which eliminates only those windows which do not contain faces. In the following stages the complexity of detectors are increased to analysis the features in more detail. A face is detected only if it is observed through the entire cascade. These detectors are constructed from integral image and Haar like features shown in figure 1.

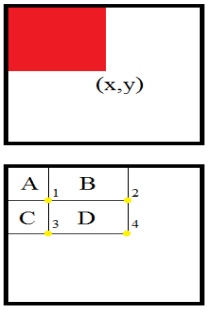
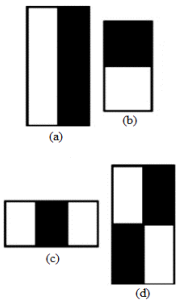


Figure 1. Viola Jones integral image construction.

The first step of this algorithm is to convert the input image into an integral image. This is done by making each pixel equal to the entire sum of all pixels above and to the left of the concerned pixel. By doing so, sum of all pixels inside any given rectangle can be calculated using only four values.



Sum of the rectangle ABCD = D - (B + C) + A

The face detector in Viola Jones method analyzes a sub-window using features. These features consist of two or more rectangles. Each feature gives a single resultant value which is calculated by subtracting the sum of the white rectangle(s) from the sum of the black rectangle(s). Different types of features are shown in Figure 2.

Figure 2. Viola Jones Haar like features.

Viola and Jones used a simple classifier built from computationally efficient features using AdaBoost [26],[27] for feature selection. AdaBoost is a machine learning boosting algorithm that constructs a strong classifier through a weighted combination of weak classifiers. Mathematical description of weak classifier is,

Where x is a sub-window, f is the applied feature, p the polarity and θ is threshold that concludes whether x should be classified as a negative (non-face) or a positive (face).

Viola-Jones face detection algorithm scans the detector several times through the same image – each time with a new size. The detector detects the non face area in an image and discards that area which results in detection of face area. To discard non face area Viola Jones take advantage of cascading. When a sub window is applied to cascading stages, each stage concludes whether the sub window is a face object or not. Sub

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

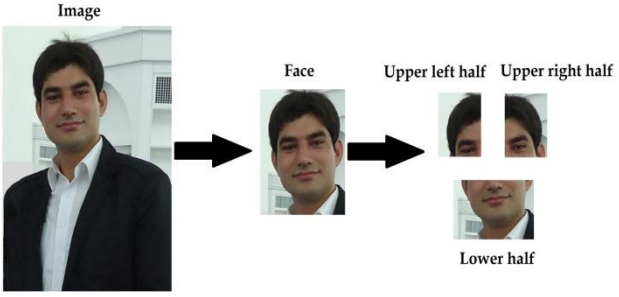
windows which contain some percentage of having faces are passed to next stage and those which are not faces are discarded. Final stage is considered to have a high percentage of face objects.

2.2 Skin color pixel detection technique

Second method used for face detection is based on skin color face detection method. In face detection approaches detection of skin color is first step. Major advantages of this technique are its robustness, non sensitivity to position and shape invariance. However for this technique to work, it is vital to use an accurate color space model. Some existing color spaces are RGB, CMY, XYZ, UVW, LSLM, L*a*b*, L*u*v*, LHC, LHS, HSV, HSI, YUV, YIQ, YCbCr[28],[29] out of which Most commonly RGB, HIS , YCbCr are used. We are using RGB color space for skin color detection as it is native representation of color images and it is widely used for processing and storing digital images. RGB color space consist of three basic colors R (red), G(green), B(blue) that can be combined to produce any resultant color. Although different people have different skin colors, studies have shown that actual difference lies between the intensities[30]. So if Brightness is removed from color representation, the difference between human skin colors can be reduced. In order to detect skin color following set of rules have been found to be more accurate than other models.

(R>95) AND (G>40) AND (B>20)
(*first condition*)

AND (max – min > 15) AND (|R-G|> 15) AND (R>G) AND (R>B) AND (R>220) AND (G>210) AND (B>170) AND (R>B) AND (G>B)
(2)



Pixels of RGB image are detected as skin if first condition holds true and rest of the conditions are used to ensure that RGB components must not be close together, that ensures grayness elimination. It also ensures that R and G component must not be together which must be true for fair complexion R is largest in R, G and B for pixel of skin color regions in RGB color space and B maybe be larger than G for pixels of skin color regions in the shadow. These rules are suitable for detection of skin color under regular illumination conditions as well as in shadow. The resultant image from this technique is a black and white image in which skin is converted to white and rest of the colors are converted to black. Face can be detected then by cutting the biggest white connected area with in black and white image.

3 Hybrid Detector Design

The hybrid design consists of three phases. In phase one face is detected. Skin color pixel detection is used to extract all the entire skin color pixels from the image. Once they are extracted Viola Jones is applied to detect face. This increases efficiency of Viola Jones techniques and decreases consumed time.

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

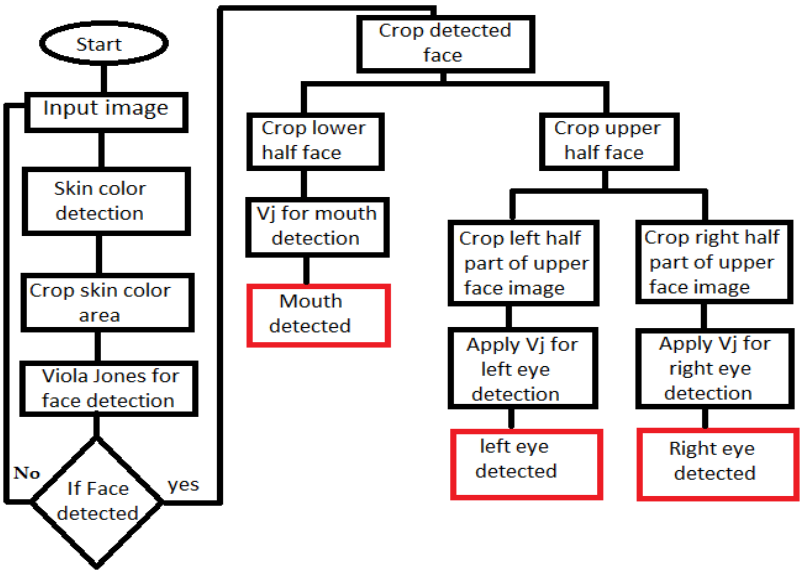


Figure 4. Hybrid design flow chart for face eyes and mouth detection

In
second phase the face detected is extracted from input image for further processing. Extracted face image is divided into three sub portions upper left half, upper right half and lower half.
Figure 3. Face image division based on physical approximation of location of eyes and mouth on face.

The divisions are made on the basis of physical approximation of eyes and mouth locations. Figure 3 shows the process done in second phase. This will help in phase three in which Viola Jones is applied to detect eyes and mouth in three sub portions. Figure 4 shows the flowchart of hybrid design.

4 EXPERIMENTAL RESULTS

The system parameters set for experimentation are shown in table 1.

TABLE 1
System parameters.

System parameters	Values and versions
Simulation software	*MATLAB R2012b*
Processor	*Intel(R) core(TM)i5-2450(2.50GHzX2)*
RAM	*4GB*
Operating System	*Windows 7(64bit)*

Keeping these system parameters constant experimentation was done on images of different sizes. The time taken by

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

standard Viola Jones and our hybrid design for some test images is shown in table 2.

TABLE 2
Time takes by test images for face mouth and eye detection.

Size of the image	Viola Jones results In seconds	Proposed system results In seconds
326 x 484	*2.6753*	*0.9805*
326 x 484	*1.3445*	*1.1932*
326 x 484	*1.3067*	*0.7347*
326 x 484	*1.2416*	*0.8560*
326 x 484	*1.2271*	*0.7707*
326 x 484	*1.2816*	*1.0485*
326 x 484	*1.5272*	*0.7581*
326 x 484	*1.4263*	*0.8748*
326 x 484	*1.3182*	*0.8208*
326 x 484	*1.6522*	*1.0261*
326 x 484	*1.3616*	*1.0477*
326 x 484	*1.5340*	*0.8044*
326 x 484	*1.4809*	*0.8297*
326 x 484	*1.4071*	*0.7449*
326 x 484	*1.2542*	*1.0041*
326 x 484	*1.1562*	*1.0164*
326 x 484	*1.1133*	*0.7435*
326 x 484	*1.2718*	*0.7943*
326 x 484	*1.2080*	*0.8182*
326 x 484	*1.3944*	*0.8821*
326 x 484	*1.6077*	*0.9095*
326 x 484	*1.7213*	*1.0352*
326 x 484	*1.4377*	*0.8471*
326 x 484	*1.5049*	*1.0034*
326 x 484	*1.5820*	*0.7852*
326 x 484	*1.4006*	*0.8631*
326 x 484	*1.5013*	*0.8671*
326 x 484	*1.6021*	*0.7621*
960 x 503	*3.2249*	*1.8728*
720 x 540	*3.4823*	*1.5320*
540 x 720	*2.9341*	*1.2763*
240 x 320	*0.8568*	*0.5804*
1006 x 714	*4.2226*	*1.7404*
1480 x 1568	*10.8782*	*5.6796*
1119 x 834	*5.3891*	*2.6488*
494 x 385	*1.5324*	*1.0302*
720 x 639	*3.6769*	*1.5638*
864 x 658	*3.5624*	*1.5163*

From table 2 we can observe that time taken by standard Viola Jones technique is higher than our hybrid design. To check accuracy of our design ten thousand sample images of size 326 X 484 were processed. Table 3 shows comparison of some images in their processing time and face, eyes and mouth detection accuracy.

In table 3 it appears clearly that accuracy of Viola Jones in detecting face is excellent. But when it comes to detection of eyes and mouth, this technique looses in accuracy as well as in time consumption. On the other hand our design consumes less time and detects mouth and eyes accurately.

International Journal of Engineering and Applied Sciences

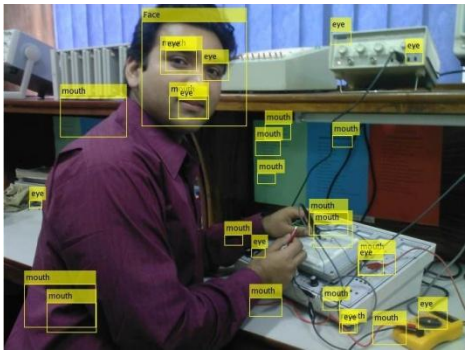
www.eaas-journal.org

Table 4 shows average time consumed by techniques and their face, eyes and mouth detection accuracy.

TABLE 4
Average time consumed and face, eyes and mouth detection accuracy of ten thousand test images.

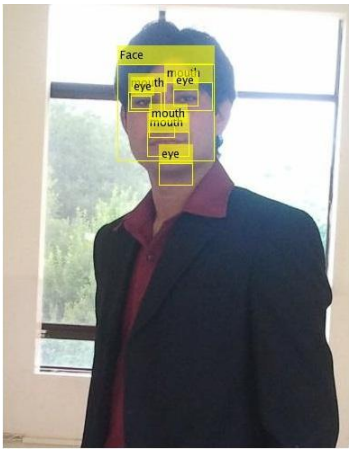
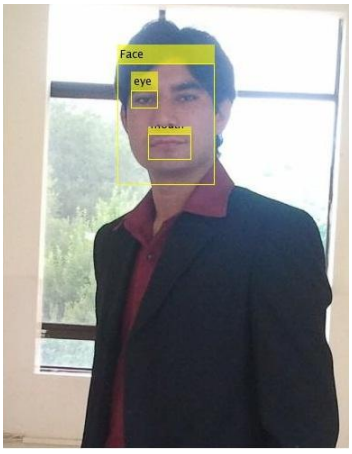
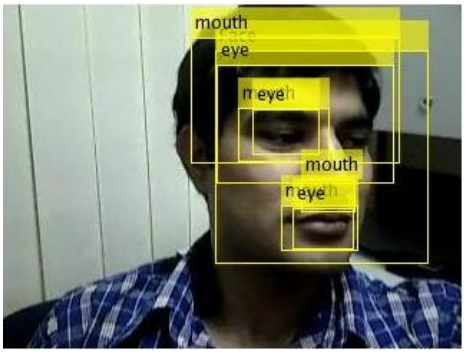
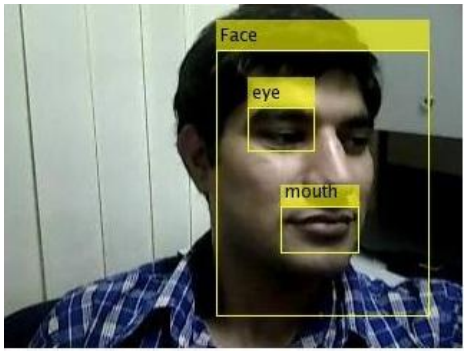
Average accuracy in image size 326 x 484	Viola Jones	Hybrid design
Average time consumed per image	*1.4674s*	*0.9067s*
Face detection accuracy	*99.4%*	*99%*
Eyes detection accuracy	*65%*	*99.5%*
Mouth detection accuracy	*57%*	*98%*

TABLE 3
Comparison in accuracy and consumed time of standard Viola Jones and proposed system for face, left eye and mouth detection.

Image size	Standard Viola Jones technique			Proposed hybrid technique		
540X720						
Processing time	2.9341s			1.2763s		
Accuracy	Face detection	Left eye detection	Mouth detection	Face detection	Left eye detection	Mouth detection
	100%	10%	8%	99%	100%	99%

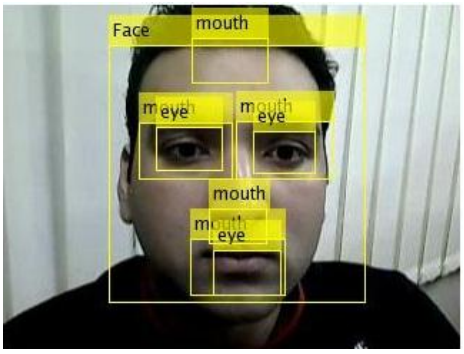
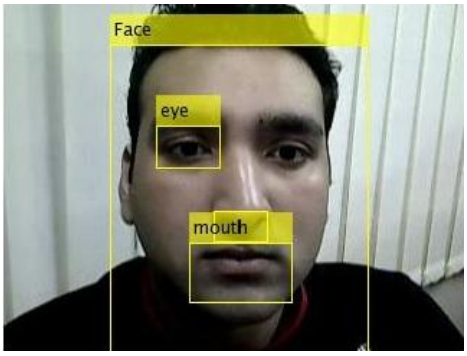
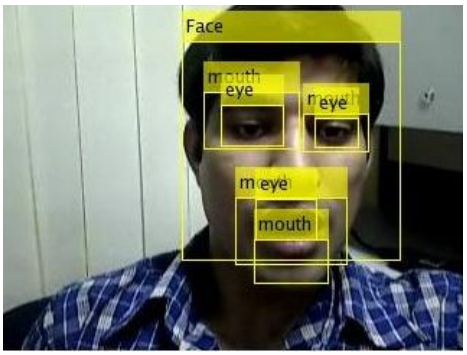
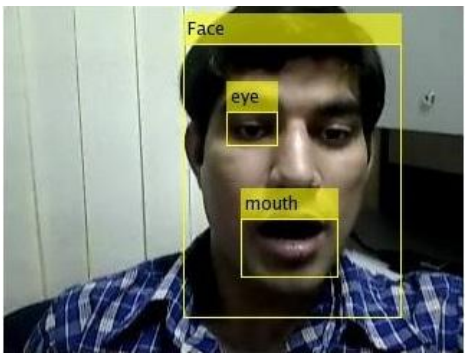
International Journal of Engineering and Applied Sciences

www.eaas-journal.org

<table>
<tr><td>494X385</td><td colspan="3"> 
</td><td colspan="3"> 
</td></tr>
<tr><td>Processing time</td><td colspan="3">1.5324s</td><td colspan="3">1.0302s</td></tr>
<tr><td rowspan="2">Accuracy</td><td>Face detection</td><td>Left eye detection</td><td>Mouth detection</td><td>Face detection</td><td>Left eye detection</td><td>Mouth detection</td></tr>
<tr><td>100%</td><td>35%</td><td>25%</td><td>99.4%</td><td>100%</td><td>100%</td></tr>
<tr><td>326X484</td><td colspan="3"> 
</td><td colspan="3"> 
</td></tr>
<tr><td>Processing time</td><td colspan="3">1.5340s</td><td colspan="3">0.8044s</td></tr>
<tr><td rowspan="2">Accuracy</td><td>Face detection</td><td>Left eye detection</td><td>Mouth detection</td><td>Face detection</td><td>Left eye detection</td><td>Mouth detection</td></tr>
<tr><td>98%</td><td>35%</td><td>25%</td><td>94%</td><td>100%</td><td>100%</td></tr>
</table>

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

326X484						
Processing time	1.3067			0.7347		
Accuracy	Face detection	Left eye detection	Mouth detection	Face detection	Left eye detection	Mouth detection
	99%	40%	20%	95%	100%	70%
326X484						
Processing time	1.4006s			0.8631s		
Accuracy	Face detection	Left eye detection	Mouth detection	Face detection	Left eye detection	Mouth detection
	99.4%	35%	20%	96%	100%	97%

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

5 CONCLUSION

Through researching and experimenting the existing face and facial feature detection methods, we proposed a technique which perform accurately while consuming less time. This technique is a combination of existing Viola Jones with skin color pixel detection. In case of eyes and mouth detection physical location approximation is made in detected face to locate eyes and mouth. This method increased the accuracy of system and decreased consumed time which can be seen in table two, which are the experimental results of ten thousand images. This technique increases the application of system in real time scenarios for face, eyes and mouth detection and most likely in system in which eyes and mouth need to be continuously monitored.

REFERENCES

[1] R. E. S.-Y. a. F. J. M.-P. V. Ayala-Ramirez, "On the Application of Robotic Vision Methods to Biomedical Image Analysis," *IFMBE Proceedings of Latin American Congress on Biomedical Engineering, pp.1160-1162, 2007.*

[2] N. A. M. Yang, "Face Detection and Gesture Recognition for Human-Computer Interaction," *The International Series in Video Computing , vol.1, Springer, 2001.*

[3] K. Parmar*, et al.*, "Facial-feature based Human-Computer Interface for disabled people," in *Communication, Information & Computing Technology (ICCICT), 2012 International Conference on*, 2012, pp. 1-5.

[4] H. L. Z. Guo, Q. Wang, and J. Yang, "A Fast Algorithm of Face Detection for Driver Monitoring," *In Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications,vol.2, pp.267 - 271,*, 2006.

[5] J. Zhichao and Q. Huabiao, "A vehicle surveillance system for face detection," in *Vehicular Electronics and Safety, 2007. ICVES. IEEE International Conference on*, 2007, pp. 1-4.

[6] G. P. Z. Zhang, M. Liu, T. Huang,, "Robust Multi View Multi-Camera Face Detection inside Smart Rooms Using Spatio-Temporal Dynamic Programming," *International Conference on Automatic Face and Gesture Recognition, pp.407-412,*, 2006.

[7] W. Y. D. K. H. Yoon, "Fast Group Verification System for Intelligent Robot Service," *IEEE Transactions on Consumer Electronics, vol.53, no.4, pp.1731-1735, Nov.2007.*

[8] R. C. Luo*, et al.*, "Face Detection and Tracking for Human Robot Interaction through Service Robot," in *Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE*, 2007, pp. 2818-2823.

[9] D. Z. Badal, "Neural network recognition of human face images stored in the database," in *Computers and Communications, 1993., Twelfth Annual International Phoenix Conference on*, 1993, pp. 552-558.

[10] H. A. a. M. S. B. Z. Ijaz Khan, " A Robust Hybrid Design for Driver FatigueDetection,"
JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES, vol. VOLUME 3, MAY 2013.

[11] J. Y. a. A.Waibel, "A Real-Time Face Tracker," *Proc. Third Workshop Applications of Computer Vision, 1996,pp. 142-147.*

[12] Y. D. a. Y.Nakano, "“Face-Texture Model Based on SGLD and Its Application in Face Detection in a Color Scence”," *Pattern Recognition,vol. 29,no. 6, 1996, pp.1007-1017.*

[13] M. C. B. T.K.Leung, and P.Perona, "“Finding Face in Cluttered Scenes Using Random Labeled Graph Matching," *Proc. Fifth IEEE Int’l Conf.Computer Vision, 1995, pp637-644.*

[14] C. J. T. A. Lanitis, and T.F. Cootes,, "An Automatic Face Identification System Using Flexible Appearance Models," *Image and Vision Computing,vol. 13,no. 5, 1995, pp. 393-401.*

[15] D. T. I. Craw, and A. Bennett,, "Finding Face Features," *Proc.Second European Conf. Computer Vision, 1992, pp. 92-96.*

[16] J. M. Shapiro, "Embedded Image Coding Using Zerotress of Wavelet

International Journal of Engineering and Applied Sciences

www.eaas-journal.org

Coefficients," *IEEE Transaction on Signal Processing Vol.41 No.12 Decemrer 1993*.

[17] S. B. H. Rowley, and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. Pattern Analysis and Matchine Intelligence,vol.20,no.1,Jan. 1998*.

[18] K. K. A. Rajagopalan, J.Karlekar,R. Manivasakan,M. Patil,U. Desai, P. Poonacha, and S. Chaudhuri,, "Finding Faces in Photographs,",*Proc. Sixth IEEE Int, l Conf. Computer Vision,1998*.

[19] Y. F. a. R. E. Schapire, "A Decision-Theoretic Generaliztion of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences, no.55, pp. 119-139, 1997*.

[20] F. Bayoumi, *et al.*, "Feature-based human face detection," in *Radio Science Conference, 2004. NRSC 2004. Proceedings of the Twenty-First National*, 2004, pp. C21-1-10.

[21] C. Yao-Jiunn and L. Yen-Chun, "Simple Face-detection Algorithm Based on Minimum Facial Features," in *Industrial Electronics Society, 2007. IECON 2007. 33rd Annual Conference of the IEEE*, 2007, pp. 455-460.

[22] H. Rowley, "Neural network-based face detection,," PhD thesis, Carnegie Mellon University, Pittsburgh (1999).

[23] K. Curran, X. Li, et al. (2005). "Neural network face detection.," *Imaging Science Journal, the 53(2): 105-115*

[24] A. Mohamed, *et al.*, "Face detection based neural networks using robust skin color segmentation," in *Systems, Signals and Devices, 2008. IEEE SSD 2008. 5th International Multi-Conference on*, 2008, pp. 1-5.

[25] P. a. M. J. Viola, "Rapid Object Detection Using a Boosted Cascade of Simple Features,," *in: Proc. IEEE Conf. Computer Vision and Pattern Recognition.*

[26] P. a. M. J. J. Viola, "Robust Real-Time Face Detection," *International Journal of Computer Vision 57(2): 137-154.*, (2004).

[27] Y. F. R.E. Schapire, P. Bartlett, W.S. Lee,, "“Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods”,," vol. The Annals of Statistics, pp. 1651-1686, 1998.

[28] H. X. Xinyu Wang, Xi Chen and Heng Li, "Fast and Robust Face Detection with Skin Color Mixture Models and Asymmetric AdaBoost," *Proc. of SPIE Vol. 7496 749618-1,* 2009.

[29] I. Hsieh, S., Fan, K., C., Lin, C.,, "A statistic approach to the detection of human faces in color nature scene," *Pattern Recognition, 35(7), 1583-1596 (2002).*

[30] V. Vezhnevets, Sazonov, V. and Andreeva, A. (2003). "A Survey on Pixel-Based Skin Color Detection Techniques.."

[31] I. K. a. M. B. Z. Hadi Abdullah, "Low cost Ultra Wide Band hybrid radar design using virtual reference tags as road safety feature in vehicles," *JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGIES,* vol. 3, APRIL 2013.