HDL Based Illumination Invariant highPerformance Face Detection SystemforMobile applications

T. Suguna, Y. Mahesh
II M.Tech,Assistant Professor, ECE, CRECTirupati, INDIA

ABSTRACT

Illumination variation is a big problem in face detection which usually requires a costly compensation prior to classification. To avoid this problem we are proposing a method for face detection irrespective of illumination variations. In this context the contribution of the work is twofold. First we introduce illumination invariant Local Structure Features for face detection. For an efficient computation we propose a Modified Census Transform which enhances the original work of Zabih and Wood [11]. Secondly we introduceanefficient face detection classifier for rapid detection to render high performance face detection rate.The Classifier structure is much simpler because we use only single stage classifier than multi-stage approaches, while having similar capabilities. The combination of illumination invariant features togetherwith a simple classifier leads to a real-time Detection results are presented on two commonlyused databases namely Bio ID set of 1526images and Yale face data base set of 15 people with 11 images for each .We are achieving detection rates of about 99.76% witha very low false positive rate of 0.18%In this paper, we are also proposing a novel hardware architecture of face-detection engine for mobile applications[1]. Here MCT (Modified Census Transform) and Adaboost learning technique as basic algorithms of face-detection engine. The face-detection chip is developed by verifying and implementing through FPGA and ASIC. The developed ASIC chip has advantage in real-time processing, low power consumption, high performance and low cost. So we expect this chip can be easily used in mobile applications.

KEY WORDS --FPGA, MCT (modified census transform), Adaboost algorithm, Bio id test, Yale test

I. INTRODUCTION

Face-detection systems carry out a major role in biometric authentication, which uses features of the face, iris, fingerprint, retina, etc. These systems are usually used in places requiring high security, such as government agencies, bank, and research institutes; it is also applied to two- or three-dimensional face detection in areas such as artificial intelligence and robots, access control Systems, cutting-edge digital cameras and advanced vehicle systems. Recently, the face-detection technology is being adopted in the mobile phone applications because of the pros in easy installation, low-cost, non-contacting method. Most of the existing face-detection engines in digital camera or mobile phone have been run by software. However, the tendency is that the technology is being developed to be run by hardware for improving the processing speed. These days, the technology is to combine hardware technique of face-detection and software technique of emotion, feeling, and physiognomy and fortune recognition. Face detection performance is known to be highly influenced by variations in illumination. Especially in mobile environment, theillumination condition is dependent on the surroundings (indoor and outdoor), time, and light reflection, etc.

The proposed face-detection method is designed to detect in the variable illumination conditions through the MCT techniques, which can reduce the effects of illumination by extracting the structural information of objects and also renders high performance face detection rate by extracting highly reliable and optimized learning data through the Adaboost learning algorithm.

Therefore we advocate the use of inherently illuminationinvariant image features for object detection which conveyonly structural object information. We propose the new featureset Local Structure Features computed from a 3 × 3pixel Neighborhood using MCT [13]. We use this feature set with a single stage classifier. Classifier is a linear classifier which consists of a set of lookup-tables of feature weights. Detection is carried out by scanning all possible analysis windows of size 22 × 22 by the classifier. In order to find faces of various size the image is repeatedly downscaled with a scaling factor of Smra = 1.25[4]. This is done until the scaled image is smaller than the sub window size.

The rest of the paper is organized as follows. Features used in the work are introduced in section 2. In section 3, basic algorithms are presented in detail. In Section 4 training the classifier for final face detection is discussed and in section 5 hardware structure for the proposed method is introduced .The experiment results and analysis are presented in section 6. Conclusions are drawn in section 7.

II. FEATURE GENERATION

The features used in this work are defined as structure kernels of size 3×3 which summarize the local spatial image structure. Within the kernel structure information is coded as binary information. The resulting binary pattern can represent oriented edges, line segments, junctions, ridges, saddle points, etc.

Fig.1 shows some examples of these Structure kernels. On a local 3×3 lattice there exists. The whole procedure can be thought of as non-linear filtering where the output image is assigned the best fitting kernel index at each location. In thenext section we show how the index of the best kernel can be obtained by modifiedcensus transform.

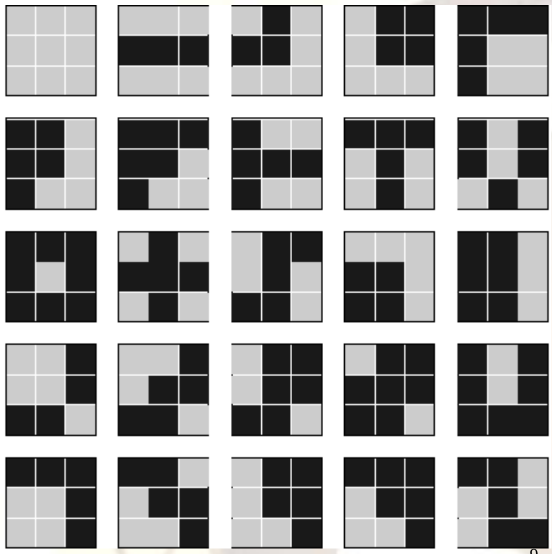

Fig.1. A randomly chosen subset of 25 out of 2^9 – 1possible Local Structure Kernels in a 3 × 3 neighborhood. Kernel is used for description.

III. BASIC ALGORITHMS

A. *MCT (Modified Census Transform)*

The modified census transform could capture all the 3 x 3 local structure kernels, as shown in the Figure1.MCT presents the structural information of the window with the binary pattern {0, 1} moving the 3×3 window in an image small enough to assume that lightness value is almost constant, though the value is actually variable. This pattern contains information on the edges, contours, intersections, etc. MCT can be defined with the equation below:

$$\Gamma(X) = \bigotimes_{Y \in N'} \zeta(\bar{I}(X), I(Y)) \quad \text{---- (1)}$$

Here X represents a pixel in the image, the 3×3 window of which X is the center is W(X); N' is a set of pixels in W(X) and Y represents nine pixels each in the window. In addition,) Ī(X) is the average value of pixels in the window, and I(Y) is thebrightness value of each pixel in thewindow. As acomparison function, ζ ( ) becomes 1 in the case of, Ī(X) <I(Y) other cases are 0. As a set operator, ⊗ connects binary patterns of function, and then nine binary patterns are connected through operations.

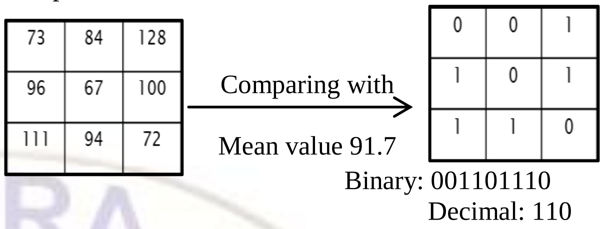

Fig.2 .Example.Converting binary patterns to binary numbers using MCT

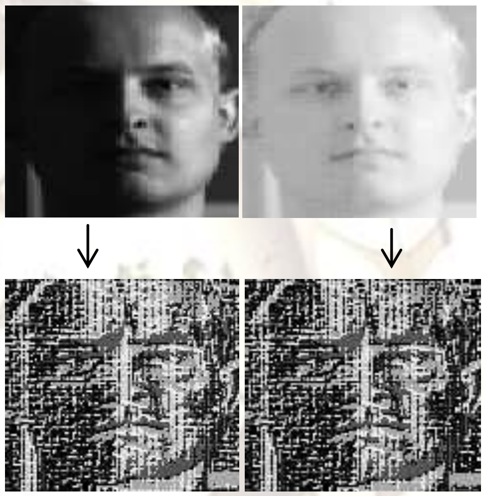

Fig.3. Example of the illumination invariance of the Modified MCT transformed test images

Figure 2 shows an example of structure kernel converting binary patterns to binary numbers by using MCT and figure 3 shows the test images on which the MCT is applied to extract the structural information of the image in different illumination conditions.As a result, a total of 511 structures can be produced, as theoretically not all nine pixel values can be 1. Thus connected binary patterns aretransformed into binary numbers, which are the values of the pixels in MCT transformed images. The main advantages of MCTisthat it is tolerant to illumination and fast in extraction if the information of the ima

B. *ADABOOST LEARNING ALGORITH*

Adaboost learning algorithm is created high-reliable learning data as an early stage for face-detection using faces data. Viola and Jones [2] have proposed fast, highperformance face-detectionalgorithm. It iscomposed of cascade structure with 38 phases, using extracted features to effectively distinguish face and non-face areas through the Adaboost learning algorithm proposed by Freund and Schapire [3]. Furthermore, Froba and Ernst [4] have introduced MCT transformed images and a face detector consisting of a cascade structure with 4 phases using the Adaboost learning algorithm.

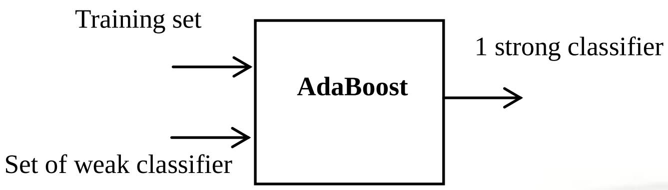

Fig 4. Basic scheme of AdaBoost

This paper consists of a face detector with a single-layer structure, using only the fourth phase of the cascade structure proposed by Froba and Ernst.AdaBoost, short for Adaptive Boosting, is machinelearning. It is used to select a set of features and train a classifierand is based on the idea that a strong classifier can be created by linearly combining a number of weak classifiers.

The algorithm has two main goals:

- Selecting a few set of features which represents as well as possible faces.
- Train a strong final classier with a linear combination of these best features.

A weak classifier $h_t(x)$ consists of a feature f_t, a threshold $\theta_t$and a parity $P_t$indicating thedirection of the inequality sign. It is given in the below equation:

$$h_t(x) = \begin{cases} 1 & if p_t f_t(x) < p_t \theta_t \\ 0 & otherwise \end{cases} \quad (2)$$

In the boosting algorithm T hypotheses areconstructed each using a single feature the final hypothesis is a weighted linear combination of the T hypotheses where the weights are inversely proportional to the training errors. Each iteration T, it will train a best weak classifier which can minimizes the training errors. After T iteration, we can obtain a strong classifier which is the linear combination of the T best weak classifiers multiplied by the weight values α as shown in the below equation:

$$H_T(X) = \sum_{t=0}^{T} \alpha_t h_t(x) \quad (3)$$

IV. RAINING OF CLASSIFIER

The face detector analyzes image patchesW of size 22×22 pixel. In the current setup for face detection each windowhas to be classified either as face or background. For a fast detection system this decision should be computable in an efficient manner. Let H (Γ) be the classifier, which classifies the current analysis window, represented by the modified census features Γ, by

$$H(\Gamma) = \sum_{x \in W} h_x(\Gamma(x)) \quad (4)$$

where**x** denotes the location within the analysis windowand W'⊆ W is the set of pixel-locations with an associatedpixel-classifier hx. The pixel-classifier $h_x$alsocalledelementary classifier consists of a lookup tableof length511 which is the number of the possible kernel indices ofthe Modified Census Transform.

This stage is trained usingthe Winnow update rule[14].It producesthesamekindof lookup-tables, where it uses the entiregivenpixellocations. The decision function for this classifier is given below:

$$H(\Gamma) = H^b(\Gamma) - H^f(\Gamma)$$

$$= \sum_{x \in W} h_x^b(\Gamma(x)) - \sum_{x \in W} h_x^f(\Gamma(x)) \text{ ----- } (5)$$

where$h_x^f$are the weights for the face class andh_x^bthoseof the background class at location x. The pixel-classifierin this case consists of the difference of the two classes -specific weight tables.

As the summation is performed inthe same domain the pixel-weights can be summed up attraining time, hx = $h_x^b - h_x^f$. Two separate weight-tablesare only necessary for training as we shall see.The two sets of weight – tables$\{h_x^b\}$ and$\{h_x^f\}$are trained using an iterative procedure. Initially all weights are set to zero. If a certain weight is addressed for the first time during training it is set to a start-value of 1. The adaptation of the weights is mistake-driven, i.e. if the current pattern is misclassified only the weights addressed by the pattern are changed. The change applies immediately (online update policy). The weights are only updated if the current training pattern is misclassified. The weight update is done with the Winnov Update Rule [14] which is a multiplicative update rule. The finaldecision is made by applying a threshold T to H (Γ),which is determined to achieve a given error ordetection rate on a database different from the training set.

The final pixel classifier h_x is the weighted sum of all weak classifiers $w_t$at location **x**. This classifier His the sum of all pixel classifiers h_x, see Equ. (4). The decision rule for a valid face atstage j is

$$H(\Gamma) \leq T \quad (5)$$

Where T is the score threshold if that particular window. The score thresholdTis tuned so that it maximizes the detection rate on atest database different from the training set.

By using this algorithm there isno need to have any a prior knowledge about face structure. The most representative features will automatically be selected during the learning.At each stage of the learning, the positive and negative examples are tested by the current classifier.The training error theoretically converge exponentially towards 0.

V. PROPOSED HARDWARE STRUCTURE

The proposed hardware structure for this face detection system has shown in Figure 5, it is composed of color conversion module to convert to gray image from color image, noise reduction module to reduce image noise, image scalar module to detect various size of the face. MCT transform module to transform image for robustness various illumination, CD (Candidate detector) / CM (Confidence mapper) to detect candidate for final face-detection, position resizermodule to resize face candidate areas detected on the scaled-down images as their corresponding points of original image size, detection areas.The data grouper module to group the duplicate area determined to be the same face prior to determining

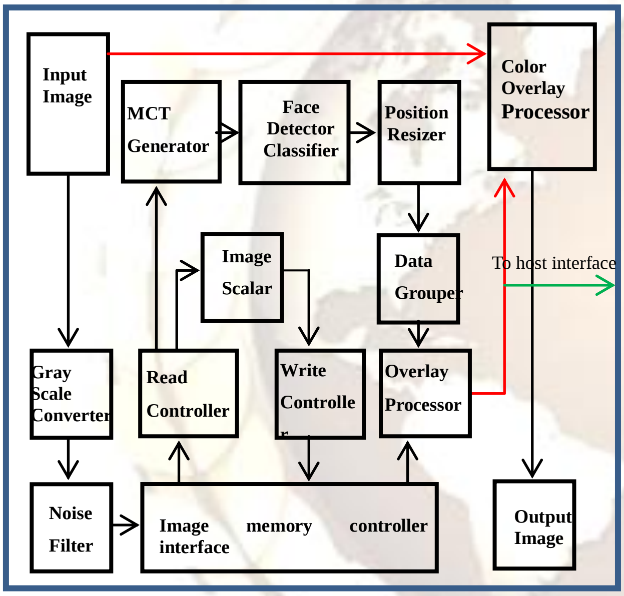

Fig.5. Block Diagram of Proposed Face-Detection System

The final face overlay processor to play in displaying an output by marking square in relation to the final face-detection area on the color-based original image from the camera or to output to transfer the information of area and size in face-detection area to embedded system through host interface.

VI. EXPERIMENTAL RESULTS

The developed face detection system has verified superb performance of 99.76 % detection-rate in various illumination environments using Yale face database [5] and BioID face database [6] as shown in Table 1, and Figure 5. And also we had verified superb performance in real-time hardware though FPGA and ASIC as shown in figure 4. First, the system was implemented using Virtex5 LX330 Board [7] with QVGA (320x240) class camera, LCD display. The developed face-detection FPGA system can be processed at a maximum speed of 149 frames per second in real-time and detects at a maximum of 32 faces simultaneously. Additionally, we developed ASIC Chip [8] of 1.2 cm x 1.2 cm BGA type through 0.18um, 1-Poly / 6-Metal CMOS Logic process. Consequently, the developed face-detection engine has verified at real-time by 30 frames per second within 13.5 MHz clock frequency with 226 mW power consumption.

TABLE 1: FACE DETECTION RESULTS

Test results	Amount used
The Maximum Operating Frequency	54 MHz
The Maximum Operating Speed	149 Frame/Sec
The Number of Simultaneous Face Detection	32 Faces
Face Detection Ratio(Yale db and Bio ID db)	99.76%(1682/ 1686)
False Positive Rate	0.18%

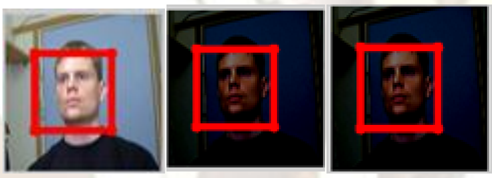

Fig 6. Fig.6. Output of Face Detector on Test Images On different illumination conditions

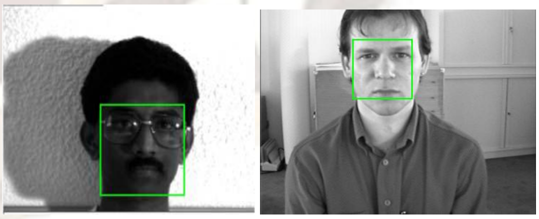

Fig 7. Face detection Results

VII. CONCLUSION

This paper has verified a process that overcomes low detection rates caused by variations in illuminations thanks to the MCT techniques. The proposed face-detection hardware structure that can detect faces with high reliability in real-time was developed with optimized learning data through the Adaboost
Algorithm. Consequently, the developed face

detection engine has strength in various illumination conditions and has ability to detect 32 various sizes of faces simultaneously. The developed FPGA module can detect faces with high reliability in real-time. This technology can be applied to human face-detection logic for cutting-edge digital cameras or recently developed smart phones. Finally, face-detection chip was developed after verifying and implementing through FPGA and it has advantage in real-time, low power consumption and low cost. So we expect this chip can be easily used in mobile applications.

REFERENCES

[1] Dongil Han, *Member*, *IEEE*, Jongho Choi, Jae-Il Cho, and Dongsu Kwak "Design and VLSI Implementation of High-Performance Face-Detection Engine for Mobile Applications",2011 IEEE International Conference on Consumer Electronics (ICCE)

[2] Paul Viola and Michael J. Jones, "Robust real-time face detection" In International Journal of Computer Vision, pp. 137-154, 2004.

[3] Yoav Freund and Robert E. Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting" in Journal of Computer and System Sciences, pp. 119-139, 1997.

[4] Bernhard Froba and Andreas Ernst, "Face detection with the Modified Census Transform", IEEE International Conference. On Automatic Face and Gesture Recognition, pp. 91-96, Seoul, Korea, May. 2004.

[5] Georghiades, A.: Yale Face Database, Center for computational Vision and Control at Yale University, http: //cvc.yale.edu /projects /yalefaces /yalefa

[6] The BioID face database: [http: //www.bioid.com /downloads /facedb /facedatabase.html]

[7] Dongil Han, Hyunjong Cho, Jongho Choi, Jaeil Cho, "Design and Implementation of real-time high performance face detection engine", SP(Signal Processing), 47th Book, 3 Issue, The Institute of Electronics Engineers of Korea Publication(IEEK), March, 2010.

[8] Seung-Min Choi, Jiho Chang, Jaeil Cho, Dae-Hwan Hwang, "Implementation of Robust Real-time Face Detection ASIC", IEEK Conference 33th Book, 1 Issue, 467470,Jeju, Korea, June, 2010.

[9] F.Crow.Summed-area tables for texture mapping.In Proceeding so SIGGRAPH, volume18 (3), pages207212, 1984.

[10] L.G Valiant. Theory of the learnable. Communications of the ACM, 27(11):1134-1142, November1984.

[11] Raman Zabih and John Wood fill. Anon-parametric approach to visual correspondence. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996.

[12] Y.Freund and R.ESchapire. Gametheory, on-line prediction and boosting In Proc.COLT, pages325332, New York, NY, 1996.ACMPress.

[13] Raman Zabih and John Wood fill A non-Parametric approach to visual Correspondence IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996.

[14] Ming-Hsuan Yang, Dan Roth, and Narendra Ahuja. A snow based face detector. In *Advances in Neural Information Processing Systems 12 (NIPS 12)*, pages 855–861. MIT Press,2000.

AUTHORS:

Author1: T.Suguna doing M.Tech in VLSI System Design in CREC in Tirupathi and done bachelor's degree in Electronics and Communication Engineering.

Author2: Y.Mahesh is working as an Assistant professor in the department of ECE, Chadalawada Ramanamma College of Engineering, Tirupati. He published one international and one national journals and attended 5 international,7 national conferences and4 workshops at his credit. He ratified by the J.N.T.U.A, Anantapur