

# OPTIMIZED ALGORITHM FOR FACE DETECTION INTEGRATING DIFFERENT ILLUMINATING CONDITIONS

Sumaya Abusaleh, Varun Pande and Khaled Elleithy  
Department of Computer Science and Engineering  
University of Bridgeport  
{abusaleh, vpande, elleithy}@bridgeport.edu  
Bridgeport, CT 06604, USA

## ABSTRACT

Face detection is a significant research topic to recognize the identity for many automated systems. In this paper, we propose a face detection algorithm to detect a single face in an image sequence in the real-time environment by finding unique structural features. The proposed method allows the user to detect the face in case the lighting conditions, pose, and viewpoint vary. Two methods are combined in the proposed approach. First, we use the components  $Y$ ,  $C_b$ , and  $C_r$  in  $YC_bC_r$  color space as threshold conditions to segment the image into luminance and chrominance components. Second, we use Roberts cross operator [1] to approximate the magnitude of the gradient of the test image and outline the edges of the face. Experimental results show that the proposed algorithm achieves high detection rate and low false positive rate.

**Index Terms**— Face detection,  $YC_bC_r$  color space, Roberts cross operator, Computer Vision, Illumination, Segmentation.

## 1. INTRODUCTION

Recently, Computer Vision has become one of the fields that have inspired a large number of researchers to develop efficient techniques for programming computers to understand the features in the images. Digital image processing technology makes the challenges of automated image interpretation more attractive and interesting. This growing interest can be attributed to useful applications such as: medical imaging, video surveillance, video coding, content-based image retrieval, movie-post processing, human computer interaction (HCI), industrial inspection, and counting people, etc [2] [3] [4]. Vision, as a source of semantic information, is one of the computer Vision themes. This means using algorithmic methods to recognize objects, people, and motions in order to understand the relationships among different components in the real world.

Face detection is the highlight of such automated applications. Detecting the faces is considered as an indispensable first step, because it is concerned with finding out whether or not there are any faces in a given image. If any face appears, it should be localized and extracted from the background.

It is also discussed in [5], that after capturing the image using a camera, some processing should be performed on the image to analyze the information on the detected faces in order to extract the features. These features are important to determine the location of the face, recognize, verify, and track its motion.

However, there are several peculiarities that makes face detection more challenging; faces are non-rigid and have different components. For instance, color, shape, size, and texture are different components of the face. Also the variations in illumination intensity and chromaticity in real-time environment affects the appearance of the skin color of the face. Moreover, the face may also be occluded by other objects such as glasses, a scarf, long hair and partial or full occlusion of faces with each other. Furthermore, facial features such as beards and mustaches may affect the appearance of the face. Orientation of faces can also be affected by facial expressions, for example, smiles, winks, and anger, etc. There are two kinds of rotations that can be also attributed as some of the major challenges associated with face detection. In-plane rotation is the orientation of the image such as frontal, upside down, and profile. Out-of-plane rotation means the angular pose of the face relative to the cameras optical axis.

The paper is organized as follows. In Section 2 we overview some related works. Section 3 illustrates the motivation for our research. Section 4 presents our proposed approach. Section 5 shows experimental results and discussions. The conclusion and future work are given in Section 6.

## 2. RELATED WORKS

Many authors have addressed the problem of face detection and developed many approaches to detect faces. Yang *et al* [6] surveyed the face detection techniques and divided the single image face detection approaches into four areas: knowledge-based methods, feature invariant approaches, template matching methods, and appearance-based methods.

Yang and Huang [7] utilized a hierarchical knowledge-based method to locate unknown human faces in black & white pictures by defining certain rules. The system consists of three levels. In the first level, all the possible candidates

of human face will be found by using a set of simple rules based on the quartet. In the second level, local histogram equalization is applied and a set of face detection rules is established on the basis of octet to detect the eyes, nose, and mouth. In the third level, the human face is detected after extracting some features such as: slope, curvature, concavity, and the coordinates of the ends of a curve.

Yow and Cipolla [8] introduced a feature-based face detection algorithm by using different types of image evidence. In the first phase, the authors design a plane that includes six facial features; eyebrows, eyes, nose, and mouth. Each facial feature is shaped by pairs of oriented edges. For detecting these feature points, the algorithm performs a second derivative Gaussian filter. The second stage is, finding out the edges around these interest points. To achieve that, Geometric and grey level constraints are used to group the interest points into a face candidate. In the third stage, a probabilistic framework is applied to propagate probabilities among the facial groups and evaluate the possibility for the candidate to be a face.

Viola and Jones proposed a new approach to real-time face detection [9]. The proposed framework includes three steps: In the first step, for fast feature evaluation, the integral image representation is used to compute simple Haar-like rectangular features. In the second step, for feature selection, AdaBoost training algorithm is performed to search the most discriminating features in order to construct an effective classifier. In the third step, for fast rejection of non-face or negative sub windows, an attentional cascade structure is prepared. This face detector achieved low false positive and high detection rates.

Among the techniques mentioned above, our proposed technique can be classified as one of the feature variant approaches which mentioned in [6]. We present a face detection algorithm in the presence of varying illumination conditions. We use the components  $Y$ ,  $C_b$ , and  $C_r$  in  $YC_bC_r$  color space as threshold conditions to segment the image into luminance and chrominance components, then we perform Roberts cross operator to approximate the magnitude of the gradient of the test image and outline the edges of the face.

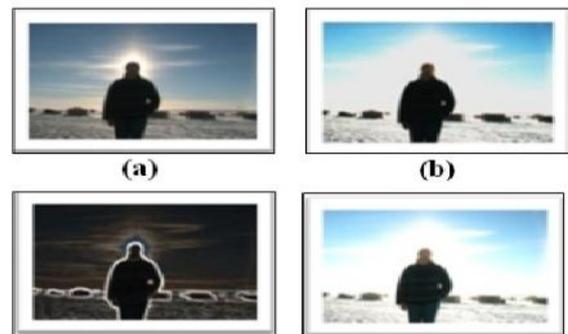
The proposed method allows the user to detect the face in case the lighting conditions, pose, and viewpoint vary in the real-time environment. The technique locates the face in video sequence by finding structural features. Skin color is an effective feature to detect the human faces. Based on the components of  $YC_bC_r$  color model, the pixel can be classified to have skin tone if its value between specific thresholds. So using  $YC_bC_r$  would save the computation time. Furthermore, Roberts cross edge detection algorithm will be applied to separate the integrated regions into the face and highlight these regions of high spatial gradients which are corresponded to edges of the face, Hence, Using Roberts cross operator will need fewer computation.

### 3. MOTIVATION

Since illumination changes over time in the real-time environment; the influences of illumination variation make the detection of the faces quite complex. The illumination includes visible light source, shadow and other illumination gradients [10].

The motivation is to detect a single face in a real-time environment by finding out an effective approach that allows the user to trade-off between accuracy and efficiency in the presence of varying illuminating conditions.

The best way to detect faces or objects in a Vision based system is to understand an image by comparing it with a real Human Vision. Computers, similar to humans, can vary the contrast perception of multimedia. It is a well-known fact that under a fixed luminosity, humans cannot detect subtle variations in color. Figure 1 represents a scenario for which the human eye fails to see the small variation in luminosity. However, the computer can successfully distinguish these variations. As seen in Figure 1.(a) the human eye cannot recognize the face of the person standing. In figure 1.(b) the computer runs a set of algorithms to calculate the luminosity of the image. In figure 1.(c) the computer differentiates the foreground from the background. Figure 1.(d) shows a better resolution for the person's face. Thus, we can conclude that the Computer Vision can see things that the Human Vision cannot perceive.



**Figure 1.** (a) Original image (b) Calculation of luminosity (c) Differentiating foreground and background (d) The new image with better Visibility.

### 4. OUR APPROACH

#### 4.1. Segmentation

Image segmentation is the identification and isolation of an image into regions that correspond to structural units. Our goal is to segment skin pixels from the background. General approaches to segmentation can be grouped into three classes: Pixel Based Methods, Continuity-based methods, and Edge-based methods [11].

## 4.2. Pixel based methods

One of the pixel based methods is thresholding in which all pixels having intensity above or below a certain value will be classified as part of the segment. The Input color image is typically in the RGB format. RGB color model consists of three components: red, green, and blue. Since the RGB color space is sensitive to the light conditions, thus the face detection may fail if the lighting condition changes. Moreover, the human eye is more sensitive to brightness than color. Hence; the human being perceives a similar image even if the color varies slightly. To overcome this issue,  $YC_bC_r$  color space is used in our proposed method. The actual color displayed depends on the actual RGB primaries used to display the signal.

$YC_bC_r$  is a digital transmission color space and it is used in digital video.  $YC_bC_r$  belongs to the television transmission color spaces. These include YUV and YIQ analog transmission color spaces for NTSC and PAL.

$YC_bC_r$  segments the image into the luminance and chrominance.  $YC_bC_r$  color space stores the Luminance information (Y) separately from the chrominance information ( $C_b$  and  $C_r$ ). Luminance (Y) can be defined as the brightness (Luma), however chrominance is usually represented as the difference of two color components,  $C_b$  and  $C_r$ . The Component  $C_b$  is blue minus the luma (B-Y), whereas the component  $C_r$  is red minus the luma (R-Y). The conversion of RGB color space into full-range  $YC_bC_r$  color space is done by the following formula [12]:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ C_b &= -0.1687R - 0.3313G + 0.500B + 128 \\ C_r &= 0.500R - 0.4187G - 0.0813B + 128 \end{aligned} \quad (1)$$

Where  $R, G, B \in [0,255]$ , and  $Y, C_b, C_r \in [0,255]$ . This linear conversion is simple and effective.

## 4.3. Edge- based methods

Edge detectors are elementary tools in Computer Vision, and they are also known as gradient operators. An edge of the image is a substantial local change in intensity of the image. The concept of image gradient means the gradation of the color from low values to high values. Thus, converting two-dimensional image into boundaries or edges is more compact than pixels and leads to the extraction of the salient features of the image, which results in reducing the amount of data to be processed. The gradient operators are used to segment the image and form the outlines of the potential object edges by enhancing the intensity change in the image. Hence, the edges will occur at points where the gradient is at a maximum which eliminates points below the maximum. To achieve this, the magnitude and direction of the gradient is computed at each pixel, so the gradient of an image  $I(x, y)$  is a vector at pixel location  $(x, y)$  with magnitude and direction.

There are four different approaches of the problem of edge detection: Gradient based operator, Template matching, Edge fitting, and Statistical edge detection [13].

Gradient based operator approach consists of two steps. First, the edge strength is measured by calculating a first-order derivative expression such as the gradient magnitude. Second, a computed estimate of the local orientation of the edge is used such as gradient direction in order to look for the local directional maxima of the gradient magnitude.

Roberts cross operator [1] is considered as one of the Gradient- and difference- based operators such as Sobel, and Prewitt. The Roberts Cross operator is used to approximate the gradient of an image through discrete differentiation. To perform that, the sum of the squares of the differences between diagonally adjacent pixels is calculated. The operator consists of a pair of  $2 \times 2$  Convolution kernels. The original image should be convolved with the following two kernels. As seen in Figure.2 the first kernel ( $G_x$ ) and the second kernel ( $G_y$ ) rotated by  $90^\circ$ .

$$\begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$$

$G_x \qquad \qquad G_y$

Figure 2. Roberts cross convolution masks.

The masks are performed separately to the original image to produce separate measurements of the gradient component in each orientation ( $G_x$  and  $G_y$ ); Let  $I(x, y)$  be a point in the original image.  $G_x(x, y)$  be a point in an image formed by convolving with the first kernel, and  $G_y(x, y)$  be a point in an image formed by convolving with the second kernel. These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient. The gradient magnitude can be defined as:

$$\nabla I(x, y) = \sqrt{G_x^2 + G_y^2} \quad (2)$$

Let  $\theta(x, y)$  represents the direction angle of  $I$  at  $(x, y)$ , then the direction of the gradient is expressed as:

$$\theta(x, y) = \arctan \left( \frac{G_y(x, y)}{G_x(x, y)} \right) \quad (3)$$

Hence, Roberts cross gradient operator is used in our proposed approach to estimate the magnitude of the gradient of the test image. The main advantages of using Roberts cross operator in our approach are: simplicity, applicability of spatial gradient measurements on two-dimensional images and computing masks of size  $2 \times 2$  which needs fewer computations. The regions of high spatial frequency will be highlighted. These regions often correspond to edges of the face.

#### 4.4. Proposed method

Several face detection approaches can detect different ethnic groups [14]. The pigments carotene, hemoglobin, and melanin involved in skin color are varying among people. Hence, skin color can be considered as a robust feature to detect human faces. Also, skin color feature allows fast processing.

Our proposed approach is shown in Figure 3. It includes two levels to detect a single face. In the first level, the system uses skin color as a feature for face detection. To achieve that, the camera captures 2 frames every 1 second. The algorithm calculates the RGB color for each pixel in the captured frame. After that, the RGB color space is converted to  $Y C_b C_r$  color space. The newly obtained  $Y C_b C_r$  color frame is decomposed into 3 separate layers of  $Y$ ,  $C_b$  and  $C_r$  components respectively.

The proposed approach has the advantage of creating an interaction between the user and the computer. The user can choose between multiple options which are Face Detection option, Environment Detection option and Luminance option, which is the face detection in varying light conditions.

In the normal lighting conditions, the algorithm detects the face depending only on the value of  $C_r$  to get more accurate results and not the combination between the three color channels. To achieve this goal, user has to choose Face detection option in the GUI. The value of  $C_r$  should be in the range between two threshold values;  $T_1$  and  $T_2$  respectively. After extensive experimentation, we found that the best threshold value for  $C_r$  is:

$$157 < C_r < 180$$

For people with really dark skin the amount of  $C_r$  when calculated on a color space is negligible thus, we set the threshold for  $C_r$  as follows:

$$45 < C_r < 150$$

In some cases the face is partially occluded by other objects such as mask, scarf, long hair, or hat as seen in figure 4.b. It is quite difficult to detect the face in such scenarios. In this case the user should choose the environment detection option in the GUI. This allows the user to have an approximate to where the face is present in the frame. To achieve this, a suitable threshold value for the  $C_b$  components has been set. The algorithm segments the image based on the  $C_b$  components, the system isolates all the skin and non skin pixels giving an outline of the environmental surroundings as a result. The resulting image will be an outline of the environment surrounding the face which it is present in. We set the value of  $C_b$  as follows:

$$157 < C_b < 377$$

In case the face appears in an environment which has low light conditions or there is a shadow, the user has to choose the luminance option in the GUI. The algorithm will perform the detection for the face depending on the value of  $Y$  for the current frame. To achieve this, a suitable threshold value for  $Y$  is set as follows:

$$441 < Y < 740$$

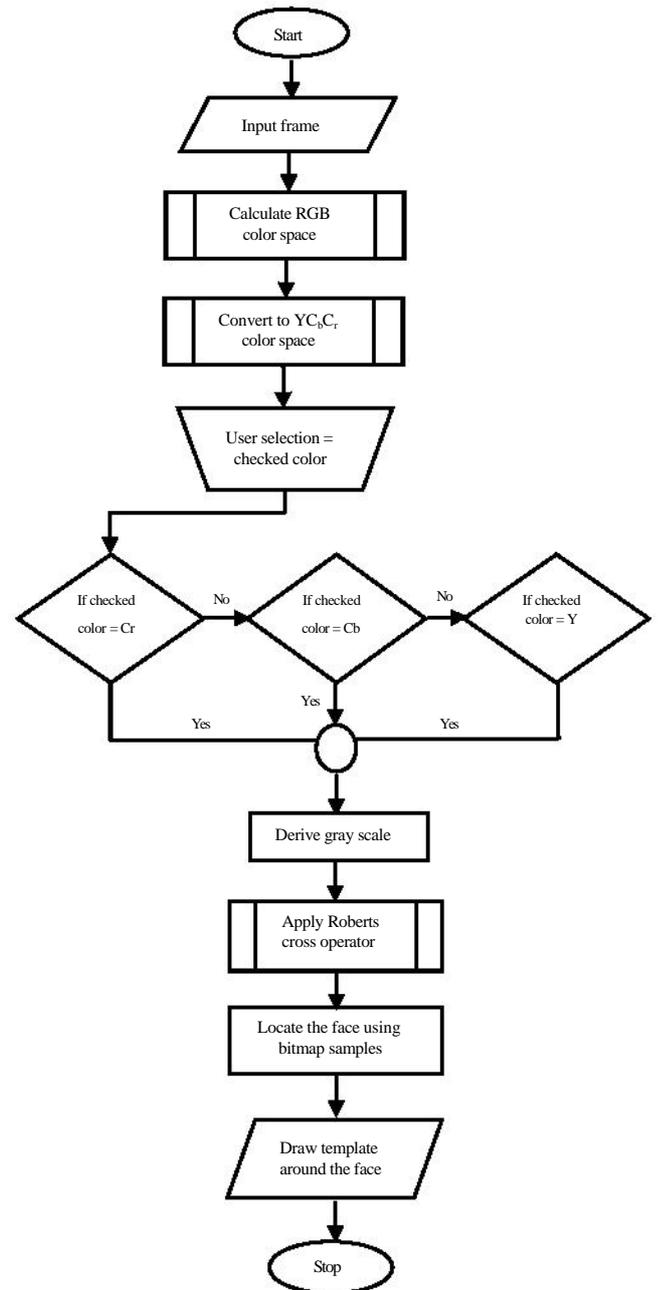
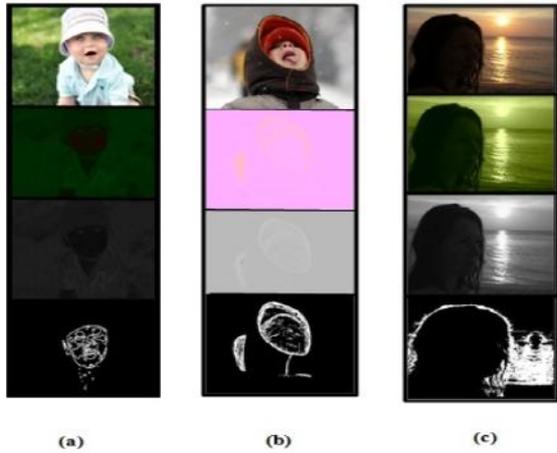


Figure 3. Face detection flow chart



**Figure 4.** (a) Decomposition based on Cr component (b) Decomposition based on C<sub>b</sub> component (c) decomposition based on Y component.

We demonstrate how the proposed method is working based on user selection, Figure 4.(a) shows color image, its C<sub>r</sub> Component, the derived grayscale of C<sub>r</sub> component, and the edge detection after performing the Roberts cross operator. Similarly, Figure 4.(b) shows the decomposition based on C<sub>b</sub> component. Figure 4.(c) shows the decomposition based on the Y component.

In the second level, the Y, C<sub>b</sub> and C<sub>r</sub> layers are converted into grayscale images. The derived grayscale image will be an input for Roberts Cross operator; which outlines the edges. The reason behind using Roberts cross operator is to increase the performance of the processing of the video frames.

It is known that all the distinguishable characteristics of a face lie in the facial features. Two Bitmap images were formed after using a training set of 25 faces for each. These bitmaps are used as samples in the proposed method. Either a small template of size 50 × 50 pixels or a large template of size 25 × 25 pixels is drawn around the face based on the distance between the camera and the face in the real-time environment.

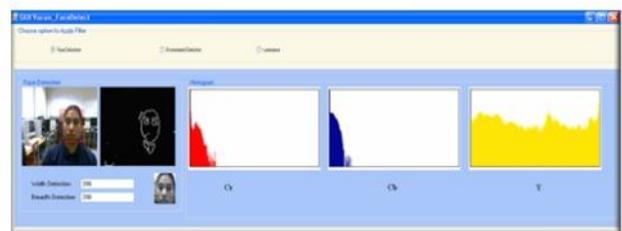
## 5. EXPERIMENTAL RESULTS

The accuracy of face detection systems can be evaluated by measuring some important factors. The first one is detection rates which mean how many faces that have been detected by the system (Computer Vision) over the actual number of faces in the whole image (Human Vision). The second factor is false positive, which means the number of regions claimed to be faces by the system (Computer Vision), but they are not. However, a false negative means the system will not detect the faces and results in lower the detection rate. We did our experiments by capturing a sequence of images for single face in real-time environment with complex backgrounds.

The system is written using C# language and operating on an Intel Core 2 duo Processor at a speed of 2.2GHz. The image resolution is 149 by 154. The graphical user interface (GUI) includes three histograms that show the user the amount of Luma (Y) in yellow color and the amount of Chrominance shown with C<sub>r</sub> in red color and C<sub>b</sub> in blue color in the real-time system as seen in Figure 5. Even though the user has difficulty in detecting the face, with the values provided in the histograms; C<sub>r</sub> is a histogram for the difference of red. C<sub>b</sub> is a histogram for the difference of blue. Y is a histogram for illumination components which appears in yellow color. The user can choose the best option to detect the face for constantly changing illuminating conditions.

As illustrated in table.1, many experiments have been done on 1100 samples. We divided the samples into two categories of 550 frames each with presence or absence of the face respectively. We compared between the Computer Vision and Human Vision in order to calculate the following values: true positive, false positive, true negative and false negative [15].

Two charts have been generated to analyze the results given in Figures 6 and 7. We have achieved 90% detection rate and 10% false positive rate. There are several advantages of the proposed system. It can detect the face regardless of how far the face is from the camera's point of view. Moreover, the proposed method can detect the face in different orientations (frontal, profile, 45 degrees) as well as different poses on still images. Furthermore, the proposed method is able to detect faces of people with different ethnicity. Finally, under variation of lighting conditions, such as cloudy or sunny weather, our approach detects a single face with low false positive rate because the luminance option in the GUI allows the detection of the face in low light condition.



**Figure 5.** Output of the proposed face detection method

**Table 1.** Computer Vision verses Human Vision

Computer Vision \ Human Vision	Positive	Negative	Total
Presence	512	38	550
Absence	54	496	550
Total	566	534	1100

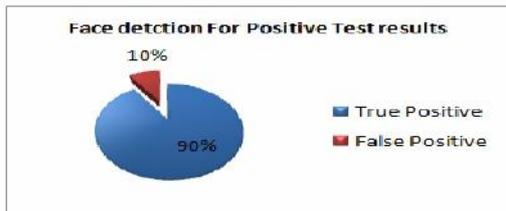


Figure 6. False positive rate versus true positive rate

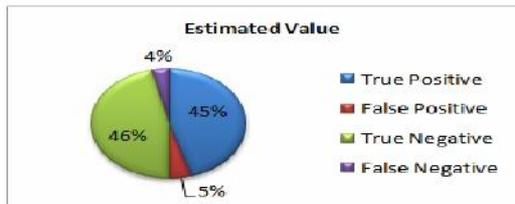


Figure 7. Accuracy of the proposed method

## 6. CONCLUSION

We have presented an effective method to detect faces in real-time environment. Our results show that the proposed algorithm achieves 90% detection rate and 10% false positive rate. One of the main issues that make face detection techniques a hard task is how to cope with different illumination conditions. Our technique detects the face under variation of lighting conditions. The proposed algorithm combined two segmentation approaches. The first technique is a Pixel-based approach by using the components  $Y$ ,  $C_b$ , and  $C_r$  in  $YC_bC_r$  color model as threshold conditions to segment the image into luminance and chrominance components. The second technique is Edge-based approaches by using Roberts cross operator to approximate the magnitude of the gradient of the test image and outlining the edges of the face.

We will be extending our research work by implementing the proposed approach using IMB400 wireless multimedia sensor network boards. In the application field of wireless sensor networks (WSNs), computational resources are very limited. Extracting features such as skin color and edges requires fewer computations which makes them effective in the field of WSNs applications.

## 7. REFERENCES

- [1] V.B. Baru, C.S. Beokar, " VLSI Implementation of Real-Time Image Segmentation", International Conference on Industrial Electronics, Control, and Robotics (IECR), pp.211-214, 2010.
- [2] A. Boyali, M. Kavakli, " 3D and 6DOF User Input platform for Computer Vision Applications and Virtual reality", International Symposium on Innovations in Intelligent Systems and Application (INISTA), pp. 258-263, 2011.
- [3] H. Meng, N. Pears, C. Bailey, "A human Action Recognition System for Embedded Computer Vision Application", Department of Computer Science, The University of York, IEEE conference on computer vision and pattern recognition, June, 2007.
- [4] W.L. Chan, Leo, S.L. Pang, C.F. Ma, "Computer Vision Applications in Power Substations", Proceedings of the 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies, DRPT 2004, Hong Kong, pp. 383-388, April, 2004.
- [5] S. Kr. Singh, D.S. Chauhan, M. Vatsa, R. Singh, "A robust Skin Color Based Face Detection Algorithm", Tamkang Journal of Science and Engineering, pp.227-234, Vol.6, No.4, 2003.
- [6] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey", IEEE Transaction on Pattern Analysis and machine Intelligence, pp.34-58, Vol. 24(1), 2002.
- [7] G. Yang, T.S. Huang, " Human face Detection in a Scene", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.453-458, 1993.
- [8] K.C Yow, R. Cipolla, "Feature-Based Human Face Detection", Image and Vision Computing, Vol.15, No.9, pp.713-735, 1997.
- [9] P. Viola, M. Jones, "Robust Real-Time Object Detection", Second International Workshop on statistical and Computational theories of Vision-Modeling, Learning, Computing, And Sampling, Vancouver, Canada, 2001.
- [10] D. Marr, E. Hildreth, "Theory of Edge Detection", Proceedings of the Royal Society of London. Series B, Biological Sciences, Vol. 207, No. 1167, pp. 187-217, 1980.
- [11] J.L. Semmlow, "Biosignal and biomedical Image Processing: Matlab Based Applications", ISBN: 0- 8247-4803-4, pp.343, 2004.
- [12] G. Wyszecki, W.S. Stiles, "Color Science: Concepts and Methods, Quantitative Data and Formulae", John Wiley and Sons, 2000.
- [13] K.S Sim, L.W. Thong, M.A. Lai, C.p. Tso, "Enhancement of Optical Images Using hybrid Edge Detection Technique", Conference on Innovative Technologies in Intelligent Systems and Industrial Applications(CITSI 2009), pp.186-191, 2009.
- [14] M.J. Jones, J.M. Rehg, "Statistical Color Models with Application to Skin Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, 1999.
- [15] P. Brennan, A. Silman, "Statistical methods for assessing observer variability in clinical measures", BMJ, Vol.309, 1992.