A Robust Algorithm for Face Detection in Color Images Based on Color Segmentation and Neural Network Techniques

Safwan R. Wshah and Ibrahim M. Mansour *

ABSTRACT

Human face detection plays an important role in applications such as biometric identification, video conferencing, intelligent human computer interface, face image database management, and face recognition. We propose a face detection algorithm for color images in the presence of varying light conditions as well as complex background based on light control, skin detection and color segmentation techniques. Our method detects the faces' rectangle that contains eyes and mouth. The algorithm constructs expected regions resulted from skin detection and color segmentation stages and search inside them for any possible face features (eyes, and mouth) and pass these expected mouth and eyes rectangle to a neural network to confirm face validation.

Experimental results demonstrate successful face detection over a wide range of facial variation in color, position, scale, orientation, 3D pose, and expression in images from several photo collections.

KEYWORDS: Face detection, Face recognition, light enhancement, skin detection, face feature extraction and neural networks.

1. INTRODUCTION

Human face detection plays an important role in applications such as biometric identification, video conferencing, intelligent human computer interface, face image database management and face recognition. Face recognition has many applications such as identification for law enforcement, authentication for banking, security system access and personal identification among others.

Face detection is not straightforward because it has lots of variations of image appearance, such as pose variation (front, non-front), occlusion, image orientation, illuminating condition and facial expression. A lot of research is going on in the area of human face detection at present (Horprasert et al., 1996:242-24). Many researchers have proposed different approaches to address face detection problem. These approaches utilize techniques such as neural networks, machine learning, (deformable) template matching, Hough transform, motion extraction, and color analysis (Hsu et al., 2001).

Ming-Hsuan Yang survey (Yang et al., 2002:34-58)

classifies the face detection methods inside single image into four categories: Knowledge-based methods, feature invariant approaches, template matching methods and appearance-based methods. The knowledge-based methods are mainly designed for face localization, the difficulty in this approach is to detect faces in different poses since it is challenging to enumerate all possible cases. The feature invariant approaches aim to find structural features that exist even when the pose, viewpoint or lighting conditions vary, and then use these to locate faces. One problem with these feature-based algorithms is that the image features can be severely corrupted due to illumination noise, and occlusion (Yang et al., 2002:34-58).

In template matching methods a standard face pattern (usually frontal) is manually predefined or parameterized by a function. The existence of a face is determined based on the correlation values computed for the face features. This approach has the advantage of being simple to implement. However, it cannot effectively deal with variation in scale, pose and shape. Appearance-based methods rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face and non-face images. The learned characteristics are in the form o f distribution models or discriminate
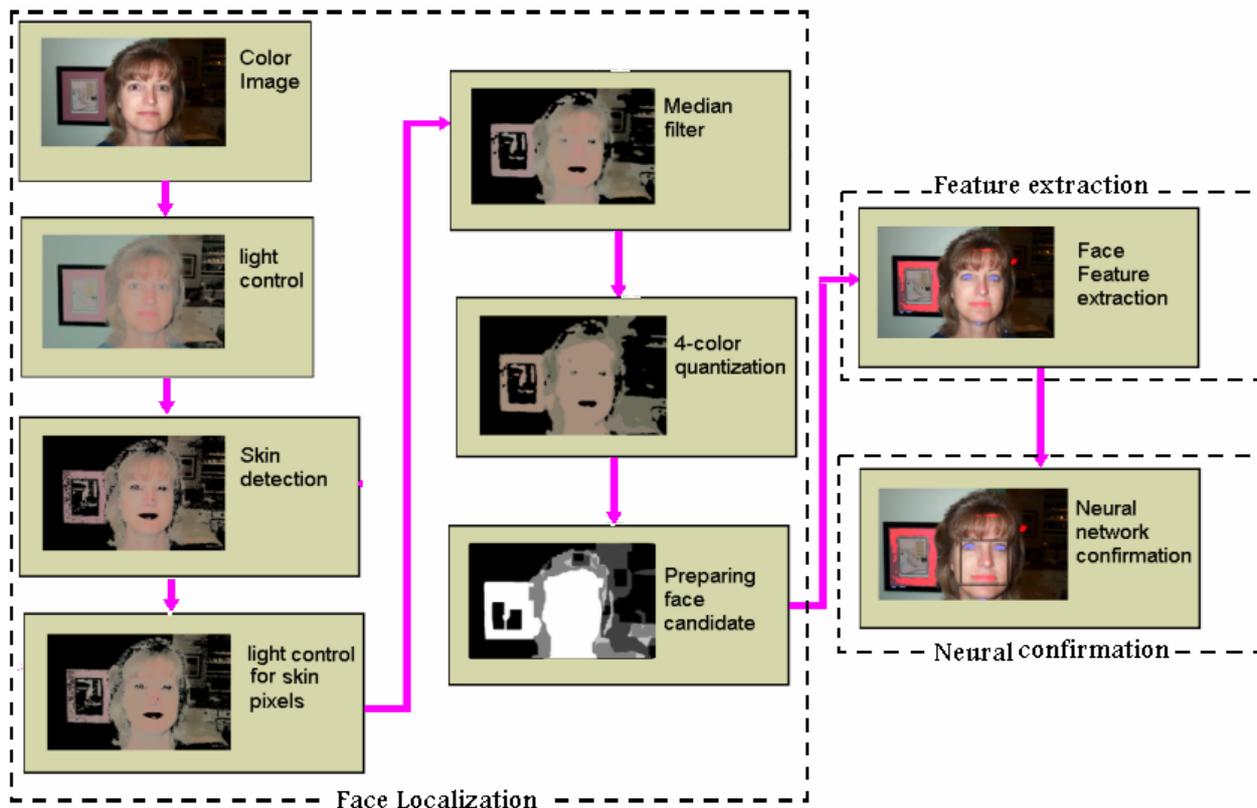
Fig. 1. Face detection algorithm.

functions that are consequently used for face detection. All Appearance-based methods are based on multi-resolution window scanning to detect faces at all scales, making them computationally expensive (Yang et al., 2002:34-58).

2. THE PROPOSED APPROACH

Our approach for face detection is designed to robust the variations that can occur in face illumination, shape, color, pose and orientation. An overview of the presented face detection algorithm is depicted in Fig. 1 which contains three major stages:

1) Face localization for finding face candidate.
2) Extraction of the expected face features from the resulted candidate.
3) Confirmation of the detected face using neural network technique.

The algorithm first uniformly distribute the brightness based on light control technique, the benefit of using the light control technique is to get uniform light distribution for all face parts. The skin tone pixels are detected based on chrominance of multi color spaces under the

assumption of discrete probability distribution of skin colors. The light control is applied another time on the skin pixels only to insure uniform brightness distribution. In the next step, the color quantized into four levels based on octree algorithm (Clark, 1996: 54-57 and 102-104), due to this quantization, the image is divided into many segments according to its colors. Because the uniform distribution of face colors produced by light control, the face is segmented in one segment.

For each separated segment resulted from the 4-levels quantization method, the face feature extraction is applied to extract possible mouth and eyes areas and pass its rectangles that may contain a face to neural network to confirm the face validation.

3. LIGHT CONTROL (UNIFORM BRIGHTNESS DISTRIBUTION)

Different light conditions cause the intensity histogram of the image to be concentrated in some different ranges, these changes in lighting condition dramatically decrease recognition system performance. Also, they cause a big problem in color segmentation results.

Many techniques have been proposed to spread out the intensity and make the image easier to analyze. Such as Histogram equalization, local histogram equalization, gamma correction, etc. These methods do not correctly improve all parts of the image, especially when the image is irregularly illuminated, some details of the results will remain too bright or too dark (Starovoitov et al., 2003).

In our algorithm, we enhanced the new light control technique proposed by Shadeed (Shadeed et al., 2003:890-893). to be suitable with the face's criteria. The light control is performed only on the luminance component of the color (i.e. Y-component of YUV space or V-component of HSV space) while keeping image chrominance unchanged (i.e. UV of YUV space and HS of HSV space). Its main idea is to map different image luminance averages to the user specified values (Shadeed et al., 2003:890-893). The averages are:

- General luminance average (G_{av}), which is the average of all histogram bins.
- Upper luminance average (U_{av}), which is the average of histogram bins between (G_{av}) and the maximum intensity (Max) exist in the histogram.
- Lower luminance average (L_{av}), which is the average of histogram bins between the minimum intensity (Min) exist in the histogram and G_{av}.

So the user should specify the Desired General average (DG_{av}), the Desired Upper average (DU_{av}) and the Desired Lower average (DL_{av}). If image's averages equal the desired values, the image will not change. Image averages are computed from image luminance histogram. The process of light control is performed as follows:

1. Convert image to HSV (or YUV) color space.

2. Compute the histogram (h) of V (or Y) component.

3. Compute the general luminance average as follows:

$$G_{av} = \frac{\sum_{i=Min}^{Max} i * h[i]}{\sum_{i=Min}^{Max} h[i]} \qquad (1)$$

Where h[i] is the histogram value at intensity i.

$$U_{av} = \frac{\sum_{i=G_{av}}^{Max} i * h[i]}{\sum_{i=G_{av}}^{Max} h[i]} \qquad (2)$$

4. Compute the upper luminance average as follows:

$$L_{av} = \frac{\sum_{i=Min}^{G_{av}} i * h[i]}{\sum_{i=Min}^{G_{av}} h[i]} \qquad (3)$$

5. Compute the lower luminance average as follows:

6. Set the desired general (DG_{av}), upper (DU_{av}) and lower (DL_{av}) averages.

7. Map the values of (Min-L_{av}) into (0-DL_{av}), (L_{av} -G_{av}) into (DL_{av} -DG_{av}), (G_{av} -U_{av}) into (DG_{av} -DU_{av}) and (U_{av} -Max) into (DU_{av} -200).

8. Convert the image back to RGB color space.

Fig. 2 shows the block diagram of the proposed light control technique.

This technique does not change the color information and yields an image within the specified lighting conditions irrespective of the image lighting conditions. It can be applied more than once on the same image to reach stable conditions. Fig. 2 shows some results of this technique:
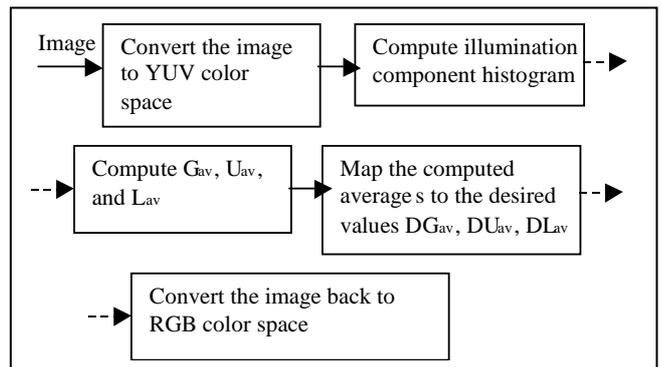

Fig. 2. Light control technique block diagram.

The benefit of using the light control technique is to get uniform light distribution for all the face parts. We found that the control of the light through Y-component gave better results than any other luminance components. Uniform brightness of all face parts is done by brightening the dark and darkening the high brightness areas. This can be controlled through average values (DL_{av}, DG_{av}, and DU_{av}). We applied different average values over 1000 of test images and we found that the best results are DL_{av} = 100(to brightening the dark areas), DG_{av}=130 (to keep the midtones light ranges in the same level), and DU_{av}=150 (to darkening the high brightness areas).

Fig. 3 demonstrates an example of the light control technique compared with other used techniques, note how the brightness is uniformly distributed through all the face parts.

(a) Original Image.   (b) Histogram Equalization.

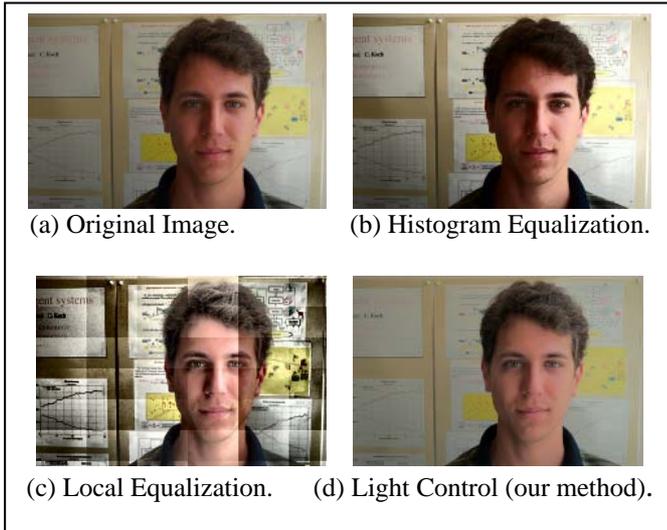(c) Local Equalization.   (d) Light Control (our method).

Fig. 3. Light enhancement (a) original image, (b) Histogram equalization effect on original image, (c) local histogram equalization effect on original image, (d) light control effect on original image.

4. COLOR SEGMENTATION

4.1 Skin Detection

The goal of skin detector is to distinguish between skin and non-skin pixels in order to reduce the search area.

When building a system that uses skin color as a feature for face detection, the researcher usually faces two main problems. First, what color-space to choose, second, how exactly the skin color distribution should be modeled (Vezhnevets et al., 2003: 85-92).

Many works on skin detection drop the luminance component of the color-space. This decision seems logical, as the goal is to model what can be thought of as "skin tone", which is more controlled by the chrominance than luminance coordinates. The dimensionality reduction achieved by discarding luminance, also simplifies the consequent color analysis. Another argument for ignoring luminance is that skin color differs from person to person mostly in brightness and less in the tone itself (Shin et al., 2002).

Inside our algorithm, we used illumination invariant skin detector based on chrominance of the different color spaces, such as HS from HSV color space, UV from YUV color space, the AB from SCT (spherical color transform) color space, and a*b* from CIE L*a*b* color space.

The main work is concentrated on building a color model that results in minimum false detection of non-skin pixels and maximum correct detection for skin pixels, this is done by selecting appropriate thresholds for each color channel that contains the skin tones.

A general color model is constructed from the generic training set using a histogram with full bins per channel. The histogram is done for each color component separately (i.e. histogram for H, histogram for S, histogram for A, histogram for B…etc). This histogram counts are converted into a discrete probability distribution (P(.)) as follows:

$$P(i) = \frac{C[i]}{T_C} \tag{4}$$

Where $C[i]$ gives the count in the histogram bin associated with the color component i and T_C is the total count obtained by summing the counts in all of the bins.

The generic training set is used to construct skin detection model that contains about 1000 samples of faces' images with different illumination conditions ranged from poor lighting to bright lighting with no control on the direction of the illumination source. These pictures were captured in both outdoor and indoor situations, and are thought to represent a large sample of the typical lighting conditions met in personal pictures captured with different digital cameras. We only considered images that were thought to be captured with a white illuminant, as a non-white illuminant changes completely the true skin color.

For each image inside the training set, we extract the face region manually. These images were prepared to include only the skin pixels used in calculating the chrominance histograms.

Figs. 4, 5, 6, and 7 show chrominance histograms for each color space, from these histogram we can find the best thresholds that give minimum false detection for non-skin pixels and maximum correct detection for skin pixels.

The selected ranges of the skin tones are combined together using AND operator as shown in Fig. 8, this figure shows the flow chart of skin detection.

The equations listed below are used in converting RGB color space to the other required spaces:

RGB to YUV (Vezhnevets et al., 2003):

$$Y = R * 0.299 + G * 0.587 + B * 0.114 \tag{5}$$
$$U = B - Y \tag{6}$$
$$V = R - Y \tag{7}$$

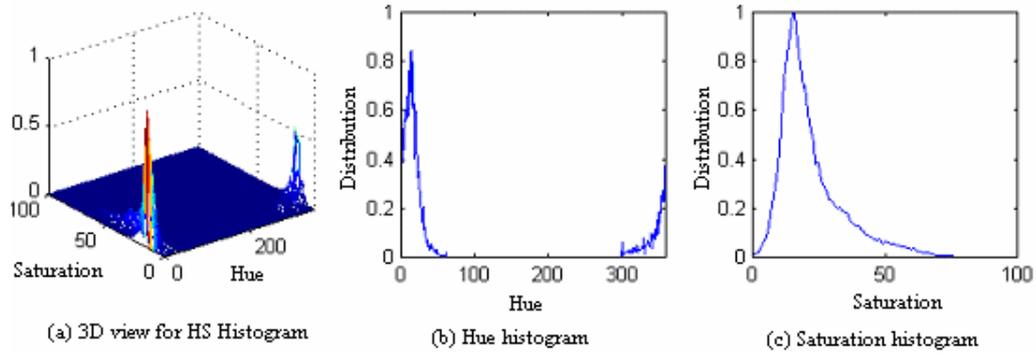**RGB to SCT** (Brainard, 1989):

$$L = \sqrt{R^2 + G^2 + B^2} \tag{8}$$

(a) 3D view for HS Histogram

(b) Hue histogram

(c) Saturation histogram

Fig. 4. HS Histograms (from HSV color space) for the chrominance of the skin pixels.

(a) 3D view for UV Histogram

(b) U histogram

(c) Value histogram
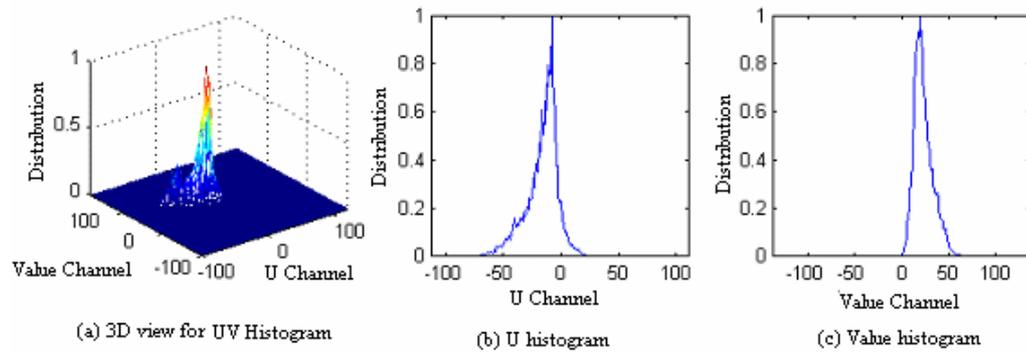
Fig. 5. UV Histograms (from YUV color space) for the chrominance of the skin pixels.

(a) 3D view for CIE L*a*b* chrominance Histogram

(b) CIE a* Histogram
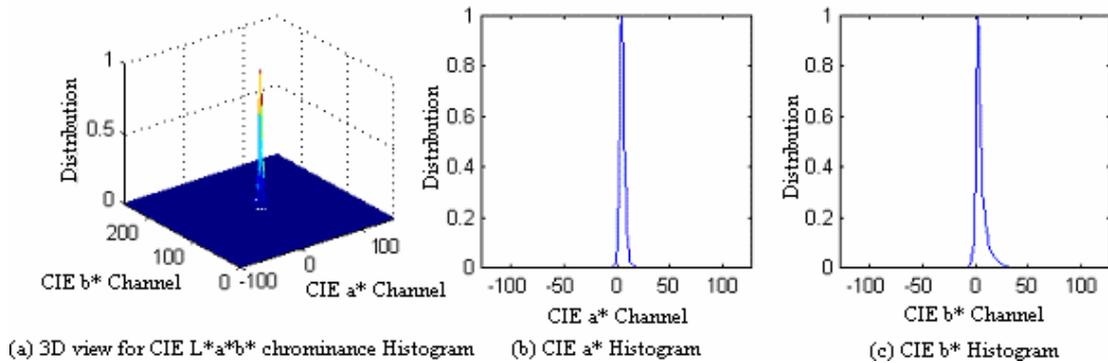
(c) CIE b* Histogram

Fig. 6. CIE a*b* Histograms (from CIE L*a*b* color space) for the chrominance of the skin pixels.

(a) 3D view for SCT chrominance Histogram

(b) A Histogram
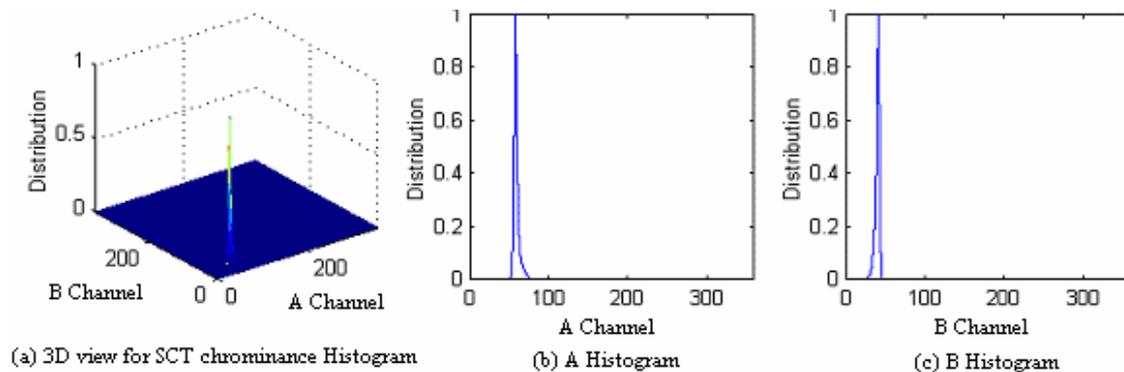
(c) B Histogram

Fig. 7. AB Histograms (from CST color space) for the chrominance of the skin pixels.

$$H = \begin{cases} \dfrac{(-B+G)*60}{V-\min(R,G,B)}, & \text{if } R=V \\[2mm] \dfrac{(-R+B)*60}{V-\min(R,G,B)}+120, & \text{if } G=V \\[2mm] \dfrac{(-G+R)*60}{V-\min(R,G,B)}+240, & \text{if } B=V \\[2mm] \text{undefined}, & \text{if } R=G=B \end{cases} \qquad (13)$$

RGB to CIE L*a*b (Brainard, 1989):

$$R,G,B = \begin{cases} \left(\dfrac{\dfrac{R,G,B}{255}+0.055}{1.055}\right)^{2.4} & if\left(\dfrac{R,G,B}{255}>0.04045\right) \\[4mm] \dfrac{R,G,B}{12.92} & else \end{cases} \qquad (14)$$

$$X = (R*41.24 + G*35.76 + B*18.05)/95.047 \qquad (15)$$
$$Y = (R*21.26 + G*71.52 + B*7.22)/100.000 \qquad (16)$$
$$Z = (R*1.93 + G*11.92 + B*95.05)/108.883 \qquad (17)$$

$$X,Y,Z = \begin{cases} (X,Y,Z)^{\frac{1}{3}} & if\,(X,Y,Z>0.008856) \\[3mm] (7.787*X,Y,Z)+\left(\dfrac{16}{116}\right) & else \end{cases} \qquad (18)$$

$$CIE\text{-}L^* = (116*Y) - 16 \qquad (19)$$
$$CIE\text{-}a^* = 500*(X-Y) \qquad (20)$$
$$CIE\text{-}b^* = 200*(Y-Z) \qquad (21)$$

4.2 (4-Levels) Color Quantization

Because of the uniform distribution of brightness in all the face segments that were caused by light control, we don't hesitate to use the color segmentation algorithms. The main idea behind the color segmentation is to separate the face segments from the other parts of the image.

We found that the use of color quantization caused a high success rate in face segmentation process, we quantize the image into 4 color levels, based on octree technique developed by M. Gervautz and W. Purgathofer (Clark, 1996: 54-57 and 102-104).

Color quantization reduces the image colors into 4 levels and caused the image to be segmented into many segments according to its colors. Because the uniform distribution of the face colors caused by light control, the face segmented in one segment. Fig. 9 shows the effect of light control on the color quantization process for different samples.

Fig.8. Skin Detection Classifier Flow Chart.

$$\angle A = \cos^{-1}\left(\frac{B}{L}\right) \qquad (9)$$

$$\angle B = \cos^{-1}\left(\frac{R}{L\sin(\angle A)}\right) \qquad (10)$$

L contains the brightness info and angles A and B contain the color info.

RGB to HSV (Vezhnevets et al., 2003):

$$V = \max(R,G,B) \qquad (11)$$

$$S = \begin{cases} \dfrac{(V-\min(R,G,B))*100}{V}, & if\ V \neq 0 \\[3mm] 0, & if\ V = 0 \end{cases} \qquad (12)$$
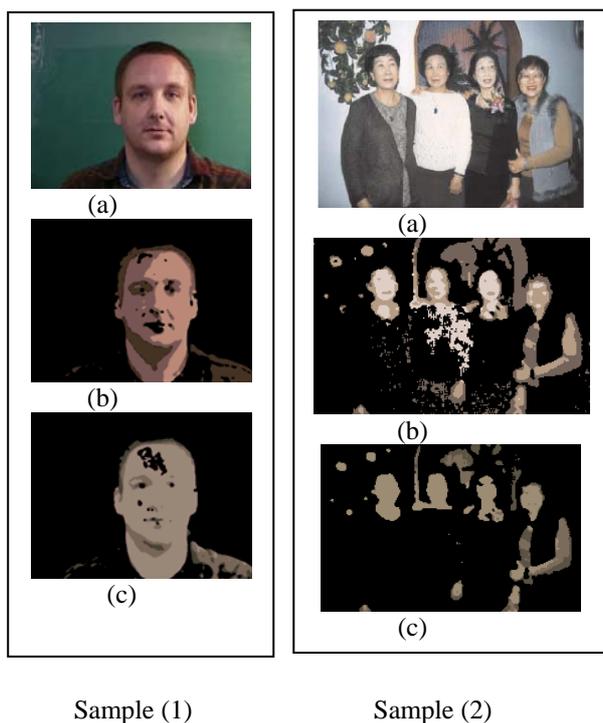
Sample (1)   Sample (2)

Fig. 9. Effect of light control on color quantization, for each sample (a) original image (b) skin region quantized without applying light control (c) skin region quantized after applying light control.

5. FACE FEATURE EXTRACTION

Among the various facial features, eyes and mouth are the most suitable features for recognition and estimation of 3D head pose (Horprasert et al., 1996:242-247). Most approaches to eye and face localization (Smeraldi et al., 2000: 323-329; Yang and Kriegman, 2002: 34-58) are template based. However, this approach has proven to be inadequate for face features detection since it cannot effectively deal with variation in scale, pose, rotations, and shape (Yang and Kriegman, 2002: 34-58).

However, we can locate possible areas for eyes and mouth regions, based on their feature maps derived from chrominance and luminance components. Our approach considers only the area covered by a mask that is built by filling the inside region of the outer layout, this can be done by applying a large size of erosion and dilation on the segmented areas. Fig. 10 shows an example of the face mask.

5.1 Mouth Detection:

The idea behind mouth detection is to find color spaces that separate mouth color from the other face region colors.

The color of the mouth region contains stronger red

component and weaker blue component than any other facial region, since the mouth region chrominance component "a*" from "CIE L*a*b*" and "V" from YUV color space is different than any other facial region color that we can adopt as a key for mouth extraction. Fig. 11 shows an example of "a*" and "V" channel for the selected face.

(a)

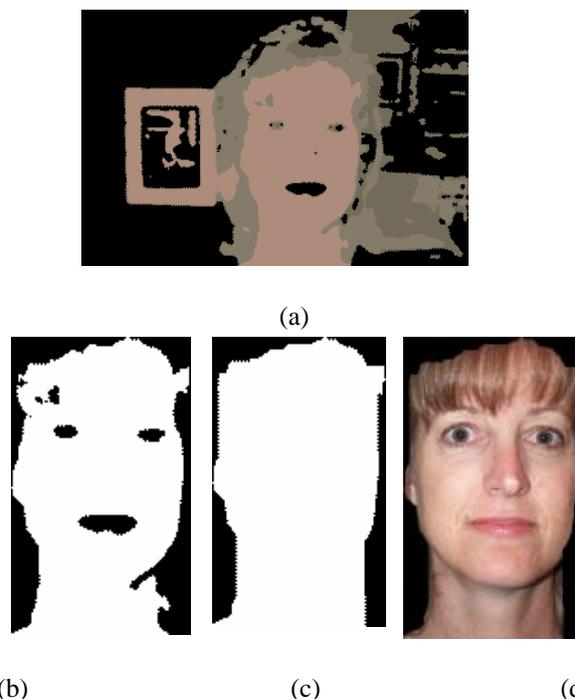(b)   (c)   (d)

Fig. 10. Face mask: (a) segments resulted from 4-level quantization (b) one of the face candidates (c) face candidates after filling the inside region of the outer layout (d) original image mask.

Skin detector passes only the skin tones, resulting in concentration of the color intensity in strict ranges inside its histogram. Fig. 12 (a, and b) shows this concentration for "a*" and "V" histograms.

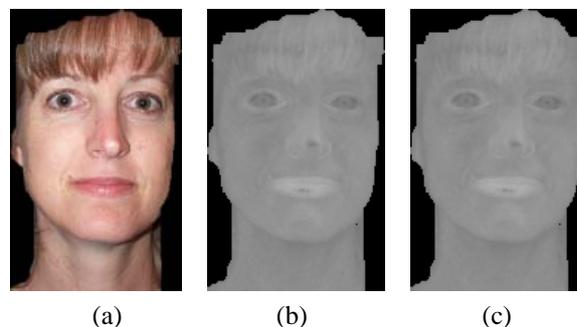(a)   (b)   (c)

Fig. 11. Color spaces used for mouth detection (a) Original image (b) a* channel from CIE L*a*b* color space (c) V channel from YUV color space.

Histogram stretching will cause the colors to be expanded and cover all histogram ranges, the stretching process will enhance the image contrast by making the high intensity colors (mouth colors) expanded toward the high intensity values, while the low intensity colors (skin colors) will be expanded toward the low intensity values as shown in Fig. 12c and Fig. 12d.

In order to get more separation between mouth colors and any other face regions; we apply a nonlinear mathematical operations to cause the colors that have high intensity values to be approximately in the same level, and the colors that have low intensity values to be darker. This can be done by normalizing the result of color multiplication by itself as follows:

$$Color = \frac{(Color)^2}{255} \tag{22}$$

Then, we apply suitable thresholds to separate high intensity values (mouth colors) from the other intensities. These operations were applied separately on "a*" and "V" channels. Applying (AND) operation between the resulted images for each channels ("a*" and "V") will minimize the number of non-mouth pixels as shown in the diagram of the mouth detection.

In last step, some of the binary filters (median, erosion and dilatations) are applied to filter out the small segments. Fig. 13 shows the expected mouth region extraction algorithm.

5.2 Eyes' Detection:

Same as mouth detection, the suitable color spaces that separate eyes colors from the other face region is the cyan channel from CMY space and U channel from YUV color space, Fig. 14 shows an example of "C" and "U" channels for the selected face.

Eyes detection algorithm is briefly shown in Fig. 15, the resulted areas from each color space are combined together by OR operation, where the C channel is used to detect the eyes in small faces and the U channel for the large face sizes.

Almost eyes' features of small faces are hidden and the only difference between the eye colors and other parts of the face is the luminance component where the eyes region is darker than the others. The benefit of using cyan space to detect eyes in small faces that contains luminance and chrominance component that can distinguish eyes colors form others region of the face. The application of light control on the face mask is used

to insure that the light is uniformly distributed over all face parts resulting in more robust eyes detection.
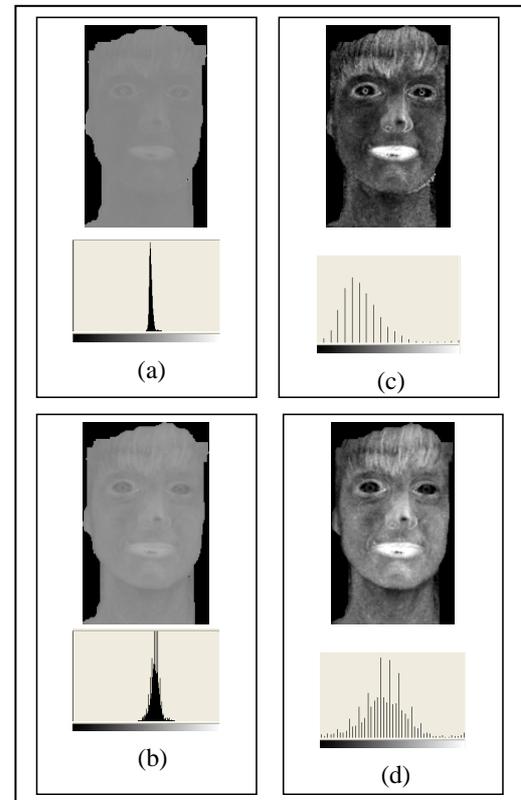

Fig. 12. Histogram stretching (a) a* channel and its histogram (b) V channel and its histogram (c) a* channel after stretching and its histogram (d) V channel after stretching and its histogram.

6. NEURAL NETWORK CONFIRMATION

This stage is designed to distinguish the face shape from other object shapes based on neural network technique. One of the well-known neural network classifiers is Multiplayer-Perceptron (MLP) model using the Back-Propagation (BP) algorithm, which consists of sets of nodes arranged in multiple layers with connections only among nodes in the adjacent layers by weights. The layer, where the inputs information is presented, is known as the input layer. The layer where the processed information is retrieved is called the output layer. All layers between the input and output layers are known as hidden layers.

In our approach, the classification takes place through two neural networks. One trained to detect face shapes and the other trained to detect non-face shapes. Each neural network is a two layer, feed-forward, back-propagation- training network. The input of each neural
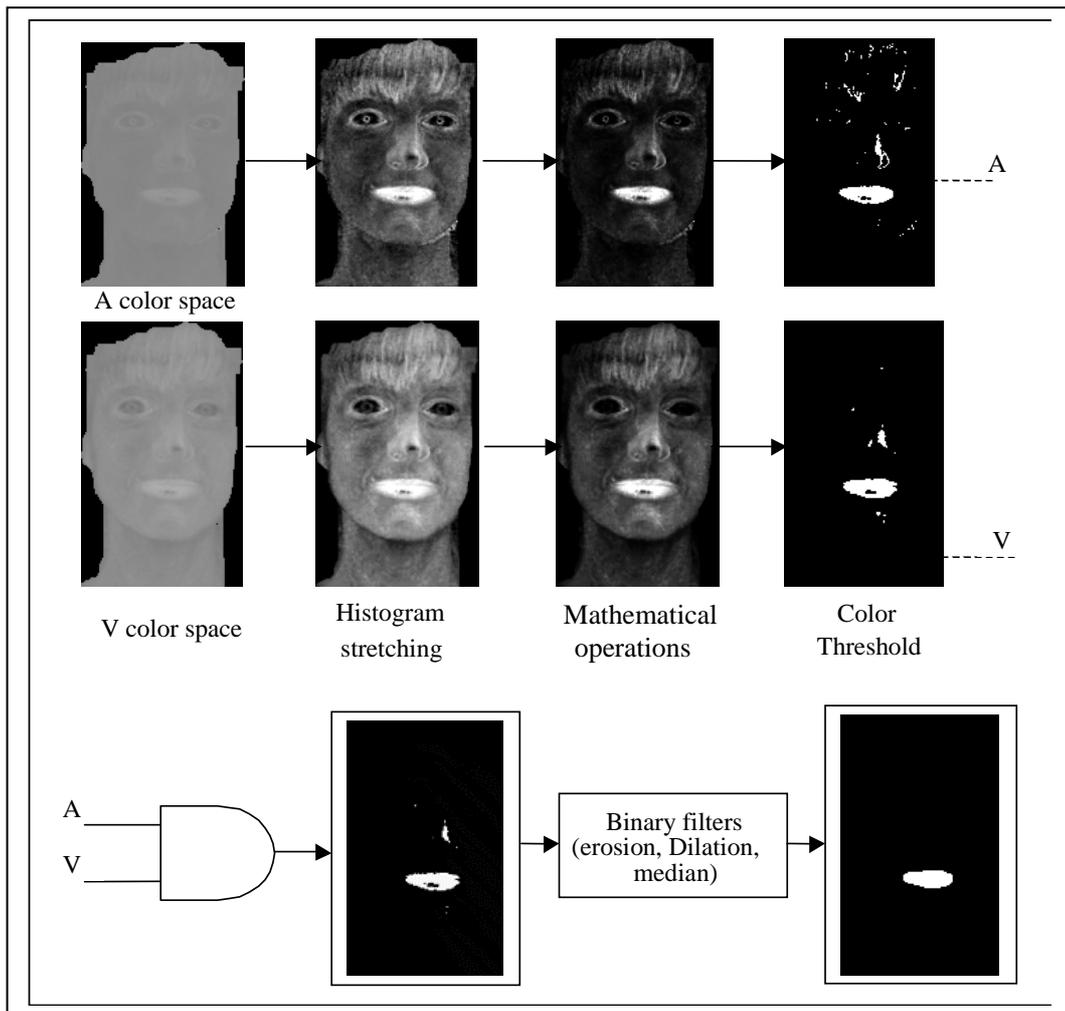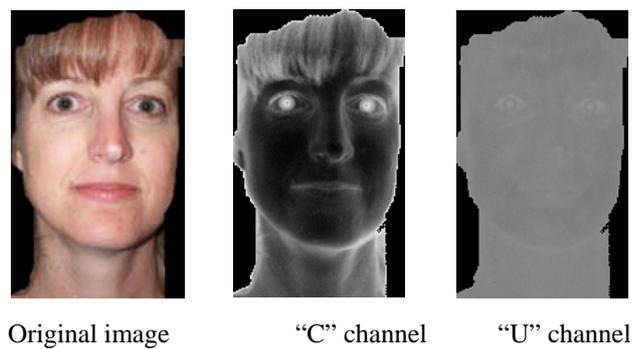
Fig. 13. Expected mouth regions extraction.

Original image   "C" channel   "U" channel

Fig. 14. Color spaces used for eyes detection (a) C channel from CMY color space (b) U channel from YUV color space.
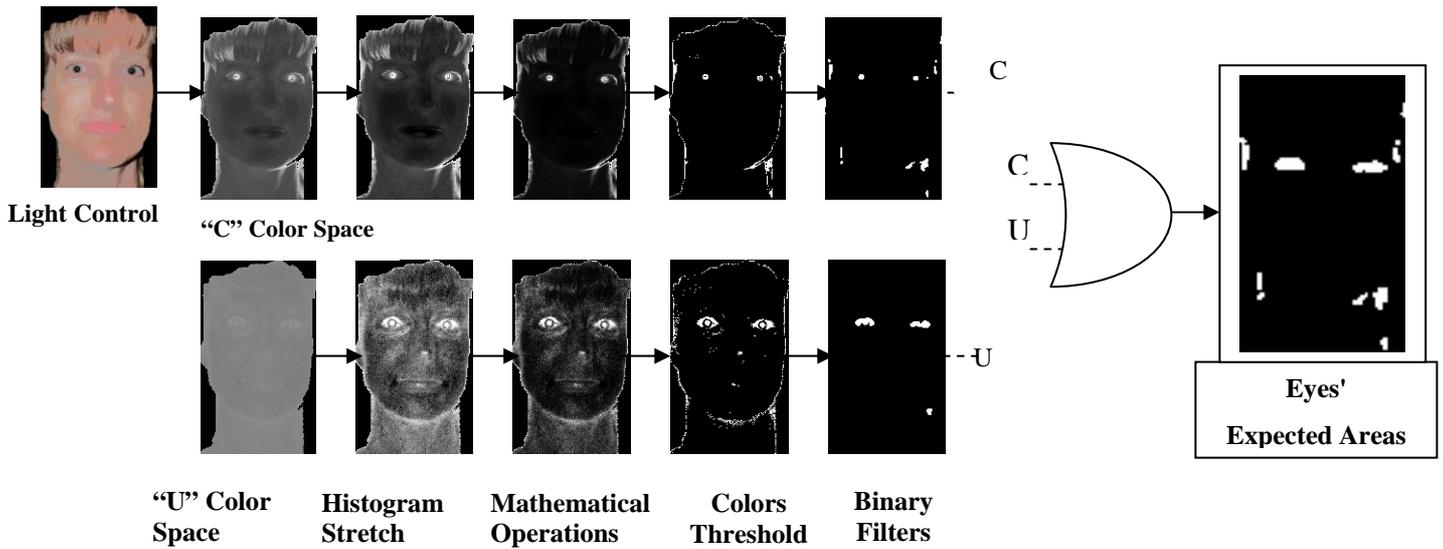
Fig. 15. Expected eyes' regions extraction.

network is a 40x40 gray image generated by adaptive threshold step with 20 hidden neurons and single output. The output of the face shape neural network: is one for face shape and zero for non-face shape and the output of the non-face shape neural network: is one for non-face shape and zero for face shape.

The rectangle regions that contain expected mouth and eyes generated from face feature extraction stage is preprocessed before applying the neural test by the following steps:

1. Rotating the face features to have a frontal face view that depends on the mouth and the eyes positions.
2. Applying face feature geometrical test (The mouth or any part of it must lie inside the selected rectangle as shown in Fig. 16, if this test failed the rectangular region will be considered as a non-face region).
3. Resize them to 40x40 pixel.
4. Enhancing its light using light control technique.

Fig. 17 shows the preprocessing to the input neural network image. If the output of face shape neural network is one and the output of non-face shape neural network is zero, then the object is a face; otherwise, it is not a face. The decision rule is:

$$Object\ shape\ is \begin{cases} Face\ shape,\ if\ the\ output\ of\ face \\ \qquad shape\ NN\ is\ true\ \&\ the\ output\ of \\ \qquad non-Face\ shape\ NN\ is\ false \\ Non-face\ shape,\ otherwise \end{cases}$$

The training was done using 1600 face and 3656 non-face images. The mean value of the sum of squares of the network errors is used as the performance function and training continues until the mean falls beneath 0.01.

7. EXPERIMENTAL RESULT

Personal photo collections usually contain color images that are taken under varying lighting conditions as well as with complex backgrounds. Further, these images may have quality variations and contain multiple faces with variations in color, position, scale, rotation, orientation, pose, and facial expression. We present detection results in Fig. 19 on the Computational Vision Group face dataset (Computational Vision Groupface Dataset), The IMM Face Database (Nordström et al., 2004), and our database that contains a lot of images with multiple faces of different sizes with a wide variety of facial variations. The algorithm can detect both dark skin-tone and bright skin-tone because of the illumination invariant skin. Varying lighting conditions do not affect our algorithm because the enhancement produced by light control.

Fig. 19 shows that our algorithm can detect multiple faces of different sizes with a wide variety of facial variations. Further, the algorithm can detect both dark skin-tone and bright skin-tone where it depends on chrominance of multi color spaces, different rotation due to face geometrical correction founded in the preprocessing stage to the neural, different orientation, and pose, with minimum false detection rate due to neural network test.

In the tables below, we consider the main three stages (skin detection, facial extraction, neural confirmation) that cause false detection alarms. In the neural confirmation stage, we consider the false for both face false detection and non-face false detection.

Table 1: Detection results on the Computational Vision Group dataset, the specification of this dataset is (Computationak vision groupface dataset):

1. 450 face images (one face per image).
2. 896 x 592 pixels, Jpeg format.
3. 27 or so unique people under with different lighting/expressions/backgrounds.

	Skin detection	Face facial extraction	Neural confirmation	All
False Number	5	1	1	10
Detection Rate	98.9%	99.7%	99.7%	98.4%

Table 2: **Detection results on the Markus Weber (Computationak vision groupface dataset) dataset, the specification of this dataset is (Nordström et al., 2004):**

1. 240 still images of 40 different human faces, all without glasses.
2. The gender distribution is 7 females and 33 males.
3. 640x480 JPEG format with a Sony DV video camera
4. Each person has 6 images with different types:
   - Full frontal face, neutral expression.
   - Full frontal face, "happy" expression.
   - Face rotated approx. 30 degrees to the person's right, neutral expression.
   - Face rotated approx. 30 degrees to the person's left, neutral expression.
   - Full frontal face, neutral expression, spot light added at the person's left side.
   - Full frontal face, "joker image" (arbitrary expression).

	Skin detection	Face facial extraction	Neural confirmation	All
Number of False	0	1	1	2
Detection Rate	100%	99.6%	99.6%	99.2%

Table 3: Detection results on our dataset, the specification of this dataset is:

1. 150 images, 489 face (3.26 face per image).
2. 640x480 JPEG format with a Sony DV video camera, and collection of images from Internet.

3. 55 or so unique people under varying lighting conditions as well as with complex backgrounds. Further, these images have quality variations and contain multiple faces with variations in color, position, scale, rotation, orientation, pose and facial expression.

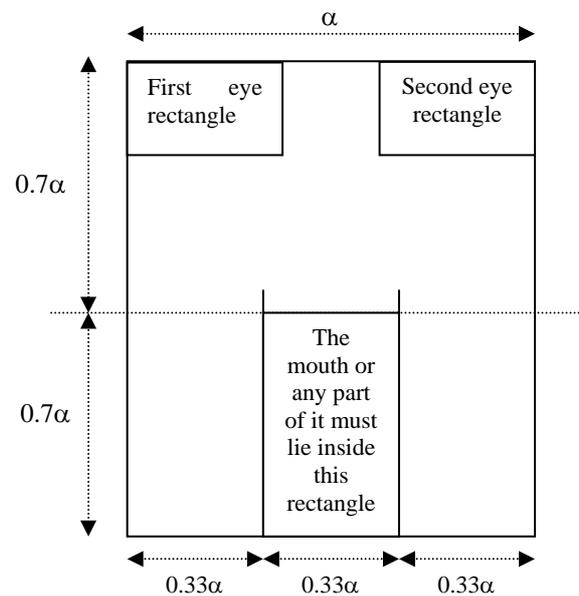|	Skin detection	Face facial extraction	Neural confirmation	All
Number of False (Per face)	15	17	11	43
Detection Rate (Per face)	96.6%	96.5%	97.7%	91.2%

Fig. 16. Geometrical face shape.

8. OTHER TECHNICAL RESULTS

There are a lot of techniques and algorithms proposed for face detection in colored images, the results of these techniques restricted on a small dataset of images with specific situations (position, scale, rotation, …etc).

The algorithm proposed by Rein-Lien Hsu (2001) is the only founded technique that evaluates his work on a wide range of images with a different situation (frontal, near frontal, profile, and half profile), this technique was evaluated on a dataset collected from World -Wide- Web, and personal photo collections. The detection rate of this algorithm on the collection of 382 family and news photos (1:79 faces per image) is 80.35% with 22.97second average time and 17.35 second standard deviation using 1.7GHz CPU.

9. CONCLUSIONS AND FUTURE WORK

We present a new algorithm for face detection designed to robust the variations that can occur in face
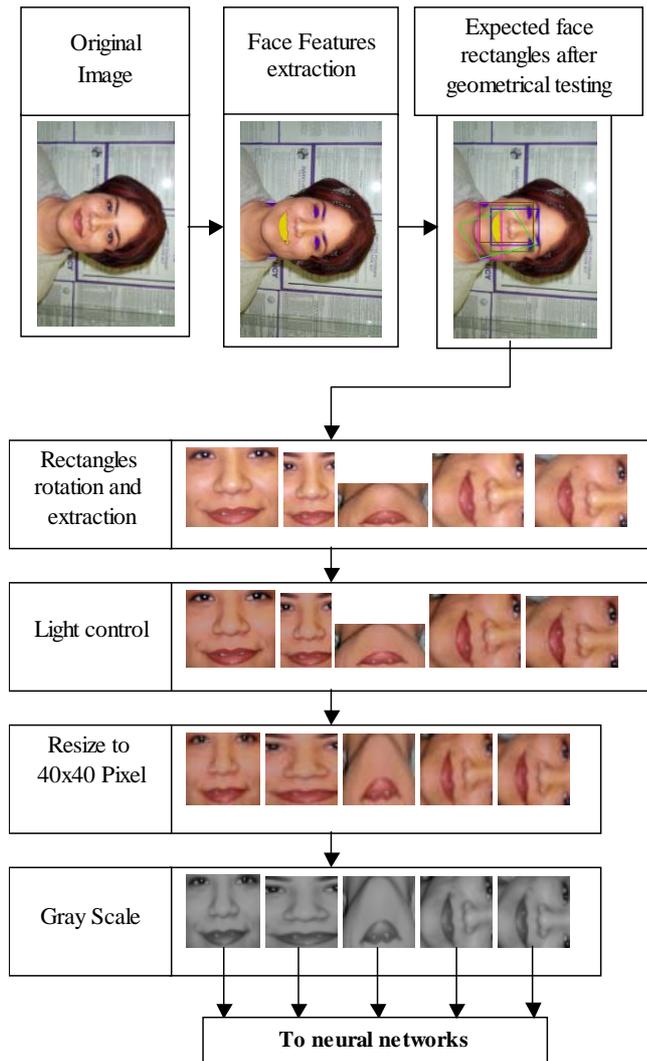
Fig. 17. **Neural network preprocessing diagram.**

Fig. 18. System flow chart.

illumination, shape, color, pose, rotations and orientation.

Our approach uniformly distributes the brightness based on light control technique to get uniform light distribution for all face parts. The skin tone pixels are detected based on chrominance of multi color spaces. The redistribution of brightness over the skin pixels only will insure uniform brightness distribution over all face parts.

The color quantization stage divides the image into four levels according to its colors based on octree algorithm (Clark, 1996: 54-57 and 102-104). Because the uniform distribution of face colors produced by light control, the face is segmented into one segment.

For each separated segment which resulted from the 4-levels quantization method, the face feature extraction is applied to extract possible mouth and eyes areas and pass their rectangles that may contain a face to neural network to confirm the face validation.

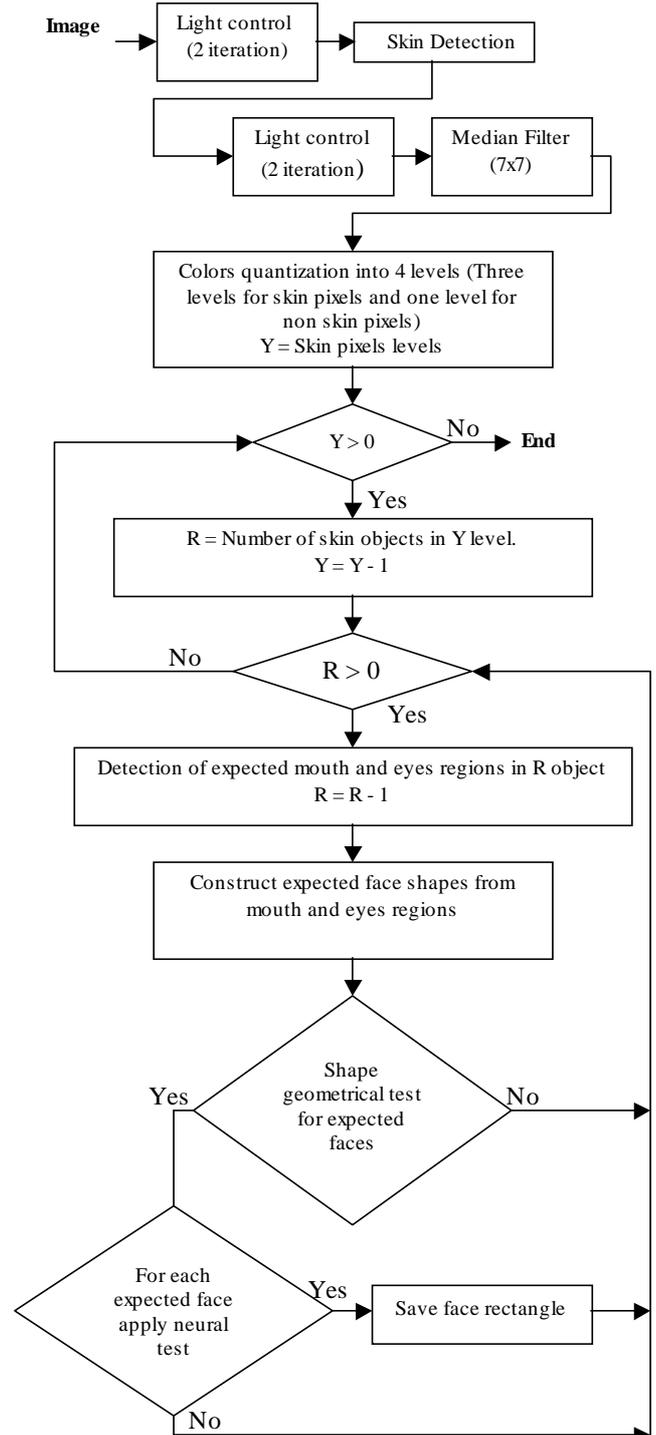Our goal is to design a system that detects faces and

facial features from color images, that allow users to edit detected faces, and uses the facial features as indices for retrieval from image and video databases.

Our algorithm proved a highly success detection rate over a large number of images with different sizes, wide variety of facial variations, illumination, shape, color, pose, rotations and orientation.

For future work, our algorithm can be extended to be evaluated on a complete face recognition system.
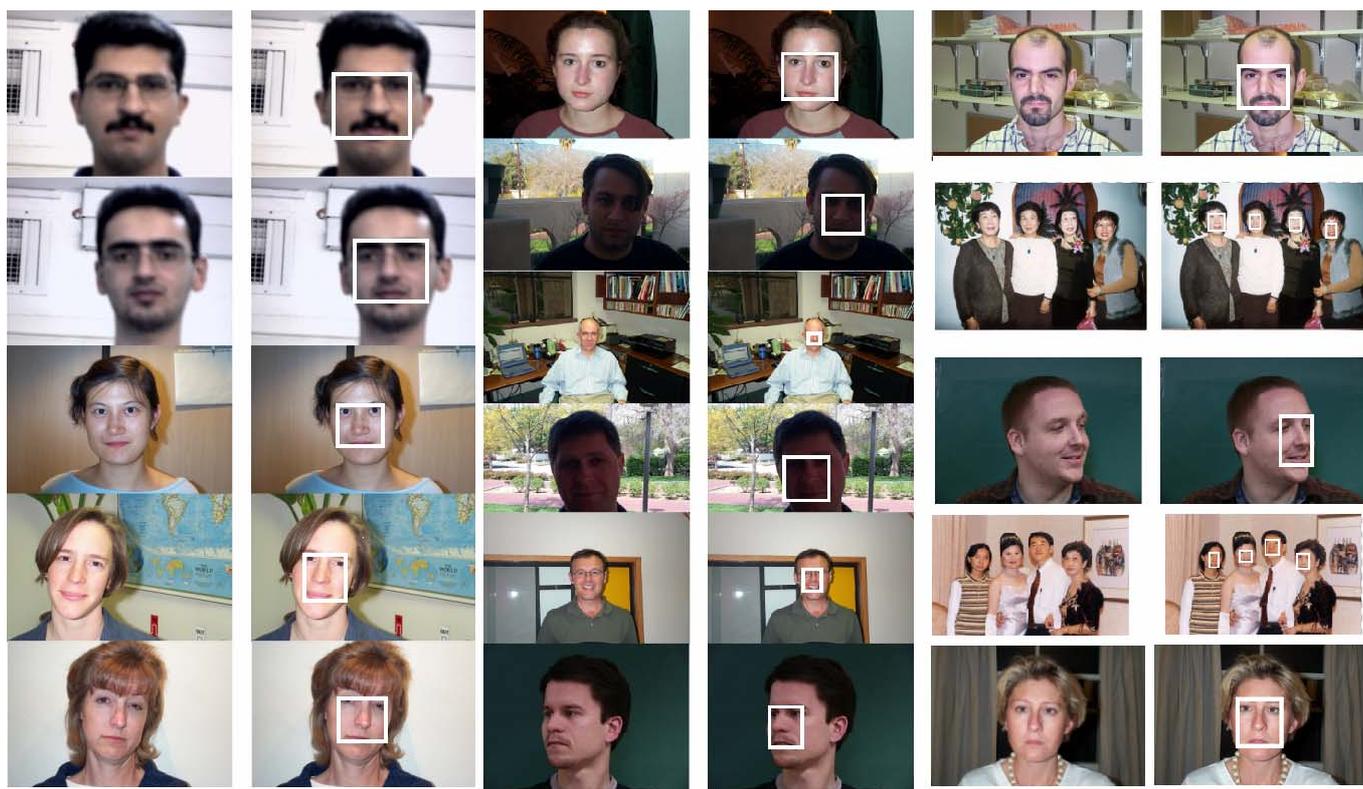
Fig. 19. Face detection result on Computational Vision Group, IMM, and our face databases (a) original image (b) detected face region represented as a rectangle.

REFERENCES

Brainard, D.H. 1989. Calibration of a Computer Controlled Color Monitor, *Color Research and Application*. 14 (1): 23-34.

Clark, D. 1996. Color Quantization Using Octrees, *Dr. Dobb's Journal*: 54-57 and 102-104.

Computational Vision Group Face Dataset. <http://www.vision.caltech.edu/Image_Datasets/faces>

Horprasert, T., Yacoob, Y. and Davis, L. S. 1996. Computing 3-D Head Orientation from a Monocular Image, Proc. Int'l Conf. *Automatic Face and Gesture Recognition:* 242-247.

Hsu, Rein-Lien, Mohamed Abdel and Anil K. Jain. 2001. Face Detection in Color Images", Michigan State University, *Technical Report*, MSU-CSE-01-7.

Nordström, Michael M., Larsen, M., Mads, Sierakowski, Janusz and Stegmann, Mikkel, B. 2004. The IMM Face Database. Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

Shadeed, W., G., Abu-Al-Nadi, D. I. and Mismar, M. J.

2003. Traffic Road Sign Detection in Colored Images, *10th IEEE International Conference on Electronics*, Circuits, and Systems (ICECS), Sharjah. UAE 14-17, 890-893.

Shin, M. C., Chang, K. I. and Tsap, L. V. 2002. Does Colorspace Transformation Make any Difference on Skin Detection? In *IEEE Workshop on Applications of Computer Vision.*

Smeraldi, F., Carmona, O. and Bigun, J. 2000. Saccadic Search with Gabor Features Applied to Eye Detection and Real-time Head Tracking, *Image and Vision Computing*, 18 (4): 323-329.

Starovoitov, V., Samal, D., Briliuk, D. and Kopendakov, A. 2003. Image Enhancement for Face Recognition. *International Conference on Iconics*, 2003, St.Petersburg, Russia.

Vezhnevets, V., Sazonov, V. and Andreeva, A. 2003. A Survey on Pixel-Based Skin Color Detection Techniques, *Proc. Graphicon*: 85-92, Moscow, Russia.

Vezhnevets, V., Sazonov, V. and Andreeva, A. 2003. A Survey on Pixel-Based Skin Color Detection

Techniques. *Graphicon-2003*, Moscow, Russia.

Yang, Ming-Hsuan, Kriegman, David and Ahuja, Narendra. 2002. Detecting Faces in Images: A Survey, *IEEE Transactions on Pattern Analysis and Machine*

Intelligence, 24 (1): 34-58.

Yang, M. and Kriegman, D. 2002. Detecting Faces in Images: A Survey, *IEEE Trans. PAMI*, 24 (1): 34-58.

*

.

.

.

.

_____