Iraqi Journal of Statistical Science (20) 2011
The Fourth Scientific Conference of the College of Computer Science & Mathematics
pp [336-350]

Face Detection Using Neural Networks

Ramadan M. Ramo* Dr.Fawziya M. Ramo**

المستخلص

اكتشف في هذا البحث وجود الوجه في الصور الرقمية باستخدام شبكة Elman بعد مرور الصورة بعدد من المراحل اذ حولت الصورة في البداية إلى المستوى الرمادي ومن ثم قطعت بالاعتماد على خصائص الجلد باعتماد خوارزمية تقطيع الجلد عن طريق استخلاص خواص جلد الإنسان باستخدام خوارزمية الجلد الثنائية ذات التلوين المتعدد ومن ثم إدخال البيانات على الشبكة العصبية. طبقت الخوارزمية على عدد من النماذج المختلفة وأثبتت كفاءتها ودقة في الاكتشاف تصل إلى 90% بالاعتماد على النتائج الظاهرة.

Abstract

This research detects the existence of the face in the digital images by using an Elman network, after a number of stages where at first it converts the image to the gray level and then segmentation of image which is based on the characteristics of the skin algorithm for segmenting the skin by extracting the properties of human skin using amulti-colored skin bilateral algorithm and then entering the data to a neural network. The algorithm was applied to a number of different models and proved to be efficient and accurate in detection reaches 90% depending on the results .

Keyword : face Recognition, skin detection, Neural network

1. Introduction

Pattern recognition is a modern day machine intelligence problem with numerous applications in a wide field, including Face recognition, Character recognition, Speech recognition as well as other types of object recognition. The field of pattern recognition is still very much in its infancy, although in recent years some of the barriers that hampered such

* Assistant Lecture\College of Administration & Economics\ University of Mosul

** Lecturer\ College of computers & Mathematic Scientific\ University of Mosul

Received:1/10 /2011 Accepted: 21 /12 / 2011

automated pattern recognition systems have been lifted due to advances in computer hardware providing machines capable of faster and more complex computation[13].

Face detection involves separating image windows into two classes; one containing faces (targets), and one containing the background (clutter). It is difficult because although commonalities exist between faces, they can vary considerably in terms of age, skin colour and facial expression. The problem is further complicated by differing lighting conditions, image qualities and geometries, as well as the possibility of partial occlusion and disguise. An ideal face detector would therefore be able to detect the presence of any face under any set of lighting conditions, upon any background. For basic pattern recognition systems, some of these effects can be avoided by assuming and ensuring a uniform background and fixed uniform lighting conditions. This assumption is acceptable for some applications such as the automated separation of nuts from screws on a production line, where lighting conditions can be controlled, and the image background will be uniform. For many applications however, this is unsuitable, and systems must be designed to accurately classify images subject to a variety unpredictable conditions.A variety of different face detection techniques exist, but all can be represented by the same basic model, depicted in figure 1[9].

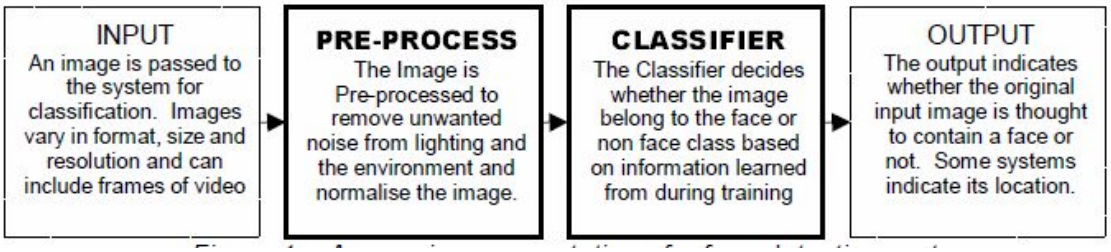


Figure 1 : Representation of a face detection system

Each technique takes a slightly different approach to the face detection problem, and although most produce encouraging results, they are not without their limitations[9].

Artificial Neural Networks (ANNs) are computing models for information processing and pattern identification. They grow out of research interest in modeling biological neural systems, especially human brains. An ANN is a network of many simple computing units called neurons or cells, which are highly interconnected and organized in layers. Each neuron performs the simple task of information processing by converting received inputs into processed outputs. Through the linking arcs among these neurons, knowledge can be generated and stored regarding the strength of the relationship between different nodes. Although the ANN models used in

all applications are much simpler than actual neural systems, they are able to perform a variety of tasks and achieve remarkable results. Over the last several decades, many types of ANN models have been developed, each aimed at solving different problems. But by far the most widely and successfully used for forecasting has been the feedforward type neural network. Figure 1 shows the architecture of a three-layer feedforward neural network that consists of neurons (circles) organized in three layers: input layer, hidden layer, and output layer. The neurons in the input nodes correspond to[4],figure (2) represents Typical Feed forward Neural Networks

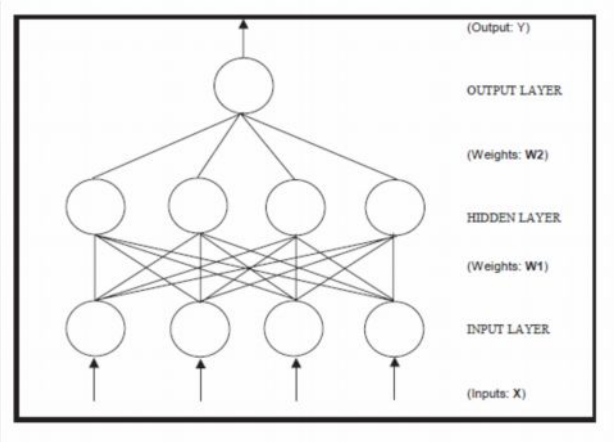


Figure 2 : A Typical Feed forward Neural Networks

A neural network represents a highly parallelized dynamic system with a directed graph topology that can receive the output information by means of a reaction of its state on the input actions. Processor elements and directed channels are called nodes of the neural network
. [7][1]

2. Biometric Pattern Recognition

Human face and human signature represent some of the most common biometric patterns that our visual system encounters daily. We present here the classification techniques of these two biometric features. A lot of interest has been generated in automated face recognition and a number of implementation approaches have been proposed [12][2].

The major strategies used in face identification are either based on features or they are based on face space, such as Eigenface or Fisherface. Most of the feature based methods extract features from front view of the face and sometimes also from side face profiles. An automatic face recognition system employing both front and side views of the face is more accurate, since it takes advantage of the explicit information inherently available in both the views of the human face. Face recognition approaches

employ diverse techniques like neural nets, elastic template matching, Karhunen-Loeve expansion, algebraic moments, iso-density lines, etc.[6][5] Each of these methods has its advantages and limitations. The feature extraction and matching techniques for face recognition are presented in the next section.[8].

3. Neural Network

The Elman Neural Network (ENN) is one type of the partial recurrent neural networks, which consists of a two-layer back propagation network with an additional feedback connection from the output of the hidden layer to its input. The advantage of this feedback path is that it allows the ENN to recognize and generate temporal patterns and spatial patterns. This means that after training, interrelations between the current input and internal states are processed to produce the output and to represent the relevant past information in the internal states. As a result, the ENN has been widely used in various fields which include classification, prediction and dynamic system identification, etc.. However, since the ENN usually uses the Back-Propagation (BP) based algorithms to deal with the various signals, it has been proved that it frequently suffers from a suboptimal solution problem. At the same time, the efficiency of the ENN is limited to low order system due to the insufficient memory capacity when Back-Propagation algorithm is employed. So, several approaches have been suggested in the literatures to enhance the performance of the BP-trained ENN with simple modifications of net structure, but not algorithms. These improved modifications attempt to add other feedback connections to the model that will increase the capacity of the memory in order to speed up the convergence and escape from the local minima [10][16].

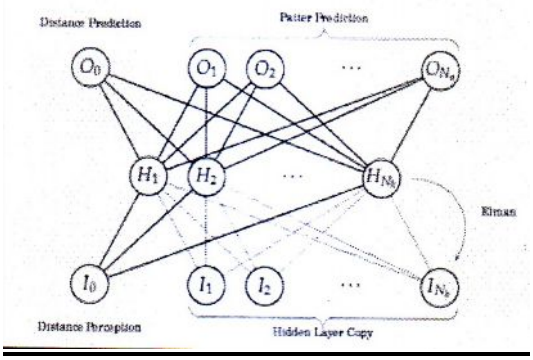


Figure 3 : The architecture of the Elman

3.1 Algorithm

Elman networks consist of NI layers using the adopted weight function. The first layer has weights coming from the input. Each subsequent layer has a weight coming from the previous layer. All layers except the last one have a recurrent weight. All layers have biases. The last layer is the network output. A adaption is done with trains, which updates weights with the specified learning function. Training is done with the specified training function. Performance is measured according to the specified performance function [11].

3.2 Network Design

Followings are some parameters set for network training:

- Number of layers : 3 Layers.
- Number of node in input layer :128x8 nodes (1024 nodes)
- Number of nodes in hidden layers : 5 nodes
- Number of nodes in output layers : 2 nodes
- Initial learing rate : 0.3
- Learing rate increment : 1.2
- Epoch : 1000
- Error goal : 0.0000006

The face images needed for this study were captured in the form of 350*280 pixel size JPEG format. Later, after converting to the Gray level and segmentation. That perform 1024 pixels which implement the number of nodes in input layer and as mentioned earlier, the number of neurons in the output layer is two ,first node represents detection and another represents nondetection [11].

4. Image Segmentation

Segmentation of an image entails the division or separation of the image into regions of similar attribute. The most basic attribute for segmentation is image luminance amplitude for a monochrome image and color components for a color image. Image edges and texture are also useful attributes for segmentation.

The definition of segmentation adopted is deliberately restrictive; no contextual information is utilized in the segmentation. Furthermore, segmentation does not involve classifying each segment. The segmenter only subdivides an image; it does not attempt to recognize the individual segments or their relationships to one another.

There is no theory of image segmentation. As a consequence, no single standard method of image segmentation has emerged. Rather, there are a collection of ad hoc methods that have received some degree of

popularity. Because the methods are ad hoc, it would be useful to have some means of assessing their performance. Haralick and Shapiro (1) have established the following qualitative guideline for a good image segmentation: “Regions of an image segmentation should be uniform and homogeneous with respect to some characteristic such as gray tone or texture. Region interiors should be simple and without many small holes. Adjacent regions of a segmentation should have significantly different values with respect to the characteristic on which they are uniform. Boundaries of each segment should be simple, not ragged, and must be spatially accurate.” Unfortunately, no quantitative image segmentation performance metric has been developed [14] [15].

4.1 Skin Segmentation

The segmentation of skin regions in color images is a preliminary step in several applications, such as video classification and retrieval in multimedia databases, semantic filtering of web contents (through the definition of medium-level features), human motion detection, human computer interaction, and video-surveillance. It is also useful in image processing algorithms, as well as in intelligent scanners, digital cameras, photocopiers, and printers. Many different methods for discriminating between skin and non skin pixel are available in the literature. These can be grouped into three types of skin modeling: parametric, nonparametric, and explicit skin cluster definition methods. The Gaussian parametric models' assume that skin color distribution can be modeled by an elliptical Gaussian joint probability density function. Non parametric methods estimate skin color distribution from the histogram of the training data without deriving an explicit model of skin color'. The simplest, and often applied, methods build what is called an"explicit skin cluster" classifier with expressly defines the boundaries of the skin cluster in certain color spaces. The underlying hypothesis of methods based on explicit skin clustering is that skin pixels exhibit similar color coordinates in an appropriately chosen color space. These binary methods are very popular as they are easy to implement and do not require a training phase. The main difficulty in achieving high skin recognition rates, with the smallest possible number of false positive pixels, is that of defining accurate cluster boundaries through simple, often heuristically chosen, decision rules. In this study we compare the performance of various explicit skin cluster methods applying the thresholds presented in the literature with that achieved when a genetic algorithm is applied to determine the boundaries of the skin clusters in multiple color spaces. To quantify the performance of these skin detection methods, we use recall and precision

scores. Classification results are assigned as True Positive (TP), False Positive (FP) and False Negative (FN). Recall is defined as the ratio between the number of skin pixel correctly classified and the total number of actual skin pixels (TP/(TP+FN)), while precision is defined as the ratio between the number of skin pixels correctly classified and the total number of pixels labeled as skin pixels by the skin detection method considered (TP/(TP+FP)).[3]

4.1.1 Binary skin classifiers

The separate skin and non skin colors using a piecewise linear decision boundary. propose a set of fixed skin thresholds in a given color space. Some color spaces permit searching skin color pixels in the 2D chromatic space, reducing dependence on lighting variation, others, such as the RGB space, address the lighting problem by introducing different rules depending on illumination conditions (uniform daylight, or flash). Working within different color spaces, we have implemented the six different algorithms analyzed in this paper. They are named for the color space adopted: YCbCr. RGB, HSV ; HSV2, HSI" and rgb. The details of their implementation can be found in the referenced paper and are summarized in the subsections here below. Examples of the skin maps obtained applying these methods to the image.[2]

4.1.2 YCbCr

A skin color map is derived and used on the chrominance components of the input image to detect pixels that appear to be skin. The algorithm then employs a set of regularization processes to reinforce those regions of skin-color pixels that are more likely to belong to the facial regions. We use only their color segmentation step here. Working in the YCbCr space find that the ranges of Cb and Cr most representative for the skin-color reference map were:[3]

$$77 \le cb \le 127 and 133 \le cr \le 173 \qquad (1)$$

4.1.3 RGB

RGB colour space deal with the illumination conditions under which the image is captured. Therefore, they classify skin color by heuristic rules that take into account two different conditions: uniform daylight and flash or lateral illumination. Uniform daylight illumination

$R > 95, G>40, B>20$

$Max(R,G,B)-Min\{R,G,B\}<15 \qquad .....(2)$

$$|R-G|>15, R>G, R>B$$

Flashlight or daylight lateral illumination:

$$R>220,\ G>210,\ B>170 \quad |R-G|\le 15, B<R, B<G \qquad (3)$$

4.1.4 HSV1

HSV color space and select pixels having skin-like colors by setting the following thresholds:

V>40

0.2<S<0.6

0^0<H<25^0 OR 335^0<H<360^0

The selected range of Hrestricts segmentation ton reddish colors and the saturation range selected ensures the exclusion of pure red and very dark red colors, both of which are caused by small variations in lighting conditions. The threshold on V is introduced to discard dark colors.[3]

4.1.5 HSV2

Starting from a training data set composed of skin color samples, the color histogram in Hue-Saturation-Value (HSV) color space, and estimate the shape of this skin color subspace.V≥40

H≤ (-0.4v +75)

10≤S≤(-H-0,1V+110)

if H ≥ 0 S≤ (0.08(100-V)H+0.5V) ……(4)

if H< 0 S≤ (0.5 H+ 35).

4.1.6 HSI

HSI colour space system to design their colour classification algorithm because it is stable for skin colour under different lighting conditions. These rules apply to the intensity 1, hue }-I and saturation S, and are detailed as follows[3]:

r>40

if 13<S<I10, 0'<H<28° and 332"<N<360° …..(5)

if 13<S<75, 309°<N<331°

The thresholds are empirically determined from the training set and the color system transformation from RGB to IISI is defined as follows :

$$I = \frac{1}{3}(R+G+B); \quad I_2 = \frac{1}{2}(R-B); \quad I_3 = \frac{1}{4}(2G-R-B).$$

$$I = I_1$$

$$S = \sqrt{I_2^2 + I_3^2} \tag{6}$$

$$H = Tan^{-}\left(\frac{I_3}{I_2}\right)$$

4.1.7 rgh

Starting with the three rgb components in a normalized form and a simple set of arithmetic operators, the authors produce a model for skin detection. The algorithm uses a Restricted Covering Algorithm (RCA) as its selective learner. The RCA searches for single rule in parallel. Among the different combination rule presented by the authors we have chosen the one with the highest precision and success rate:

$$\frac{r}{g} > 1.185, \frac{r.b}{(r+g+b)^2} > 0.107 and \frac{r.g}{(r+g+b)^2} > 0.112 \tag{7}$$

where rgb are the normalized coordinates obtained as:

$$r = \frac{R}{R+G+B} \qquad g = \frac{G}{R+G+B} \qquad b = \frac{B}{R+G+B} \tag{8}$$

5. Experimental Results

The suggested algorithm was applied at a number of images which have shown the accuracy of results in face detection algorithm based on neural networks, where we note that results vary much depending on the selected image when there is convergence to the color of the face in any part of the picture is taken into account. Where we note from the application of the algorithm, the first images to determine there was a part of the neck as a result of Taqarib color of the neck of the color of the face. While we find that in the image number (2) was discovered face accurately, despite our reliance on taking the image upside down in order to ensure the accuracy of the algorithm to locate the face. Also images (3,4,5,6), which has identified detect face accurately. As for the picture (7,8) have taken part of the hand, due to the convergence of the color of the hand tip of the color of the face. And shapes following algorithm describe the work of a different set of images with the extension Jbj.

Image (1)

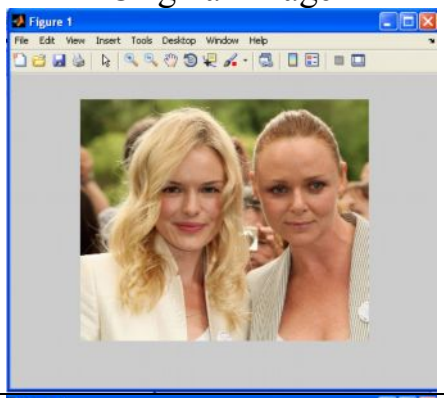
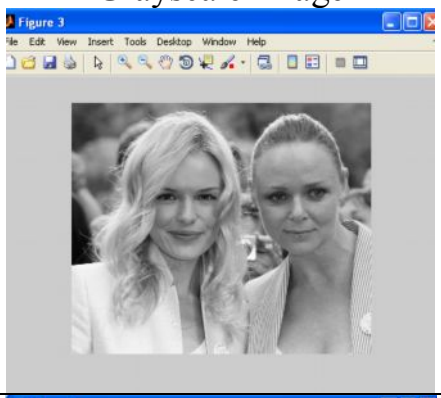
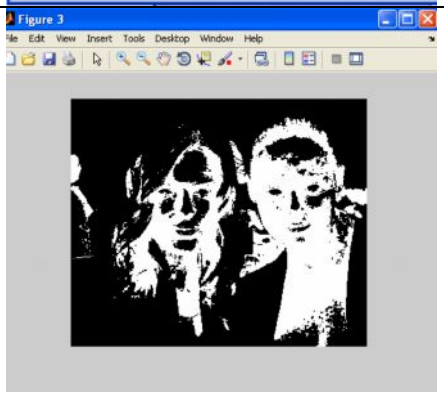
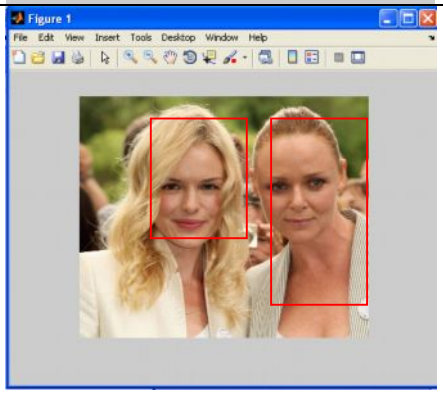
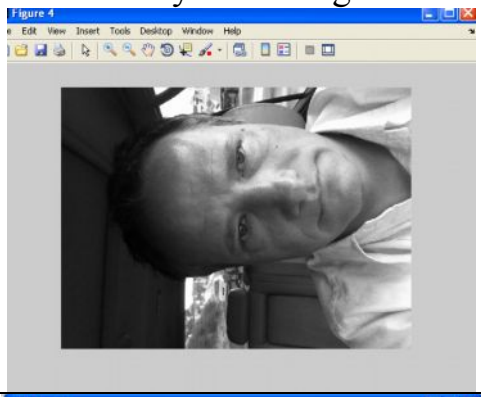
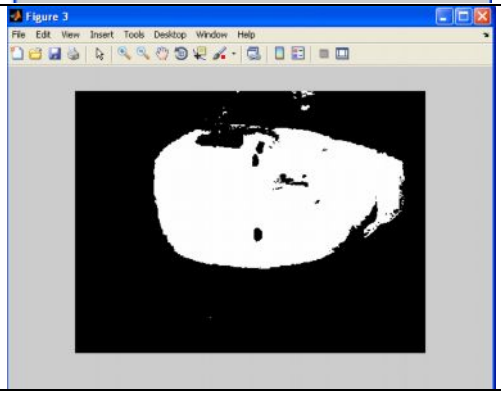
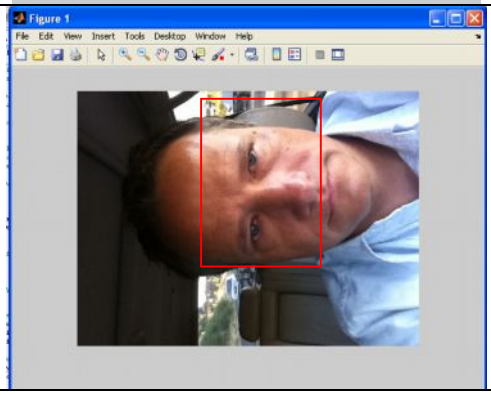
Original Image	Grayscale Image
	
	
Skin Detection	Suggested method

Image (2)

Original Image	Grayscale Image
	
	

Skin Detection	Suggested method

Image (3)

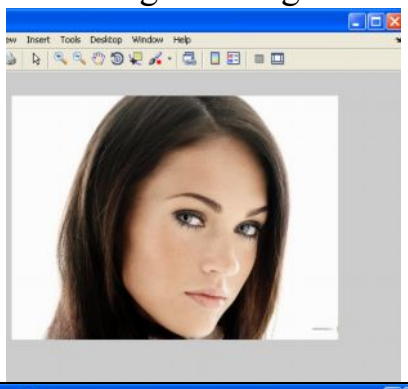
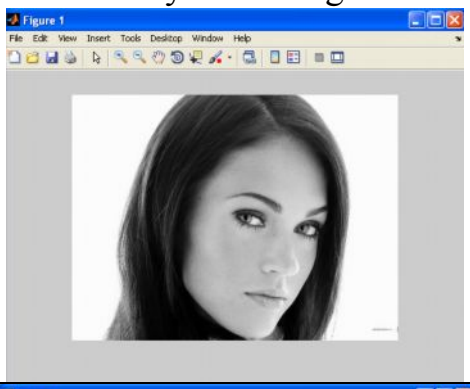
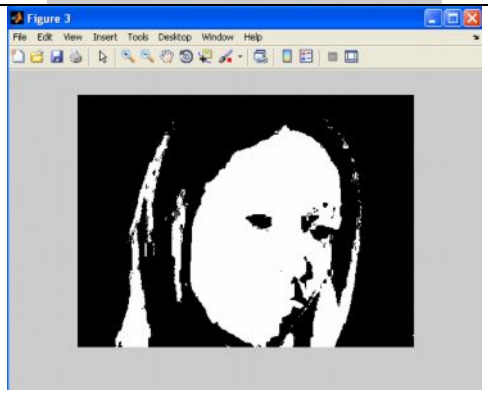
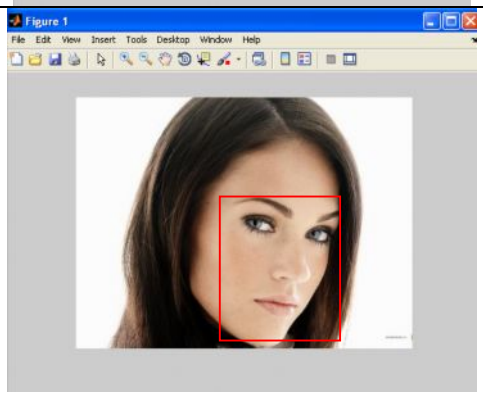
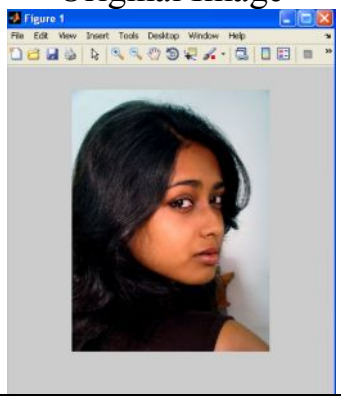
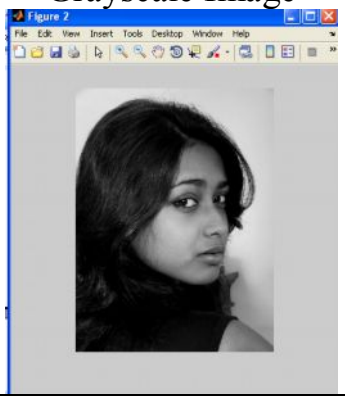
Original Image	Grayscale Image
	
	
Skin Detection	Suggested method

Image (4)

Original Image	Grayscale Image
	

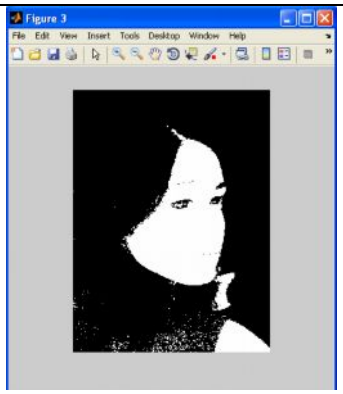
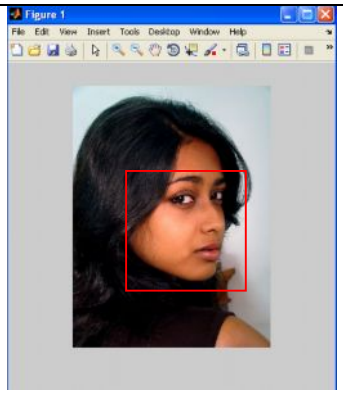
	
Skin Detection	Suggested method

Image (5)

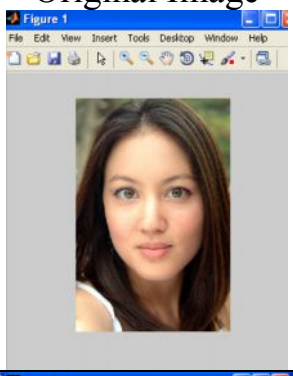
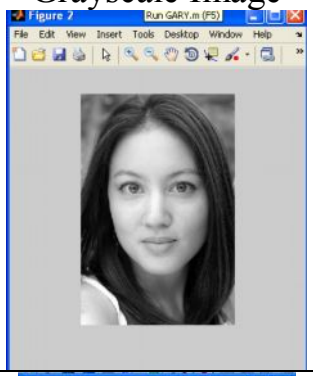
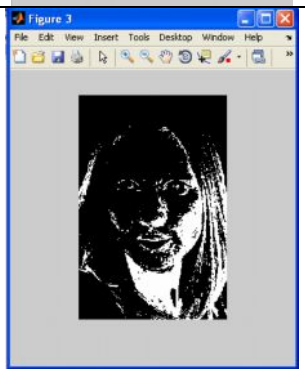
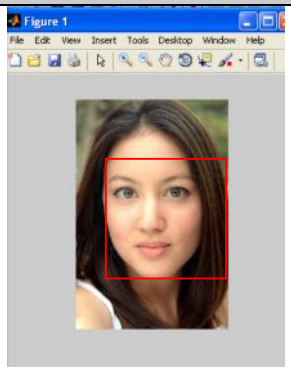
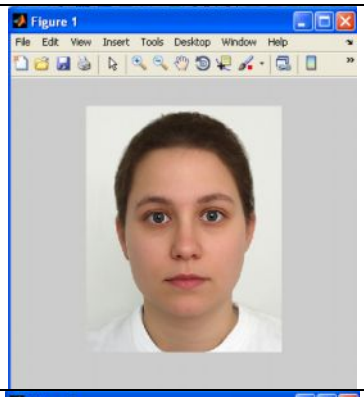
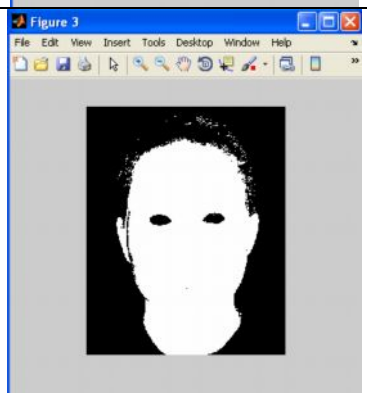
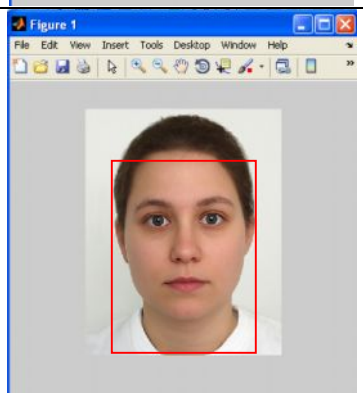
Original Image	Grayscale Image
	
	
Skin Detection	Suggested method

Image (6)

Original Image	Grayscale Image

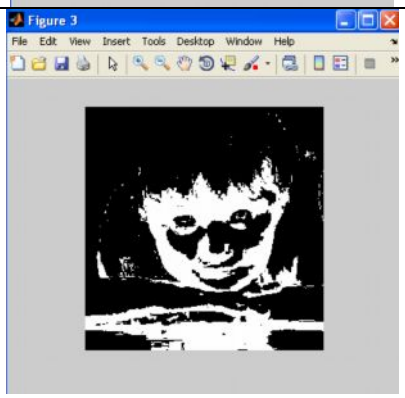
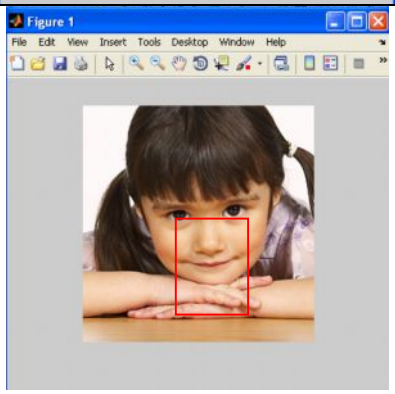


Skin Detection | Suggested method

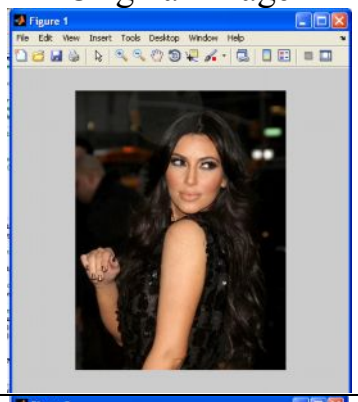
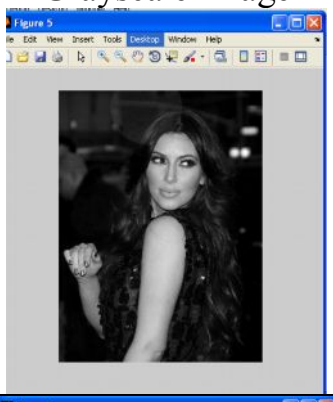
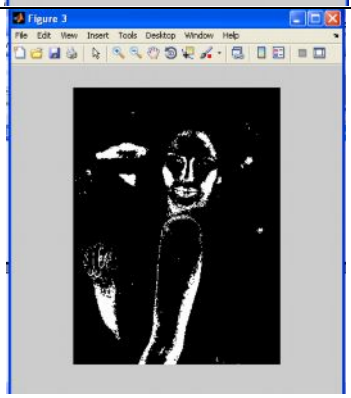
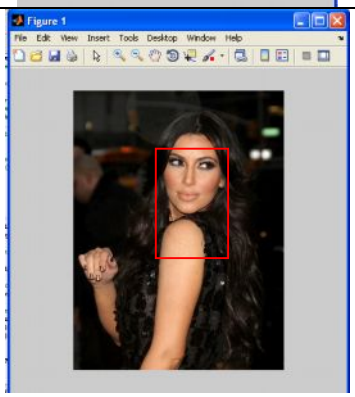
Image (7)

Original Image | Grayscale Image

Skin Detection | Suggested method

Image (8)

Original Image	Grayscale Image
	
	
Skin Detection	Suggested method

Reference

1. Alexander I. Galushkin," Neural Networks Theory", Moscow Institute of Physics & Technology Department Neurocomputers Institutskiy, Springer-Verlag Berlin Heidelberg, 2007.
2. D. S. Taubman, "high-performance scalable image compression with ebcot", IEEE transactions on image processing, july 2000, 1158-1170.
3. Francesca gasparini, rairnondo schettini,"skin segmentation using multiple thresholding", university degli studi di milano-bicocca, via bicocca degli arcimboldi, milano, Italy,2002.
4. G. Peter Zhang," Neural Networks in Business Forecasting", Georgia State University, USA, Published in the United States of America by Idea Group Publishing (an imprint of Idea Group Inc.), 2004.
5. Iso/iec 14492-1, “lossy/lossless coding of bi-level images,” 2000.
6. Iso/iec 15444-1, “information technology-jpeg2000 image coding System, part 1: core coding system,” 2000.
7. Kevin gurney,"an introduction to neural networks", university of sheffield, ucl press limited is an imprint of the taylor & francis group this edition published in the taylor & francis e-library, 2005.
8. Mi. B. Pennebaker, j. L. Mitchell, g. G. Langdon, jr., and r. B. Arps,

"an overview of the basic principles of the q-coder adaptive binary Arithmetic coder", Ibm journal of research and development, November 1988, 717-726.
9. Phil brimblecombe,” face detection using neural networks", university of surrey, electronic engineering school of electronics and physical sciences,2002.
10.Stefano zappacosta, Gianluca baldassarre, and Stefano nolfim, "Elman neural networks and time integration for object recognition", 2004.
11.Shaymaa m. Al-mashahdany , "human iris recognition using elman neural network", department of computer sciences ,college of computer sciences & math., university of mosul,2009.
12.T. Acharya and p. Tsai, "jpeg2000 standard image compressron: concepts, algorithms and vlsi architectures", wiley, hoboken, 2004.
13. Tinku charya& ajoy k. Ray," image processing principles and applications", avisere, inc.tucson, arizona and department of electrical engineering arizona state university tempe, arizona, avisere, inc.tucson, arizona and electronics and electrical communication engineering department indian institute of technology kharagpur, india,2005.
14.Torsten seemann, " digital image processing using local segmentation", school of computer science and software engineering faculty of information technology, monash university, australia. Submission for the degree of doctor of philosophy, 2002.

15.William k. Pratt," digital image processing : piks scientific inside, fourth edition", copyright © 2007 by john wiley & sons, inc.
16.Zhiqiang zhang, shangce gao, gang yang, fangjia li and zheng tang," an algorithm of supervised learning for elman neural network", faculty of engineering toyama university gofuku 3190, toyama shi, 930-8555, japan, dalaosha@hotmail.com, received june 2008.