C# IMPLEMENTATION OF A FACE DETECTION SYSTEM USING TEMPLATE MATCHING AND SKIN COLOR INFORMATION

Adesola Olaoluwa Anidu [1], Fasina Ebunoluwa Philip [2]

[1] Department of Mathematical Sciences, Redeemer's University, Mowe, Ogun, Nigeria
[2] Department of Computer Science, University of Lagos, Lagos, Nigeria

ABSTRACT: Face detection is the first step for any automatic face recognition system, Human Computer Interaction systems, surveillance systems and also a step towards Automatic Target Recognition (ATR) or generic object detection/recognition.
This paper presents a face detection system using template matching and skin color information as methods for detecting a face / faces in an image. The system is developed using C#.net programming language. The stages of development are categorized into three namely: the pre-processing, normalization and the face detection stages. The system was tested using pictures taken from a digital camera and stored on the computer system and an accuracy of 80% on the average was achieved for the tested images.
KEYWORDS: Template, normalization, template matching, skin color

INTRODUCTION

As computers become faster and faster, new applications dealing with human faces become possible. Examples of these applications are face detection and recognition applied to surveillance systems, gesture analysis applied to user-friendly interfaces, or gender recognition applied to reactive marketing. Human face detection by computer systems has become a major field of interest. In fact, detection of faces in a digital image has gained much importance in the last decade, with application in many fields such as law enforcement and security. Although facial detection is an extremely simple task to the human eye, automating the process to a computer requires the use of various image processing techniques.
Face detection and localization is usually the first step in many biometric applications like, face recognition, video surveillance, human computer interface etc. Face detection in general terms is defined as to isolate human faces from their background and exactly locate their position in an image. A face is naturally recognizable to a human being despite its many points of variation (e.g. skin tone, hairstyle, facial hair, glasses etc.).
Automating the process to a computer requires the use of various image processing techniques.
According to [MBV05] face detection schemes can be classified into four different categories, although some methods can belong to more than one category.
(1) Knowledge-based methods, where some rules or relationships between features are encoded.
(2) Feature-invariant approaches, where the idea is to detect the facial features first, such as eyes, mouth, eye brows, and group them into candidate faces.
(3) Template-matching methods, where there is a predefined face pattern that is correlated with the image. Point distribution models (PDMs) have also been used for this purpose.
(4) Appearance-based methods, where the goal is to train a classifier that learns the features of the faces from a training set with face and non-face images. Many classic techniques such as principal component analysis, Gaussian mixture models, neural networks, hidden Markov model, support vector machines, and probabilistic models have been applied.
Template matching is a technique used in classifying objects. Template matching techniques compare portions of images against one another. The sample image may be used to recognize similar objects in source image.
The aim of this project is to detect as many faces as possible in a picture that is either taken through a web camera or a digital camera.

1. FACE DETECTION

Face detection is an important step in any automatic face recognition system. Given an image of arbitrary size, the task is to detect the presence of any human face appearing in the image. Detection is a challenging task since human faces may appear in different scales, orientations, and with different head poses. The imaging condition, such as illumination, also acts the appearance of human faces considerably.
Moreover, human faces are non-rigid objects. There are a lot of variations due to varying facial expressions. In addition, the presence of other objects or facial features such as glasses, make-up or

beards contributes substantially to the variation of facial appearance in an image [PW00].

According to [PW00] they classified face detection methods broadly into two classes: template-based and feature-based detection. In the first class, the pixel set in an image window is classified to see if there is a human face at that location. In the second approach, faces are detected by grouping facial features according to their geometric configuration in a model of a face, or by segmenting candidate facial regions based on color information and further verifying these regions based on their shapes and pixel values.

The classification of [PW00] was further classified by [PS99]. The few distinct approaches they mentioned are namely:

1. **The top-down model-based approach** assumes a different face model at different coarse – to - fine scales. For efficiency, the image is searched at the coarsest scale first. Once a match is found, the image is searched at the next finer scale until the finest scale is reached. In general, only one model is assumed in each scale (usually in the frontal-parallel view) and thus it is difficult to extend this approach to multiple views.
2. **The bottom-up feature-based approach** searches the image for a set of facial features and groups them into face candidates based on their geometric relationship. Though this approach can be easily extended to multiple views, it is unable to work well under different image conditions because the image structure of the facial features vary too much to be robustly detected by the feature detectors.
3. **In texture-based approach** faces are detected by examining the spatial distribution of the gray-level information in the sub-image (using Space Gray Level Dependency (SGLD) matrices). This is again not easily extensible to multiple viewpoints.
4. **The neural network approach** detects faces by sub-sampling different regions of the image to a standard-sized sub-image and then passing it through a neural network filter. In general, the algorithm performs very well for frontal-parallel faces but performance deteriorates when extended to different views of the face.
5. **The color-based approach** labels each pixel according to its similarity to skin color, and subsequently labels each sub-region as a face if it contains a large blob of skin color pixels. It can cope with different viewpoint of faces but it is sensitive to skin color and the face shape.
6. **Motion-based approaches** use image subtraction to extract the moving foreground from the static background. The face is then located by examining the silhouette or the color of the difference image. This approach will not work well when there are a lot of moving objects in the image.
7. **At depth-based approach** primary facial features are localized on the basis of facial depth information. In the first step, pairs of stereo images containing frontal views are sampled from the input video sequence. Then point correspondences over a large disparity range are determined using a multi-resolution hierarchical matching algorithm. Finally, the facial features are located based on depth information.

As face detection is the first step of any face processing system, it finds numerous applications in face recognition, face tracking, facial expression recognition, facial feature extraction, gender classification, clustering, attentive user interfaces, digital cosmetics, biometric systems, . In addition, most of the face detection algorithms can be extended to recognize other objects such as cars, humans, pedestrians and signs etc

1.1 Template matching

Template matching is a technique in digital image processing for finding small parts of an image which match a template image. It can be used in manufacturing as a part of quality control, a way to navigate a mobile robot, or as a way to detect edges in images.

Template matching can be subdivided between two approaches: feature-based and template-based matching. The feature-based approach uses the features of the search and template image, such as edges or corners, as the primary match-measuring metrics to find the best matching location of the template in the source image. The template-based, or global approach, uses the entire template, with generally a sum-comparing metric that determines the best location by testing all or a sample of the viable test locations within the search image that the template image may match up to.

Feature-based approach

If the template image has strong features, a feature-based approach may be considered; the approach may prove further useful if the match in the search image might be transformed in some fashion. Since this approach does not consider the entirety of the template image, it can be more computationally efficient when working with source images of larger resolution, as the alternative approach, template-based, may require searching potentially large amounts of points in order to determine the best matching location.

Template-based approach

For templates without strong features, or for when the bulk of the template image constitutes the matching image, a template-based approach may be effective. As aforementioned, since template-based template matching may potentially require sampling of a large number of points, it is possible to reduce the number of sampling points by reducing the resolution of the search and template images by the same factor and performing the operation on the resultant downsized images (multi-resolution, or pyramid, image processing), providing a search window of data points within the search image so that the template does not have to search every viable data point, or a combination of both.

Template matching has various different applications and is used in such fields as face recognition and medical image processing. Systems have been developed and used in the past to count the number of faces that walk across part of a bridge within a certain amount of time. Other systems include automated calcified nodule detection within digital chest X-rays.

1.2 Skin color information

Skin detection plays an important role in a wide range of image processing applications ranging from face detection, face tracking, gesture analysis, content-based image retrieval (CBIR) systems and to various human computer interaction domains. Recently, skin detection methodologies based on skin-color information as a cue has gained much attention as skin color provides computationally effective yet, robust information against rotations, scaling and partial occlusions. Skin color can also be used as complimentary information to other features such as shape and geometry and can be used to build accurate face detection systems. Skin color detection is often used as a preliminary step in face recognition, face tracking and CBIR systems. Skin-color information can be considered a very effective tool for identifying/ classifying facial areas provided that the underlying skin-color pixels can be represented, modelled and classified accurately.

Most of the research efforts on skin detection have focused on visible spectrum imaging. Skin-color detection in visible spectrum can be a very challenging task as the skin color in an image is sensitive to various factors such as:

• **Illumination:** A change in the light source distribution and in the illumination level (indoor, outdoor, highlights, shadows, non-white lights) produces a change in the color of the skin in the image (color constancy problem). The illumination variation is the most important problem among current skin detection systems that seriously degrades the performance.

• **Camera characteristics:** Even under the same illumination, the skin-color distribution for the same person differs from one camera to another depending on the camera sensor characteristics. The color reproduced by a CCD camera is dependent on the spectral reflectance, the prevailing illumination conditions and the camera sensor sensitivities.

• **Ethnicity:** Skin color also varies from person to person belonging to different ethnic groups and from persons across different regions. For example, the skin color of people belonging to Asian, African, Caucasian and Hispanic groups is different from one another and ranges from white, yellow to dark.

• **Individual characteristics:** Individual characteristics such as age, sex and body parts also affect the skin-color appearance.

• **Other factors:** Different factors such as subject appearances (makeup, hairstyle and glasses), background colors, shadows and motion also influence skin-color appearance [KMB07]

2. REVIEW OF RELATED JOURNALS

In visual object recognition using template matching [CA04], a two-tier hierarchical representation of images for each object is created, reducing the set of images to match against. The reduced set is generated via a pre-processing task of extracting areas of interest (blobs) and then reducing the number of blobs found by clustering.

The results are based on a raw database of 90 classes and a total of 140; 000 images each of size 680x480. From the 91 classes, over 100; 000 blobs were extracted as the training set which then was reduced by over 90%. For a recognition rate of 90%, only 0.59% of this training set was examined, giving us an execution time to identify a new image in 6.7 seconds.

For visual object recognition, a large number of views of each object are required due to viewpoint changes and it is necessary to recognise a large number of objects, even for relatively simple tasks. Clearly, a large number of images must be gathered in order to correctly classify new objects. The training was based on two-dimensional images taken with a digital camera. However, the world is three-dimensional. All of the images are projections of the three-dimensional object and do not contain information such as depth and geometry. For robotics applications, it is not acceptable to pause for long periods trying to recognise an object. Hence the focus of the work is on real-time methods and scaling to large sets of objects, more than accuracy.

The recognition procedure is closely tied to the approach used above to reduce the number of images that need to be compared for recognition.

A two-tier hierarchical classification approach is used. The first step is to make an approximate

classification in the form of recording possible matching classes.
In face detection using color thresholding and eigen image for template matching [MP03], the goal of the project is to take a color digital image with over 20 faces and indicate the location of the centre of each face in that image.
The approach used in the paper follows one similar to a rejection scheme algorithm [EHK02]. The first step is to reject regions in the image that are not faces based on color thresholding and skin segmentation. The next step is to use binary image processing to create clearer delineations in these regions. Template matching with both training image faces and eigenfaces using the Sirovich-Kirby method is used to detect faces from the non-faces, and also used to find the position of faces. Other sub-routines including the separation of overlapping faces and removal of high aspect ratio and low standard deviation non-face images are iterated against the main detection scheme before a final output is created. The algorithm was implemented in MATLAB.
The detection rate of the system is 95% with 4.3% false positives. Moreover, the result demonstrates that the system works extremely well for the images whose faces are full, upright, and facing towards the front. The false positives primarily occurred on large non-face body parts such as arms and legs. The false negatives were typically due to an obstructed face or variations that caused separations in the face such as sunglasses.
Using color information in images is one of the various possible techniques used for face detection. The novel technique used in this project was the combination of various techniques such as skin color detection, template matching, gradient face detection to achieve high accuracy of face detection in frontal faces. The objective of the work was to determine the best rotation angle to achieve optimal detection. Also eye and mouse template matching have been put to test for feature detection.
A face detection algorithm for color images in the presence of varying lighting conditions as well as complex backgrounds was proposed [HAJ02]. The method detects skin regions over the entire image, and then generates face candidates based on the spatial arrangement of these skin patches. The algorithm constructs eye, mouth, and boundary maps for verifying each face candidate. Experimental results demonstrate successful detection over a wide variety of facial variations in color, position, scale, rotation, pose, and expression from several photo collections. The use of color information can simplify the task of face localization in complex environments. The algorithm contains two major modules: (i) face localization for finding face candidates; and (ii) facial feature detection for verifying detected face candidates.
The appearance of the skin-tone color can change due to different lighting conditions. A lighting compensation technique that uses reference white to normalize the color appearance was introduced. Modelling skin color requires choosing an appropriate color space and a cluster associated with skin color in this space. Since the skin-tone color depends on Luminance, the YCbCr color space to make the skin cluster luma-independent was nonlinearly transformed. This also enables robust detection of dark and light skin tone colors.
[LS03] presented an approach to automatic detection of human faces in color images. The proposed approach consists of three parts: a human skin segmentation to identify probable regions corresponding to human faces; adaptive shape analysis to separate isolated human faces from initial segmentation results; and a view-based face detection to further identify the location of each human face. The human skin segmentation employs a model-based approach to represent and differentiate the background colors and skin colors. To further refine the initial segmentation results, the adaptive shape analysis uses a series of morphological operations based on the prior shape knowledge of upright human faces. The view-based face detection is constructed based on principal component analysis and neural network classification. The face detector has been applied to several test images, and satisfactory results have been obtained.
Color is a prominent feature of human faces. Using skin color as a primitive feature for detecting face regions has several advantages. In particular, processing color is much faster than processing other facial features. Furthermore, color information is invariant to face orientations. However, even under a fixed ambient lighting, people have different skin color appearance. In order to effectively exploit skin color for face detection, there is a need to find a feature space, in which human skin colors cluster tightly together and reside remotely to background colors.
In summary, a hybrid approach to automatic detection of human faces in color images is presented and the efficacy of our face detector by several test images is demonstrated.
[J+07] proposed a face detection approach. The luminance-conditional distribution model of skin-color information is used to detect skin pixels in color images, the idea is to segment eye-pair candidates using skin-color information; then, morphological operations are used to extract skin-region rectangles and finally template matching based on a linear transformation is used to detect

faces in each skin-region rectangle. The proposed algorithm is shown to be effective and efficient in detecting frontal faces in color images. The skin color information is particularly important for face detection as skin color differs from race to race, and depends on the lighting conditions. Modelling skin color requires choosing an appropriate color space and identifying a cluster associated with skin color in this space. The luminance-conditional distribution model of skin color information is utilized to detect skin pixels in color images. The skin pixels detected may form several skin regions in the color image. There are often some false skin pixels detected because of noise in the color image. The binary morphological operations can be utilized to remove the false skin pixels. Then all the skin regions can be extracted.

Experimental results show that the proposed algorithm to detect faces by using template matching and skin-color information is effective and efficient in detecting frontal faces with different races and under different lighting conditions.

3. SYSTEM DESIGN

3.1 Use Case Diagram

There are two people involved in the Use case diagram; they are the user and the system. As seen from bellow, the user is responsible for selecting the testing photograph while the system performs the face detection, displays the face detection result, load system setting and label face in image.

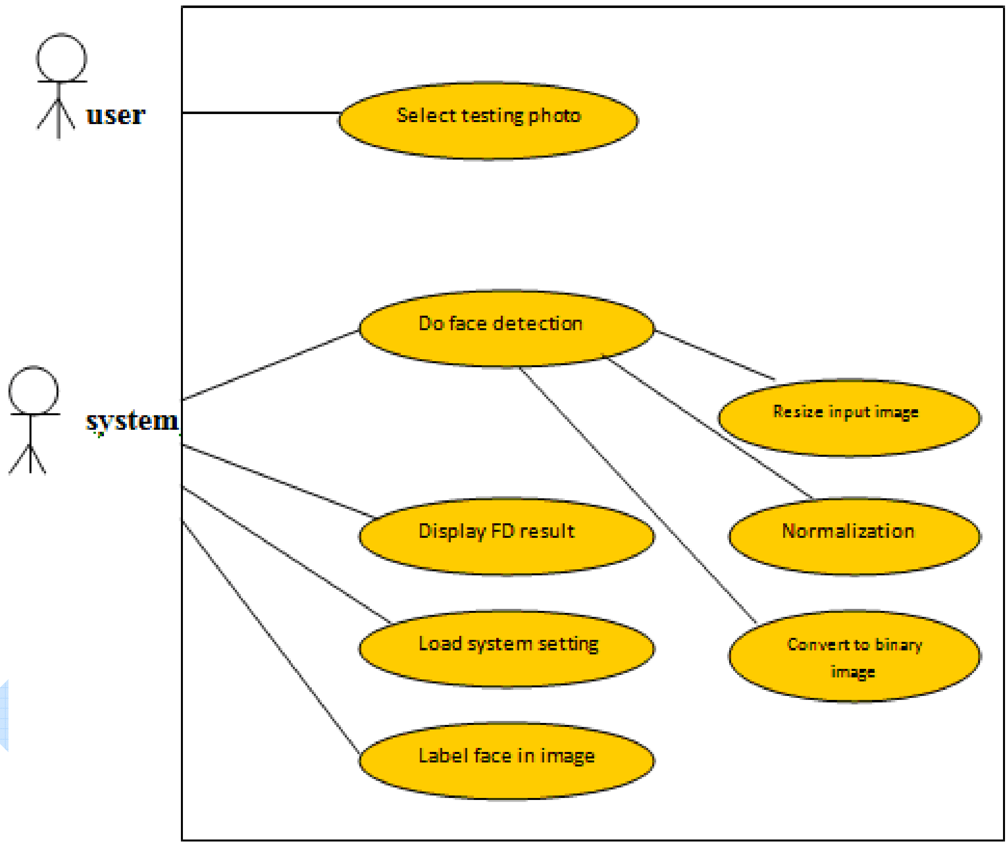

Figure 1: Use Case Diagram

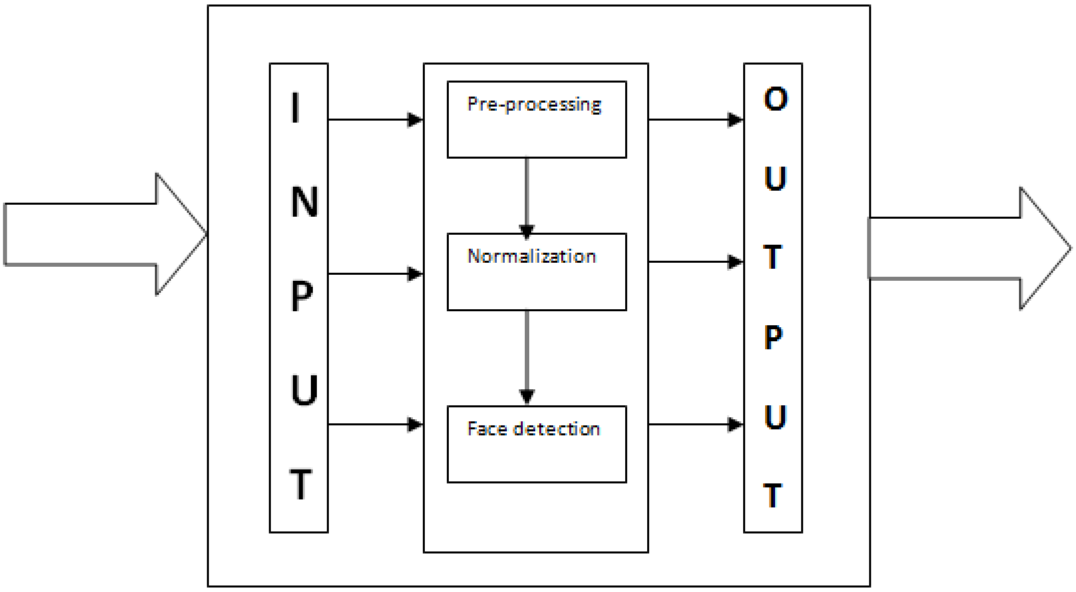

Figure 2: System Architecture

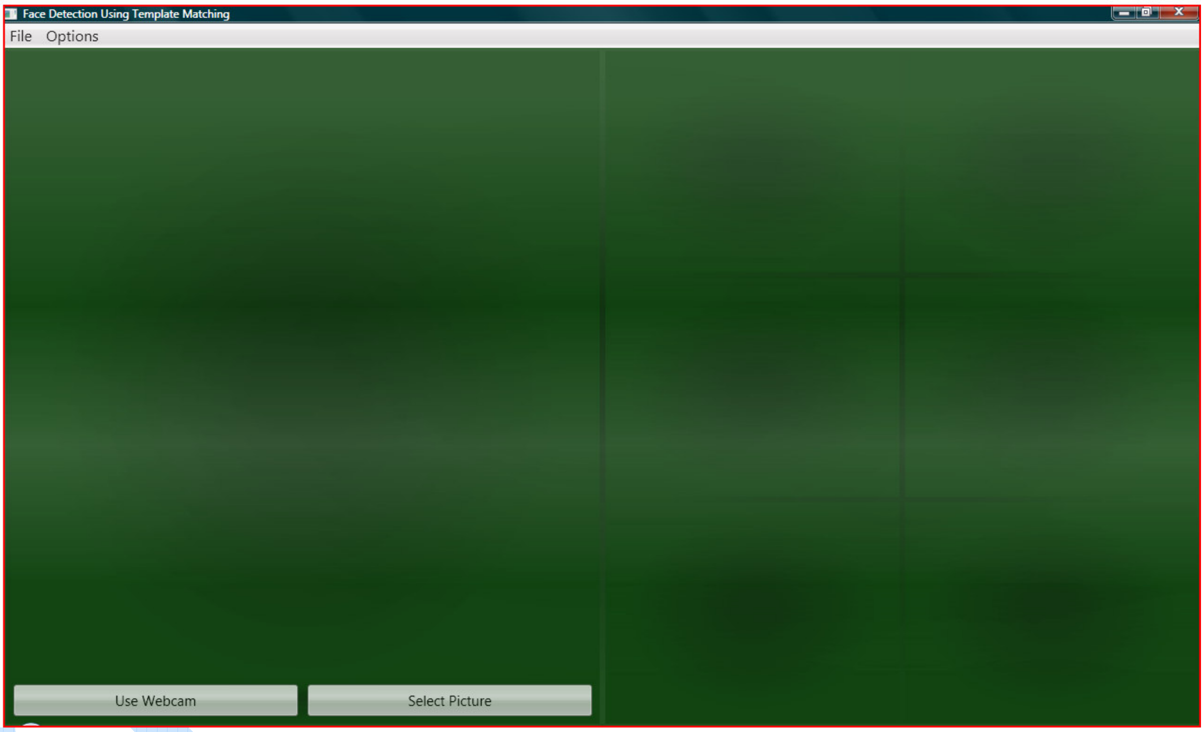

Figure 3: The user interface

3.2 Architecture of the face detection system

The input to the face detection system is the picture either taken by a web camera or saved on a folder on the system for easy accessibility.
The processing stage involves three main processes namely:
- Pre-processing stage is the first stage to be carried out by the system. It involves getting the image, conversion of image to binary image and the loading of the face template.
- Normalization stage involves normalization of binary image i.e. is a process that changes the range of pixel intensity values.
- Face detection stage involves the comparison between the face template and the extracted region of the binary image. This is the last stage of processing and it ensures that the entire binary image is searched for as many faces as can be located.

The output stage displays the result of the image selected with rectangles drawn around regions that look like faces.

3.3 User Interface

Under the file menu we have the following:

i. Load skin region samples which enable the skin calibration to de done.

ii. Repeat detection is used when parameters have been changed and the face detection is performed again on the selected image/picture.
iii. Exit is used to put an end to the program.

Under the options menu we have the following:

i. The "threshold" option in percentage which specifies the percentage match (with face template) that a skin region should have before it is considered a face.
ii. "Use Maximum threshold" option which tells the program to determine the skin region(s) with the maximum match and select those regions as faces.
iii. "Merge intersecting face regions" option which tells the application to merge face regions that overlap into a single region.

3.4 Results

The face detection system was tested with a computer system of 32 bit operating system (Microsoft Vista), 1 GB RAM, 70Gb Hard drive, Intel Celeron M processor and a processor speed of 1.86Ghz. The face detection system was tested by selecting a picture from those some pictures saved on the system. Once the picture has been selected, a skin calibration is done (depending on whether it is a White skin, African skin or Nigerian skin, after which the system goes on to do the face detection.

An important point to note is the use of the following parameters to get the desired result. The parameters are adjusted until a near perfection result is got. The parameters are:

i. The "threshold" option in percentage which specifies the percentage match (with face template) that a skin region should have before it is considered a face.
ii. "Use Maximum threshold" option which tells the program to determine the skin region(s) with the maximum match and select those regions as faces.
iii. "Merge intersecting face regions" option which tells the application to merge face regions that overlap into a single region.

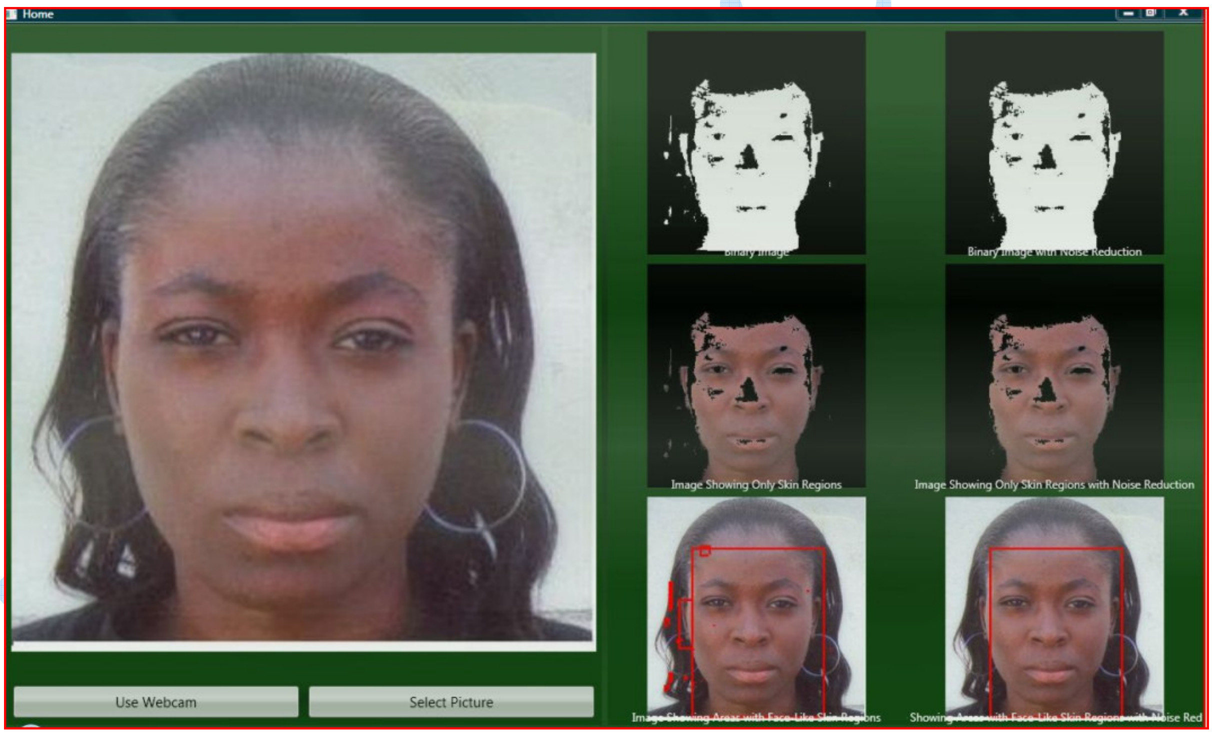

Figure 4: A face detected by the program

CONCLUSIONS

The face detection system can detect frontal faces in color images. One of the major shortcomings of template-based methods is that they are very sensitive to both rotation and scaling. To detect faces of different scales, the input image or the template can be resized to an appropriate value so that classification or matching can be carried out.

Another limitation of the system is that it does not perform well when the size of the picture is more than 200kb. It becomes very slow and takes a very long time for the result to be generated. Also, in case the desired result is not got when a particular threshold is used, the threshold value can be changed until it presents the actual result wanted. But in all, template matching using C# programming language worked well in the detection of faces.

REFERENCES

[CA04] **L. Cole, D. Austin** - *Visual Object Recognition using Template Matching*, by Australasian Conference on Robotics, 2004

[CB07] **C. Caleanu, C. Botoca C** - *C# Solutions for a Face Detection and Recognition System*, Facta Universitatis (NIS) Elec. Energy, Vol. 20, No. 1, pp. 93-105, 2007

[EHK02] **M. Elad M., Y. Hel-Or, R. Keshet** - *Rejection Based Classifier for Face Detection,* Pattern Recognition Letters (23), pp. 1459-1471, 2002

[Fai08] **A. Faizi -** *Robust Face Detection using Template Matching Algorithm*, 2008, at https://tspace.library.utoronto.ca/bitstream/1807/17164/1/Faizi_Amir_200811_MASc_thesis.pdf

[HAJ02] **R. Hsu, M. Abdel Motalleb, A. Jain -** *Face Detection in Color Images,* Pattern analysis, 2002

[J+07] **Z. Jin, Z. Lou, Y. Yang, Q. Sun** - *Face Detection using Template Matching and Skin Color Information,* Neurocomputing 70, pp.794-800, 2007

[Kin03] **A. King** - *A Survey of Methods for Face Detection*, 2003, at http://www.sfu.ca/~adk7/papers/face.pdf

[KMB07] **P. Kakumanu, S. Makrogiannis, N. Bourhakis** - *A Survey of Skin-color Modelling and Detection Methods*; Journal of Pattern Recognition Society 40, pp. 1106- 1121, 2007

[KSR**] **A. Kharpade, R. Y. Sowmya, S. Ravipudi** - *Face Detection using Color-based Segmentation and Morphological Processing: A Case Study*, n.d.

[LS03] **W. Lu, S. Sun** - *EE368 Project: Face Detection in Color Images*, 2003, at http://www.stanford.edu/class/ee368/Project_03/Project/reports/ee368group03.pdf

[MP03] **D. Marius, S. Pennathur** - *Face Detection using Color Thresholding and Eigen Image template Matching,* Digital Image Processing project, 2003

[MBV05] **D. Masip, M. Bressan, J. Vitria** - *Feature Extraction Methods for Real time Face Detection and Classification,* EURASIP Journal on Applied Signal Processing 13, pp. 2061-2071, 2005

[PS99] **P. Peer, F. Solina** - *An Automatic Human Face Detection Method*, Proceedings of Computer Vision Winter Workshop, 1999

[PW00] **T. V. Pham, M. Worring** - *Face Detection methods: A Critical Evaluation*, ISIS Technical Report Series, Vol. 11, Department of Computer Science, University of Amsterdam, Netherlands, 2000

[Tsa**] **T. M. Tsang** - *Image based Face Detection System*, B.Sc, Oxford Brookes University, n.d.

[VJ04] **P. Viola, M. J. Jones** - *Robust Real Time Face Detection*, International Journal of Computer Vision 57 (2), pp. 137-154, 2004

[Xio**] **Z. Xiong** - *An Introduction to Face Detection and Recognition*, Department of Electrical and Computer Engineering, University of Illinois, n.d.