AUTOMATIC MORPHING OF FACIAL IMAGES

A Project Report

Submitted By:

EKTA SAGAR (061050)
HEMANT SHARMA (061054)
ANGAD ABROL (061150)

In partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

ELECTRONICS AND COMMUNICATION ENGINEERING

under the Supervision of

DR. VINAY KUMAR

विद्या तत्व ज्योतिगमः

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

Jaypee University of Information Technology
Waknaghat, Solan - 173 234, Himachal Pradesh

MAY 2010

CERTIFICATE

This is to certify that the project report entitled "Automatic Morphing of Facial Images", submitted by Ekta Sagar(061050), Hemant Sharma(061054) and Angad Abrol(061150) in partial fulfillment for the award of degree of Bachelor of Technology in Electronics and Communication Engineering to Jaypee University of Information Technology, Waknaghat, Solan has been carried out under my supervision.

Signature

Date : 24-05-2010

Dr. Vinay Kumar

Sr. Lecturer

Certified that this work has not been submitted partially or fully to any other University or Institute for the award of this or any other degree or diploma

Signature

1. **EKTA SAGAR**
Roll No.061050

Signature Hemant Sharma

2. **HEMANT SHARMA**
Roll No.061054

Signature

3. **ANGAD ABROL**
Roll No.061150

ACKNOWLEDGEMENT

The zeal to accomplish the task of formulating the project report on "Automatic Morphing of Facial Images" could not have been realized without the support and co-operation of the members of the faculty of ECE department.

First and foremost, we express our gratitude and sincere thanks to Dr. Vinay Kumar for providing us the opportunity to undertake the project under his able guidance. He has helped us to develop novel solution to every problem and has always encouraged us to put in our best efforts and deliver a quality and professional output.

Secondly, we thank Dr D.S.Saini, Mr Vipin Balyan, Mr Tapan Jain for their patient hearing of our ideas and opening up our minds to newer horizons by pointing out our flaws, providing critical comments and suggestions to improve the quality of our work and appreciating our efforts.

Date: 24th May, 2010

Hemant Sharma(061054)
Ekta Sagar (061050)
Angad Abrol (061150)

ABSTRACT

Image morphing is an image processing technique used to generate a transition from one image to another. Morphing is a common technique in generating special effects for entertainment, but is becoming important in the fields of computer vision and psychology. Most image morphing techniques require the knowledge of significant image features, which is used to perform realistic morphing and are often obtained by user interaction. In this report an automatic feature based technique for morphing algorithm is discussed which eliminates the need for user interaction. The algorithm has been used to generate morphing between still images of faces of different people as well as between different images of the face of an individual. This algorithm automatically extracts feature points on the face, and based on these feature points images are aligned using the basic warping and linear interpolation techniques. The warped images are combined using a technique of cross fading to get intermediate images. The algorithms used are described in detail including an analysis of the number of critical points that must be identified for a successful morph and the assumptions taken about the nature of images as well as the possible problems that could be encountered.

Table of Contents

Chapter 1

INTRODUCTION

Image morphing is a popular method used to create dramatic special visual effects. Image morphing is basically a method for transformation of the shape and color of an image to another image creating a convincing transition, or metamorphosis. Although the resulting transition can be quite convincing and impressive, the amount of work required to produce the animation is often extensive. Different methods for producing image morphing exist, but common for most methods is that they require some sort of feature correlation between the source and destination images. In most applications this feature correlation data is specified by hand, and thus require extensive work. To manually input the required data might be feasible when image data is limited, but the amount of work becomes unrealistic for some applications.

Morphing involves the image processing techniques of cross-fading and warping. It is basically achieved by coupling image warping with color interpolation. As the morphing proceeds, the first image is gradually distorted and is faded out, while the second image starts out and is faded in. The middle image is fifty percent of both images. There are various techniques in implementing the morphing and they depend on the way of establishing the correspondence between the two images.

1.1 APPLICATIONS

1. Face recognition and human perception
2. Morphing has been used as an element in many recent motion pictures, music videos and, television commercials
3. It can enhance many multimedia projects, presentations, education and computer based training.

1.2 SOFTWARE USED

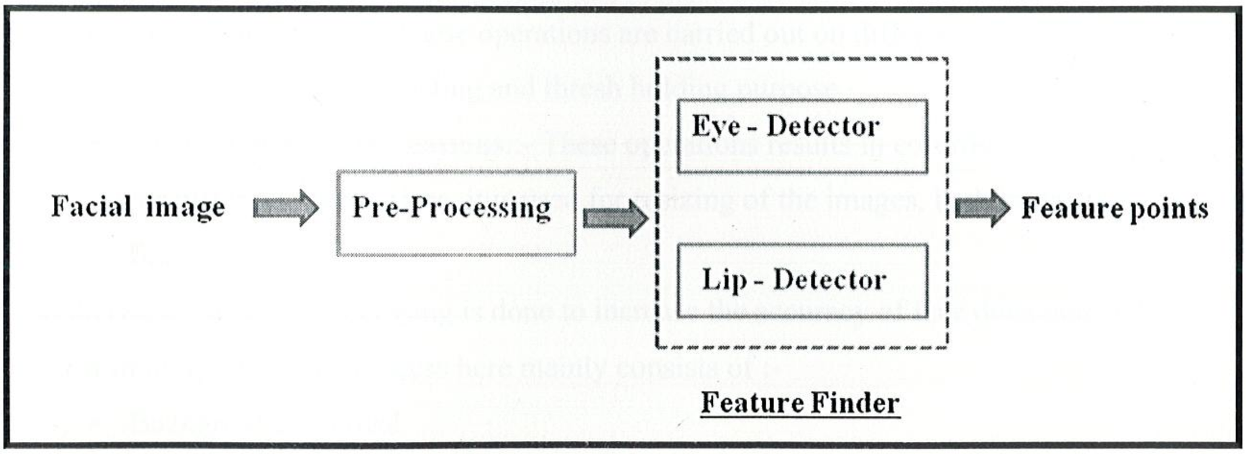
- MATLAB 7.1 :- It is a high-level language with interactive environment that enables to perform computationally intensive tasks faster than with traditional programming languages such as C, C++, and Fortran.

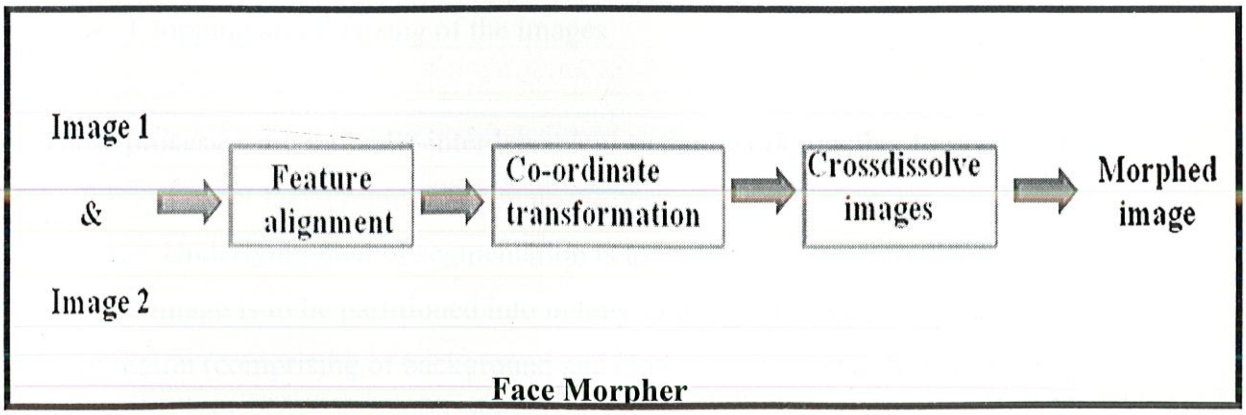
1.3 OUTLINE OF THE PROCEDURE

The proposed algorithm in this report basically consists of three steps.

- First, pre-processing scheme is applied to remove background of the initial and final images, scaling of images and detect face in the images. This results in a binary image representation in the form of face mask which facilitates the feature extraction and image warping procedures.
- Secondly, a feature extraction scheme is applied. The scheme finds key feature points comprising of eyes and lips present in the images.
- The third and final step performs warping and cross fading between the initial and final image. It uses the method of point-to-point mapping and linear interpolation to generate intermediate images.

The algorithm consists of a feature finder and face morpher. The following figures illustrates the procedures.

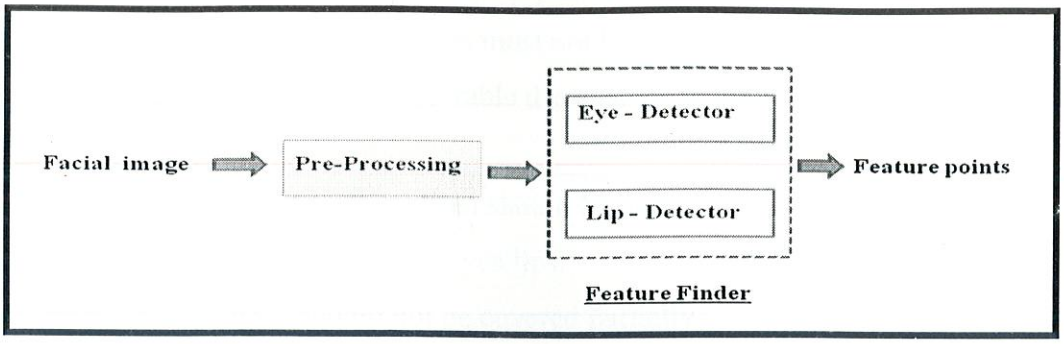




The details for the implementations will be discussed in the following chapters.

Chapter 2

PRE-PROCESSING



Pre-processing consists of the operations that occur at the beginning of the processing chain. It adds no image content – in fact, usually removes information. Aim of image pre-processing is to suppress information that is not relevant, and enhance information that is particularly relevant.

Basic pre-processing methods used in this algorithm are:-

- **Point operations**::- These operations are carried out on different pixels. Here, it is used for differentiating and thresh holding purpose.
- **Geometric transformations**::- These operations results in co-ordinate change for different pixels. Here, it is used for resizing of the images, both initial and final.

In this algorithm, pre-processing is done to increase the accuracy of face detection and perform morphing. This process here mainly consists of :-

- Background removal
- Face mask generation
- Skin region segmentation
- Cropping and Resizing of the images

These processes are mutually interdependent as they work together to give output in the form of exact facial region. The main purpose of pre-processing is segmentation of image. Underlying goal of segmentation is to partition image into different regions. Here the image is to be partitioned into mainly two regions—facial region and non-facial region (comprising of background and clothes (if present)). It helps the regions to be analyzed in a more meaningful manner.

Assumptions made for the considered images(initial and final) are:-

1. The images should have at least 3 pixel background with uniform or slight intensity transition in order to avoid presence of edges in the background.
2. The faces present in the images must not be tilted by more than 15 degree.
3. The faces should be of comparable dimensions in both the images to give optimum result.
4. The key features (eyes and lips) should be properly visible.
   Example - No hairs on the eyes/lips,
   Face should not be covered partially.

2.1 BACKGROUND REMOVAL

This step of pre-processing is done in order to decrease the region of concentration for finding face in the given image. The process of background removal involves computing intensity difference between pixels at a gap of 3 pixels and applying threshold.

Thresholding, a commonly used point-operation is used here to separate objects of interest from the background image.Trivial way for thresholding is to set pixels below certain value to 0 (*background*) and all others to 1 or 255 (*object*). In this algorithm thresholding is applied on the intensity difference between two pixels and not on the individual pixel intensity value.

The intensity difference gives the information about the intensity transition and threshold helps us decide about the presence of edge. As the background has slight intensity transition, the difference between neighboring pixels will be very less whereas at the outer boundary of facial region the difference between neighboring pixel will be more.

The computation of intensity difference is carried out in all the four directions(left to right, right to left, top to bottom and bottom to top) in order to increase the accuracy and remove maximum of the background. Different intensity difference threshold is applied in different direction to get optimum result. The co-ordinates of the average of the pixels whose intensity difference exceeds threshold is stored in a matrix. A binary image is generated using the stored co-ordinates which is used in the following steps.

2.2 FACE MASK GENERATION

This step of pre-processing uses the binary image generated during background removal process to create a binary face mask which approximately consists of whole of the facial region.

To generate face mask, maximum and minimum co-ordinates (in both X and Y directions) of the object present in the binary image are calculated. The facial region is then partitioned into 20 equal horizontal slots based on the height of the binary image.

In each horizontal slot, maximum and minimum points in horizontal direction is calculated and pixel co-ordinate is stored in a matrix. The matrix will contain 40 pixel points which are used to generate a 40 point polygon which will surround the face.

The polygonal mask is stored as a binary image which is used in further steps to define facial region.

For accurate detection of face, it is assumed here that facial region is well differentiated from the background and clothes(if present) with respect to color change.

The steps can be diagrammatically represented as:-

- Reading image from a specified location and storing it as a matrix.

- Colour format conversion (RGB to GRAYSCALE).
- Making image matrix 2D.

- Calculating pixelwise intensity difference in both x and y direction.

- Plotting those points where intensity difference is greater than threshold.

- Storing the points in a matrix.

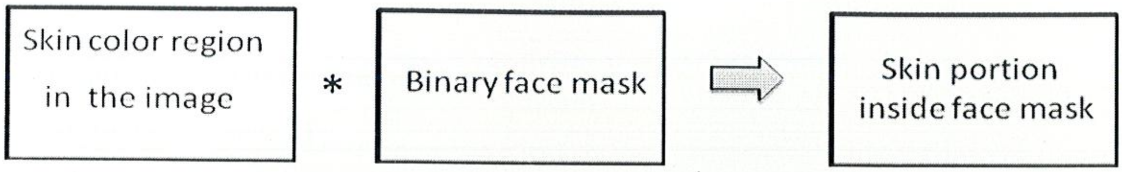
- Dividing image into sections (each of height 5% of the original) in horizontal direction.

- Calculating the minimum and maximum co-ordinates of the pixels lying in that region(as calculated and stored in previous step).

- Storing these co-ordinates in a matrix and plotting a polygon based on these points.

2.3 SKIN SEGMENTATION

This step of pre-processing is carried out to generate accurate information about the facial region. Skin color is detected in whole image and the face mask generated in the previous step is then applied to remove skin color region present in the background. The output image consists only of the skin region present inside the face.



Presence of skin color in the background can create problem if it is comparable to the skin color of the face else it will be removed during background removal.

The color range used for detecting skin region is

$$0.35 < r < 0.55$$

$$0.2 < g < 0.7$$

$$200 < b1 < 740$$

where,

$r1=(R)$

$g1=(G)$

$b1=(R)+(G)+(B)$

$r=r1/b1$

$g=g1/b1$

R= Red layer of image (RGB format)

G= Green layer of image (RGB format)

B= Blue layer of image (RGB format)

Image in RGB format can be described as 24-bit image which is formed by combining three different layers- namely red layer, green layer and blue layer. Each layer is represented in 8-bit format that is intensity value ranges from 0-255 in each layer.

2.4 CROPPING AND SCALING IMAGES

In this step of pre-processing facial region of both initial and final images are cropped and resized to same scale to make co-ordinate mapping easier.

In this algorithm the cropping is done by extracting the region inside the face mask and then taking objects from both the images and resizing them with dimensions a=max(a1,a2) and b=max(b1,b2).The resized objects(faces) are centered in two separate images whose dimensions are [a+6,b+6].

Images	Dimension
Initial_Image	(a1, b1)
Final _Image	(a2, b2)

Scaled Images	Dimension
Initial_Image	(max (a1, a2) +6, max (b1, b2) +6)
Final _Image	(max (a1, a2) +6, max (b1, b2) +6)

Resizing of the images are done by using the function

B = imresize(A, scale)

B = imresize(A, scale) returns image B that is scale times the size of A. The input image A can be a grayscale, RGB, or binary image. If scale is between 0 and 1.0, B is smaller than A. If scale is greater than 1.0, B is larger than A.

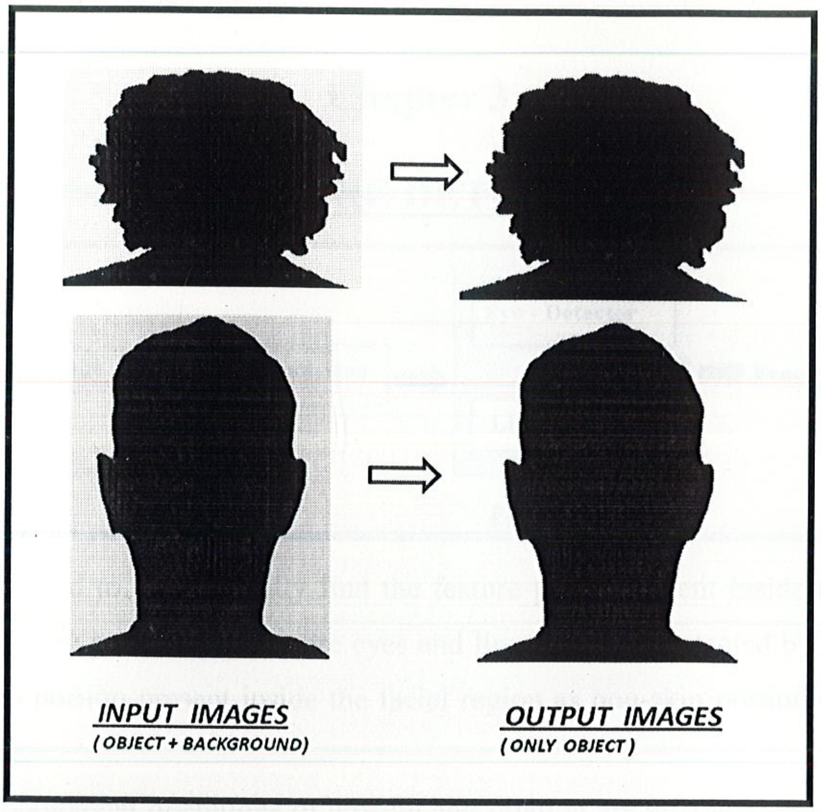


Figure: **Cropping**

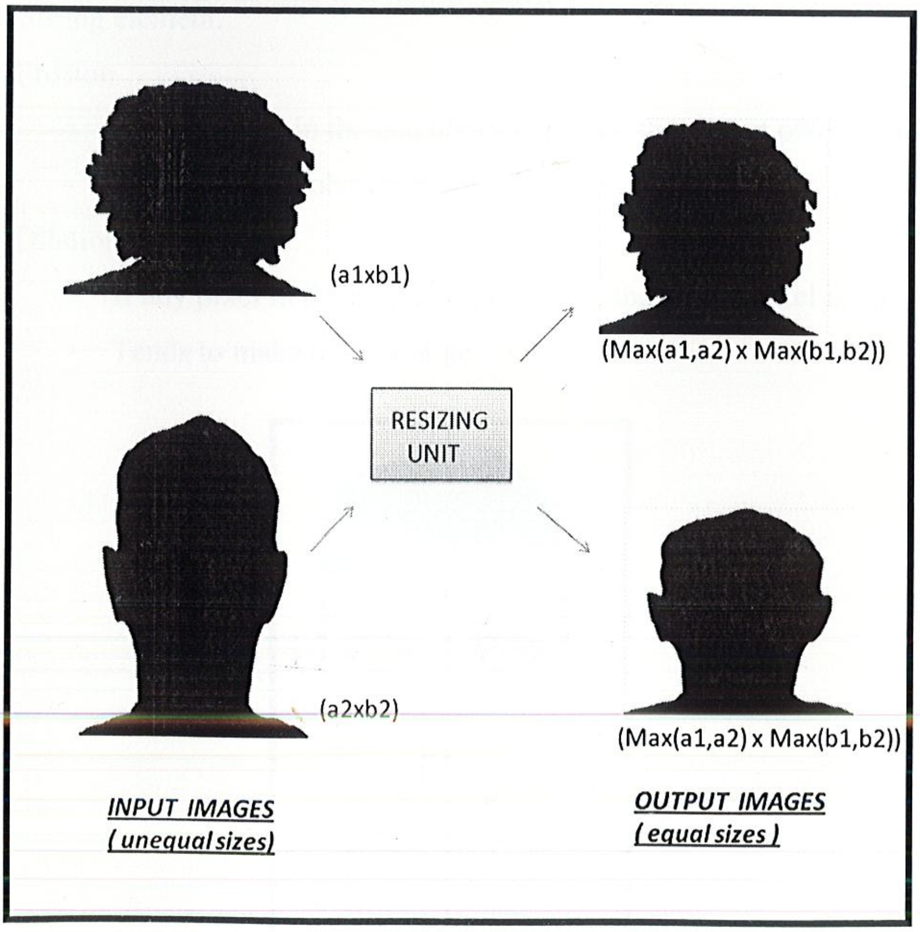
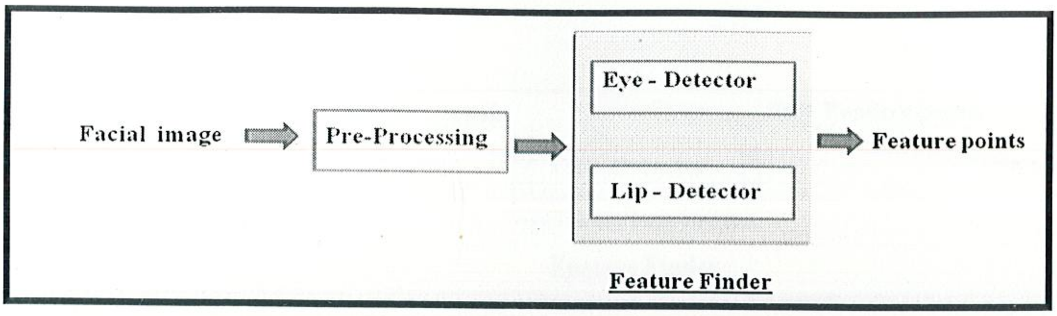


Figure: **Resizing**

Chapter 3

FEATURE DETECTION



This step is used to automatically find the feature points present inside the detected face. The feature points of a face are eyes and lips. They are detected by recognizing the non-skin portion present inside the facial region as non-skin portion comprise of eyes and lips.

Image morphological operations of erosion and dilation are also used here to remove noise. For each pixel in the input image, the neighborhood is examined as specified by the structuring element.

- Erosion
  - If every pixel in the neighborhood is on, the output pixel is on
  - Tends to make objects smaller.
- Dilation
  - If any pixel in the neighborhood is on, the output pixel is on
  - Tends to make objects large.

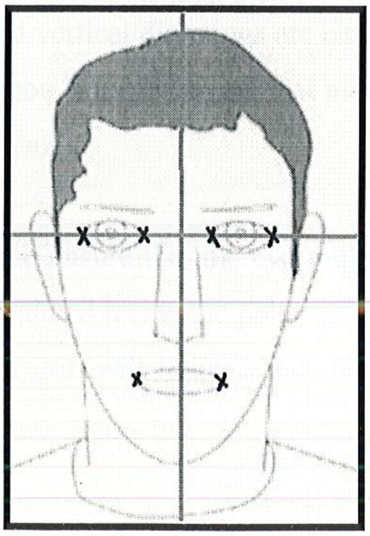
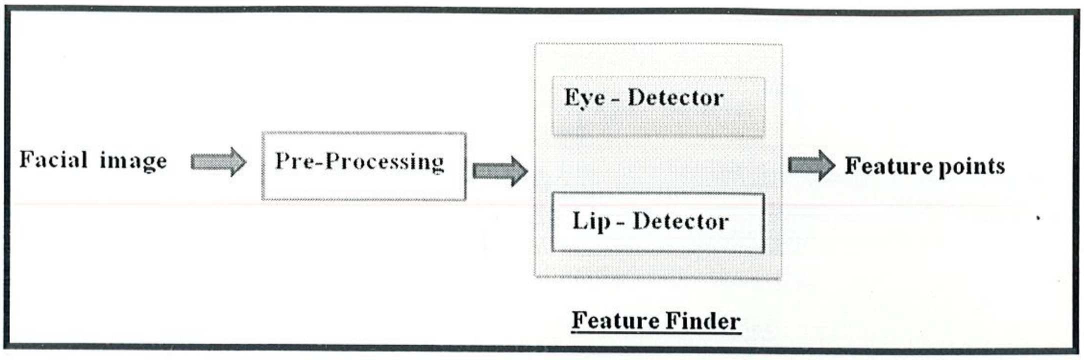
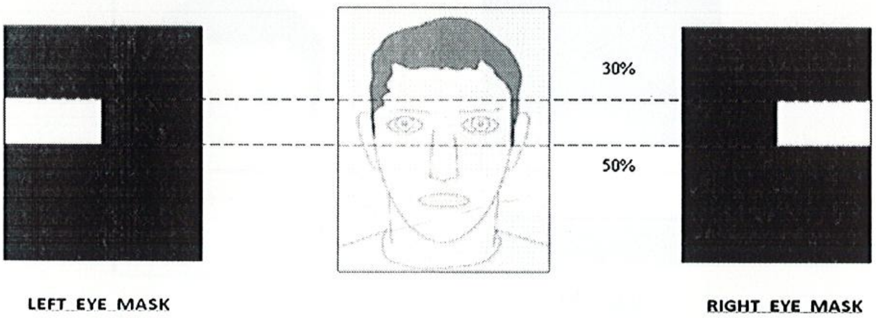


Figure: **Feature Points of Face**

3.1 EYES DETECTION



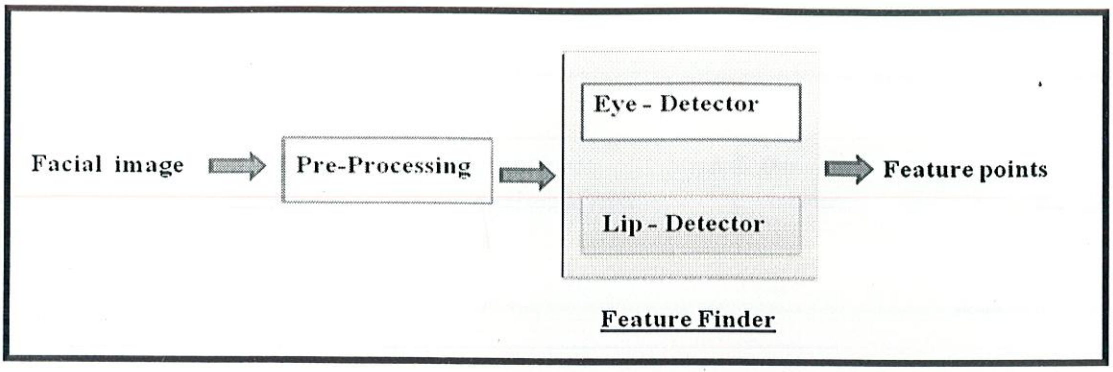
A rectangular binary adaptive eye mask whose dimension and position depends on the width and height of the facial region is generated separately for both the eyes. It is assumed that the eyes will lie inside 30-50% of the face region as calculated from topmost point.



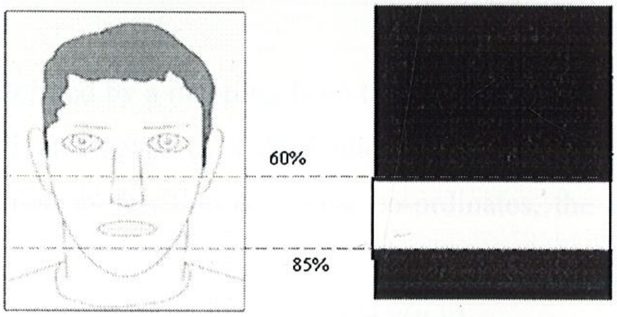
Inside each mask (for left and right eye) mean of the co-ordinates of the non-skin pixels in both horizontal and vertical directions are calculated to give the approximate position of eyes and a hexagon is made around that mean (whose size also depends on the height and width of the face).

To increase the accuracy and ensure that the eye region is well approximated by the polygon, mean is again calculated inside the polygon and final polygon is constructed around that mean. This polygon will indicate the eye region inside their respective masks.

3.2 LIPS DETECTION



A rectangular binary adaptive lip mask whose dimension and position depends on the width and height of the facial region is generated. It is assumed that the lips will lie inside 60-85% of the face region as calculated from topmost point.



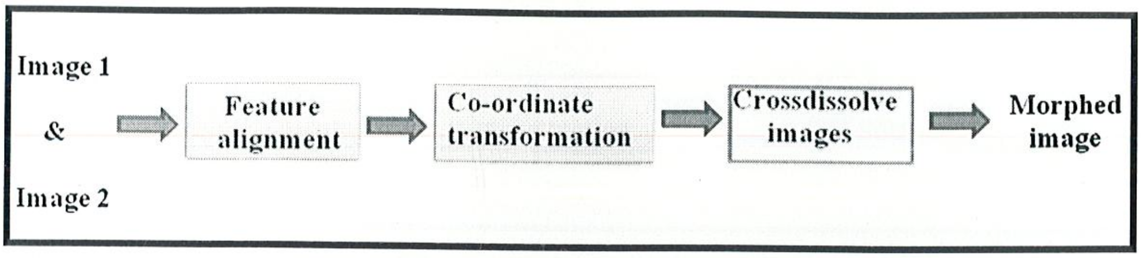
LIPMASK

Inside lip mask mean of the co-ordinates of the non-skin pixels in both horizontal and vertical directions are calculated to give the approximate position of lips and a hexagon is made around that mean (whose size also depends on the height and width of the face).

To increase the accuracy and ensure that the lip region is well approximated by the polygon, mean is again calculated inside the polygon and final polygon is constructed around that mean. This polygon will indicate the lip region inside the mask.

Chapter 4

WARPING



The term "image warping" refers to the process of geometrically transforming two-dimensional pictures or images. Image warping encompasses the whole range of transformation from simple ones such as scaling or rotation to complex, irregular warps. In morphing, image warping technique is combined with cross-fading to create a convincing illusion of one object transforming smoothly into another.

An image warp is defined by a mapping from the co-ordinate space of a source image (u,v) to the co-ordinates space of a destination image (x,y). if the destination co-ordinates are specified as function of source co-ordinates, the mapping is called a forward mapping:

$$x = x(u,v), \qquad y = y(u,v)$$

If the source co-ordinates are specified as functions of the destination co-ordinates, the mapping is called a reverse mapping:

$$u = u(x,y), \qquad v = v(x,y)$$

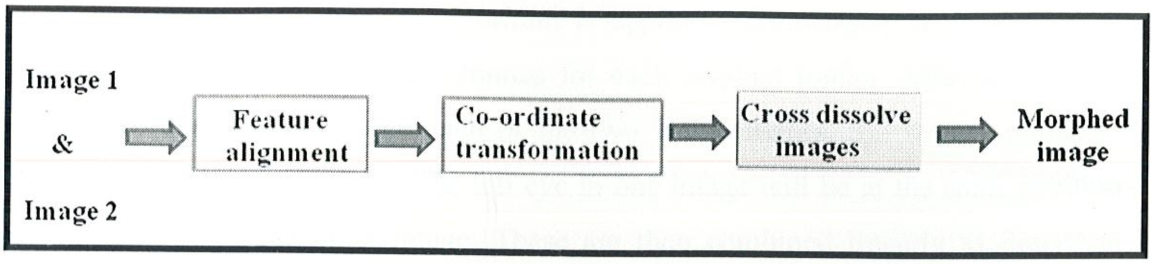
4.1 POINT-TO-POINT MAPPING

In this algorithm forward mapping is used to perform morphing. A warp described by a forward mapping is performed by scanning the source image pixel by pixel, calculating the corresponding location in the destination image by evaluating the mapping function and copying the destination image with the intensity value of the source pixel. Here, the source image is initial image and destination image is the final image.

In forward mapping the color intensity of a single source pixel has to be distributed over several destination pixels.

Chapter 5

CROSS-FADING



It involves creating a series of frames composed of linear combinations of the two original images.

Before the development of morphing, image transitions were generally achieved through the use of cross-dissolves/ cross-fading, e.g., linear interpolation to fade from one image to another. The result is poor, owing to the double-exposure effect apparent in misaligned regions. This problem is particularly apparent in the middle frame, where both input images contribute equally to the output. As the features are misaligned, the intermediate images just look like two faces superimposed, rather than one face as in goal. It has limited application and can only be applied on similar objects.

The following figure shows the results of applying a simple cross-fade to two images of faces. Despite the original images being quite similar, the features are misaligned and the intermediate images just look like two faces superimposed, rather than one face as is the goal. One can imagine that the results would be considerably worse were the original images of radically different shape or size.

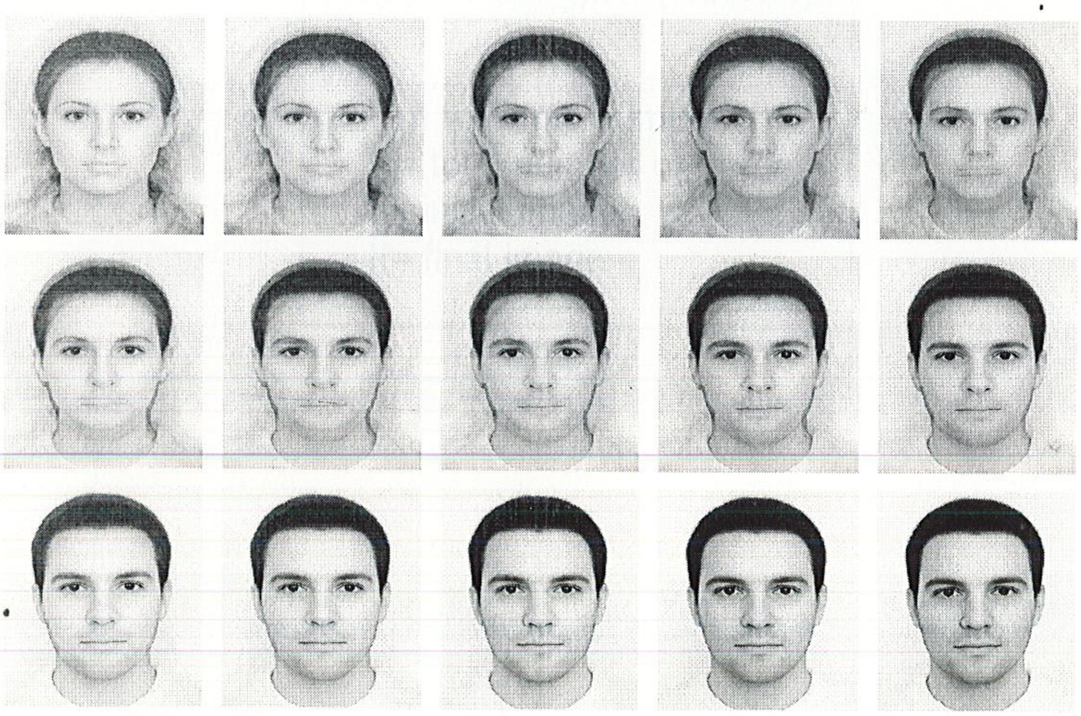


Figure: Crossfading

5.1 LINEAR INTERPOLATION

In this algorithm, the warping algorithm is applied incrementally to both images, creating a series of intermediate frames for each original image. After performing coordinate transformations for each of the two facial images, the feature points of these images are matched. i.e., the left eye in one image will be at the same position as the left eye in the other image. These are then combined linearly as described above and compiled to create the final morph sequence. The result is a transition where the key features of the original object are seen to move to match those of the final one.

Morphing achieves a fluid transformation by incorporating warping to maintain geometric alignment throughout the cross-dissolve process. To complete face morphing, we need to do cross-dissolving as the coordinate transforms are taking place. Cross-dissolving is described by the following equation,

Working formula:

$$F(n)=(n-1/N-1)If + (N-n/N-1)Ii$$

Where,

F(n) = Frame number n

N= total number of frames.

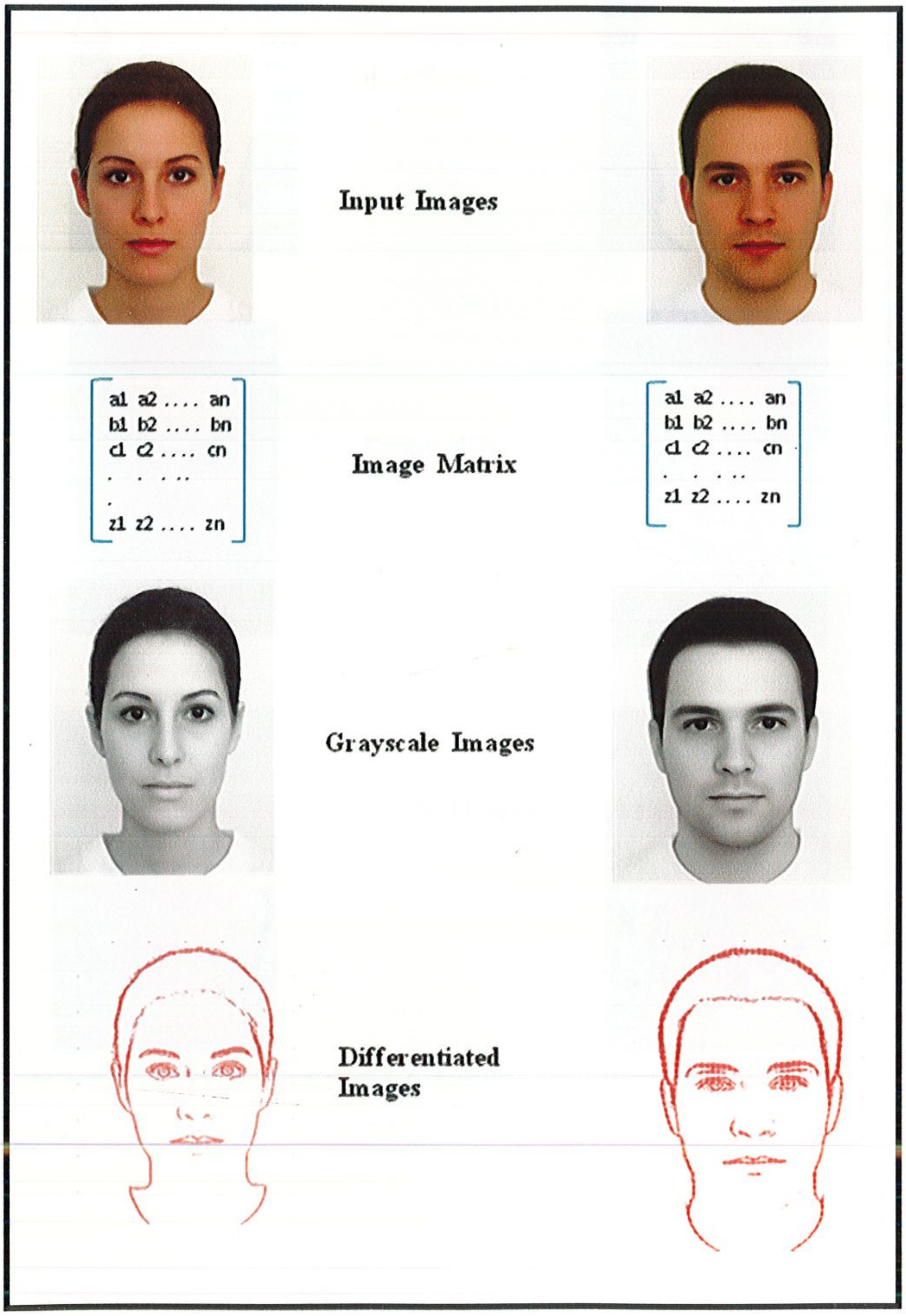
Ii= initial image

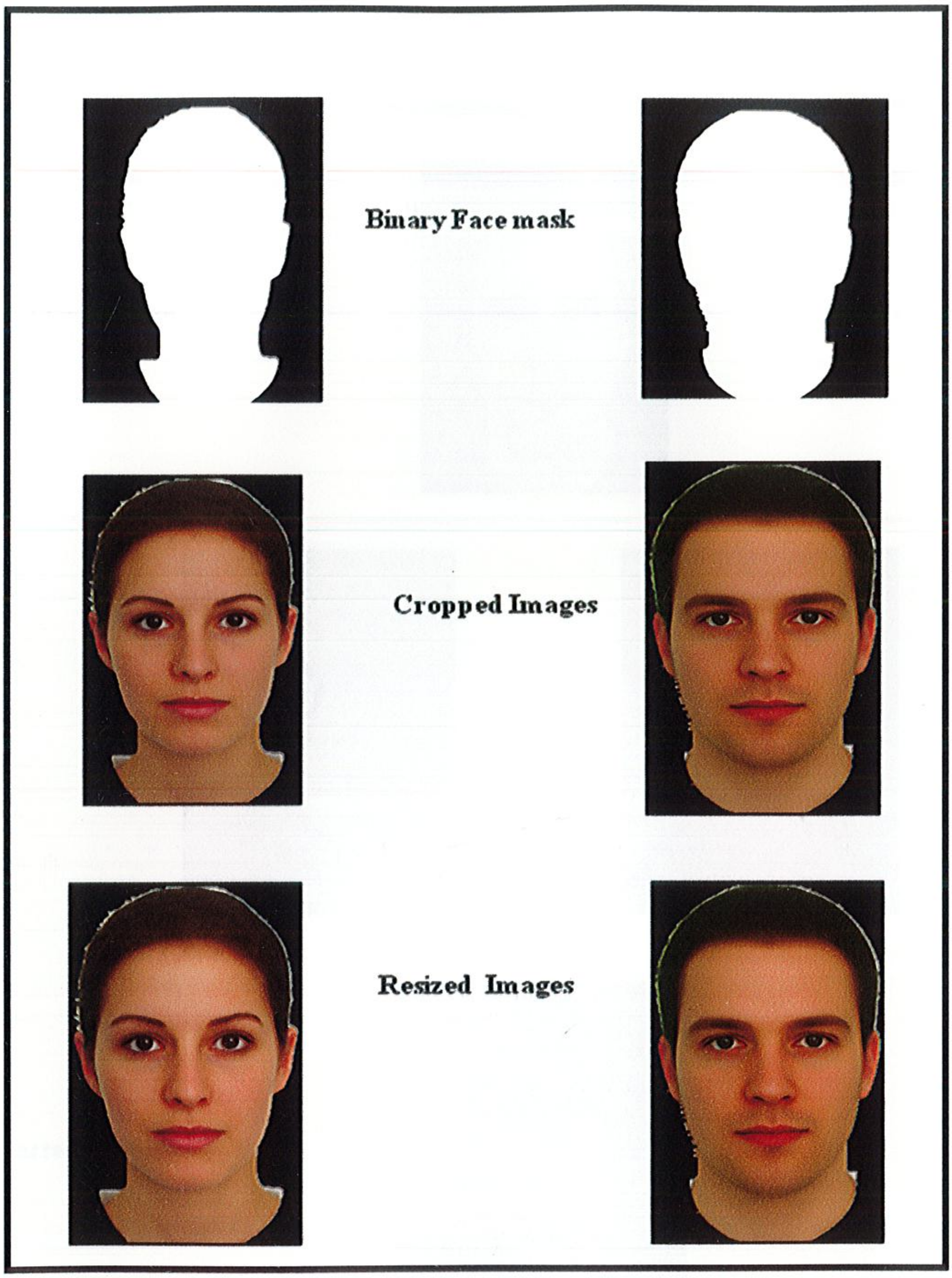
If= final image

RESULTS

Example 1:

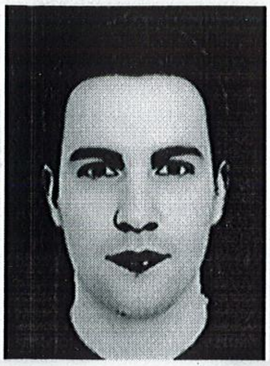
Pre-Processing

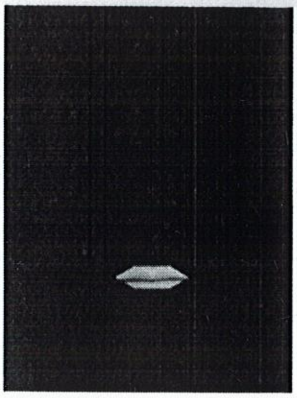


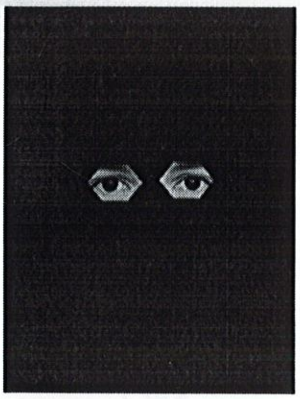

Binary Face mask
Cropped Images
Resized Images

Feature Extraction

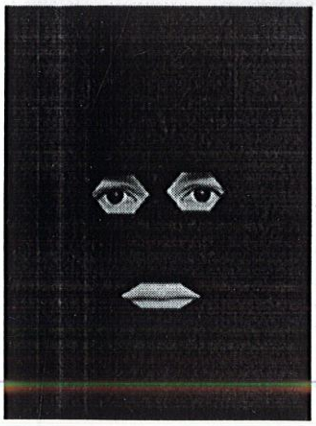
Skin Region

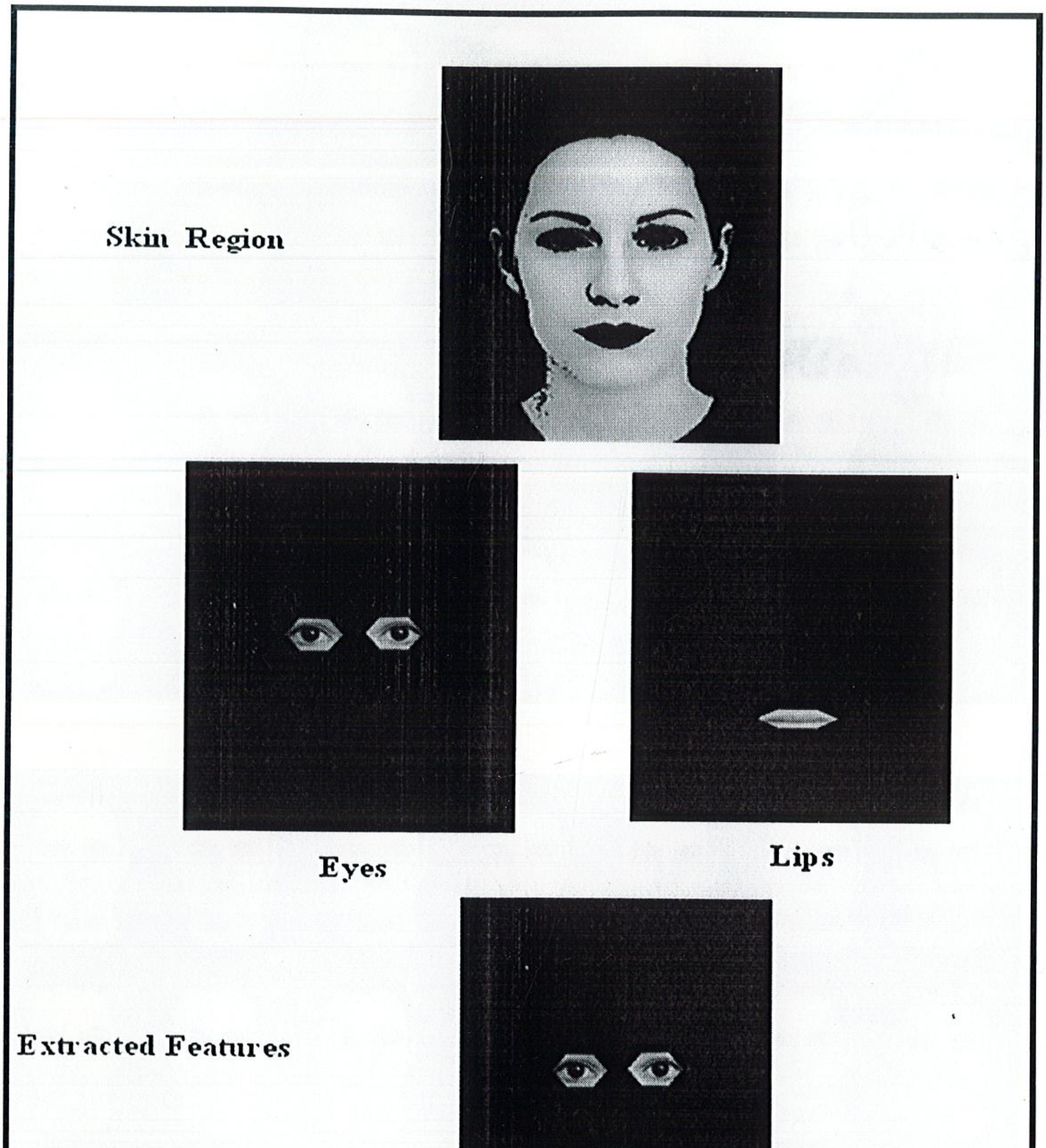




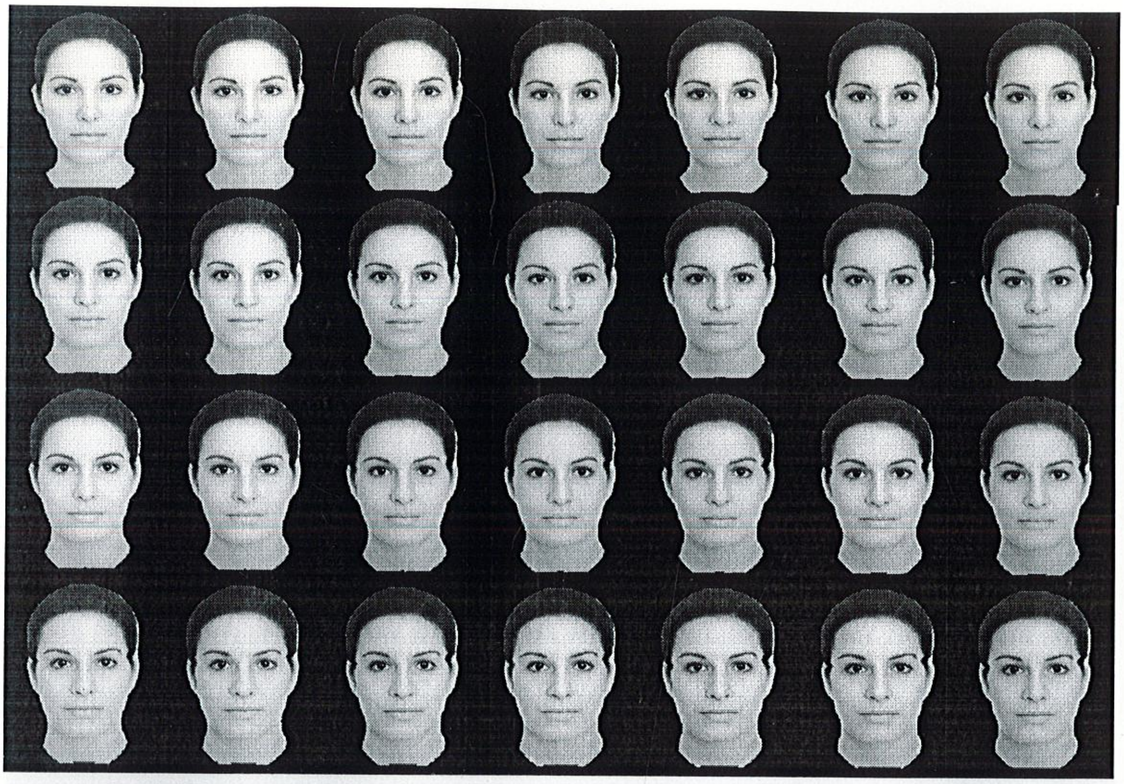


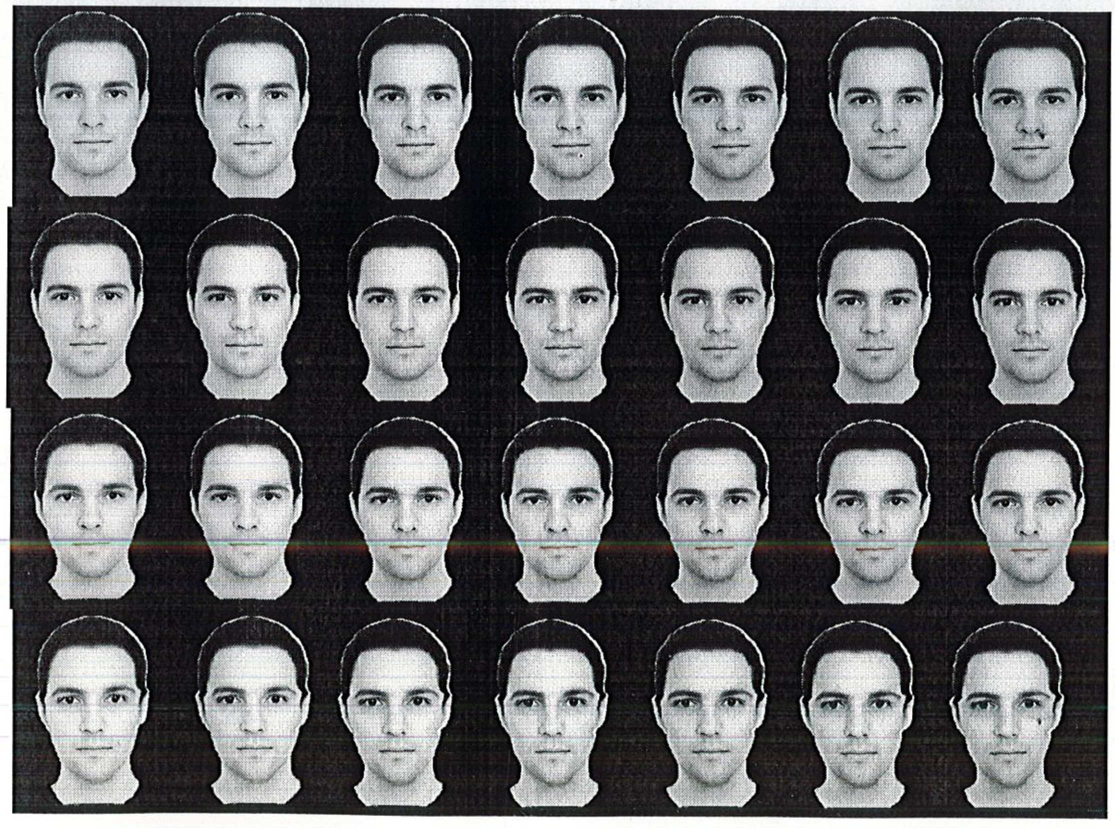
Extracted Features



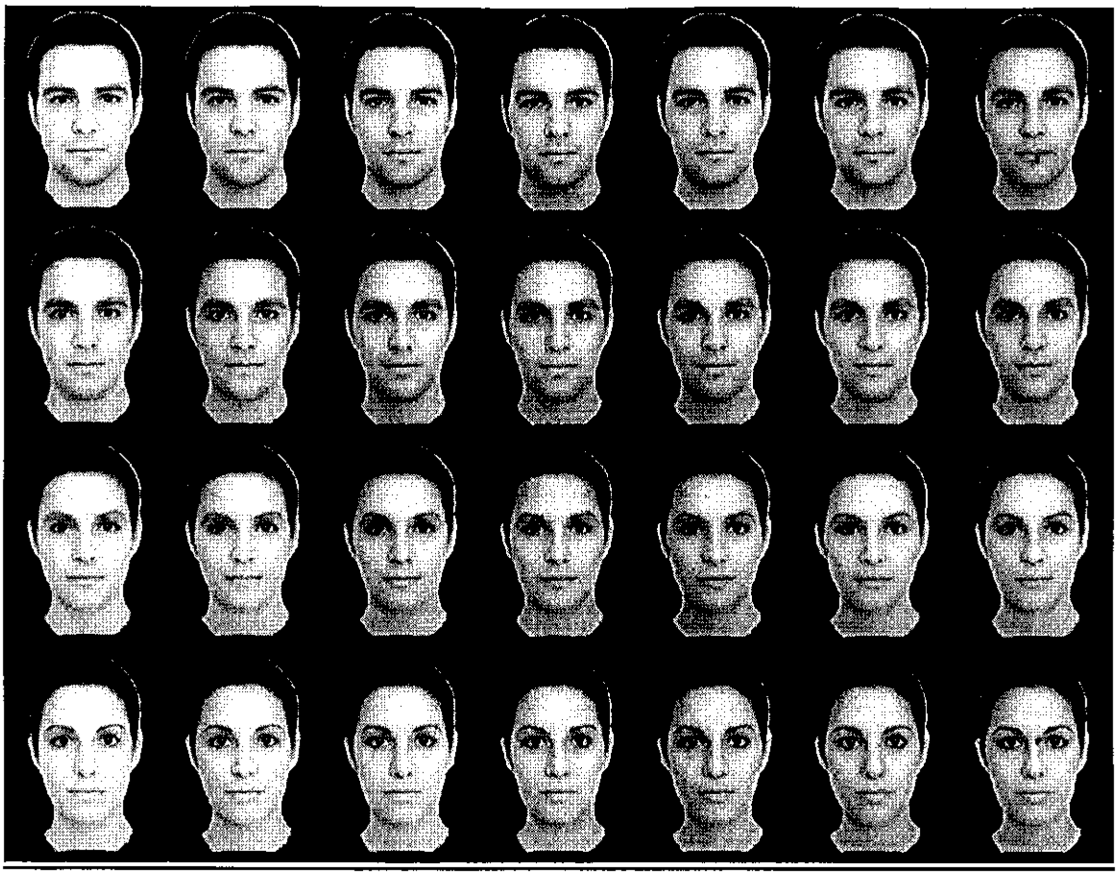

Skin Region
Eyes
Lips
Extracted Features

Warping



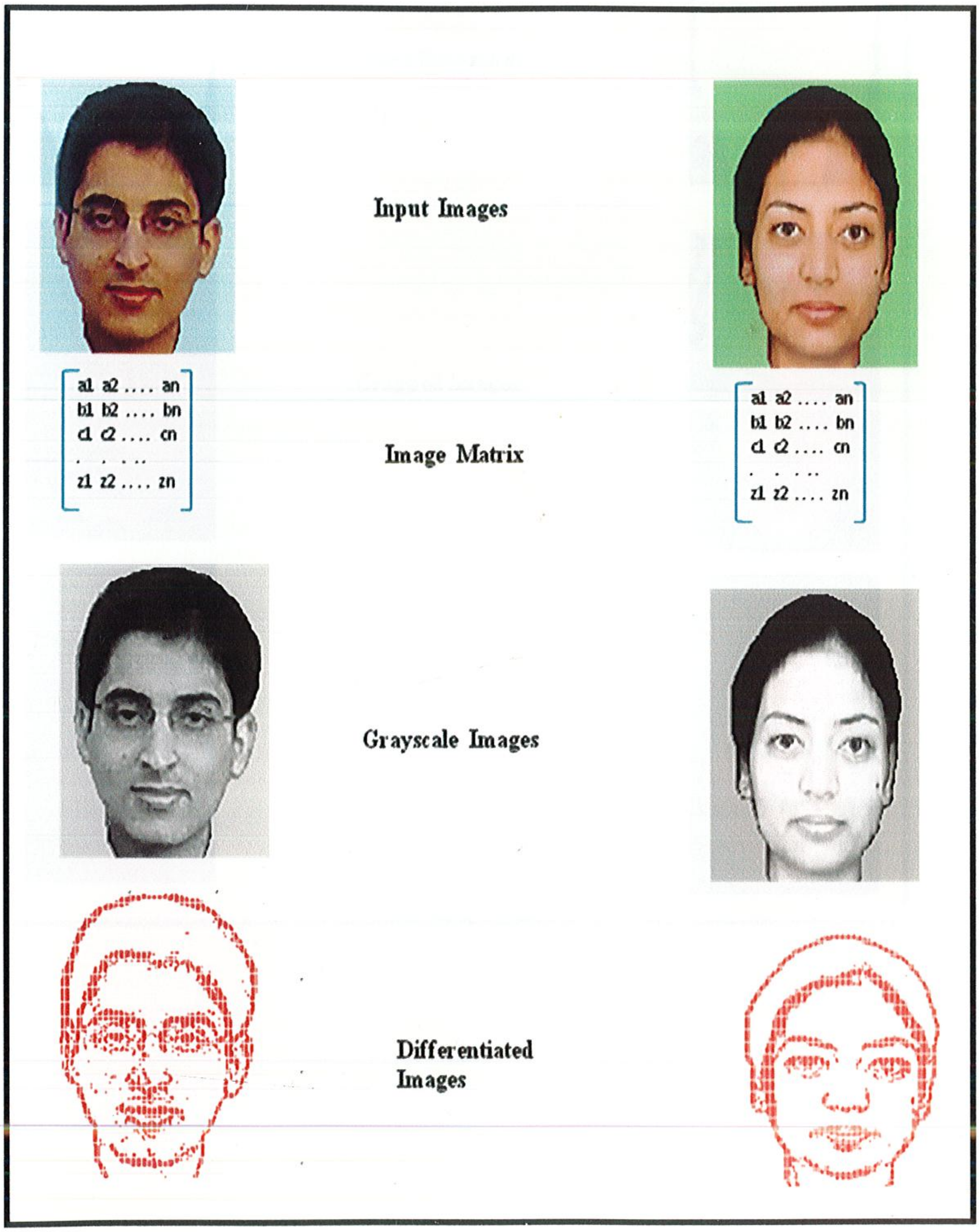


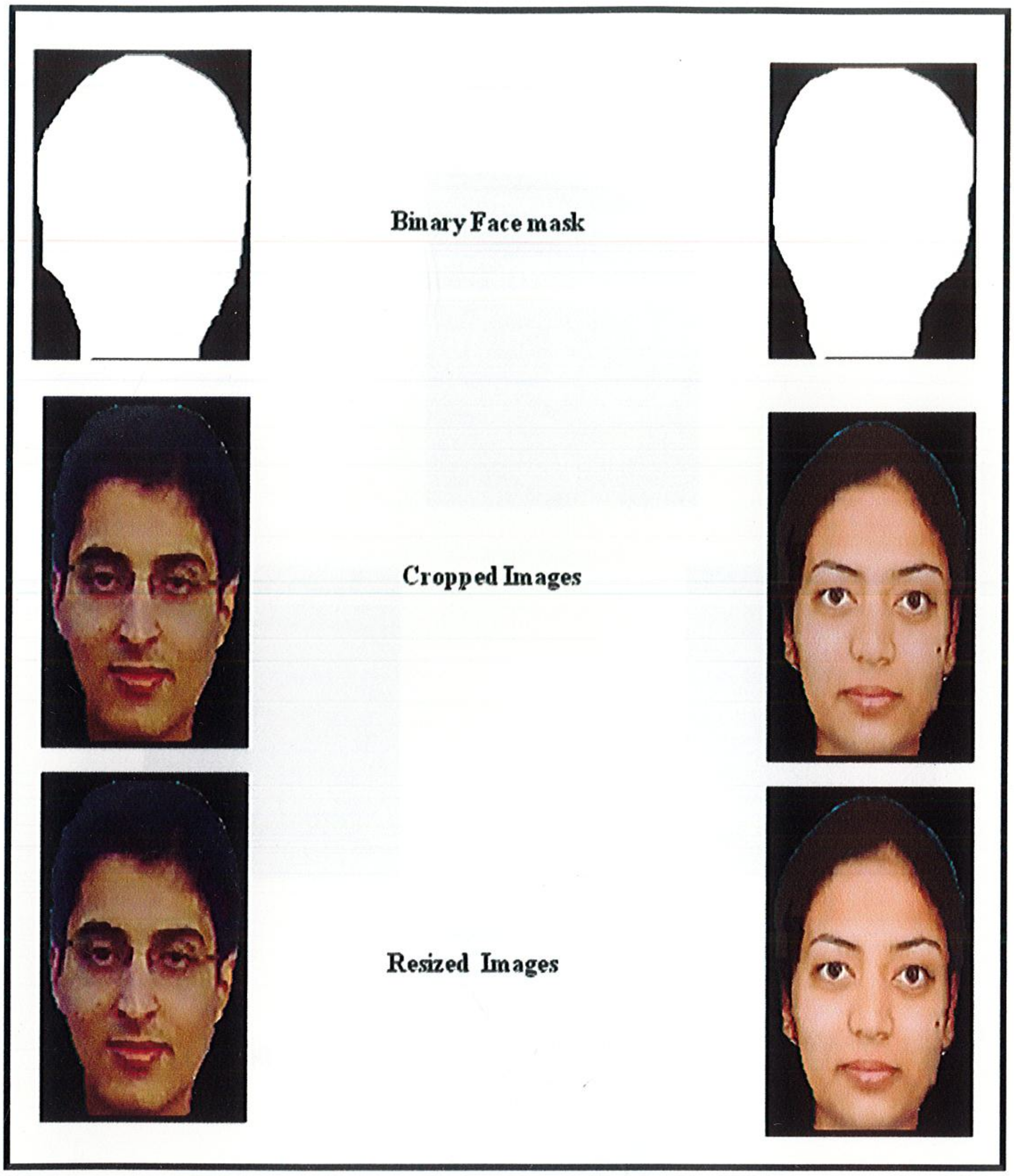
Warping and Crossfading



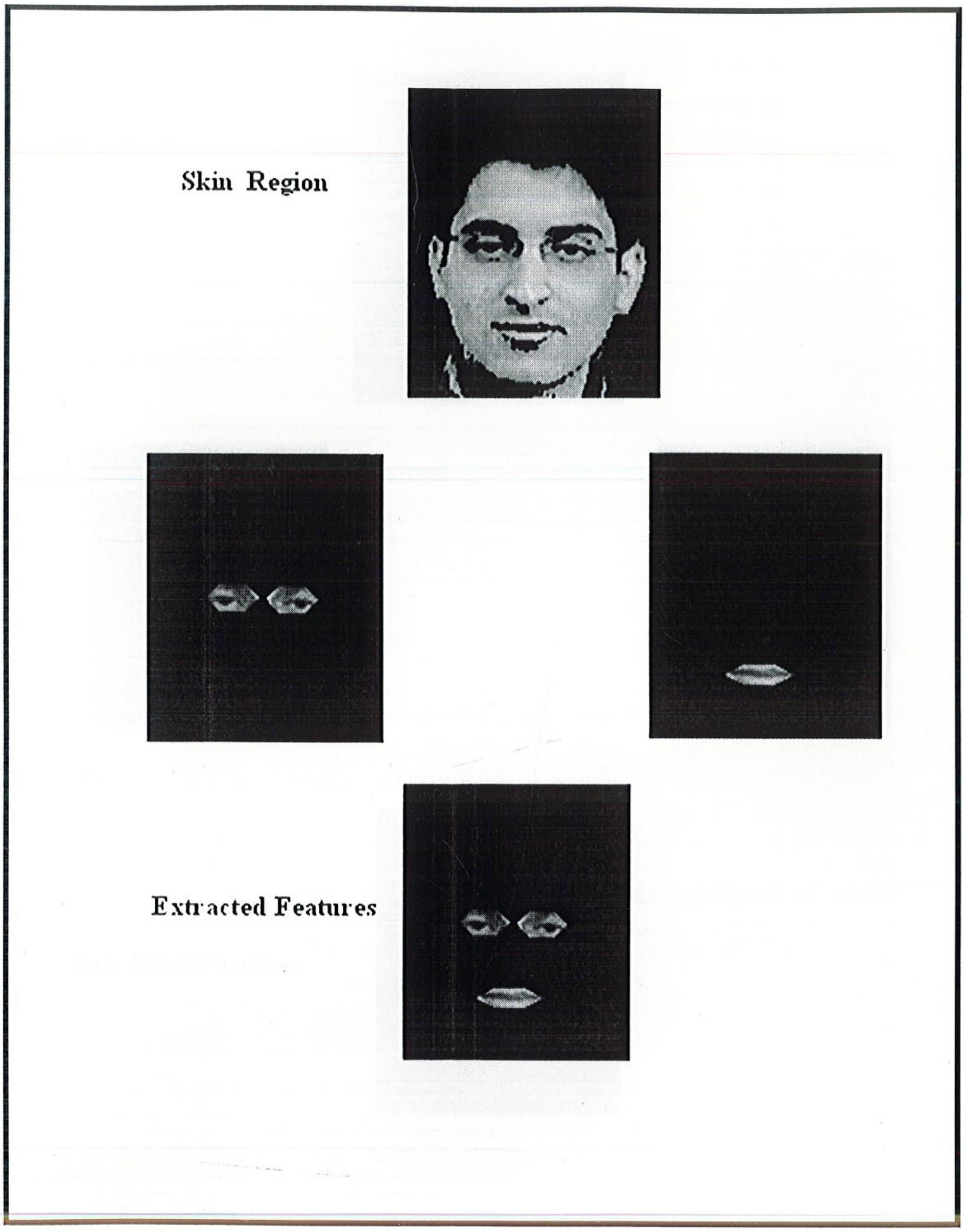
Example 2:

<u>Preprocessing</u>

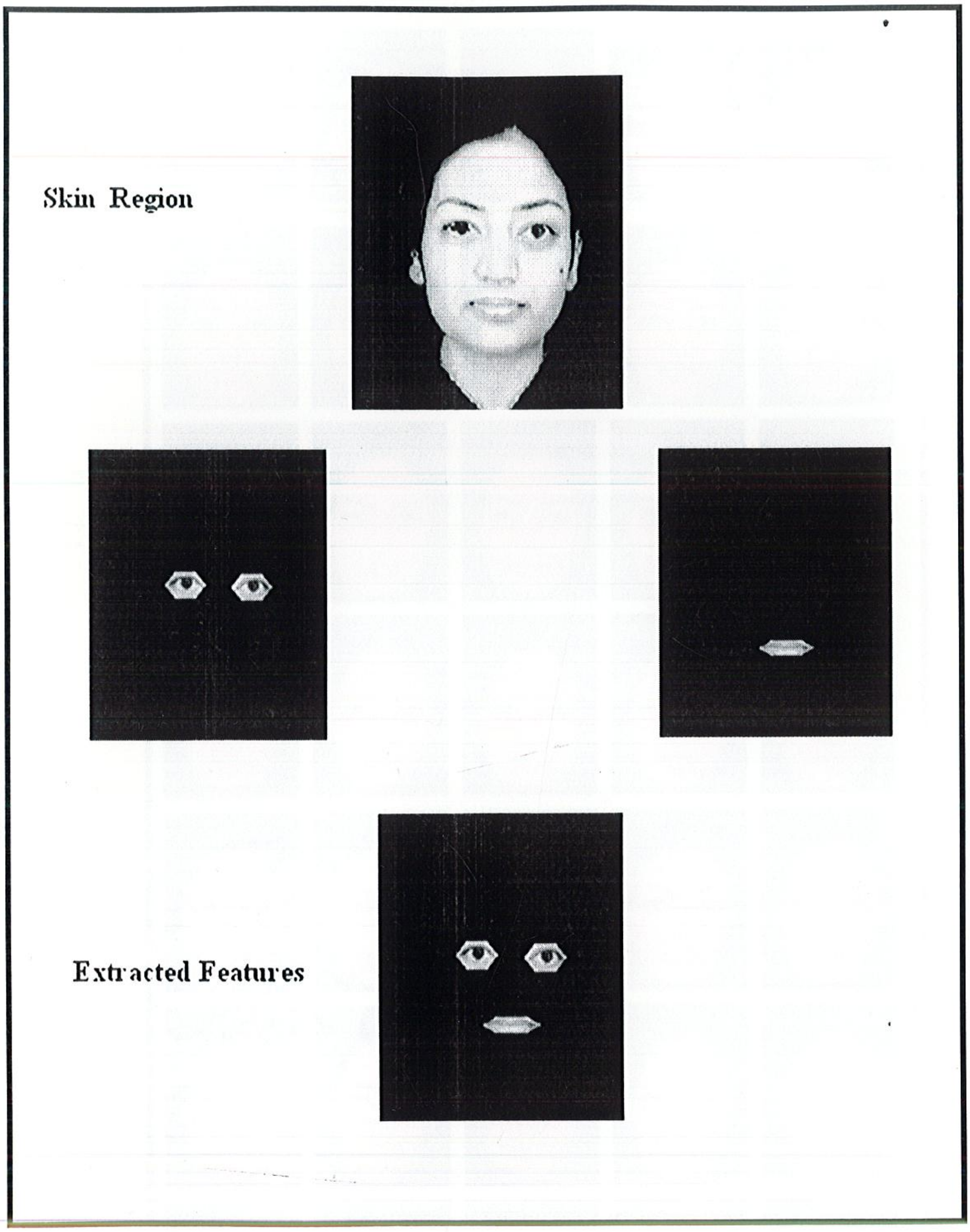



Binary Face mask
Cropped Images
Resized Images

Feature Detection

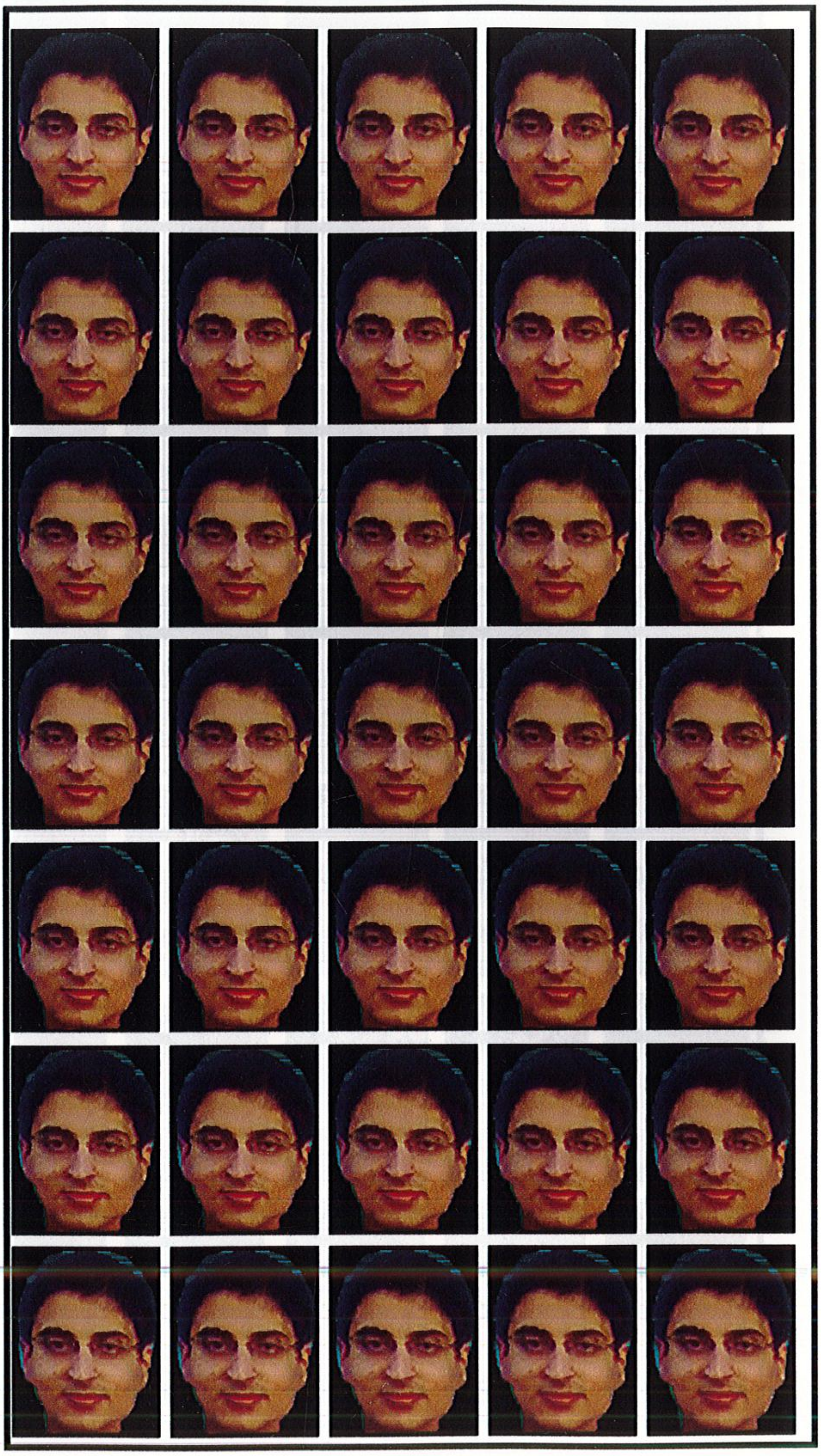


Feature detection

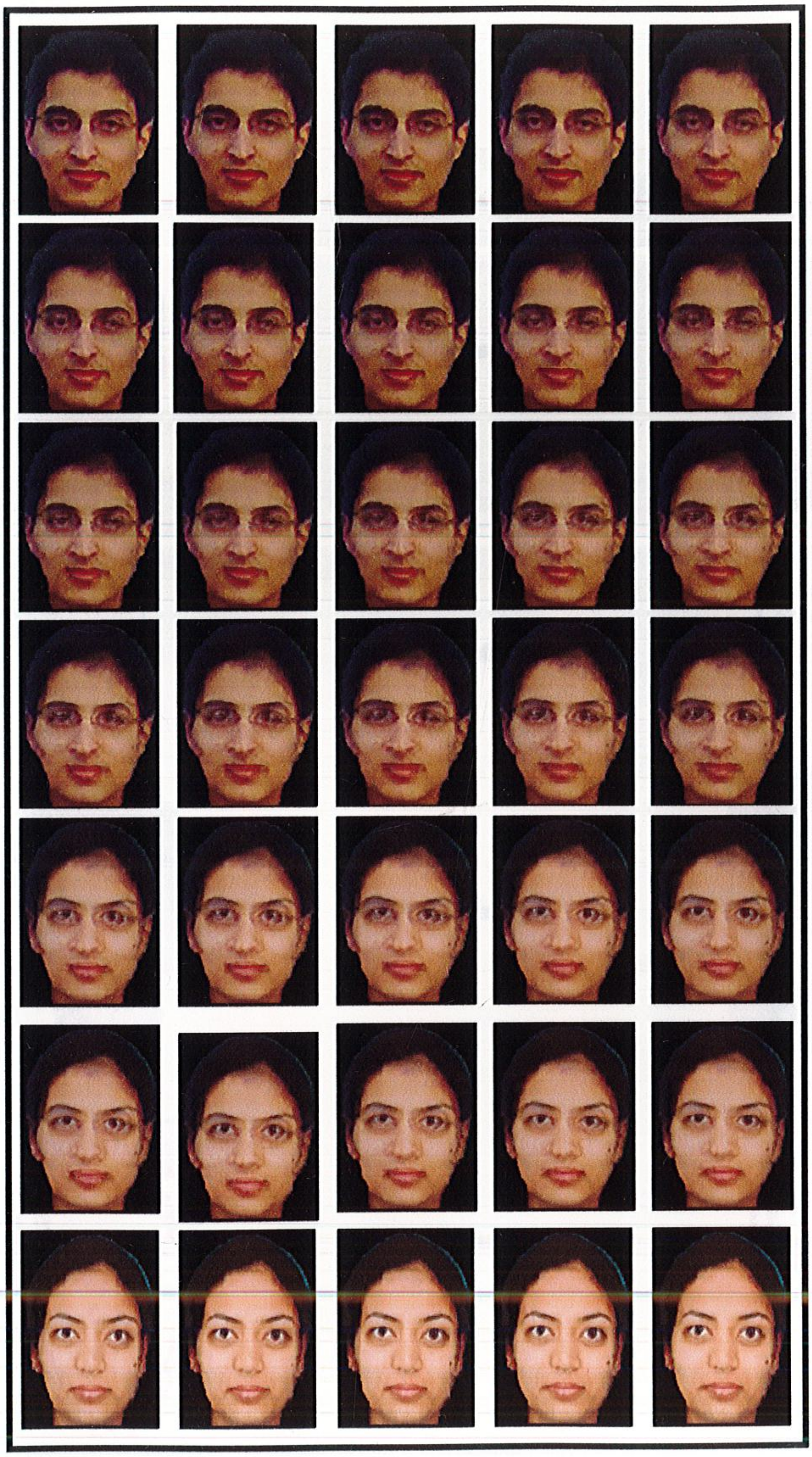


Warping

Warping

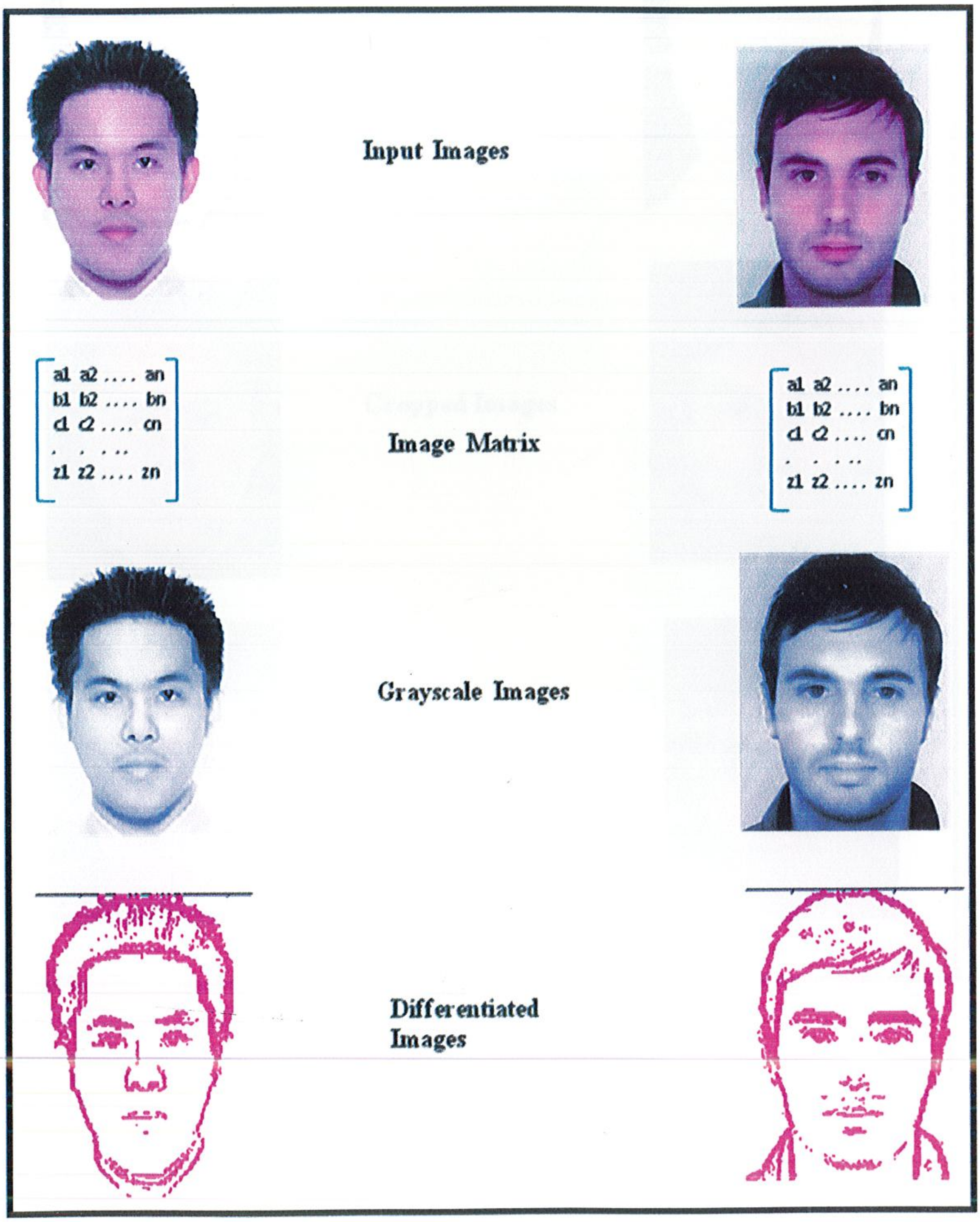


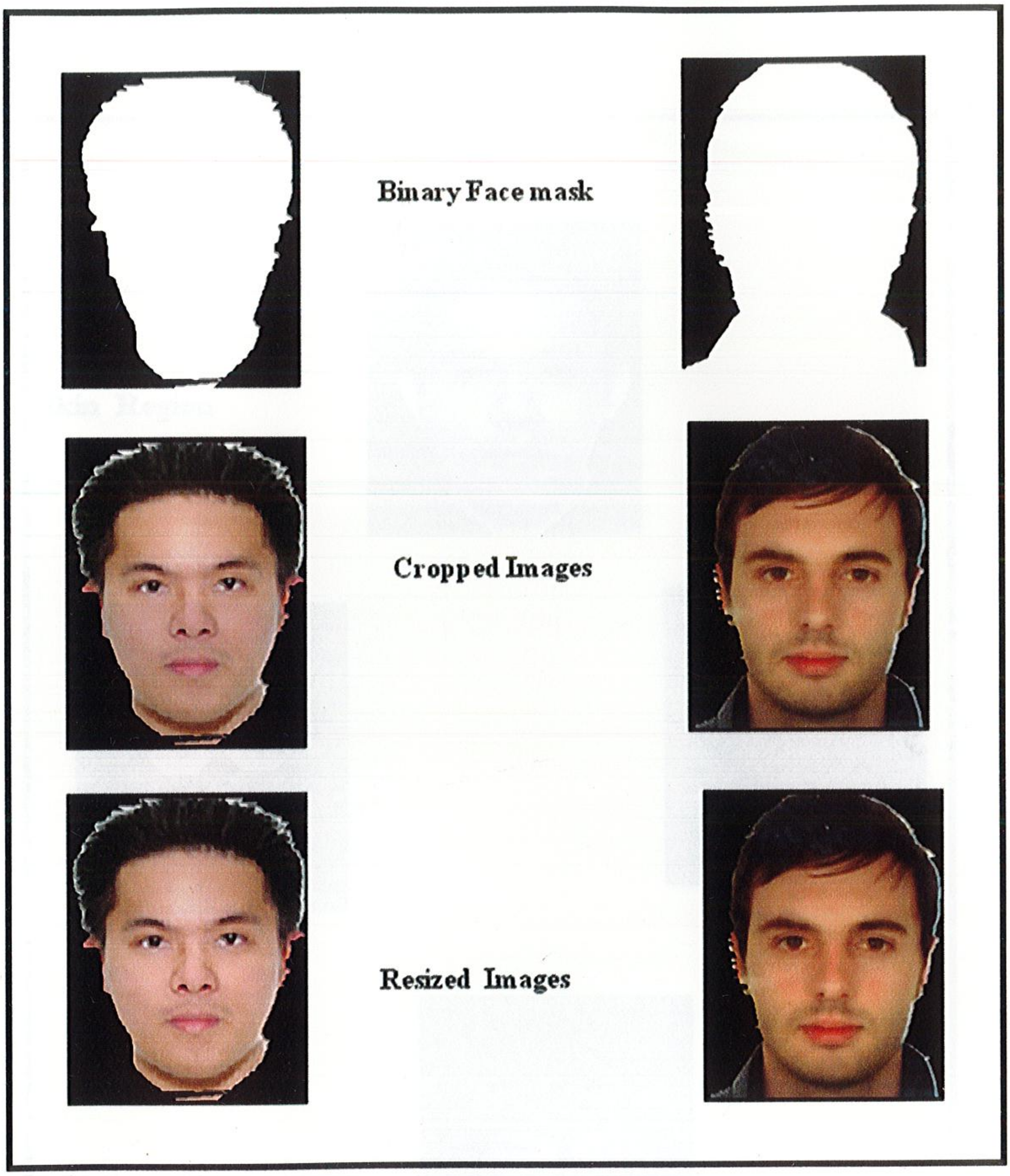
Warping and Crossfading



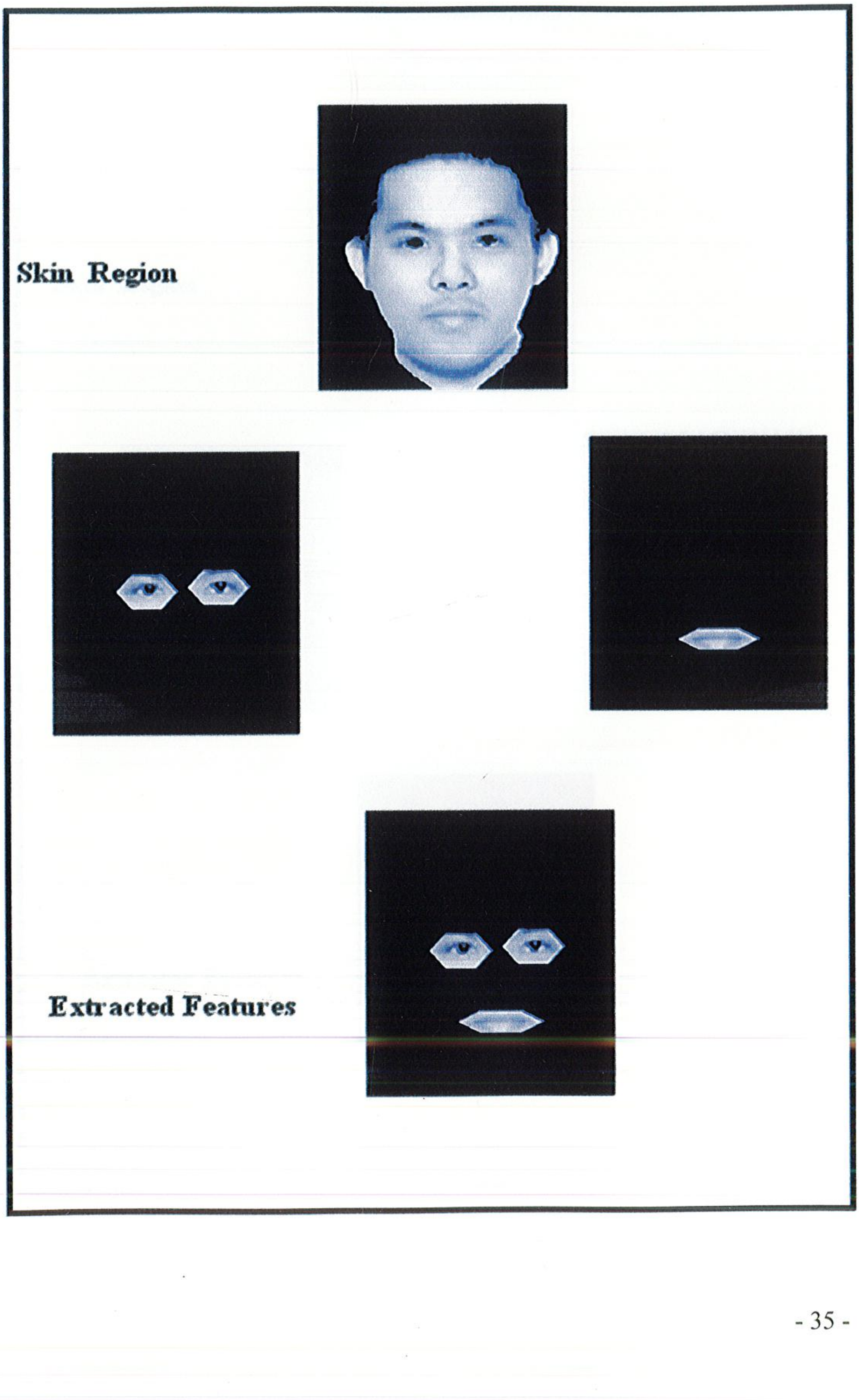
Example 3:

Preprocessing

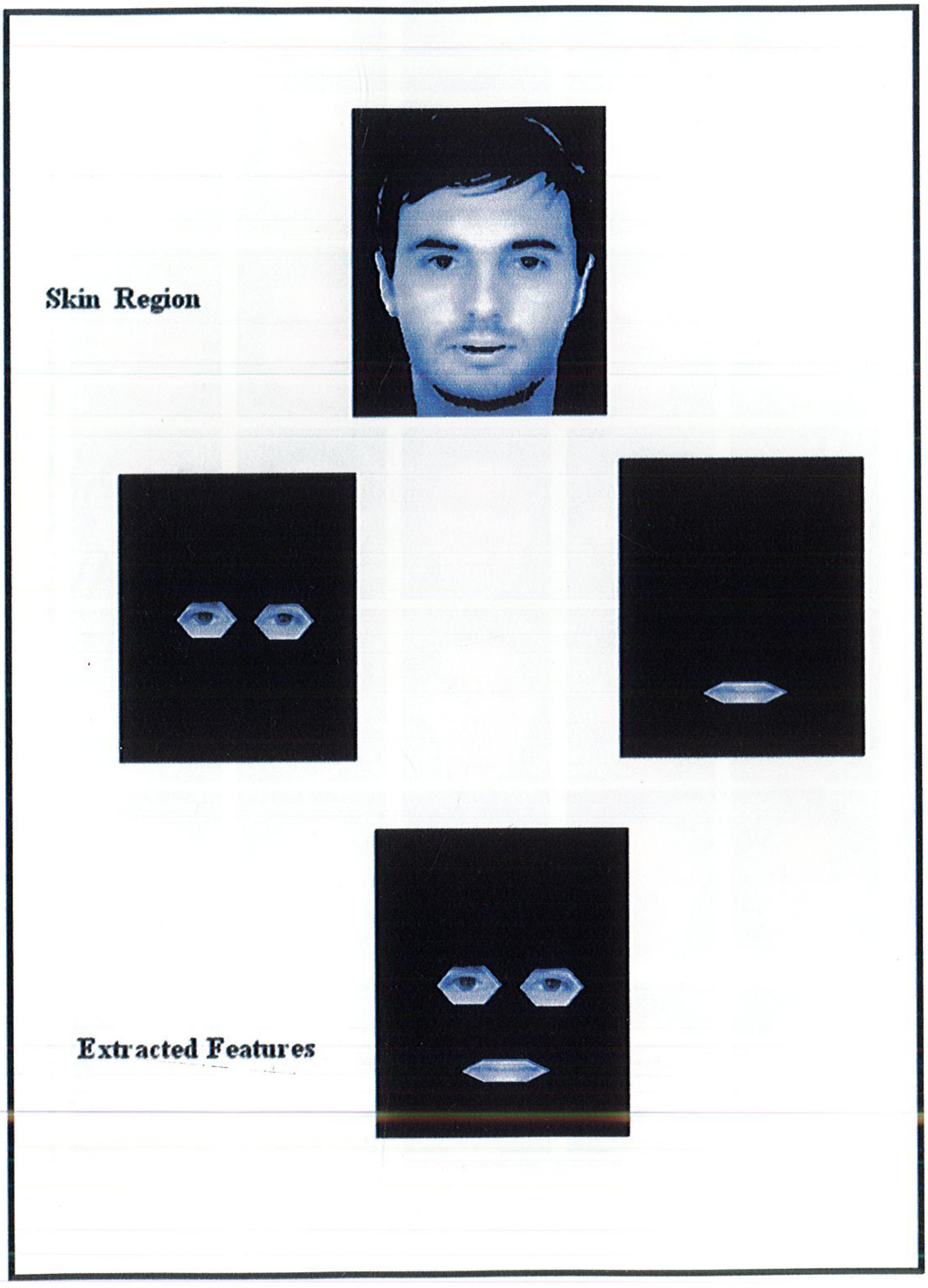



Binary Face mask
Cropped Images
Resized Images

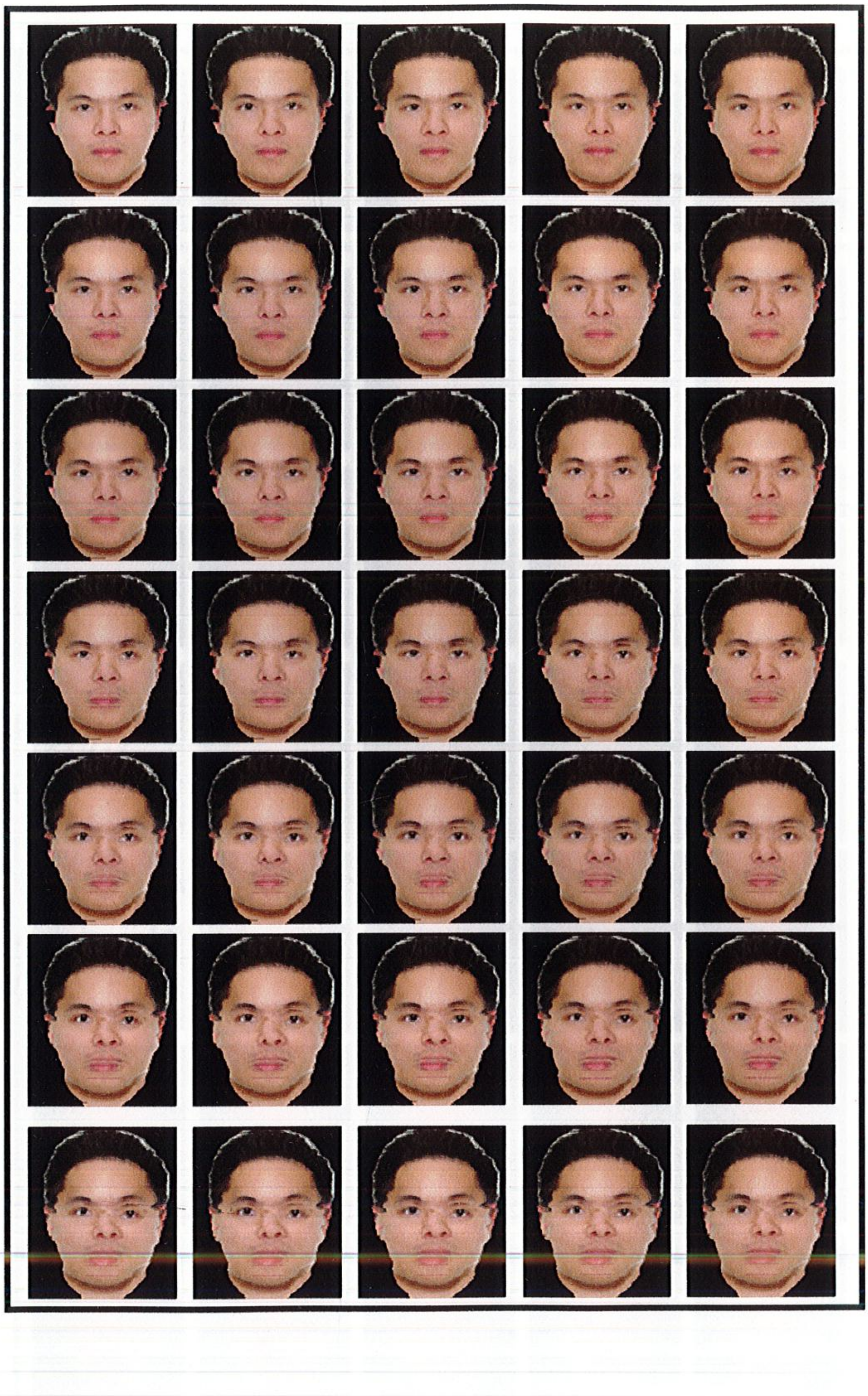
Feature detection



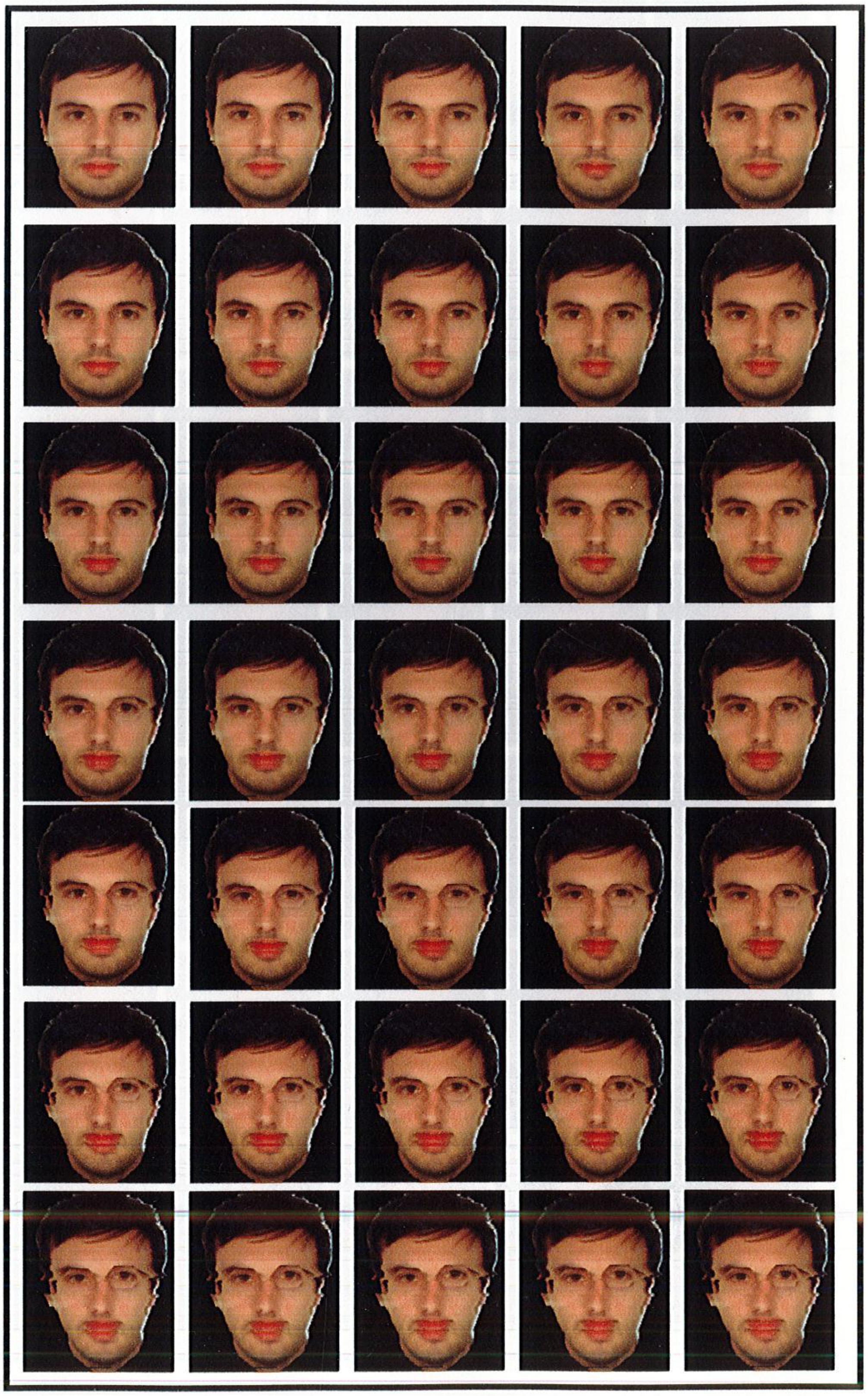
Feature detection



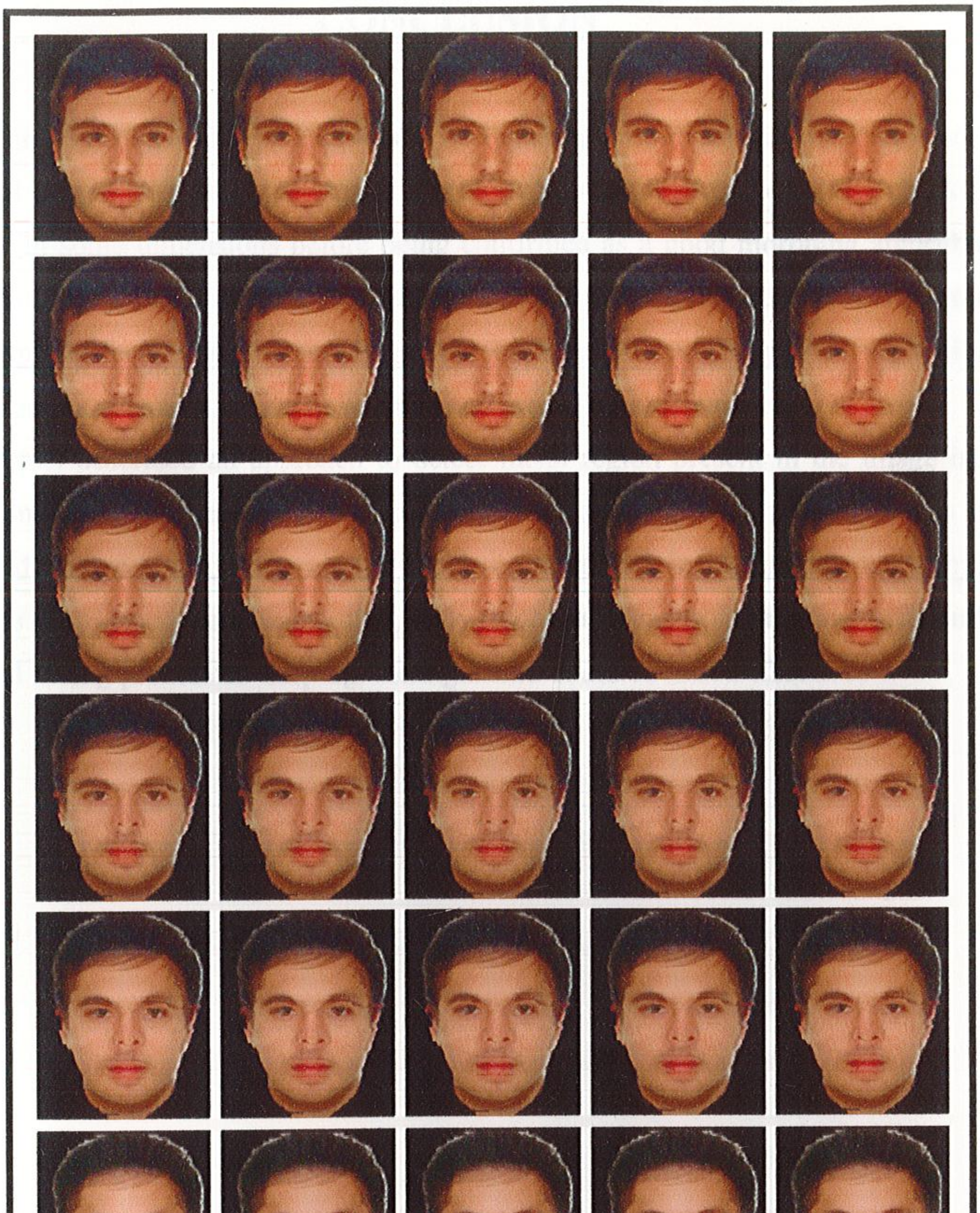
Warping



Warping



Warping and Crossfading



CONCLUSION

The main objective of our project was to design an algorithm for automatic morphing of facial images. We have seen that a morphing algorithm should carefully check its validity in some dimensions before being confirmed as a good morphing algorithm. The following parameters need to be considered while designing automatic algorithm for morphing facial images namely, proper background and foreground separation, feature points detection, efficient warping and cross fading techniques.

The project described an approach to detect facial region present in the image in a better and efficient manner using skin segmentation method. The method has an added advantage of also being able to decide on the region inside the face where eyes and lips lie. Hence concluding that, only when all feature points are efficiently matched during process of warping, the algorithm could be called effective.

FUTURE EXTENSION

We contemplate following future features which can be incorporated into this project:-

- **Dealing with various image formats:**

The implemented algorithm in this report incorporates morphing of only bit map images, so the other type of images can also be used like jpeg, gif etc. and algorithm should be flexible enough to handle the various formats.

- **Improvement in the morphing algorithm:**

This algorithm can be extended to deal with facial images with non-uniform illumination and different face orientation in addition to frontal facial images. Non-uniform background removal algorithm can be added to increase generality.

BIBLIOGRAPHY

[1] Beier, T., Neely, S., Feature-Based Image Metamorphosis. In "Proc. SIGGRAPH'92 "(Chicago, July 26-31, 1992). Published as "Computer Graphics", 26(2) (July 1992), pp. 35-42.

[2] Autocorrespondence: Feature-based Match Estimation and Image Metamorphosis. Published in the Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, October 22-25, 1995. Vancouver, Canada.

[3] Kumar, R.T., Raja, S.K. & Ramakrishnan, A.G. 2002, "Eye detection using color cues and projection functions", *Proceedings of International Conference on Image Processing*, vol. 3.

[4] J. Canny A Computational Approach to Edge Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, Nov. 1986.

[5] Digital Image Processing, R. C. Gonzalez and R. E. Woods, Addison-Wesley, 1992.