A Facial Caricature Generation System using Adaptive Thresholding

Upasna Dal
MSCI,
Mumbai, India
upasnadal@yahoo.com

Divyata Dal
Invenio Business Solutions,
Mumbai, India
divyata.dal@gmail.com

Siby Abraham
Department of Mathematics and Statistics,
Guru Nanak Khalsa College, University of Mumbai
Mumbai, India
sibyam@gmail.com

Automatic Facial Caricature Generation involves extracting the feature points and emphasizing the distinctive features of a particular face. The input face is compared to an 'Average Face' based on their respective facial distance parameters. The deviations are then scaled by an 'Exaggeration Rate', thereby elevating the peculiarity of the input face. A novel approach of Adaptive Thresholding has been used for the extraction of feature points from the input face to manage the non-uniform illuminations of the input face.

Keywords: Facial Caricature, Adaptive Thresholding, Average Face, Facial Feature Distances

I. INTRODUCTION

A Caricature is either a portrait that emphasizes the distinctive features of a particular face to create an easily identifiable visual likeness, or in literature, a description of a person using exaggeration of some characteristics and oversimplification of others. The term 'caricature', originally derived from the Italian word 'caricare', means to charge or load. Caricature is therefore, a 'loaded portrait'. The main intent of a caricature is to convey humor, widely used to entertain people through media viz. newspapers, internet, cartoon strips, etc [1].

Caricature is formalized as the process of exaggerating the differences from a mean face [2]. Psychologists suggested that human beings have an 'average face' recorded in their brain, which is an average of faces they encounter in life. A caricaturist compares one's face with this average face and draws caricatures by exaggerating the distinctive facial features. Therefore, a caricature is basically an exaggeration of the facial features, with the "norm" or "average" face as the reference point, which may represent an ideal face or a statistical average.

In the proposed work, we use a host of Digital Image Processing techniques to enhance the input image, threshold the image, fetch the facial coordinates and lastly render the image using Computer Graphics. The highlight of the work is the use of Adaptive Thresholding for the generation of the facial caricature.

The paper is organized as follows: Section II discusses the related works. Section III refers to the acquisition and preprocessing of the input image. The face localization and facial feature extraction techniques are elaborated in Section IV. Section V deals with finding the average face features, whereas Section VI illustrates how the facial features are exaggerated. Section VII exhibits the sketching of the caricature using Computer Graphics. Lastly, Section VIII presents the concluding remarks and anticipated future work related to the current research.

II. RELATED WORKS

J.J.F. van der Pol and F.M.W. Kanters [3] discuss the automatic facial feature extraction using Simulated Annealing to find the global positions of the eyes and the mouth, thereby extracting the eyes and the eyelids and Sequential Simple Deformable Templates (SSDT) Algorithm to gain these features. Image warping and morphing techniques have been used by Lav Varshney [4] to generate the facial caricature. Burcu Kepenekci [5] and Michael Lyons, Andre Plante et al [6] in their papers use Gabor Wavelet Transform for Face Recognition and Avatar Generation respectively. A 2D Caricature face method is implemented by Tun-Wen Pai and Chih-Yao Yang [7] in their work using facial feature extraction and matching techniques in transformed feature space. Pei-Ying Chiang, Wen-Hung Liao and Tsai-Yen Li [8] have combined effective feature analysis along with a simple image warping process for automatic caricature generation. Various methodologies have been used for feature extraction like SSDT algorithm, Gabor Wavelets transform, Simulated Annealing, etc. In our work, we have used a novel approach of Adaptive Thresholding for feature extraction to generate a Caricature.

III. IMAGE ACQUISITION AND ENHANCEMENT

Before we embark on creating a caricature, certain preliminary prerequisites must be met and preprocessing be performed on the input image for better results. We realize these using the following steps:

A. Image Acquisition

Image Acquisition involves acquiring the input image to be processed. Certain assumptions with respect to the input image are required to be fulfilled for best results. The project assumes a well-illuminated background for the input image.

978-1-4673-0126-8/11/$26.00 ©2011 IEEE

Also, the image of the face should ideally have a resolution of 180 x 200, since it affects the efficiency of the feature extraction process. The input image must be in JPEG format.

B. Preprocessing the Input Image

It is mandatory that the input image be preprocessed before any operations are carried out on it. The image is enhanced using various digital image processing techniques.

- To begin with, the input image is converted to grayscale. This is done because our feature extraction algorithm, like most others, operates only on grayscale data [Figure 1].

Figure 1. Converting input image to grayscale

- In order to obtain best results for feature extraction, it is desirable that the facial image occupies a standard intensity distribution. This is achieved using ‘Intensity Normalization’. We have ensured this by using Contrast Stretching, a linear transformation function which stretches the gray levels of the image to occupy the entire spectrum of 256 levels.
- The input face is then de-noised to remove any random noise from the image. For this purpose, Median filters are used since they provide excellent noise-reduction capabilities, with considerably less blurring than linear smoothing filters of similar size. Median Filters are particularly effective in the presence of impulse noise, also called salt-and-pepper noise because of its appearance as white and black dots superimposed on an image. Therefore, to remove the noise better, we first add salt and pepper noise to the image, then apply median filtering [Figure 2].

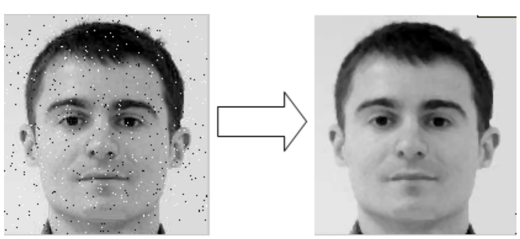

Figure 2. Removing the salt and pepper noise from image using Median Filter

IV. Face Localization And Facial Feature Extraction

For caricature generation, it is necessary to determine the individualistic features of a person. For this, the features of the face must be first extracted, given the frontal photograph of the subject. The output of this stage is a set of distances pertaining to the various features of the face.

The steps to the proposed approach for the process of Feature Extraction are illustrated in the flowchart [Figure 3].

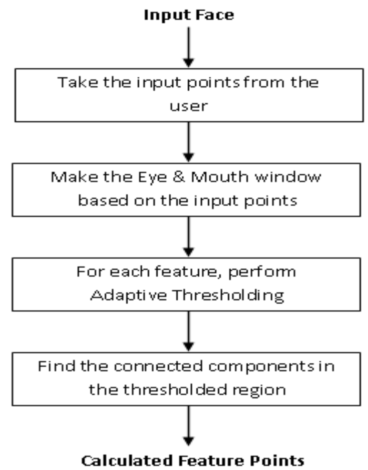

Figure 3. Steps for Facial Feature Extraction

A. Taking the input points from the user

The process of Feature Extraction primarily requires a method to detect the face i.e. we need to extract the facial region from the background. The facial region is then broken into its constituent features.

For this task of identifying the facial region, we ask the user to click on 4 predefined points which are used to determine the Eye and Mouth Windows. These windows assist in the extraction of features viz. eyebrows, ear, hair, chin, etc from the input image. These predefined points are listed as follows:

- Pupil of the Left Eye
- Pupil of the Right Eye
- Tip of the Nose
- Lowest Point of the Lip

B. Make the Eye and Month Windows Based On Input Points

The points obtained from the above step are used to build an Eye Window, comprising of the eyes and eyebrows; and a Mouth Window, containing the lips and the nose region [9].

In order to isolate the Eye Window, we have standardized the width and the height of the Eye Window. The width of the window is given in equation (1).

$$(width)_{eye} = \left[\frac{R_x - L_x}{2} \right] . 3 \quad (1)$$

Where, R_x = x-coordinate of the pupil of the right eye
L_x = x-coordinate of the pupil of the left eye

Similarly, the height of the eye window is given by equation (2):

$$(height)_{eye} \quad = \quad \left\lceil \frac{R_x - L_x}{2} \right\rceil \quad . \; 2.5 \qquad (2)$$

where, all the symbols hold the same meaning.

The localization of the mouth window starts from where the eye window ends. The width of this window is given by equation (3):

$$(width)_{mouth} \quad = \quad (\, R_x - 2 \,.\, L_x) \quad . \; 15 \qquad (3)$$

The height of the mouth window extends from the end of the eye window to the (lip point + 10).

After experimental analysis, the constants for the width and height of the eye window have been fixed to 3 and 2.5 respectively. Similarly, constants for the width and height of the mouth window have been fixed to 15 and 10 respectively.

These results hold true for all the faces input to the system. A sample result is as shown in [Figure 4]:

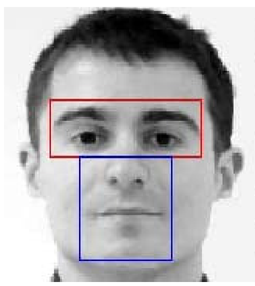

Figure 4. Eye and Mouth Windows

C. *Perform Adaptive Thresholding*

Thresholding enjoys a central position in the applications of image segmentation. Adaptive Thresholding is most useful when the imaging factors consist of uneven illumination or a histogram that cannot be partitioned effectively by a single threshold.

Adaptive Thresholding divides the original image into sub-images and then utilizes a different threshold to segment each sub-image. There are 2 important concerns in this approach, namely:

- How to sub-divide the image?
- How to estimate the threshold for each resulting sub-image?

Since the threshold used for each pixel depends on its location in the sub-image, this type of thresholding is termed as 'adaptive'. Adaptive Thresholding is carried out by following the below listed steps:

- Performing 2D filtering operations using an Average Filter of varying window size 'W'.
- Subtracting the original image and a factor 'C' from the filtered image.
- Erosion and morphing to smoothen the resulting image, thereby enhancing the feature extraction process.

The values of 'W' and 'C' are determined after extensive experimental analysis and vary for the various facial features. The constants for each of the features are summarized in the table (I) given below:

TABLE I. ADAPTIVE THRESHOLDING FACTOR 'C'

Adaptive Thresholding Factors 'W' and 'C'		
Facial Feature	***Value of 'W'***	***Value of 'C'***
Eye Mask	(Average threshold of Eye Mask) * 10 * 1.75	(Average threshold of Eye Mask) / 10
Mouth Mask	(Average threshold of Eye Mask) * 10 * 4.5	(Average threshold of Eye Mask) / 50
Nose Mask	(Average threshold of Eye Mask) * 10 * 2	(Average threshold of Eye Mask) / 150
Ear Mask	(Average threshold of Ear Mask) / 15	0.015
Chin Mask	(Average threshold of Face Mask) * 10 * 3	(Average threshold of Face Mask) / 10
Hair Mask	(Average threshold of Face Mask) * 10 * 3	(Average threshold of Face Mask) / 10

Adaptive Thresholding is now performed and the following facial feature points are extracted:

- Eye Localization
- Eyebrow Localization
- Eye Contour Localization
- Lip Points Localization
- Nose Localization
- Ear Localization
- Chin Localization
- Outer Hair Localization
- Inner Hair Localization

For example, the eye window thresholding is depicted in Figure 5:

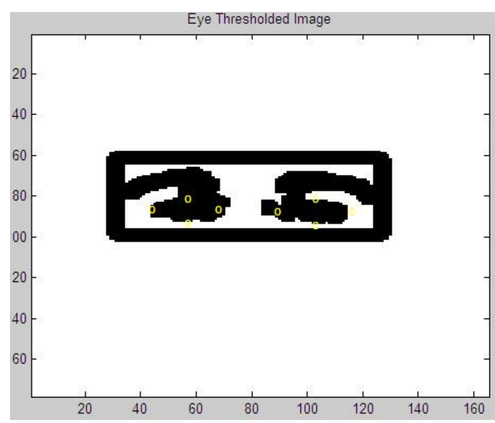

Figure 5. Eye Contour Localization

D. *Find the Connected Components in the Thresholded Region*

The next task would be to find the connected regions of the areas where the threshold has resulted. This has been achieved using 4-connected pixel connectivity [Figure 6]. In this algorithm, we say 2 pixels are connected if their edges touch, that is, they must be connected either vertically or horizontally.

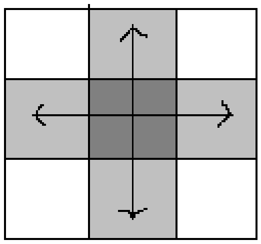

Figure 6. 4-Connected Pixel Connectivity

In addition to this, we need to locate the outlines of features like the lips, hair contour, etc. This is accomplished using the Boundary Extraction technique; an example of this for the hair contour is illustrated in [Figure 7].

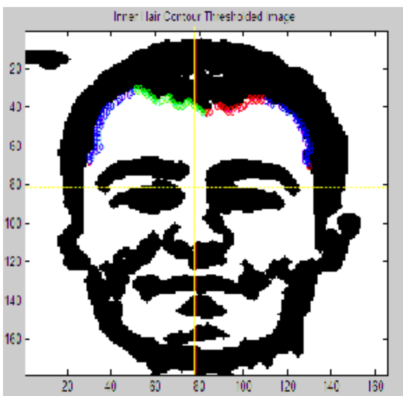

Figure 7. Hair Contour Boundary Extraction

A combination of the above techniques gives the final feature points of the input face. These feature points are also called Feature Vectors which comprise of the coordinates of the feature points [Figure 8].

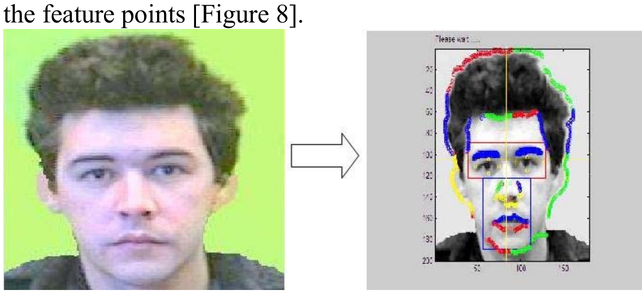

Figure 8. Facial Feature Extraction

V. FINDING THE AVERAGE FACE FEATURES

The average face concept previously discussed is widely used in the generation of caricatures. It mimics the caricaturists' drawing approach of comparing the subject's face with the average face and exaggerating/deviating the distinct facial features.

We have calculated the average male and female face using 50 different faces of each. Some of the faces used are as shown in [Figure 9] and [Figure 10] respectively.

Figure 9. A Part of the Male Face Database

Figure 10. A Part of the Female Face Database

The average face is calculated using Computer Graphics methods developed by the Face Research Organization, U.K. [10]. The Face Research Group performs research on various aspects like finding average faces, obtaining the child faces from an adult's face, etc. We have calculated the average male and female face coordinates from the average faces illustrated in [Figure 11].

Figure 11. Average Male and Female Face

VI. EXAGGERATING THE FACIAL FEATURES

We now know that a caricature not only eliminates useless information (which may only serve to hinder the recognition process), but also amplifies the individuating features thereby aiding recognition. So, the caricature appears more 'like' the face, than the face itself.

The exaggeration of the input face is carried out in the following order:

- Computation of the distances of the input face
- Computation of the distances of the average face
- Calculation of the differences between the distances of the input face and the average face
- Exaggeration of the distances

A. Computing the Exaggeration Distances

The coordinates of the various facial points obtained from the feature extraction phase are converted into distances which can then be compared against the average face. These distances are obtained by studying the caricaturists' work, that is, the areas the caricaturist stresses upon while drawing a caricature. Based on this, we define 6 distances which in turn are used for exaggeration of the facial features [9] [Figure 12] [Table II].

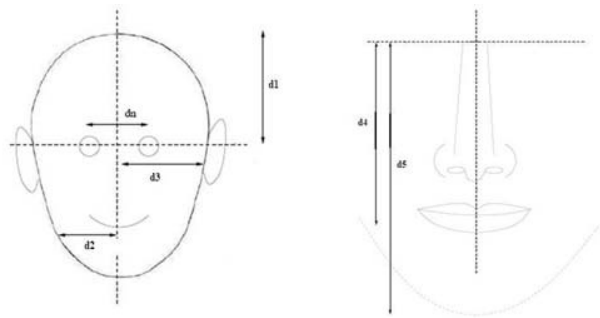

Figure 12. The Facial Distances

TABLE II. FACIAL DISTANCES

Facial Distances	
Feature Code	***Feature Description***
d1	Distance of Eye Line From the Top of the Head
d2	Distance of Face side from Central Axis at the Lip Line
d3	Distance of Face side from Central Axis at the Eye Line
d4	Distance of Bottom End of Lower Lip from Eye Line
d5	Distance of Chin from Eye Line
dn	Inter – Iris Distance

B. Exaggerating the differences between the Input and Average Facial Distances

The facial distances thereby obtained are compared to those belonging to the average face. There are different perceptions regarding the ideal face which vary with time, race and culture.

The difference between the facial parameters of the subject and the norm is exaggerated by a factor, b, to yield the caricature.

The caricature formula [9] used in the proposed work is given by the formula (4)

$$Q_{2D} = P_{2D} + b * (P_{2D} – S_{2D}) \tag{4}$$

where,
P_{2D} represents the input image
Q_{2D} represents the generated caricature
b is the degree of exaggeration
S_{2D} represents the mean face
$P_{2D}, Q_{2D}, S_{2D} = (x_i, y_i), i = 1, 2, 3, \ldots N$
(x_i, y_i) represents the points on the contours that define the face

The process of caricature generation can be better explained with the help of [Figure 13] [9].

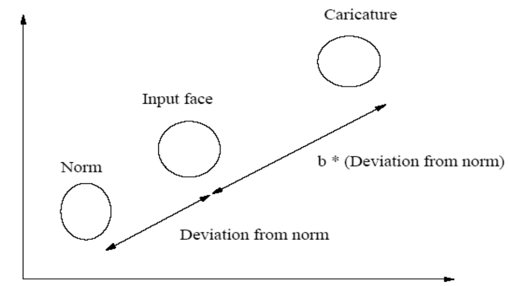

Figure 13. General Idea of Facial Caricature

The exaggeration rate is of prime importance while determining the deviation of the features from the "norm" face. In our work, the exaggeration rate depends on the differences between the distances. If the difference is negative, that is, the input distance is less than the average distance, then we decrease that distance to individuate that feature. Similarly, if the distance is positive, that is, the input distance is more than the average distance, then we increase that distance to bring out the distinguishing features of the face. To summarize, we exaggerate by a factor that depends on the differences [11].

This exaggeration process is carried out one feature at a time. For example, exaggerating the facial distance d_2 affects the chin, face as well as the hair contour.

The most important point to keep in mind while carrying out the exaggeration is to maintain the inter-iris distance d_n which plays a critical role in identifying a person. If this distance is modified, the caricature would fail to look similar to the input face.

After exaggerating all the facial distances, the deviations from the input face can be distinctly seen in [Figure 14].

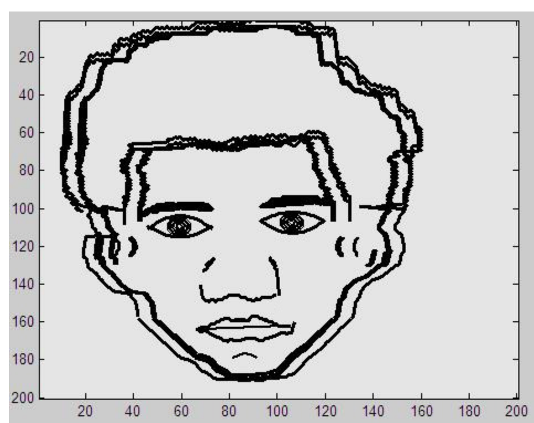

Figure 14. Exaggerated Feature Points

VII. Sketching The Final Caricature

The input to this final step of Caricature Generation is the exaggerated feature points of the input image. Using these as input, the caricature is sketched using Computer Graphics techniques of translation, scaling, etc.

Since we exaggerate features individually, B-Spline curves are used to draw smooth curves joining the feature points viz. the chin, ears, hair, lips, etc. Spline Interpolation is used to sketch the intermediate points that remain undefined during the process of exaggeration.

The collective outcome of the digital image processing and computer graphics techniques can be merged with interesting backdrops to add more worth to the comic character [Figure 15].

Figure 15. Final Caricature

VIII. Conclusion

The primary objective of the proposed work is to bring out the uniqueness of the face. The sole intention of using a grayscale image to sketch the caricature is to imitate the caricaturists' style of emphasizing their pencil strokes. Adding color is a mere formality. The highlight of our work is the use of Adaptive Thresholding combined with the concept of facial distances which speeds up the generation of the caricature. Moreover, Adaptive Thresholding is a simple and reasonably accurate method of detecting and extracting features.

We have successfully accomplished the task of 2D Caricature Generation. We plan to extend our work to generating a 3D Facial Caricature that takes into consideration the 3rd dimension of the human face, that is, the depth which would result in a more realistic-looking caricature. We also intend to extend caricature generation from merely facial caricature to caricaturing the entire body. This, in turn, will add a professional touch to the art of drawing caricatures.

References

[1] Caricature, http://en.wikipedia.org/wiki/Caricature

[2] Susan E. Brennan, Caricature Generator: The Dynamic Exaggeration of Faces by Computer, Source: Leonardo, Vol. 18, No. 3, (1985), pp. 170-178, Published by: The MIT Press.

[3] J. J. F. van der Pol (433603) and F. M. W. Kanters (429434), Automatic Facial Feature Extraction using Simulated Annealing and Sequential Simple Deformable Templates, Tehcnical Report Eindhoven University of Technology, February 2002.

[4] Lav Varshney, Face Metamorphosis and Face Caricature: A User's Guide, School of Electrical and Computer Engineering, Cornell University, March 16, 2004.

[5] Burcu Kepenecki, Face Recognition Using Gabor Wavelet Transform, The Graduate School of Natural Sciences of The Middle East Technical University, September, 2001.

[6] Michael Lyons, Andre Plante , Sebastien Jehan, Seiki Inoue , and Shigeru Akamatsu, Avatar Creation using Automatic Face Recognition, Proceedings, ACM Multimedia 98, Bristol, England, Sept. 1998, pp. 427-434.

[7] Tun-Wen Pai and Chih-Yao Yang, Facial Model Analysis for 2D Caricature Generation Systems, Dept. of Computer Science, National Taiwan Ocean University.

[8] Pei-Ying Chiang, Wen-Hung Liao and Tsai-Yen Li, Automatic Caricature Generation by Analyzing Facial Features, National Chengchi University, Taipei, Taiwan.

[9] Partha Kulkarni, Interactive Caricature Generation Using Frontal Photographs, Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur, May, 2002.

[10] Face Research Database, U.K., http://www.faceresearch.org.

[11] Lin Liang, Hong Chen, Ying-Qing Xu, Heung-Yeung Shum, Example-based Caricature Generation with Exaggeration, Microsoft Research, Asia, Xi'an Jiatong University, China.

[12] Gonzalez, Woods, Digital Image Processing, 2nd Edition.

[13] L. C. Jain, U. Halici, Intelligent Biometric Techniques in Fingerprint and Face Recognition, CRC Press.