EasyChair Preprint

№ 8425

Caricature Generation Using Dobot Magician

Arpit Ojha

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

July 10, 2022

Caricature generation using Dobot Magician

Arpit Ojha

Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru – 560064, India

Abstract

A caricature is a piece of art depicting persons or sociological conditions in a non-verbal way. A caricature is usually created by the reference models of a physical human being. The deviations from the reference model are the characteristic feature of the caricature. Basically, a caricature is used to either insult or compliment a human being or to depict an event involving human beings pictorially. The aim of caricaturist visualization is an illustrative depiction of the characteristics of a given data by exaggerating deviations from the reference model. This paper proposes a model to automate the process of generating caricature based on input images. This model uses adaptive thresholding and SVM algorithm for the cause.

Keywords: Caricature; face-shape, GAN; Exaggeration; Image Processing; Gaussian-Filter; face-shape Bilateral Filtering, SVM, Adaptive Thresholding

1. Introduction

A caricature is a rendered image showing the features of its subject in a simplified or exaggerated way through sketching, pencil strokes, or through other artistic drawings [10].

A caricature can be defined as an exaggeration likeness of a person made by emphasizing all the features that make the person different from everyone else. The key is that an exaggeration is not a distortion. “Exaggeration is the overemphasis of truth. Distortion is a complete denial of truth” [11][13].

Fundamentally a caricature is usually drawn by the artists on drawing board based on physical attributes and to a certain extent behavior. There were many attempts to automate this process of generating caricatures based on images using the computer. The process of drawing the caricature manually involves the artist to be aware of the physical traits of the human subject’s face with the grand artistic mindset and visualization. To encompass all of these attributes of an artist along with his visualization and to come up with a model that gives a close approximation to the actual drawings of the artist is a daunting task.

A support vector machine is a machine learning algorithm that is used to model both classification and regression problems. Support vectors are nothing, but individual observation and a vector machine works on these support vectors to decide on the best way to classify the observations by drawing a hyperplane [13]. In order to identify the boundaries

and facial points of the pictures adaptive threshold paradigm can be used which manipulates the pixel values based on a threshold parameter into the foreground and background pixels [14].

The upcoming sections of this paper is organized as follows, section 2 discusses the related work with respect to the context of this paper, section 3 briefs the different components in the system, section 4 explains algorithm, the flowchart of the method is given in section 5, the results of the experiments are detailed in the

2. Related Work

There were many attempts to synthesize facial caricature in the past. Akleman et al. [15] developed a model that makes caricature based on the interactive morphing tool. Tominaga et. al. [16] build a template-based facial caricature system PICASSO and Web-PICASSO. But all these works don't model the artistic patterns and observations made by an actual artist for the automation models.

"Facial Caricature Drawing based on Subjective Impressions" [1] - a paper from Japan Linguistic expressions used to reflect personal impressions of a face which includes cheerfulness, gentleness and many more. The concept of exaggeration is considered by the line thickness and exaggeration rate. The validity of the process is very essential since it is very difficult to maintain objectivity in the arts. This distinction can be very well explained by the difference of opinion for Van Goh and Leonardo da Vinci. Emotional information lacks objectivity, universality which is very essential in scientific research work instead it exhibits vagueness, ambiguity and situational dependence. It not only deals with the facial caricature drawing with image processing but also introduces emotions into the output which must be perceived by others. Facial expressions are the only way to express emotions, which is a cognitive or a mental state, or a faculty within the brain, and is not amenable to direct physiological examination.

[2] Features of a face - eyes, nose, mouth, eyebrows, ears, hair profile and position, each feature has its own element of interest which introduces the uniqueness as per the index and extent of such parameter. Mathematical modeling of the facial features has been discussed in this paper.

[3] Subjective evaluation and objective evaluation have been done which addresses the model verification as well as justification. Subjective evaluations have been done using a set of questions that led to the development of the correlation table and various modes were tested by analyzing the matching rate.

[4] "Simpson Style Caricature based on MLS (Moving Least Squares)" - a paper from Korea Moving Least Squares were deployed with exaggeration rules from caricature artist, Justin. Unique characteristics were detected using the Active Appearance Model.

The other horizon where the researchers worked on was building automated computers which generates the stylistic face sketch by observing the drawings of artists. The emphasis was concentrated on generating stylistics images rather than the caricature.

There are two broad approaches to synthesize a caricature using the computer, one relates to methods of creating regarding lightness and darkness changes in a real facial picture image. The caricature generation is done using Generative Adversarial Network (GAN) and later warping the input image and caricature generated image from the GAN called warpGAN. The purpose of these methods is to create a picture like caricatures, and it is difficult to incorporate facial feature emphasis procedures in these methods.

The other approach relates to methods of drawing a caricature over the result of extracting facial features from real picture images and emphasizing these features. Methods of emphasizing methods of obtaining differences of edge features between an object face and an average face and emphasizing these differences.

These methods have problems of pictures being destroyed with extreme exaggerations and a low degree of freedom associated with the shape and arrangement features being emphasized collectively, There are also methods of not only emphasizing feature extraction results but also replacing with illustrations closest to the extraction results selected from a database of illustrations of face parts obtained in advance and illustration suitable for the facial parts of an input facial image not necessarily being available in the database.

3. The System Framework

The framework for the testbed comprises of four stages as depicted in fig 1, they are,

- Feature Extraction stage
- Face Exaggeration stage
- Face Rendering stage
- Hair Segmentation and rendering stage

In the feature extraction stage, a set of predefined facial features are used to train a model which will act as an initial reference for the training. The features extracted by applying the SVM algorithm to the input image are tested with the trained model to obtain the face component features which will be exaggerated to obtain the caricatured picture. The system also models the head hair separately by applying the segmentation method and compared with the output of the trained model on hair structure. The face reference points are extracted out from the sample picture using the ASM algorithm which is followed by the Adaptive threshold algorithm to predict an outlined facet. This output cumulated with an exaggerated face along with hair structure is used to obtain the final caricature.

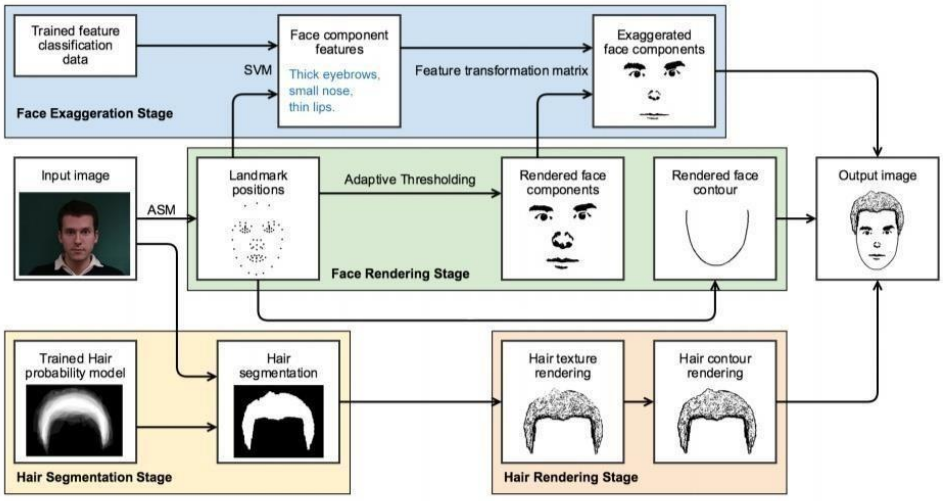

Figure 1: System Methodology

Our methodology starts with face images on which a collection of landmarks which are located. Our tests are with manual location; automatic location of landmarks is possible, most notably using active shape models or optic flow techniques but, we avoid the confusion which would arise if incorrectly located landmarks were used in subsequent coding. The restriction to having only one target in the gallery means we are not concerned with learning, based on a set of faces, leading to a sophisticated matching strategy; here, we concentrate purely on coding issues. Relationship Exaggeration

Most popular approaches exaggerate the difference from the mean, although the results of this approach have often been criticized. In fact, caricaturists tend to emphasize only one or two salient features in a caricature. The proportion description of the features is introduced Normalizing the proportion differences of a feature by using its mean is viewed as an expression of the feature distinctiveness. Caricatures think that all the facial features relate to one another fundamentally, and we cannot make a change to one feature without it affecting the others. This is one of the few constants you can rely on with respect to drawing.

3.1 Caricatures

Action and Reaction. The T-shape rule is utilized to exaggerate the relationships between the facial features, which can be stated as follows: if the eyes move apart from each other, the nose should be shortening; whereas, if the noise is lengthening, the eyes should move closer to each other. It proves both simple and intuitive.

3.2 Likeness

In existing methods, "likeness" was seldom considered for caricature synthesis due to the lack of a "likeness" metric. We introduce the Modified Hausdorff Distance (MHD) to measure the visual similarity, based on this metric the likeness is incorporated into the integral caricature by optimizing the configuration of the facial elements.

3.3 Texture Transformation

In the texture transformation process, our aim is to get the grayscale sketch with a comic look. Although the generation is easy, the resulting sketch looks acceptable.

4. Algorithm

Algorithm for finding the unexaggerated grayscale sketch by taking the feature shape of the original photo is briefed in this section.

Input: S: feature shape of the original photo I: The original photo

Output: C: The unexaggerated grayscale sketch

Start

Step 1: Transform I into a grayscale image I'.

Step 2: Stretch intensity values of I' to 0~255.

Step 3: Perform Sobel edge detection on I' to obtain edge image E1. (r>25).

Step 4: Perform Sobel edge detection on I' to obtain edge image E2. (r>25) Step 5:

Paste the facial region in E2 to the corresponding region in E1.

Step 6: Fill the dark part and eyebrow regions in E1 with a dark color.

Step 7: Fill in the inner mouth regions with white color.

Step 8: Plot the contour lines. C=E1+Contour lines.

End

To make the facial area cleaner, we compute the ratio r of non-edge points and edge points. Higher ratio means higher demands of reducing the edge points, which makes the image clearer. We take the facial region with the result of a higher threshold while filling in other regions with the result of the lower threshold as shown in step 4~5 in the algorithm. Filling the dark part in the image can make it easier to recognize and more artistic. Considering making the facial features stand out, we plot contour lines as the final procedure. To make the facial area cleaner, we compute the ratio r of non-edge points and edge points. Higher ratio means higher demands of reducing the edge points, which makes the image clearer. We take the facial region with the result of a higher threshold while filling in other regions with the result of the lower threshold as shown in step 4~5 in the algorithm. Filling the dark part in the image can make

it easier to recognize and more artistic. Considering making the facial features stand out, we plot contour lines as the final procedure. **5. Flowchart**

The algorithm steps of our system can be summarized by the following flowchart fig 2.

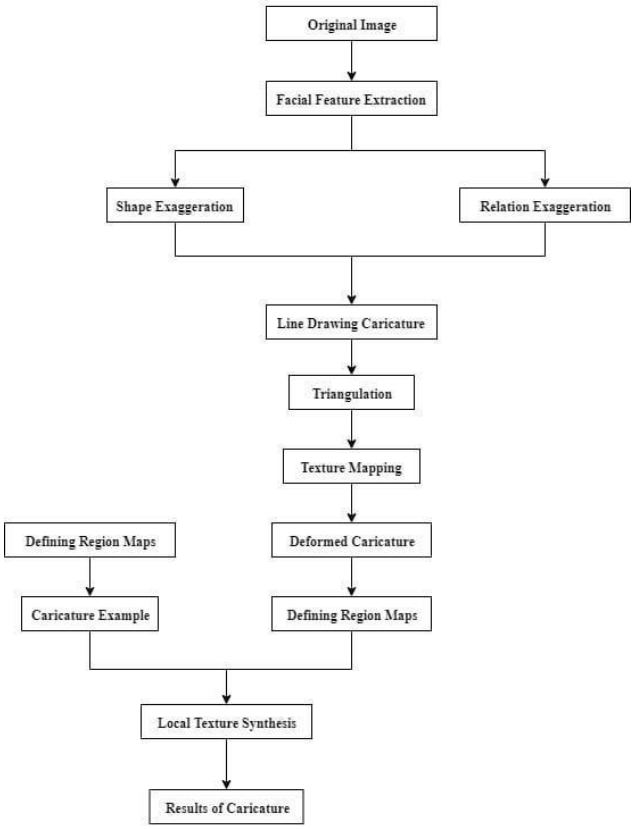

Fig 2: Flowchart

6. Results & Discussion

The paper proposed an automatic face caricature generation, which can be used to produce cartoons or inspire artists. The algorithm used recommends adaptive thresholding to eliminate the illumination and use the caricature generator to effectively detect the face component features. Hair is segmented and rendered in the same style to other components, and it has a pretty cool artistic effect. Results show that the generator proposed performs well in most cases and the synthesized caricature has a good artistic effect. The outcome of the experiment is given int the figure 2. It shows the original image which is taken as input, the intermediate smoothened image, and the final edges image. The proposed method is validated over several different images format and it is proved to be efficiently generating the caricatures of the actual image input.

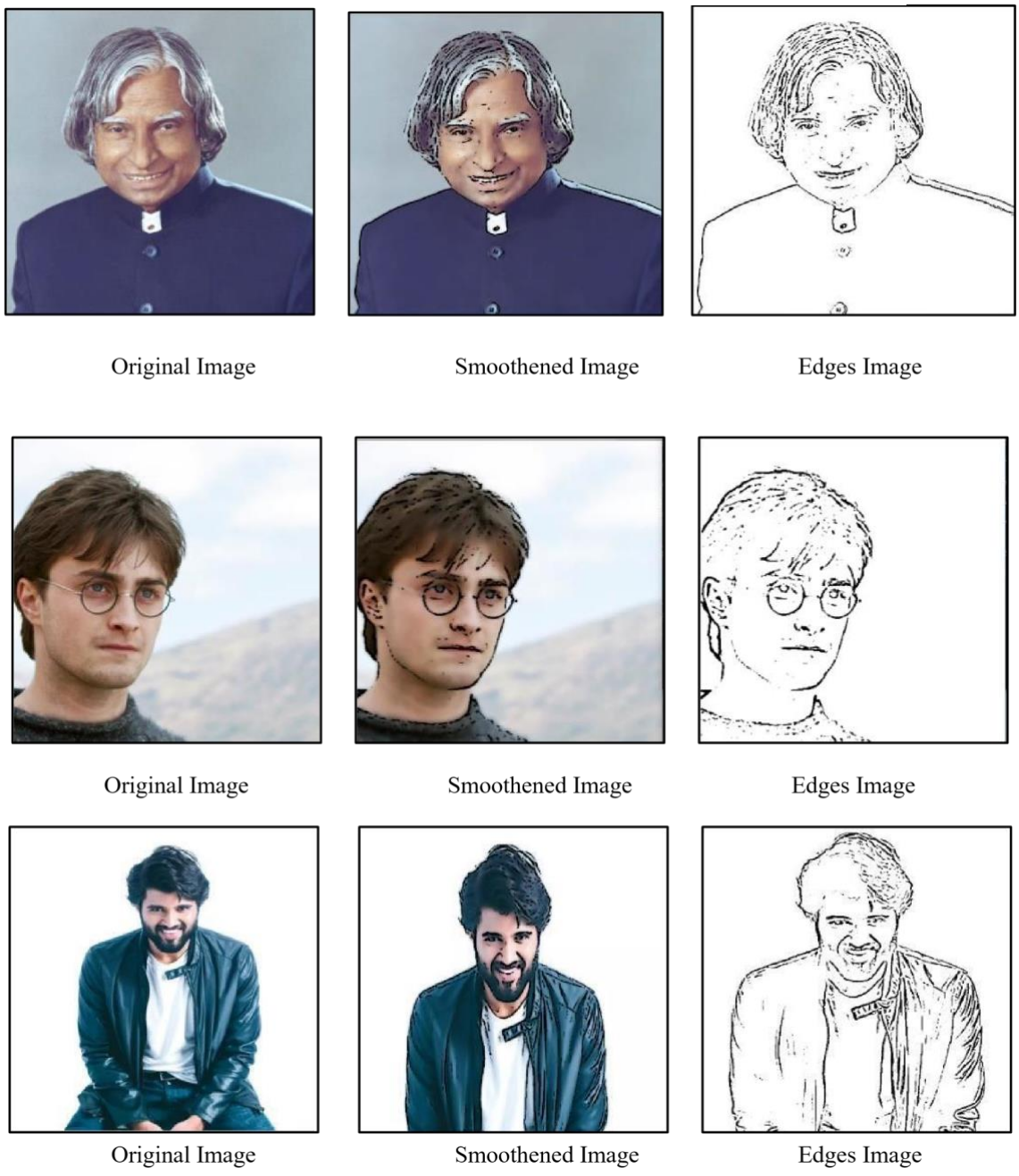

Figure 3: Results of experiment

7. Conclusion

This paper proposed the method to automate the process of generating caricature based on input images. We have demonstrated the use of adaptive thresholding and SVM algorithm for generating the caricature from the input image.

Initially the original image is converted to smoothened image, and the edges are detected and rendered to generate the final image resembling the caricature.

References

[1] Brennan, S. E. (1985). Caricature generator: The dynamic exaggeration of faces by computer. Leonardo, 18(3), 170-178.

[2] Gooch, B., Reinhard, E., & Gooch, A. (2004). Human facial illustrations: Creation and psychophysical evaluation. ACM Transactions on Graphics (TOG), 23(1), 27-44.

[3] Chen, H., Zheng, N. N., Liang, L., Li, Y., Xu, Y. Q., & Shum, H. Y. (2002, December). PicToon: a personalized image-based cartoon system. In Proceedings of the tenth ACM international conference on Multimedia (pp. 171-178). ACM.

[4] Hardy, C., Le Merrer, E., & Sericola, B. (2019, May). MD-GAN: Multi-discriminator generative adversarial networks for distributed datasets. In 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS) (pp. 866-877). IEEE.

[5] Hosseini-Asl, E., Zhou, Y., Xiong, C., & Socher, R. (2018). A multi-discriminator cyclegan for unsupervised non-parallel speech domain adaptation. arXiv preprint arXiv:1804.00522.

[6] HANAOKA, G. (1997). Facial caricature by computer based on the style of individual human caricaturist. IEICE Trans., 80(8), 21102118.

[7] Iwashita, S., & Onisawa, T. (1997, July). A study on facial caricature drawing by fuzzy theory. In Proceedings of 6th International Fuzzy Systems Conference (Vol. 2, pp. 933-938). IEEE.

[8] Nagamachi, M. (2010). Kansei Engineering, 2 Volume Set. CRC Press.

[9] Sobel, I. (1990). An Isotropic 3× 3 Image Gradient Operator, Machine Vision for Three-Dimensional Scenes (H. Freeman editor).