Paper Review 2

Semi-Supervised Learning in Reconstructed Manifold Space for 3D Caricature Generation
Junfa Liu, Yiqiang Chen, Chunyan Miao, Jinjing Xie, Charles X. Ling, Xingyu Gao, Wen Gao

Introduction

In the current work, the authors explore the application of automatic 3D caricature generation. The 3D caricatures are generally generated manually using various softwares such as 3D Max, but the problem with this methodology is that it is very time consuming and requires professional expertise. The initial explorations in automatic 3D caricature generation had two major drawbacks: being not automatic enough for practical applications; not being able to provide complex caricature styles. The authors explore a machine learning based approach to generate the 3D caricatures from the 2D images of the object.

The authors propose a method to build 3D models from 2D facial photographs, using a manifold regressive model. This methodology would take advantage of the neighbourhood of the points in the 2D face images to build the 3D models, by using the 3D caricatures reconstructed from manually modeled 2D caricatures on the basis of a PCA-based technique. These 3D caricatures are used as the ground truth in the regression problem. To construct the manifold structures of the 2D real faces and 3D caricatures, the non-linear dimensionality reduction technique Locally Linear Embedding (LLE) is employed.

Data Preparation & Feature Extraction

The dataset preparation is one of the major components of this research work. In all, three classes of data is prepared: one with 2D faces; one with 2D caricatures; and one with 3D caricatures. In the first part of the dataset, 2000 photographs containing human faces are collected. In the second part of the dataset, artistic 2D caricatures are collected, and a subset (100) of them are matched with real 2D faces, so as to generate 100 2D real face - 2D caricature pairs. In the third part, the 3D caricatures are created on a subset (100) of 2D caricatures collected by PCA subspace modeling, and remaining from the PCA subspace.

One of the primary precursor to feature extraction would be to align all the 2D images. To do so, the authors use the affine transformation technique, which comprises a combination of translation, rotation and scaling. This is ignored for the 3D caricatures, beacuase each of them have been based on a standard MAYA model, due to which it's scale remains unchanged since the beginning. From the 2D face images, features are obtained using a tool developed by the authors called ShapeExtractor, using which the points of facial shapes are extracted. From the 3D caricatures, the vertex coordinates are extracted as features, which in case turns out to be a 4011 dimensional vector.

Furthermore, the 3D caricatures, which are distributed in a manifold, are reduced to a 2D manifold map using Locally Linear Embedding technique. The 2D manifold map for the 2D real faces are similarly extracted as well.

Experimentation

The experiments around 3D caricatures consist of the following parts: reconstruction, generation and evaluation. In the 3D caricature reconstruction, the authors set the parameter R as 98% for PCA, owing to which, 40 principal components are calculated. Since 100 of the 1000 2D caricatures in the dataset were used as prototypes, 900 2D caricatures are used for reconstruction. Based on evaluation metrics like Similarity, Artistry and Exaggeration, 810 of these reconstructed images are preserved, owing to which, a total of 910 (810 + 100) samples are converted into the 2D manifold.

For LLE dimensionality reduction, 6 nearest neighbours are used for both 2D real faces and 3D caricatures. 56 dimensions are used for the representation of the 2D faces, and 2 dimensions for the 3D caricatures. Since most parts of the neighbourhood remains intact, the reduction in dimensions from 4011 to 2 doesn't lead to much loss in information.

For the objective evaluation of all results, the evaluation metrics such as Similarity, Artistry and Exaggeration were employed. The authors used the scores from ten persons, who assigned scores on a scale of 5, wherein the scorers were unaware of the information as to which of them were Manifold-Regression based on Traditional methods. Furthermore, to verify the similarity of the results, a recognition experiment were conducted by using 23 persons as evaluators, on 5 female and 15 male facial photographs and their 3D caricatures generated. The scores for all the evaluation metrics from both the techniques are tracked in Table 1, while the average recognition rate and recognition time separately on male and female are recorded in Table 2. It was observed that the average recognition rate of the female caricatures was higher than that of the male caricatures. The authors attributed to this enhanced performance to the possibility of more distinctive features existing in female faces than the male faces, for example, such as a sharp chin.

Strengths & Weaknesses

The research work proved that both the 2D real faces and the 3D caricatures can be represented effectively in a low dimensional manifold feature space. Furthermore, the results in Table 1 also proves that the manifold regressional technique performs considerably better than the traditional technique. One of the major weaknesses observed in this study is that not reconstructed caricatures can be used for training, and had to be discarded from the training set manually. Also, the evaluation measures such as artistry and exaggeration which were used to compare manifold regression method and the traditional method have not been defined quantitatively in the study.

Further Work

Han, et.al explore a deep learning based sketching system for 3D caricature modeling in [1], which was found to have a better performance as compared to a simple regressive manifold technique. They used a novel CNN based approach for inferring 3D caricature models from the 2D sketches, which helped in achieving a significant performance. Also, one of the weaknesses of the current research study is the complicated training set generation method followed, which was alternated with a better technique in [2] by Qianyi et.al, which allowed them to work with more standard datasets.

References

[1] Han, Xiaoguang and Gao, Chang and Yu, Yizhou, *DeepSketch2Face: A Deep Learning Based Sketching System for 3D Face and Caricature Modeling*, ACM Trans. Graph, July 2017

[2] Qianyi Wu, Juyong Zhang, Yu-Kun Lai, Jianmin Zheng, Jianfei Cai; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 7336-7345