Face Transferring on Webcam images using StyleGAN

J. Prasanthi 1, Dr. G. Anuradha 2
1 M. Tech Student, CSE Dept, email: prasanthi12121@gmail.com
2 Associate professor, CSE Dept, email: ganuradha@vrsiddhartha.ac.in
1,2 Velagapudi Rama krishna Siddhartha Engineering, Vijayawada, AP, India

Abstract

In image processing technology, face transfer is broadly used for privacy protection, picture enhancement, and entertainment applications. Face transfer is the domain that maps one image into another image and extracts several features of the face from one person to morph that face to another person. This face transfer will carry the facial expressions also. This is also called face morph, face swap, etc. Here we propose StyleGAN technology using face transfer with the image to get high quality. In this StyleGAN contribute the bilinear interpolation and affine transformation. Bilinear interpolation is to remove the noise and increase the quality of images. Affine transformation is to supply the images with 2d warping to improve the image quantity. To upgrade the quality of the images with face transfer is adopted to increase the accuracy of the image quality after image transfer.

Keywords: StyleGAN, Affine Transformation, Deep learning.

1. Introduction

From the selection of two images the face swapping or face transfer refers to changing the face identity from the input image to the earmark image, without losing the quality of the image. From the past many years several algorithms have been applied for better face transfer. This approach is most widely used in several domains such as entertainment [1][2], privacy protection [3][4], and theatrical industry [5]. Several methods show the huge impact by using various face transfer applications. Re-rendering is utilized to merge with the actual image to get the improved face transfer. In Li et al., 2012 [6] the data driven approach is applied on various face transfer images. This approach is mainly focused on extracting the frames by using the similarity parameters that are used for the optical flow as advent and calculate the velocity and search the K-nearest neighbors (KNN) based on timestamps and flow distance.

Generative Adversarial Network Model (GAN) is a new method which is instigated to manage the huge quality face images [5]. In [6][4][7] various face image synthesis using GANs has been implemented. One of the advantages of GAN is that it will generate optimized and high resolution data. The inputs of the real data distribution are generally proposed by the GAN. The overall characteristics are transformed while protecting the identity. This approach is based on two factors to implement this task. In the first stage, the source image is selected as input, and the face image is mapped to the target face image based on the related facial expression and head pose. Secondly, the task is based on a high quality.

2. Related work

Vlasic et al. 2005 [7] explained about the multi-linear model which can morph the given input image very efficiently based on the geometric variations because of different facial expressions. In Dale et al. 2011 [8] performance is calculated on videos that are selected. Based on the face expressions the face template is identified and re-rendered by using various metrics. Garrido et al. 2014 [9] introduced 3D geometry to analyze the performance of input and output. Thies et al., 2016 [10] proposed the new face tracking model with expression modeling which is used to transfer the facial expressions to get the real time output.
Several issues are identified in the image processing, computer graphics and vision of computers and designed the image-to-image translation task by Isola et al. 2016 [11]. Various tasks such as scenes mapping with aerial view, day and night images that convert the grayscale to color. Many issues are identified with face synthesizing which is used in this approach. Several face images are modified based on the hair style, person age, face expressions, black and white beard with glasses is proposed by (Kim et al. 2017b) [12] (Shen and Liu 2016) [13]. Recently for 3 years GAN has given more attention to this technique. Various domains are used for this GAN according to the requirement. With the GAN models the quality and resolution plays the major role [14] and also seen many improved techniques [15, 16, and 17]. The generators are used to handle the black boxes and in spite of recent efforts [18], will focus on several aspects based on the process of image synthesis, for example, there is lack of stochastic features. Based on the drawbacks the latent space interpolations [19, 20, and 21] are very poorly understood by the system. This will not provide any quantitative approach which compares several generators against each other.

From the computer vision point of view the high-quality photo realistic images are generated that includes editing of photo, computer based design, and image synthesis. For the text to image transitional models the Attention GAN (AttnGAN) [22] is used regularly. Another method which is proposed in [23] is Stacked GAN for translation of text to image which is used to select the samples of text via Conditional GAN [24]. In this two stages of GANs are created such as Improved-GAN generates the image of Low Resolution (LR) on the specific text data, and stage II GAN improves the resolution of the created image.

3. Proposed methodology

Here we introduce face transfer by using StyleGAN technology to transfer the one image to another image.

3.1 Style based Generative Adversarial Network (StyleGAN)

This model controls the image synthesis through scale-specific modifications to the styles without compromising the high quality image that generates continuously by using PGGAN. StyleGAN is an advanced version of PGGAN. This model is divided into three elements namely in (i) Coarse features such as pose, hair, face, shape, (ii) middle features such as facial constitutes, eyes, and (iii) fine features color scheme. These are StyleGAN1, StyleGAN2 are two new GAN techniques that improve the image quality after the image swapping.
Hereby using StyleGAN, the main contributions are: i) Bilinear interpolation and ii) Affine transformation. Bilinear interpolation provides the resampling technique i.e resize the images to provide quality and adequately map the images. We provide interpolation to remove the latent space between the images. Affine transformation is to map the data with warping the images for quantity and remove the noise in the images as shown in Figure.1.

4. Implementation
The dataset is collected from various sources which have images for testing and training. We train some images to transfer one face to another face and use bilinear interpolation to remove noise of the images and provide proper face transfer to those particular images. We supply affine matrix transformation to get the color correction by using warping and deliver two-dimensional (2d) warping to morph the faces and provide better quality to those images. Here we need to provide the Face 1, Face 2 is the input images and transfer image is output images. We train and test 25 images to provide face transfer with better quality as shown in Figure.2.
5. Result Analysis

The performance of StyleGAN is calculated by using the bilinear interpolation and the two linear interpolations in the x-direction (horizontal) are represented as \((x_1, y_1)\) then at \((x_2, y_2)\) plots.

The function \(f\) of linear interpolation at \((a, b_1)\) using the values of \(f\) at \((a_1, b_1)\) and \((a_2, b_1)\) which are \(Q_{11}\) and \(Q_{21}\) respectively as shown in equations:

\[
R_1 = \frac{x-x_1}{x_2-x_1}Q_{11} + \frac{x-x_1}{x_2-x_1}Q_{21} - 1
\]

Bilinear interpolation formula derivation step 1

\[
R_2 = \frac{x-x_1}{x_2-x_1}Q_{12} + \frac{x-x_1}{x_2-x_1}Q_{22} - 2
\]

The linear interpolation function \(f\) at \((x, y_2)\) using the values of \(f\) at \((x_1, y_2)\) and \((x_2, y_2)\) which are \(Q_{21}\) and \(Q_{22}\) respectively:

\[
P = \frac{x-x_1}{x_2-x_1}R_1 + \frac{y-y_1}{y_2-y_1}R_2 - 3
\]

Bilinear interpolation formula derivation step 1
The linear interpolation finally at \((x, y)\) using the interpolated values of \(f\) at \((x, y_1)\) and \((x, y_2)\): The linear interpolation in the y-direction (vertical): the values at \((x, y_1)\) and \((x, y_2)\) to extract the interpolation at the final point \((x, y)\) as shown in Figure.3.

![Figure 3: Graph of bilinear interpolation](image)

6. **Affine Transformation**

It is used for scaling, mapping the data and to provide the structure of images. This is generally coordinates at \((x, y)\) directions of the values \((a_1, b_1)\), \((a_2, b_2)\) as shown in Figure.4 below along with graph and equation 4:

\[
\begin{vmatrix} a_1 & b_1 \end{vmatrix} = X \times \begin{vmatrix} a_2 & b_2 \end{vmatrix} + Y - 4
\]

![Figure 4: Graph of Affine transformation](image)

**Table 1: performance of StyleGAN**

| Algorithms | Bilinear Interpolation | Affine Transformation |
|------------|-----------------------|----------------------|
| GAN        | 95.6                  | 96.5                 |

7. **Conclusion**

The StyleGAN which is used in this paper shows the huge quality of an image. Image transfer is the more relevant feature by using the StyleGAN. From the various domains and issues, many sub-domain techniques of StyleGAN are discussed by the researchers. This paper focused on developing the quality of images, improving the quantity, based on testing and training that can provide better quality for images in the StyleGAN. We provide affine transformation and bilinear
interpolation in stylegan to improve the accuracy and high quality of images. The future scope can improve the face transfer for specific eyes, nose and lips by using stylegan.

References

[1] Alexander, Rogers, Lambeth, Chiang, Debevec creating a photo real digital actor the digital emily project in 2009 conference for visual media production (pp. 176-187). IEEE.

[2] Kemelmacher-Shlizerman, I Transfigure portraits. 2016 ACM Transactions on Graphics (TOG), 35(4), 1-8.

[3] Blanz Scherbaum K, Vetter, Seidel, H. P. Exchanging faces in images. In Computer graphics forum 2019 Oxford, UK and Boston, USA: Blackwell Publishing, Inc Vol. 23, No. 3, pp. 669-676.

[4] Mosaddegh, Simon, Jurie F. photo realistic face de-identification by aggregating donors face components in 2019 Asian Conference on computer vision springer cham pp. 159-174.

[5] Naruniec, Helminger, Schroers, Weber, R. M. high resolution neural face swapping for visual effects. In 2020 computer graphics forum Vol. 39, No. 4, pp. 173-184.

[6] Wang J., Li, Xu, Dai Q., and Liu, Y A data-driven approach for facial expression synthesis in video. In computer vision and pattern recognition (CVPR), 2012 IEEE 57–64.

[7] D. Brand, Vlasic, M. Pfister H. and Popovic, J. Face transfer with multi linear models. In ACM transactions on graphics (TOG), volume 24, 426–433.

[8] K. Sunkavalli, Dale, Johnson, Vlasic, D. Matusik, and Pfister H. 2011 video face replacement. ACM Transactions on Graphics (TOG) 30(6):130.

[9] Valgaerts, Garrido, Thormahlen, Rehmsen, Perez, P. and Theobalt, C Automatic face reenactment. In IEEE conference proceedings on computer vision and pattern recognition 2014, 4217–4224.

[10] Zollhofer, Thies, Stamminger, Theobalt, and Nießner, face2face: Real time face capture and reenactment of rgb videos. In 2016 proceedings of the IEEE conference on computer vision and pattern recognition, 2387–2395.

[11] Isola, Zhou, Zhu and Efros image to image translation with conditional adversarial networks. 2014 arXiv preprint arXiv:1611.07004.

[12] Kim, B. Cha, M Unsupervised visual attribute transfer with reconfigurable generative adversarial networks 2017 arXiv preprint arXiv:1707.09798.

[13] Liu and Shen Learning residual images for face attribute manipulation in 2016 arXiv preprint arXiv:1612.05363.

[14] Pouget Abadie, Goodfellow Mirza, Xu, B. Darrell Farley, Ozair, S. Courville, and Bengio, Generative adversarial nets. In 2014 advances in neural information processing systems, 2672–2680.

[15] Karras, Aila, Laine, and Lehtinen Progressive growing of GANs for improved quality, stability, and variation. 2017 CoRR, abs/1710.10196.

[16] Kataoka, Miyato, Koyama, and Yoshida Spectral normalization for generative adversarial networks. 2018 CoRR, abs/1802.05957.
[17] Donahue, Brock, and K. Simonyan Large scale GAN training for high fidelity natural image synthesis 2018 CoRR, abs/1809.11096.

[18] Strobelt, Zhou, Zhu, Tenenbaum, and Torralba GAN dissection: Visualizing and understanding generative adversarial networks. In Proc.ICLR, 2019.

[19] Springenberg, Dovsovistskiy and Brox Learning to generate chairs with convolutional neural networks. 2014 CoRR, abs/1411.5928.

[20] Thielk, Theilman, Sainburg, Migliori, and Gentner Generative adversarial interpolative auto encoding: adversarial training on latent space interpolations encourages convex latent distributions. 2018 CoRR, abs/1807.06650.

[21] S. Laine feature based metrics for exploring the latent space of generative models 2018 ICLR workshop poster.