Volume 5, No. 4, April 2014 (Special Issue)
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
REVIEW ARTICLE
Available Online at www.ijarcs.info

The Comparative Study of Automated Face Morphing Methods for Images and Video

Harrmesh Sanghvi
SCET College
Surat, India
sanghviharmesh@yahoo.com

D. R. Kasat
SCET College
Surat, India
dipali.kasat@scet.ac.in

Sanjeev Jain
MITS
Gwalior, India
infomitsgwl@gmail.com

V. M. Thakare
Amravati University
Amravati, India
vilthakare@yahoo.co.in

Abstract: For entertaining purposes, a computerized special effect referred to as "morphing" has enlarged huge attention because of the attraction of scenes persons saw around the display screen. Although various morphing algorithms have been developed that can be categorised as 2D and 3D object morphing and generally applied in animation and morphing, but there are few mechanisms to spread out these techniques to handle images and video sequences automatically, specially morphing of an expressive moving part in the video like human face. This paper surveys the growth of this field and describes the latest advances in Automatic face morphing in terms of feature specification, warping techniques, and various different scenarios. These areas relate to the simplicity of use and features of results.

Keywords: Face detection; Face morphing; Morph-able Model

I. INTRODAUCTION

Morphing signifies the transform with variety and often behaviors of the individual during typical improvement after the developing phase. For that reason, the idea accommodates wonderfully addressing computer animation methods managing the style involving algorithms modifying one object into an additional after a while. In addition to purposes within the entertainment industry, morphing methods get benefits within areas that require helping image modification or the change of one object into an additional, as an example, visualization of the development of human beings, the modeling of facial movements using example images, and the interpolation between different views of a single face.

Image morphing concerning a pair of photographs begins through an animator creating their correspondence using frames of feature primitives, e.g., points, curves, line segments, or mesh nodes. Each primitive identifies an image feature or landmark. The feature correspondence is then used to compute mapping functions that define the spatial correlation among all points in both images. Since mapping functions are central to warping, it can be referred as warp functions. It can be used to interpolate the positions of the features across the morph sequence. Once both images have been warped into alignment for intermediate feature positions, ordinary color interpolation (i.e., cross-dissolve) generates in between images.

Feature specification is the most wearisome phase of morphing. Although the choice of allowable primitives may vary, all morphing approaches need careful attention to the specific placement of primitives. Given feature correspondence constraints between both images, a warp function over the whole image plane must be derived. This process, which it refers to as warp generation, is basically an interpolation problem. Another interesting problem in image morphing is transition control. If transition rates are allowed to vary locally across in between images, more interesting animations are possible. Plenty of morphing algorithms have been proposed which resolve these kinds of problems but feature specification process is manual. Certain methods have been proposed which provide completely automatic face morphing. The following section surveys different face morphing techniques with various scenarios.

II. FACE MORPHING TECHNIQUES

A. *Face Morphing in images :*

In 2004, V. Zanella et al. [1] proposed method to perform the morphing of face images in frontal view with uniform illumination automatically, using a generic model of a face and evolution strategies to find the features in both face images. They used a model of 73 points based on a simple parameterized face model. The model does not rely on color or texture; it only uses information about the geometrical relationship among the elements of the face. For example, they use the fact that the eyes are always at the same level and above the mouth in a face in frontal view.

In this work, automatic fitting of this model to source image and target image has been proposed using genetic operation. Once the model has been adjusted to the images, it performs image deformation, or warping, by mapping each feature in the source image to its corresponding feature in the target image. Right here, Inverse Distance Weighted Interpolation Method is utilized for warping. The run time average is 30 seconds to perform the morphing on a 2.0 GHz Pentium IV machine with 128 MB of RAM. This method performs the automatic morphing of face images using evolutionary strategies and a generic face model. It does not need a training set of faces to obtain the model because it uses a model based on a simple parameterized face model. The results are good although it worked with a simplified problem using only images of faces in frontal

CONFERENCE PAPER
Two day National Conference on Innovation and Advancement in Computing
Organized by: Department of IT, GITAM UNIVERSITY Hyderabad (A.P.) India
Schedule: 28-29 March 2014

view and with uniform illumination.

The described method works only for frontal view face morphing; otherwise this face morphing technique tends to generate blurry intermediate frames when the two input faces differ significantly. In 2012, F. Yang et al. [2] proposed a new face morphing approach that deals explicitly with large pose and expression variations. It recovers the 3D face geometry of the input images using a projection on a pre-learned 3D face subspace. The geometry is interpolated by factoring the expression and pose and varying them smoothly across the sequence. Finally, it poses the morphing problem as an iterative optimization with an objective that combines similarity of each frame to the geometry-induced warped sources, with a similarity between neighboring frames for temporal coherence.

In this system, it fits a 3D shape to both the input images. A 3D shape contains two sets of parameters: external parameters describing the 3D pose of the face, and intrinsic parameters describing the facial geometry of the person under the effect of facial expression. Then, it linearly interpolates both the intrinsic and external parameters of the two input faces, and generates a series of interpolated 3D face models. In each frame, the warped faces are blended together. It takes about 10 minutes to create 8 intermediate frames for face regions about 200 by 200 pixels, on an Intel CPU of 2.40 GHz. A large portion of the time is taken by the repeated optical flow computations.

This method focus on the problem of morphing faces and it does not have a sophisticated model for the background. Another direction for future work is in improving the model of the facial structure, such as explicitly modeling the locations of the pupils and ensuring they properly interpolate when the input images have large differences in gaze direction.

Certain methods also allowed for automatic face replacement of people in single image [3, 4]. For example, in 2004 the method by V. Blanz et al. [3] fits a Morph able model to faces in both the source and target images and renders the source face with the parameters assessed from the target image. Finally, it replaces the target face with source face in the target image. Morph able model is built from a statistical analysis of human faces, obtained from a large database of 3D scans, which can be morphed by adjusting parameters. It can estimate the 3D shape of a human face, its orientation in the space, and illumination conditions in the scene. Thus the reconstructed face extracted from 2D image can be manipulated in 3D [5]. In 2008, D. Bitouk et al. [4] described another system for automatic face swapping using a large database of faces. Though it is hard for user's to find a candidate face to match the target face in appearance and pose from their images, the system allowed de-identification automatically by selecting candidate face images from a large face library that is similar to the target face in appearance and pose. Lastly, it replaces the target candidate with selected candidate from the library image using image based method.

In 2013, S. Gao et al. [6] introduced simple 3D face model, which is known as Morph able guidelines. They proposed a system which allows morphing specific part of face like nose, mouth, cheek etc. in single image. This Morph able guideline is a 3D model structured similar to the ball and plane method. This model consists of simple curves like a circle, line etc. which is controlled by the 3D Vertices. Individual shape can be changed by changing the parameter. In this paper they have applied this model to reshape human shape parameter such like nose, mouth, cheek etc. But model fitting process to the human face in image is manual.

B. Face Morphing in video:

In 2009, Y. Liang [7] proposed the system which provides plausible face replacement in video. This system allows replacing target human face from target video with source face in source video. This replacement algorithm has three main stages. First, given an input video, it detects all faces that are present, and align such detected faces. Second, it analyses facial expressions of each detected faces and select candidate face images from source video that are most similar to the target face in pose and expression. Finally, it blends candidate replacements to target video.

For face detection here Bayesian Tangent Shape Model (BTSM) is used. At that moment system contains two models. One indicates the prior shape distribution in tangent shape space and the other is a likelihood model in image shape space. Based on these two models, the posterior distribution of model parameters can be derived. To replace the face, it requires outlines of the facial profile and structures. For face alignment 2D landmark points are used from BTSM (as mentioned above). To identify appropriate source pose and expression, system allows user to select the best candidate for each frame independently in the target video. For that it splits the target video into several partitions. This process is referred to as clustering. After clustering, user has to select appropriate frame for each cluster. The best candidate face must have the most similar pose and expression.

Once the user selected the best candidate face for each cluster, it will synthesize in-between frames. To concatenate the clusters, it requires interpolating the frames between clusters. A smooth image interpolation can be done by warping two images to the same positions and then dissolve the image textures together. After producing the smooth image sequence, the final step is to warp the facial features to appropriate positions to match the expressions exactly. It warps the smoothly interpolated image sequence according to the difference vectors. The warped images are about to be blended into the target video.

Sometimes, the size of the target character's face and the source character's face does not correspond with each other due to scale. Face concealment can be easily done by clearing the gradient field to zero in the facial region then reconstructed by integrating the modified gradient. It can obtain a smooth and de-identification face by assigning null gradient to the pixels within the facial region. While system paste up the source faces on the target face, the pasting region may exceed the boundary of the target face. The irrelevant background would be blended together such that the smudge effect will occur. To seamlessly clone the source patch into a target image, the operation is typically carried out by solving a large linear system. Instead of solving the large linear system of Poisson equation, it uses image cloning.

In 2009, Y. T. Cheng et al. [8] proposed the system that replaces the target subject face in the target video with the source subject face, under similar pose, expression, and illumination. This approach is based on 3D morph-able model [5] and an expression model database to deal with expressions and the input information of the source subject face is reduced to one to two images. The system takes a target video and one source image as input, and the output is the video with the target subject face replaced with the source subject face.

Given the source image, it reconstructs the 3D model of the source subject face using 3D Morph able model [5]. The 3D face synthesizer derives a Morph able face to fit the input image, and map the texture from the image to the derived 3D face model. A face alignment algorithm is applied to the target video to detect the detailed facial features and outlines of the target subject face [7]. A pose estimator exploits the face alignment results to estimate the head pose parameters of the target subject face. Here method employs a 3D face expression database to clone the expressions to the source face model. To fit the expressions to the target video, Y. T. Cheng et al. [8] proposed an algorithm to extract the expression parameters. In some videos, directly rendering the source subject face model onto the target frame results in illumination inconsistency. A relighting algorithm relights the rendered source subject face for illumination consistency.

Finally, it seamlessly composites the rendered source model with the target frame using Poisson equation, proposed in 2003 by P. P´Erez, et al. [9]. The output is a video with the target face replaced by the source face, with similar pose, expression, and lighting.

In 2010, F. Min, et al. [10] proposed an automatic face replacement approach in video based on 2D Morph able model. This approach includes three main modules: face alignment, face morph, and face fusion. Given a source image and target video, the Active Shape Models (ASM) is applied to source image and target frames for face alignment. Then, the source face shape is warped to match the target face shape by a 2D Morph able model. The color and lighting of source face are adjusted to keep consistent with those of target face, and seamlessly blended in the target face. This approach is fully automatic without user interference, and generates natural and realistic results.

This approach includes three main modules: face alignment, face morph, and face fusion. In face alignment, the ASM algorithm is used to the target frame and source image to detect the detailed facial features. According to these facial features, face alignment is implemented. In face morph, a 2D Morph able model is fitted to the target and source face. Adjusting the parameter of the Morph able model, the shape of source face is warped to match that of target face. In face fusion, a relighting and recolor algorithm is applied to the source face to keep illumination and color consistency with the target face. The source face is seamlessly blended in the target face using Poisson equation, proposed in 2003 by P. P´Erez, et al. [9].

In 2011, K. Dale, et al. [11] proposed the method which allows replacing facial performances in video. It also provides face replacement in target video from source video. The system tracks both the faces in source and target video using multilinear model [12]. Using this tracked 3D Geometry, source face is warped to target face in every frame of video. It is sometimes important that the timing of the facial performance matches exactly in the source and target video; this is done by retiming algorithm. After tracking and retiming, it blends the source performance in the target video to produce the final result. They computed optimal seam through the video volume that maintains temporal consistency in the final composite.

This System consists of following Major steps: 1) Face Tracking 2) Face alignment 3) Blending

a. ***Face tracking:*** To change the default, adjust the template as follows. To track a face across a sequence of frames, the method of D. Vlasic et al. [12] computes the pose and parameters of the multilinear face model. Initialization is critical, as errors in the initialization will be propagated throughout the sequence. Therefore, System provides a simple user interface that can ensure good initialization and can correct tracking for troublesome frames. The interface allows the user to adjust positions of markers on the eyes, eyebrows, nose, mouth, and jaw line, from which the best-fit pose and model parameters are computed. This initial face mesh is generated from the multilinear model [12] using a user-specified set of initial attributes corresponding to the most appropriate expression, vise me, and identity.
b. ***Face Alignment:*** To align the source face in the target frame, it uses face geometry from the source sequence and pose from the target sequence. It also takes texture from the source image. Automatic retiming algorithm then compares the average minimum Euclidean distance between the first partial derivatives with respect to time that is used to match the facial performance in video. Retiming applied on sequence of images till first partial derivatives with respect to time will match in both videos.
c. ***Blending:*** A truly photo-realistic composition of two images is possible by the Poisson blending algorithm proposed in 2003 by P. P´Erez, et al. [9] but it requires specifying the region from the aligned video that needs to be blended into the target video. In addition, this seam needs to be specified in every frame of the composite video, making it very tedious for the user to do.

This method incorporates these requirements in a novel graph cut framework that estimates the optimal seam on the face mesh. For every frame in the video, It computes a closed polygon on the having constructed this graph. The construction of the graph confirms that, in every frame, the graph-cut seam forms a closed polygon that separates the target vertices from the source vertices. Finally, it blends the source and target videos using gradient-domain fusion.

C. Face Morphing in real time video:

In 2003, K. Timm [13] described a real-time virtual camera application based on view morphing. This system takes video input from multiple cameras aimed at the same subject from different viewing angles. It performs real-time pattern matching and generates artificial views for a virtual camera that can pan between the real views. The main steps of the algorithm can be summarized as follows. First, the images from two video streams must be acquired and the subject segmented from the background. Next, most important piece of the view morphing process is creating a correspondence between the images. After the correspondence has been created, each image is warped to an in-between view, blended with each other, and then displayed to the screen.

In 2009, Y. Yang et al. [14] proposed a system which provides an automatic morphing of face in real time video. In this, they have used efficient AdaBoost face tracking algorithm which gives facial feature points. In this proposed scheme using feature point, it finds a center point of a face and according to that point it made a square around the face and applied morphing algorithm. Here, they have proposed two morphing algorithm which provides two different views of face, face bulging and face shrinking. They have solved two challenges. First, described method will work only for frontal face. They have solved this by introducing wangling angle. And, second to maintain spatial and temporal coherence they have used buffer scheme. They have used five frames and find averaging of that and finally they warped the image.

Table 1. Performance Analysis of Different Techniques

Face morphing Techniques	Ref. Works	Pros	Cons
Face morphing in images	An Approach to Automatic Morphing of Face Images in Frontal View [1]	Comparatively less time consuming	Works only if face images in frontal view with uniform illumination
	Face morphing using 3D aware appearance optimization [2]	Deals explicitly with large pose and expression variations in face images	Time consuming
	Exchanging Faces In Images [3]	It can be used to replace face; Also works for non-frontal view	Does not give pleasant result if large illumination difference
	Face Swapping: Automatically Replacing Faces In Photographs [4]	It works even large illumination difference in both images	Does not work in case of any Occlusion
	Morph able Guidelines For The Human Head [6]	Allow to change face parameter like nose, cheek, jaw, chin, mouth and eyes etc.	Model fitting process is manual
Face morphing in video	Image Based Face Replacement in Video [7]	Less time complexity because of Image based method; It can be used for entertainment purpose	Facial expression and pose would be same; It requires to shoot target video with facial expression and pose similar to the source video; It requires manual inputs in clustering process; The tolerance to pose variance is limited by the robustness of face alignment algorithm
	3D Model Based Face Replacement in Video [8]	Here source is reduced to single image; It takes care of facial expression and pose of target face	Time consuming process because of 3d model based method; The tolerance to pose variance is still limited by the robustness of face alignment algorithm
	Automatic Face Replacement in Video Based On 2D Morph able Model [10]	This approach is fully automatic without user interference; Less time consuming process because of 2D model based method	It does not take care about facial expression; The tolerance to pose and expression variance is limited by the robustness of ASM; Sharp lighting and violent movement in videos may affect the final result
	Video Face Replacement [11]	It gives plausible results; This approach is fully automatic with less user interference; It takes care about facial expression	Tracking is based on optical flow, which requires that the lighting change slowly over face; Tracking often degrades beyond the range of poses outside 45^0 from frontal; Lighting must also be similar between source and target.
Face morphing in real time video	Real-time view morphing of video streams [13]	Morphing effect in real time; It can be used to create virtual environment	Required Higher configuration of hardware
	Entertaining video warping [14]	It morphs face area only instead of complete frame	It provides limited warping function like face bulging, face shrinking

III. CONCLUSION

In this paper, we have surveyed the growth of face morphing and described recent advances in the field. Table I above depicts the performance analysis of the different face morphing techniques surveyed in the paper. We surveyed various face morphing techniques, including those based on model based and image based morphing for images and video sequences. For those cases where the input images are appropriately similar, feature specification can actually be automated. So the feature extraction is the key technique toward building entirely automatic face morphing algorithms. In Zanella V. et al. [1] have introduced the problem of generating low quality face morphing animation given two faces of difference in pose and expression, which have been resolved in [2]. In a model based face morphing methods [2,3,8,11,12], they used a 3D subspace model learned from a large collection of human faces exhibiting realistic expression, constrains well the space of possible face deformations for interpolating casual face images. Traditional warping methods used for morphing that requires accurate correspondence between the two source faces, but in these methods they warp the two faces independently and only roughly using a 3D model based flow. Model-based vision has been examined to exploit knowledge about the relative position of these features and automatically locate them for feature specification. The same automation applies to morphing between two video sequences [7, 8, 10, 11, 13, 14], where time-varying features must be tracked. Its real-time looks coupled with the interesting face warping results reveal that these algorithms have a realistic view to be applied in many entertainment applications such as editing special effects on video including faces.

IV. REFERENCES

[1] V. Zanella and O. Fuentes, "An Approach to Automatic Morphing of Face Images in Frontal View", In: Monroy, R., Arroyo-Figueroa, G., Sucar, L.E., Sossa, H. (eds.) MICAI LNCS, vol. 2972, pp. 679–687, Springer, Heidelberg, 2004.

[2] F. Yang, E. Shechtman, J. Wang, L. Bourdev and D. Metaxas, "Face morphing using 3D-aware appearance optimization", GI '12 Proceedings of Graphics Interface, Pages 93-99, 2012.

[3] V. Blanz, K. Scherbaum, T. Vetter, and H.P.Seidel, "Exchanging Faces In Images", Computer Graphics Forum, (Proc. EUROGRAPHICS) 23, 3, Pages 669–676, 2004.

[4] D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and S. K. [10] F. Min, N. Sang, Z. Wang, "Automatic Face Replacement in Nayar, "Face Swapping: Automatically Replacing Faces in Photographs", In SIGGRAPH '08: ACM SIGGRAPH 2008 Papers, Pages 1-8, New York, USA, 2008.

[5] V. Blanz and T. Vetter, "A Morph able Model For The Synthesis Of 3D Faces", In SIGGRAPH '99: Proceedings of The 26th Annual Conference On Computer Graphics and Interactive Techniques, ACM Press/Addison-Wesley Publishing Co. Pages 187-194, New York, USA, 1999.

[6] S. Gao, C. Werner and Amy A. Gooch, "Morph able guidelines For the Human Head", Proceedings of the Symposium on Computational Aesthetics. ACM New York, USA Pages 21-28. 2013.

[7] Y. Liang, "Image Based Face Replacement in Video", Master's Thesis, CSIE Department, National Taiwan University, 2009.

[8] Y. T. Cheng, V. Tzeng, Y. Liang, C. C. Wang, B. Y. Chen, Y. Y. Chuang, and M. Ouhyoung,"3D Model Based Face Replacement in Video", In SIGGRAPH 2009 Poster, ACM. Video Based On 2D Morph able model", Proceeding ICPR '10 Proceedings Of The 2010 20th International Conference On Pattern Recognition, IEEE Computer Society Washington, Pages: 2250-2253, DC, USA 2010.

[9] P. P´Erez, M. Gangnet, and A. Blake, "Poisson Image Editing", ACM Trans. Graphics (Proc. SIGGRAPH) 22, 3, Pages 313– 318, 2003.

[10] K. Dale, K. Sunkavalli, M. K. Johnson, D. Vlasic, W. Matusik, and H. Pfister, "Video Face Replacement", ACM Transactions on Graphics (Proc. SIGGRAPH Asia) 30, 6, 2011.

[11] D. Vlasic, M. Brand, H. P fister and J. Popovi´C, "Face Transfer with Multilinear Models", ACM Trans. Graphics (Proc. SIGGRAPH) 24, 3, Pages 426–433, 2005.

[12] K. Timm, "Real-time view morphing of video streams", ACM SIGGRAPH 2003 Sketches & Applications, San Diego, California, July 27-31, 2003.

[13] Y. Yang, Y. Zhu, C. Liu, C. Song and Q. Peng, "Entertaining video warping", Computer-Aided Design and Computer Graphics, CAD/Graphics '09, 11th IEEE International Conference, Pages 174-177.Aug. 2009.