ANALYSIS OF FACIAL EXPRESSION TRANSFER TECHNIQUES: A SURVEY

[1]Jigna Patel, [2]Prof.D.R.Kasat, [3]Dr. Sanjeev Jain, [4]Dr. V. M. Thakare
[1]PG Student, Comp.Engg Dept, [2]Associate Professor,[3]Director,[4]Professor and Head –CSE
[1]Dept.,SCET College, SURAT, [2]SCET, Surat, [3]MITS, Gwalior, [4]Amravati University, Amravati
Email: [1]patel11jigna@gmail.com, [2]dipali.kasat@scet.ac.in

Abstract— Facial Expression transfer system is designed for expression transfer from one source subject to another. This paper summarizes theoretical approaches used in published work and also describes their strengths, weaknesses, and relative performance. Facial Expression transfer system is adopted to support computer vision and in movie industry. First, Facial expression transfer system was performed on source image to target image then performed on existing video. Yet less work are done on Real time video. Un-natural expression which is often more noticeable and severe, is the bad expression of the subject, such as closed eyes, half-open mouth, etc. However, how to make the synthesized facial expression realistic remains a challenge due to the complexity of human facial anatomy and inherent sensitivity to facial expression. Motion capture and the transfer of facial expressions from an actor to a CGI-generated movie character have long been a focus of much research in the movie industry and computer graphics community. In this paper, Several famous facial expression transfer approaches, such as Expression ratio images based approach, Morphing based approach, Multi-linear model based approach, Deformable Face Model based approach and Bilinear model based approach will also be explained.

Index Terms — AAM, ASM, ERI, Morphing.

I. INTRODUCTION

Everyone who has the experience to take photographs / video of family members and friends that she knows how hard it is to capture the perfect moment. One, the camera may not be properly set at the right time. In addition, a delay between the time a perfect smile in the viewfinder and the image / video is really slow response, especially for the low-end cell phone camera has captured the looks that the time is always there. For this reason, face images captured by amateur photographers often contains various imperfections. Generally speaking, there are two types of imperfections. The first type of photometric errors, incorrect camera settings, so even in the face, grainy, dark or might appear to be blurry. The second type, which is often more substantial and serious such as eyes closed, half-open mouth, the subject is a bad expression. Recent interest in facial modeling and animation is spurred by the increasing appearance of virtual characters in film industry and video, inexpensive desktop processing power for human-computer interaction.

With recent advances in image editing, photometric imperfections can be largely improved using modern post-processing tools. For instance, the personal photo enhancement system provides a set of adjustment tools to correct global attributes of the face such as color, exposure, and sharpness. Compared with photometric imperfections, expression artifacts are much harder to correct. Given a non-smiling face photo, one could simply find a smiling photo of the same person from his/her personal album, and use it to replace the whole face using different methods. Unfortunately, this global swap also replaces other parts of the face which the user may want to maintain. Local component transfer among face images is thus sometimes more preferable.

Transfer of Facial expression has long been an important topic in computer graphics and vision, driven applications in character animation, and more recently, healthcare. However, how to make the synthesized facial expression realistic remains a challenge due to the complexity of human facial analysis and our basic sensitivity to facial expression. Here proposed works related to facial expression transfer of an actor in video are based on exiting video. There are few works done on real time video which is related to morphing [6, 7]. It is difficult to transfer facial expression in real time because humans are very sensitive to any changes on face of the actor and if that modification/s are not perfect then it is very easy to recognize the unnatural change/s that have been done. For some reasons, face images captured by photographers often contain various imperfections. Like, which is often more noticeable and severe, is the bad expression of the subject, such as closed eyes, half-open mouth, etc. Editing special effects on actor face such as open/close eyes, pleasant smile in real time could be useful in many applications like in computer graphics and vision, driven by applications in character animation, and more recently, healthcare.

The paper is organized as follows: Section II presents the facial performance transfer related work. In section III, concludes this paper.

II. REALATED WORK

Recently, various approaches such as [1]-[2]-[3]-[4]-[5] [8] have been proposed for facial performance transfer. A method for creating photorealistic 3D facial models from images of human subjects was proposed in [1]. A 3D morphable model (3DMM)-based approach has been used for animating novel faces by transferring mouth movements and expressions. It achieves good accuracy for both pose and illumination, but at the expense of higher computational complexity. They employ manual facial pose marking and use a 3D shape morphing algorithm among the face models, but the computational cost is high due to expensive 3D calculations. Others have used an Expression Ratio Image (ERI) [2], the limitation of which is in dealing with illumination changes. In a different approach to transfer just the lip movements [5], a multidimensional morphable model, trained on a large data set of a speaker, is adapted to animate the lips of another speaker that requires only a small data set for training. Motion capture and the transfer of facial expressions from an actor to a CGI-generated movie character have long been a focus of much research in the movie industry and computer graphics community. The goal of this study is to provide the survey of facial expression transfer different techniques and to identify the research for future work. Followings are some expression transfer based approaches:

A. *Expression ratio images based approach*

Details of the changes of light is ignored, while the traditional expression attribute mapping techniques used to determine the speed. In 2010, Zicheng Liu et al. [2] approach, a novel technique for mapping facial expression represents. That capture the light of changes in the expression of a person's expression ratio image (ERI) is known. Along with the geometric warping, more expressive facial expressions to generate an image of the face of any other person ERI map. With harsh lighting conditions rather than changing geometry geometric deformations under constant lighting handles. The other person on the face of a person's facial expression is the first technique that is capable of aligning details.

		
Left: neutral face	**Middle: result from geometric warping**	**Right: Result from ERI method**

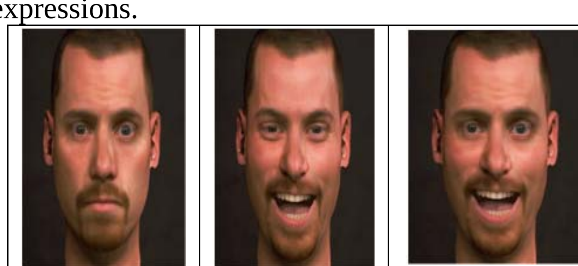
Fig.1 Expression mapping expression details[2]

Figure 1 shows an example of an expression with and without the expression details. The left image is the original neutral face. The one in the middle is the expression generated using the traditional expression mapping method. The image on the right is the expression generated using the ERI method. The feature locations on the right image are exactly the same as those on the middle image, but because there are expression details, the right image looks much more convincing than the middle one. Expression ratio image is an effective technique to enhance facial expression mapping with illumination changes. An expression ratio image can capture subtle but visually important details of facial expressions. The resulting facial expressions are significantly more expressive and convincing than the traditional expression mapping based on geometric warping. This method is very simple. One limitation of this method is difficult to deal with different lighting conditions. It can be used to animate 2D drawings and images, as well as textured or non-textured 3D face models.

B. Morphing based approach

This approach based system is used for creating realistic face models and also for performing realistic transitions between different expressions. In 1998, Frederic Pighin et al. [1] shows 2D morphing techniques can be combined with 3D transformations of a geometric model to automatically produce specifics 3D facial expressions with a high degree of naturalness. In this process consists of few basic steps. First, using cameras at arbitrary locations to capture multiple views of a human subject. Next, digitize these photographs and manually mark a small set of initial correlative points on the face in the different views (corners of the eyes and mouth, tip of the nose, etc.). These points are then used to automatically recover the camera parameters (position, focal length, etc.) corresponding to each photograph, as well as the 3D positions of the marked points in space. The 3D positions are then used to deform a generic 3D face mesh to appropriate the face of the particular human subject. At this stage, additional corresponding points can be marked clearly fit. Finally, extract one or more texture maps for the 3D model from the image. A view of the texture map can be extracted either independent, or view images based on the original size of the texture can be used to accomplish. With many different facial expressions of the entire process is repeated for the human subject.

To produce facial animation, interpolate between two or more different 3D models build in this way, while at the same time merge the textures. All the 3D models are constructed from the same generic mesh, morph as to achieve a natural correspondence between all the geometrical points. Thus, the transitions between expressions can be produced fully automatically once the different face models have been constructed, without having to specify pair wise correspondences between any of the expressions.



Left: "Neutral" expression	**Middle: "Happy" expression**	**Right: "Fake smile"**

Fig. 2 Morphing based facial expression transfer. Combining the upper part of a "neutral" expression (left) with the lower part of a "happy" expression (center) produces a "fake smile" (right) [1].

Figure 2 shows the example of morphing based facial expression transfer where combining the upper part of a "neutral" expression (left) with the lower part of a "happy" expression (middle) produces a "fake smile" (right). This modeling approach is based on photogrammetric techniques in which images are used to create precise geometry. The earliest such techniques applied to facial modeling and animation in place grids that were picked directly on the human subject's face. This system gives the user complete freedom in specifying the correspondences, and enables the user to able the initial fit as needed. Another advantage of this technique is its ability to handle fairly arbitrary camera positions and lenses, rather than using a fixed pair that are precisely oriented. And it is able to perform of extracting texture maps of higher resolution. This is the need for user interference in the modeling process.

Another advantage of this technique, a texture map model compared with more traditional animation, it is deformed face, which is as light and textures (e.g., facial creases) can capture subtle changes in. The degree of complexity and realism because the skin fold inter-reflections and self-shadowing same difficulty, but it is difficult to achieve with physical-based model. It's a very real face models and have the ability to produce natural looking animations.

C. Multi-linear model based approach

This approach incorporates all such information through multi-linear analysis, which naturally accommodates variations with multiple attributes. In 2005, Daniel Vlasic proposed a method for face Transfer is based on a multi-linear model [8] of 3D face meshes that separably parameterizes the space of geometric variations due to different attributes (e.g., identity, expression, and viseme). Separability means that each of these attributes can be vary independently. A multi-linear model can be estimated from a Cartesian product of examples (identities × expressions × visemes) with techniques from statistical analysis, but only after careful preprocessing of the geometric data set to secure one-to-one correspondence, to reduce the cross coupling of the artifacts, and sufficient in any missing examples. In this technique, Face transfer offers new solutions to this problem, expression, pose and viseme parameters to extract a face tracking algorithm connects with the estimated model. Example of facial expression transfer with multi-linear models where combine facial attributes from several videos as shown in Figure 3, where the pose, expressions, and visemes are mixed from three different input videos. Figure 3 shows Face Transfer with multi-linear models gives animators decoupled control over facial attributes such as identity, expression, and viseme.

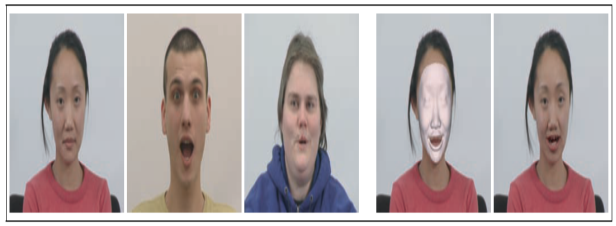


Fig. 3 Face Transfer with multi-linear models.

In this example, combine pose and identity from the first video, expressions from the second, and visemes from the third one to get a composite result blended back into the original video [8].

A convenient control of Facial features support by multi-linear models. And all the benefits of the system performance, change lifted from the video, and are easily return to the video, where video rewriting programs are combined. The disadvantage of this system is roughly a thousand vertices to track, process (e.g. around wrinkles above the lip) are less sensitive to changes in the local intensity. Multilinear model is applicable for editing of identity, performance, and facial texture in video, enabling video rewrite applications such as performance animation (puppetry) and actor replacement. Also the model offers a rich source of synthetic actors that can be controlled via video.

D. Deformable Face Model based approach

In 2012, Akshay Asthana [4] proposed a deformable model based approach; Facial operations in real-time using the AAM framework focused on the transfer issue. And also propose a new approach of learning the mapping between the parameters of two completely independent AAMs, using them to facilitate the facial performance transfer in a more real manner. The main advantage of modeling this correspondence parametric is that it allows a “meaningful” transfer of both the non-rigid shape and texture across faces whatever of the speakers’ gender, shape, and size of the faces, and illumination conditions. In that explore linear and nonlinear methods for modeling the parametric correspondence between the AAMs and show that the sparse linear regression method performs the best. Also show the utility of the proposed work for a cross-language facial performance transfer that is an area of interest for the movie dubbing industry.

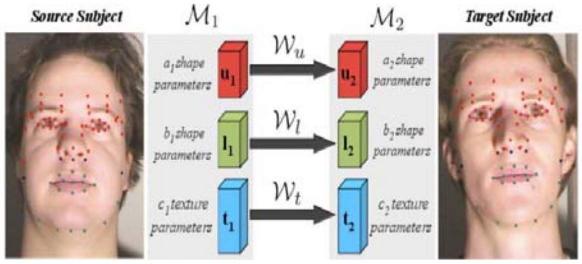


Fig. 4 Active appearance model based Facial performance transfer. (points in red—upper face shape, points in green—lower face shape, and points in blue — shared by both) [4].

Figure 4 shows the facial Performance transfer based on AAM. AAM works on shape and texture of the face. This approach describes the task of facial performance transfer by directly modeling the parametric correspondence between shape and texture of two completely independent AAMs. Regression-based approach is used to model the relationship between the shape and texture parameters of two completely independent face models, enabling “meaningful transfer” of the variation in both shape and texture. The “Meaningful” means taking the personal characteristics into account. Example, if the source subject exhibits large facial movements while the target subject normally shows little, it would look unrealistic if the source’s large movements were transferred verbatim to the target face. This approach facilitates realistic facial performance transfer between two subjects, irrespective of their gender, shape of their face or their skin tone. It also shows good generalization capability and works irrespective of the subject’s gender, ethnic background and language. And it requires only a small video corpus (4s videos) for modeling the parametric correspondence. Overall, the Sparse Linear Regression Method is best suited for the task of learning the parametric correspondence. The results and the rating provided by the human participants are very encouraging.

Active shape models (ASMs) are statistical models of the shape of objects that iteratively deform to fit to an example of the object in a new image, developed by Tim Cootes and Chris Taylor in 1995. Generate a suggested shape by looking in the image around each point for a better position for the point. This is commonly done using what is called a "profile model", which looks for strong edges or to match a model template for the point uses the distance. Usually in the context of a "shape model" is known as a good distribution model to suit the shape of the notification. The technique has been widely used to analyze images of faces, mechanical assemblies and medical images (in 2D and 3D).It is closely related to the active appearance model. It is also known as a “Smart Snakes” method that is an analog to an active contour model which would respect explicit shape constraints.

Deformable face facial performance model-based approach for the transfer of facial performance in computer vision and graphics community has gained tremendous interest.

E. Bilinear model based approach

In 2014, Chen Cao et al. [5] present Face Warehouse which is a 3D facial expression database for visual computing applications. The database contains the facial geometry and texture of 150 subjects, each with 20 expressions. The raw data set is used to construct 47 expression blend shapes for each person, capable of representing most expressions of human faces. Then these entire blend shapes are assembled into a rank-3 tensor, which is decomposed to build a bilinear face model. This model can be used to accurately estimate face expression and identities for facial images and videos. It is applied in a wide range of applications, such as facial image manipulation, face component transfer, real-time performance-based facial image animation, and facial animation retargeting from video to image. Facial component transfer application performs face component copy-and-paste to improve the expression in a facial image. It takes two images as input of the single person: an unhealthy expressions such as smiling inserted and the other as a necessary expression that includes a reference image that is the target photo. 3D face models adapted to synthesize the input images using bilinear face model. After that use the two face meshes to calculate a 2D expression flow in the target image, which warps the target face to match the desired expression. A 2D alignment flow is also calculated from these two meshes to warp the reference face to an appropriate size and position for transferring. Finally, Warp reference image, select the crop region, and transfer the result to produce a Warp combination of the target image.

Table 1 illustrates the summary of different model based facial expression transfer approaches.

Table 1: Summary of different model based approach

Sr. No.	Approach	Pros	Cons
1.	Expression ratio images based approach [2]	To enhance facial expression mapping with illumination changes	Difficult to deal with different lighting conditions
2.	Morphing based approach [1]	Capture the subtle changes in illumination and appearance that occur as the face is deformed Ability to generate highly realistic face models and natural looking animations	Difficult to deal with different lighting conditions
3.	Multi-linear model based approach [8]	Provide a convenient control of facial attributes It is an inexpensive	less sensitive to localized intensity changes
4.	Deformable Face Model based approach [4,9]	Applicable for Real time application	For more accuracy, require longer computation times
5.	Bilinear model based approach [5]	Applicable in a wide range of applications Capable of representing most expressions of human faces	Require more calculation time.

III. Conclusion

Facial expression transfer is a both challenging and important technique, which is its user-friendliness. In this paper, we have given an introductory survey for the facial expression transfer technology. We hope this paper can provide the readers a better understanding about facial expression transfer, and we encourage the readers who are interested in this topic to go to the references for more detailed study. Facial expression transfer techniques concern how to make facial expression changes. This can be categorized into the Expression image ratio based approach, Multi-linear model based approach, Bi-linear model based approach, and deformable model based approach and the Morph-based approach. This paper addresses the facial performance transfer technique scheme on images or existing video sequence which allows expression transfer on specific parameter such as eyes, lips etc of the face using different models. These system is adopted to support computer vision and in movie industry. Moreover, limitations and advantages of each algorithm is also discussed.

REFERENCES

[1] Frederic Pighin Jamie Hecker Dani Lischinskiy Richard Szeliskiz David H. Salesin, “Synthesizing Realistic Facial Expressions from Phoogtraphs”, ACM, 1998, pp.75-84.

[2] Zicheng Liu, Ying Shan and Zhengyou Zhang, “Expressive Expression Mapping with Ratio Images”, ACM, August 2001, pp.12-17.

[3] Sen Wang, Xianfeng David Gu and Hong Qin , “Automatic Non-rigid Registration of 3D Dynamic Data for Facial Expression Synthesis and Transfer”,IEEE,2008,pp.978-986.

[4] Akshay Asthana, Miles Delahunty, Abhinav Dhall, and Roland Goecke, “Facial Performance Transfer via Deformable Models and Parametric Correspon-dence”, IEEE, September 2012,pp.1511-1519.

[5] Chen Cao, Yanlin Weng, Shun Zhou, Yiying Tong, and Kun Zhou, “FaceWarehouse: A 3D Facial Expression Database for Visual Computing”, IEEE, 2014, pp.413-425.

[6] A Jain, T. Thormahlen, H.P. Seidel, and C.Theobalt, “Movie reshapes: Tracking and reshaping of humans in videos.”ACM Trans. Graph, 29(5), 2010.

[7] Michal Richter, Kiran Varanasi, Nils Hasler, “Real-time Reshaping Of Humans”, In Proc. IEEE Transaction, 2012, pp. 340-347.

[8] Daniel Vlasic, Matthew Brand, Hanspeter Pfister, Jovan Popovic, “Face Transfer with Multi-linear Models”, ACM, 2005, pp.426-433.

[9] Jaewon Sung, Takeo Kanade and Daijin Kim, “ Pose Robust Face Tracking by Combining Active Appearance Models and Cylinder Head Models”, springer, 2008, pp.260-274.

[10] Thomas Albrecht, Marcel Lüthi, Thomas Vetter, “Deformable Model”, 2009, pp.1-7.