An approximative and semi-automated method to create MPEG-4 compliant human face models

Ákos Tóth, Roland Kunkli

Abstract: In this paper, we introduce our method to facilitate the process of creating an MPEG-4 compliant face model based on a simple 3D mesh. The presented method is semi-automatic, and the user needs to choose 56 points on the model as a preprocessing step. We use a cage based deformation technique to approximate an input model with a generic one which already contains the required MPEG-4 parameters. In the paper, we also show how the cage can be constructed to surround the model completely. The resulting model can be used in any MPEG-4 based facial animation player.

Keywords: cage based deformation, MPEG-4, talking head, facial animation, automation

Introduction

The usage of virtual avatars — also called talking heads in the case of eliminating the body part — is widespread in HCI (Human-Computer Interaction), and they may play an essential role in the future as well.

Methods for creating a talking head based on the face of a real or a fictive person typically consist of two main steps. The first one is the creation of the model itself, while the second step is to prepare our model for further facial animations.

Some of the main challenges are the portability and the reusability of the prepared model. Their principles can be very different; therefore, in most cases, the conversion between the outputs is nearly impossible. For solving this problem, many earlier systems [1, 2] support the well-known MPEG-4 facial and body animation standard [7]. Using MPEG-4, we can guarantee that the desired animation can be attached to any standard model automatically, but unfortunately, the usual creation of an MPEG-4 compliant face model requires a lot of manual user interactions. With all this in mind, our goal in this paper is to give a possible solution to this problem by reducing the amount of required interaction in a semi-automatic way.

In the next section, we discuss some previous works related to the area of MPEG-4 facial animation. We highlight their main advantages and disadvantages as well. In Section 3, we present our method, which can provide a solution to some of these mentioned disadvantages, based on a prepared generic model and a cage based mesh deformation technique. Then, in the last section, we demonstrate the results of our algorithm and our future ideas as well.

Previous work

The face model adaptation (parameterization) heavily depends on the topology of the considered mesh. In the case of simple models, which consist of few hundred vertices, it can be executed easily manually. However, if the models are highly detailed, relatively complex, and include thousands of vertices, then the calibration is not practical. Unfortunately, the adaptation of an existing parameterization to a new facial model is not possible. To solve some of these problems (e.g., the parameterization of the models), Sheng et al. have been released a PDE based method [9] but unfortunately, it does not support the MPEG-4 standard.

MPEG-4 compliant deformation-based methods have been also published. Escher et al. have been proposed a solution [3] which can create standard faces by using a generic model. Lavagetto et al. introduced a method [5] which can help in the calibration process. But due to the limitations of the used deformation algorithm (i.e., RBF and Free Form Deformation [10]), the techniques cannot provide a satisfactory deformation of the generic model, especially in the case of the nose, the eyes, and the chin (see Figure 1).

As we discussed in Section , our primary goals are to ease the adaptation and avoid possible errors caused by the users. We suggest a semi-automatic method based on a generic model which has to be deformed in order to approximate the input one. MPEG-4 parameters are already specified on the generic model. Thus, the face model calibration can be skipped. For the manipulation of the generic

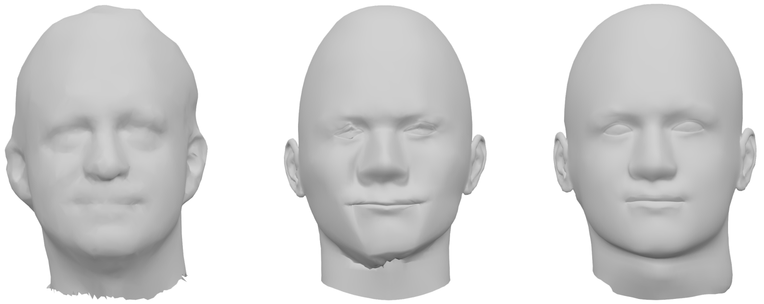

Figure 1: Input model (left), the results of the deformation using Surface-Oriented Free-Form Deformation [10] (middle) and cage based deformation technique (right) which is the basis of our approach.

model, we used a deformation method called cage based deformation technique. One of the main advantages of these methods against other deformation solutions is that using them we can work in real-time, and we have an easy to use, smooth, and intuitive control over the mesh with the defined cage [6].

Our method for automating the face model calibration

As we mentioned in Section , the creation of an MPEG-4 compliant face model requests a lot of time if we do it manually because we need to define all of the 84 feature points and the zone of influence for each FP. Thus, we suggest a semi-automatic solution to create a standard face if its simple 3D mesh is given. At first, we consider a specified generic model together with all its predefined standard points, and then a cage based method is used to approximate the input model with the generic one. Therefore, the whole procedure of the face model adaptation does not have to be executed.

Our deformation method

The necessary modifications of the generic model are achieved by a cage based deformation technique called harmonic coordinates [4]. As we mentioned earlier, a topologically flexible cage (or also known as control mesh) is used to control the deformation of the interior object. If we move the cage vertices C_i to the new positions C_i', an interior point p moves to the new location p', and it is computed as

$$p' = \sum_i h_i(p) C_i',$$

where $h_i(p)$ is the harmonic coordinate of p respect to C_i.

Therefore, at first, we define simple control cages (shown in Figure 2) surrounding the generic and the input models in the same way, based on pre-marked facial feature points on the heads. These cages are responsible for the deformation.

Then, we modify the generic model's cage by translating each point to the corresponding position on the input model's cage separately. As a result, the generic model sufficiently approximates the input one, and it still remains MPEG-4 compliant.

Conclusions and future works

In this paper, we have proposed a robust and semi-automatic method for creating an MPEG-4 compliant face model. Only the definition of the feature points needs user interaction, but it takes just a few minutes. After applying our method, the resulting talking heads can be used in any MPEG-4 compatible facial animation player [8] immediately.

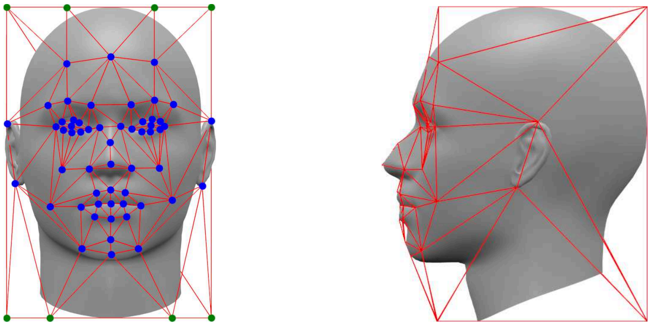

Figure 2: The control cage which has been created for the generic model. Blue dots are the facial feature points of the model, which need to be marked manually, while the green ones are the auxiliary points.

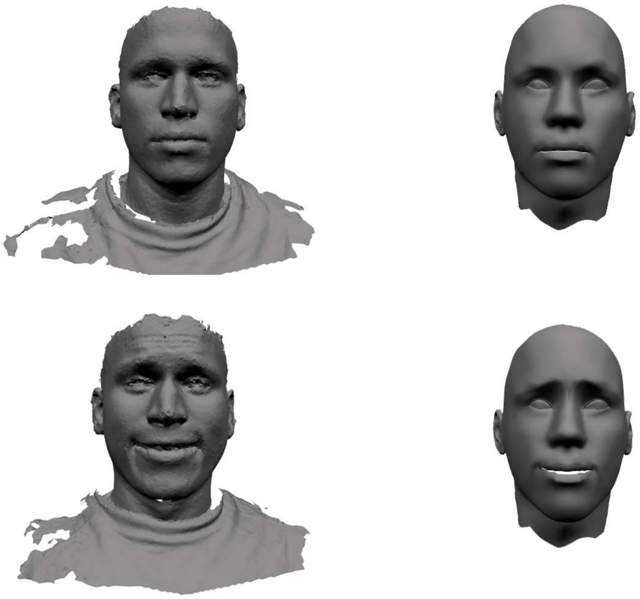

Figure 3: Left: 3D reconstructed human facial expressions from [11]. Right: Our resulting model using the generated FAPs file.

Our solution uses a cage based deformation method called harmonic coordinates to approximate the input model with the generic one. The main difference from earlier systems is that we do not have to define the MPEG-4 parameters. Therefore, calibration errors do not influence the quality of the animation of the model.

To validate our resulting models, we generated FAPs files using 3D reconstructed human facial expressions from the BU-3DFE (Binghamton University 3D Facial Expression) Database [11]. Then these FAPs files were played on our resulting face models to get the same expressions. The original reconstructed model and our deformed model with different facial expressions can be seen in Figure 3.

Additionally, we would like to attempt to minimize the number of user interactions by using a face tracking method which can mark the necessary facial feature positions for us.

Acknowledgement

The first author was supported by the construction EFOP-3.6.3-VEKOP-16-2017-00002. The project was supported by the European Union, co-financed by the European Social Fund.

The second author was supported by the ÚNKP–17–4 New National Excellence Program Of The

Ministry Of Human Capacities.

References

[1] Koray Balci. Xface: MPEG-4 Based Open Source Toolkit for 3D Facial Animation. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 399–402, 2004.

[2] Fiorella De Rosis, Catherine Pelachaud, Isabella Poggi, Valeria Carofiglio, and Berardina De Carolis. From Greta's mind to her face: Modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human Computer Studies*, 59(1–2):81–118, 2003.

[3] M. Escher, I. Pandzic, and N. M. Thalmann. Facial deformations for MPEG-4. In *Computer Animation 98. Proceedings*, pages 56–62. IEEE, 1998.

[4] Pushkar Joshi, Mark Meyer, Tony DeRose, Brian Green, and Tom Sanocki. Harmonic Coordinates for Character Articulation. *ACM Trans. Graph.*, 26(3):71, 2007.

[5] F. Lavagetto and R. Pockaj. The facial animation engine: toward a high-level interface for the design of MPEG-4 compliant animated faces. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(2):277–289, 1999.

[6] Jesús R. Nieto and Antonio Susín. *Cage Based Deformations: A Survey*, pages 75–99. Springer Netherlands, Dordrecht, 2013.

[7] Igor S. Pandzic and Robert Forchheimer, editors. *MPEG-4 Facial Animation: The Standard, Implementation and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 2003.

[8] Roland Rácz, Ákos Tóth, Ildikó Papp, and Roland Kunkli. Full-body animations and new faces for a WebGL based MPEG-4 avatar. In *CogInfoCom 2015: 6th IEEE International Conference on Cognitive Infocommunications*, pages 419–420, 2015.

[9] Yun Sheng, Phil Willis, Gabriela Gonzalez Castro, and Hassan Ugail. *PDE-Based Facial Animation: Making the Complex Simple*, chapter Poster, pages 723–732. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[10] Karan Singh and Evangelos Kokkevis. Skinning Characters using Surface Oriented Free-Form Deformations. In *Proceedings of the Graphics Interface 2000 Conference*, pages 35–42, 2000.

[11] Lijun Yin, Xiaozhou Wei, Yi Sun, Jun Wang, and M. J. Rosato. A 3D Facial Expression Database for Facial Behavior Research. In *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pages 211–216, 2006.