Pacific Graphics (2014)
J. Keyser, Y. J. Kim, and P. Wonka (Editors)

Extraction and Transfer of Facial Expression Wrinkles for Facial Performance Enhancement

Il-Kyu Shin[1], A. Cengiz Öztireli[2], Hyeon-Joong Kim[1,3], Thabo Beeler[4], Markus Gross[2,4], Soo-Mi Choi[1,*]

[1]Sejong University, Seoul, Korea [2] ETH Zürich, Zürich, Switzerland
[3] 3D Systems, Inc, Seoul, Korea [4] Disney Research Zürich, Zürich, Switzerland

Abstract

Creating realistic models and animations of human faces with adequate level of detail is very challenging due to the high complexity of facial deformations and the sensitivity of the human visual system to human faces. One way to overcome this challenge is using face scans captured from real facial performances. However, capturing the required level of detail is only possible with high-end systems in controlled environments. Hence, a large body of work is devoted to enhancing already captured or modeled low resolution faces with details extracted from high fidelity facial scans. Details due to expression wrinkles are of particular importance for realistic facial animations. In this paper, we propose a new method to extract and transfer expression wrinkles from a high resolution example face model to a target model for enhancing the realism of facial animations. We define multi-scale maps that capture the essential details due to expression wrinkles on given example face models, and use these maps to transfer the details to a target face. The maps are interpolated in time to generate enhanced facial animations. We carefully design the detail maps such that transfer and interpolation can be performed very efficiently using linear methods.

1. Introduction

Modeling human faces has been one of the main challenges in many fields including computer graphics. This is due to the importance of faces for all types of communication, the subtle details and complex expressions present, and the very well trained human visual system to capture such facial features. A diverse set of emotions can be conveyed via a face by moving various muscles, which are manifested as wrinkles on the facial surface. Insufficient modeling of these wrinkles often result in emotionless and dull digital face models that human observers cannot relate to. The problem becomes even harder for facial animations where the facial expressions and details dynamically change.

One way to obtain realistic face models and animations is using digital authoring tools to model the facial surface geometry, textures, and materials. The facial geometry is then rigged, generating intuitive parameters for expressive animations. However, even with significant efforts from a well-trained artist, the generated animations can easily become unrealistic due to the well-known uncanny valley effect, as some of the details are very difficult to model and yet perfectly perceivable by human observers. Another path to realistic facial animations is capturing high quality scans of human faces. With the current high-end systems, even subtle details can be captured. However, these systems can be expensive to setup and impractical to use for some applications. Instead, many capturing pipelines use methods such as motion capture from videos or depth maps that can efficiently generate a coarser version of the facial geometry. However, the resulting geometry lacks the perceptually important details and wrinkles. Hence, it is essential to enhance the captured facial geometry with the necessary details and in particular expression wrinkles.

One method for enhancing coarse facial geometry is transferring wrinkles and details from already captured face models [GMP*06, MJC*08, BLB*08, BBB*14]. In this paper, we present an approach specifically designed for transferring expression wrinkles for enhancing facial animations. Given example face models portraying different expressions of a person, we extract and transfer expression wrinkles to a target face. The transferred details are then interpolated for en-

* Corresponding author (smchoi@sejong.ac.kr)

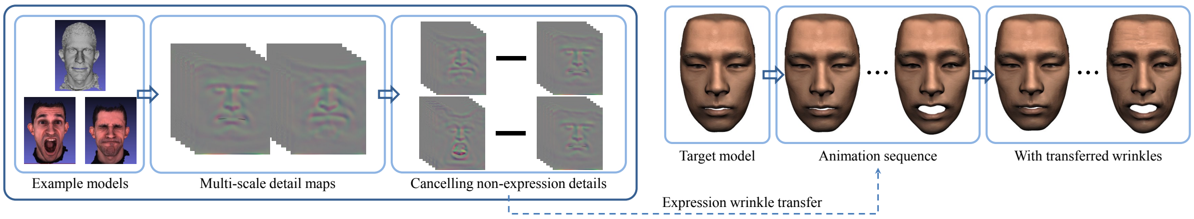

Figure 1: *Overview of our method. Given example face models at different poses, we extract multi-scale detail maps that capture normal changes only due to expression wrinkles. These maps are then used to enhance a coarse output animation sequence.*

hancing a target animation. In order to isolate the high frequency details due to expression wrinkles, we utilize multi-scale detail maps. These maps capture the surface normal perturbations due to expression wrinkles. Our technique offers an easy and efficient way of quick prototyping of facial animations with expressive details.

Overview An overview of our technique is given in Figure 1. The example face models for different expressions and the target model are first parametrized and normal difference maps that capture the changes in the surface normal at different scales are extracted. These maps contain fine to middle scale details due to wrinkles and pores. Next, we cancel the details of the facial geometry in the neutral pose to keep only the normal changes due to expression wrinkles. To retarget these details to another face, we first transfer the large scale deformation of a chosen expression of the example face to the target neutral face. We then augment the resulting deformed target face with the computed detail maps. Finally, the detail maps are interpolated in time in synch with the large scale deformation to generate an animation of the chosen expression.

2. Related Work

Our method is related mainly to facial detail transfer techniques and also to other methods of synthesizing face models and wrinkles. Below, we review the most relevant works.

Facial detail transfer Transferring details from a source to a target shape is a well studied problem for general digital 3*D* models. The details can be extracted from a single model via local vertex displacements and coordinates [SCOL*04], or can be encoded with respect to a reference pose [SP04]. These methods can be specialized for faces by coupling the details on a face to a low frequency representation of the same face. Explicit smoothing can be used to generate a database of such pairs [BBB*14], which can then be used to enhance captured low detail face models. It is also possible to relate details to facial rigs [ARL*10], skin strain [BLB*08], or motion capture data [MJC*08]. In order to cancel the dependency of details on facial features, specialized local statistics can be utilized [GMP*06]. Our method also relies on the idea of representing details as a function of facial geometry. However, we develop a technique tailored to expression wrinkles that allows efficient transfer of relevant details for fast animation prototyping.

Facial deformation transfer In addition to details, low frequency deformations can also be transferred from an input to an output face to reuse existing facial animations. Deformation transfer can be performed by utilizing various tools such as deformation gradients [SP04], vertex motion vectors [NN01], or radial basis function based parametrization [PKC*03]. Although these methods work very well for large scale deformations, they are not suited for transfer of details due to expression wrinkles.

Wrinkle generation Enhancing a coarse face model can also be achieved by synthesizing wrinkles in a manual or semi-automatic way. A common approach used for modeling wrinkles is allowing the user to specify the wrinkle center-lines, profile shapes, and other parameters [BKN02, LC04, LYKL07, LLLC11, JEOG11]. To facilitate physically correct behavior of expression wrinkles, models of internal structures such as facial muscles and elastic fibers can also be utilized [Wat87, TW90, WMTT94, ZPS04]. These methods can generate detailed wrinkles on a given face but are not designed for transfer and automatic synthesis of such facial details.

3. Method and Results

3.1. Extraction of the Detail Maps

Facial details such as bulging due to wrinkles and pores can be extracted by creating an approximate version of a face without such details via smoothing, and using a notion of difference between the original and the smoothed versions of the face [GMP*06, BBB*14]. Previous works represented this difference via displacement maps [GMP*06] or deformation gradients [BBB*14]. Although representation by deformation gradients allows operations such as linear combinations easily and accurately, it is computationally more expensive to especially reconstruct the face model from the modified gradients and requires care to accurately represent the fine-scale details [BBB*14].

Instead, we use a normal map based representation, which allows transfer and linear interpolation of details efficiently. In order to use normal maps, we first parametrize all example faces E_i and the target face F (each of which contain around 1 million vertices). We then establish correspondences among them via topology templates that segment the facial surface into semantically meaningful regions (Figure 2, a). To this end, for each facial model, the locations of the vertices of the topology template are coarsely aligned to the facial geometry manually. The connectivity of the template is the same for all faces and only the vertex locations change, providing a natural correspondence for template vertices. Barycentric interpolation is utilized at each patch of the topology template to extend these correspondences to the in-patch vertices of the facial models. Hence, we obtain parametrized normal maps for all faces (Figure 2, b) in correspondence. Note that we mask out the nose, eyes, and lips regions in the topology template since they do not contain any expression wrinkles.

To generate the maps capturing the details in an example face, we smooth the normal map $\mathbf{N}$ of the example face by convolving it with Gaussians of increasing standard deviations to obtain $\mathbf{N}_s$, $s = 0 \cdots n$ with $\mathbf{N}_0$ representing the original normal map (we used $n = 3$ scales in this paper). Next, we compute the difference between normal maps at consecutive scales to extract the details at different scales. This difference is represented via the quaternions $\mathbf{q}_s(\mathbf{x})$, $s = 1 \cdots n$ that measures how much the normal $\mathbf{n}_{s-1}(\mathbf{x}) \in \mathbf{N}_{s-1}$ at a point $\mathbf{x}$ is rotated with respect to $\mathbf{n}_s(\mathbf{x}) \in \mathbf{N}_s$. (Figure 3, left). We illustrate some of the generated detail maps in Figure 3, right. As shown, the scales capture overall wrinkle shapes at different detail levels.

Canceling non-expression wrinkle details The extracted maps contain details due to expression wrinkles as well as other face specific details such as pores and aging wrinkles. In order to cancel this dependency on the example face, we subtract the maps $\mathbf{q}_s^0$ of the neutral example face E_0 from those $\mathbf{q}_s^i$ of the example faces E_i with expressions. Subtraction in this context means applying the inverse of the rotations that are due to the non-expression wrinkles, which are stored in $\mathbf{q}_s^0$. Hence, the new detail maps are given by $\mathbf{q}_s^i \leftarrow \mathbf{q}_s^i \mathbf{q}_s^{0*}$, with $*$ denoting the conjugate. Figure 4 illustrates the effect of canceling these details. With the non-expression details canceled, the aging wrinkles in the forehead, around the eyes, or under the mouth are not transferred to the output model, as expected.

3.2. Detail Transfer and Interpolation

In order to transfer an expression from an example face E_i to the target face F, we start by transferring the large scale deformation that captures the overall facial shape [BLB*08]. Although the resulting deformations correctly mimic the example expressions, they lack the important details due to the wrinkling of the face as illustrated in Figure 4.

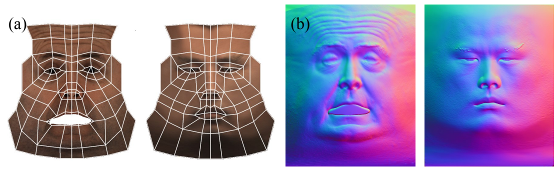

Figure 2: *(a) Topology templates overlaid for two faces, and (b) their normal maps.*

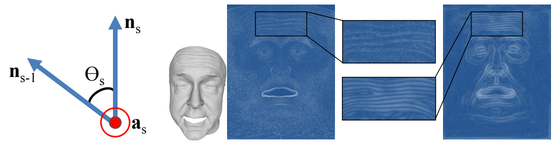

Figure 3: *(Left) Detail maps capture differences between normals at consecutive scales given by the quaternions* $\mathbf{q}_s = \cos(\theta_s/2) + \mathbf{a}_s \sin(\theta_s/2)$. *(Right) A face and generated maps for two different scales (only the angle θ_s of the quaternion $\mathbf{q}_s$ is plotted).*

Detail transfer We transfer such details via the extracted detail maps $\mathbf{q}_s^i$. This is achieved by using the computed quaternions at multiple scales to rotate the surface normals of the target face F. An example result of such a transfer is illustrated in Figure 4. Note that due to the canceling of the face specific details not related to the expression wrinkles, and also the semantic correspondence between the example and target faces given by the topology templates, the expression wrinkles are accurately retargeted. Further examples are shown in Figure 5. Transferring only the large scale deformation results in dull expressions. In contrast, our method is able to transfer wrinkles essential for conveying the desired expressions and realism.

Animation via interpolation The quaternions from the detail maps can also be interpolated linearly to get in-between details, which are transferred to F to generate facial animations, similar to previous methods that blend wrinkle maps [Oat07]. We illustrate frames from such animations where the expression is transitioning from surprise to frowning and surprise to smiling in Figure 5 (please also refer to the accompanying video for animations). A natural transition is obtained between the two expressions.

3.3. Implementation

The example and target faces are obtained by utilizing a high resolution scanning system [BBB*10]. It is especially important to have high quality example face geometries, since they need to contain the required details. After all faces are parametrized and the topology templates are aligned as explained in Section 3.1, all of the following steps (except large scale deformation transfer) are performed based on 2D maps using GPU shaders. In a pre-processing step, for all

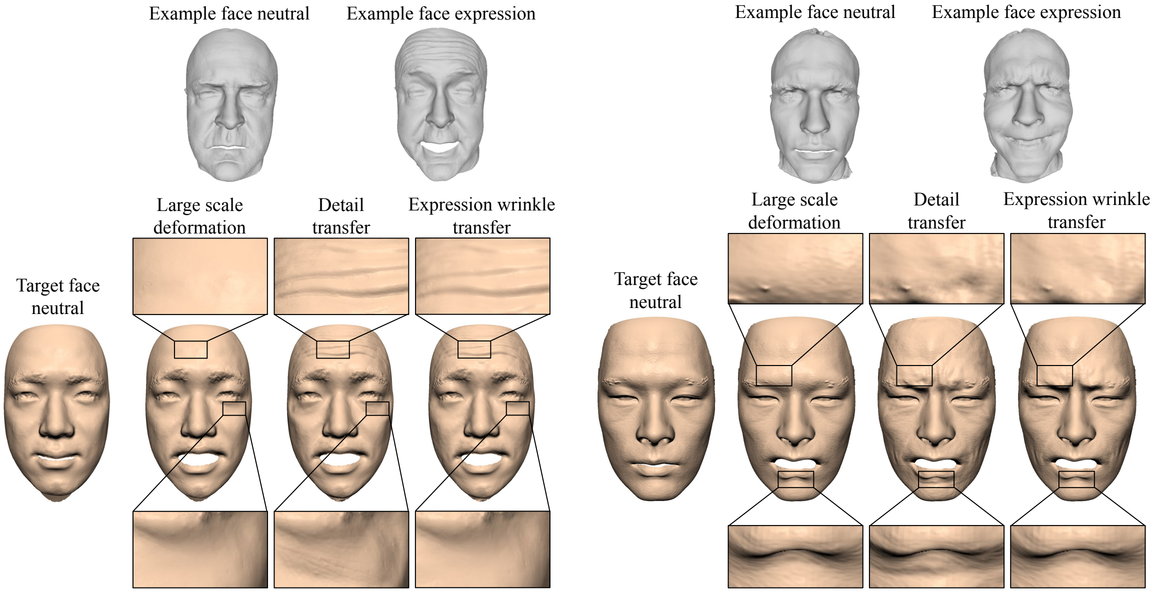

Figure 4: *Transferring only the large scale deformation cannot preserve the expression wrinkles, while trying to retarget the missing details via raw detail maps results in also transferring the aging wrinkles and other face specific details. Our method transfers only the expression related details required for realism and vividity of the expressions.*

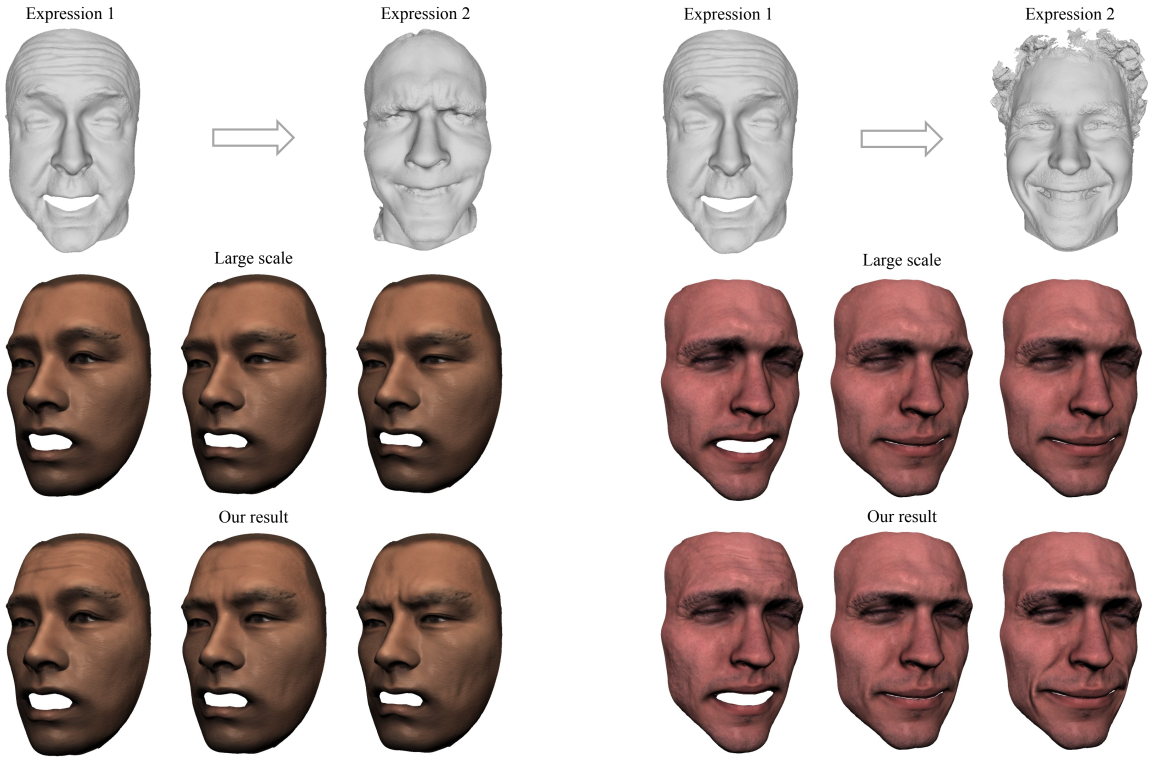

Figure 5: *Frames from animations, where the target faces are transitioning from expression* 1 *to expression* 2*.*

example faces, the in-patch vertex correspondences are first computed by interpolation, and quaternion maps that capture expression wrinkles at multiple-scales are computed and saved. The user can then select a parametrized target face and desired expressions for different time frames. The system first transfers large scale deformations to the target face for all time frames including the in-between frames. Finally, expression wrinkles extracted from the indicated example faces and interpolated in time are transferred.

4. Conclusions and Future Work

In this paper, we presented an approach for transferring facial detail resulting from expression wrinkles to enhance a coarse animation of a target face. Our approach is based on representing such details via quaternions measuring rotations in surface normals at different scales. This allow efficient operations for transfer and interpolation of details for fast prototyping of realistic animations. However, there are several limitations, open problems and directions to pursue.

Adapting details for facial characteristics We cancel the dependency of the details of a face on the face geometry by computing differences between different scales, omitting non-expression related details, and establishing semantically meaningful correspondences among faces. However, as observed in previous works [GMP*06, BBB*14], transfer of some fine-scale details might require further measures to make the details better conform to the target facial geometry. We will investigate statistical measures and texture synthesis methods that incorporate our detail maps. We believe that this will also help to eliminate the loss of some small scale details of the transferred expression wrinkles in the current method.

Selecting and exploiting wrinkle scales In this paper, only three scales that can be easily set to roughly capture the small, middle, and large scale details are used. Utilizing more scales would require a more careful scale selection possibly adapted to the facial geometry and expressions. We used the multi-scale nature of the detail maps mainly when canceling the non-expression related details. This multi-scale representation can be further utilized to, for example, selectively transfer, amplify, or interpolate details at particular scales.

Controlling the transfer and animations Similar to previous works, our method does not allow full artistic control over the transferred details. We would like to investigate intuitive editing of details and timing of animations for more controllable transfer.

Acknowledgments

This work was supported by Korean NRF and Swiss SER under the Korean-Swiss Cooperative Program (2013K1A3A1A14055180).

References

[ARL*10] ALEXANDER O., ROGERS M., LAMBETH W., CHIANG J.-Y., MA W.-C., WANG C.-C., DEBEVEC P.: The digital emily project: Achieving a photorealistic digital actor. *IEEE Computer Graphics and Applications 30*, 4 (2010), 20–31. 2

[BBB*10] BEELER T., BICKEL B., BEARDSLEY P., SUMNER B., GROSS M.: High-quality single-shot capture of facial geometry. *ACM Trans. on Graphics (Proc. SIGGRAPH) 29*, 3 (2010), 40:1–40:9. 3

[BBB*14] BERMANO A. H., BRADLEY D., BEELER T., ZUND F., NOWROUZEZAHRAI D., BARAN I., SORKINE-HORNUNG O., PFISTER H., SUMNER R. W., BICKEL B., GROSS M.: Facial performance enhancement using dynamic shape space analysis. *ACM Trans. Graph. 33*, 2 (Apr. 2014), 13:1–13:12. 1, 2, 5

[BKN02] BANDO Y., KURATATE T., NISHITA T.: A simple method for modeling wrinkles on human skin. In *Proc. of the 10th Pacific Conference on Computer Graphics and Applications* (Washington, DC, USA, 2002), PG '02, IEEE Computer Society, pp. 166–175. 2

[BLB*08] BICKEL B., LANG M., BOTSCH M., OTADUY M. A., GROSS M.: Pose-space animation and transfer of facial details. In *Proc. SCA 2008* (Aire-la-Ville, Switzerland, Switzerland, 2008), SCA '08, Eurographics Association, pp. 57–66. 1, 2, 3

[GMP*06] GOLOVINSKIY A., MATUSIK W., PFISTER H., RUSINKIEWICZ S., FUNKHOUSER T.: A statistical model for synthesis of detailed facial geometry. In *ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), SIGGRAPH '06, ACM, pp. 1025–1034. 1, 2, 5

[JEOG11] JIMENEZ J., ECHEVARRIA J. I., OAT C., GUTIERREZ D.: *GPU Pro 2*. AK Peters Ltd., 2011, ch. Practical and Realistic Facial Wrinkles Animation, pp. 15–27. 2

[LC04] LARBOULETTE C., CANI M.-P.: Real-time dynamic wrinkles. In *Proc. CGI* (Washington, DC, USA, 2004), CGI '04, IEEE Computer Society, pp. 522–525. 2

[LLLC11] LI L., LIU F., LI C., CHEN G.: Realistic wrinkle generation for 3d face modeling based on automatically extracted curves and improved shape control functions. *Computers & Graphics 35*, 1 (2011), 175 – 184. Extended Papers from Non-Photorealistic Animation and Rendering (NPAR) 2010. 2

[LYKL07] LI M., YIN B., KONG D., LUO X.: Modeling expressive wrinkles of face for animation. In *Proc. of the Fourth International Conference on Image and Graphics* (Washington, DC, USA, 2007), ICIG '07, IEEE Computer Society, pp. 874–879. 2

[MJC*08] MA W.-C., JONES A., CHIANG J.-Y., HAWKINS T., FREDERIKSEN S., PEERS P., VUKOVIC M., OUHYOUNG M., DEBEVEC P.: Facial performance synthesis using deformation-driven polynomial displacement maps. *ACM Trans. Graph. 27*, 5 (Dec. 2008), 121:1–121:10. 1, 2

[NN01] NOH J.-Y., NEUMANN U.: Expression cloning. In *Proc. of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2001), SIGGRAPH '01, ACM, pp. 277–288. 2

[Oat07] OAT C.: Animated wrinkle maps. In *ACM SIGGRAPH 2007 Courses* (New York, NY, USA, 2007), SIGGRAPH '07, ACM, pp. 33–37. 3

[PKC*03] PYUN H., KIM Y., CHAE W., KANG H. W., SHIN S. Y.: An example-based approach for facial expression cloning. In *Proc. SCA 2013* (Aire-la-Ville, Switzerland, Switzerland, 2003), SCA '03, Eurographics Association, pp. 167–176. 2

[SCOL*04] SORKINE O., COHEN-OR D., LIPMAN Y., ALEXA

M., RÖSSL C., SEIDEL H.-P.: Laplacian surface editing. In *Proc. SGP 2004* (New York, NY, USA, 2004), SGP '04, ACM, pp. 175–184. 2

[SP04] SUMNER R. W., POPOVIĆ J.: Deformation transfer for triangle meshes. *ACM Trans. Graph. 23*, 3 (Aug. 2004), 399–405. 2

[TW90] TERZOPOULOS D., WATERS K.: Physically-based facial modelling, analysis, and animation. *The Journal of Visualization and Computer Animation 1*, 2 (1990), 73–80. 2

[Wat87] WATERS K.: A muscle model for animation three-dimensional facial expression. In *Proc. of the 14th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1987), SIGGRAPH '87, ACM, pp. 17–24. 2

[WMTT94] WU Y., MAGNENAT THALMANN N., THALMANN D.: A plastic-visco-elastic model for wrinkles in facial animation and skin aging. In *Proc. of the second Pacific conference on Fundamentals of computer graphics* (River Edge, NJ, USA, 1994), Pacific Graphics '94, World Scientific Publishing Co., Inc., pp. 201–213. 2

[ZPS04] ZHANG Y., PRAKASH E. C., SUNG E.: A new physical model with multilayer architecture for facial expression animation using dynamic adaptive mesh. *IEEE Trans. on Visualization and Computer Graphics 10*, 3 (May 2004), 339–352. 2