

Facial Aging Simulation by Patch-Based Texture Synthesis with Statistical Wrinkle Aging Pattern Model

Akinobu Maejima, Ai Mizokawa, Daiki Kuwahara
and Shigeo Morishima

Abstract We propose a method for synthesizing a photorealistic human aged-face image based on the patch-based texture synthesis using a set of human face images of a target age. The advantage of our method is that it synthesizes an aged-face image with fine skin texture such as spots and pigments of facial skin, as well as age-related facial wrinkles without blurs (such as those resulting from lack of accurate pixel-wise alignments as in the linear combination model) while maintaining the quality of the original image.

Keywords Automatic · Facial aging · Image patch · Texture synthesis · Wrinkle · Agind pattern · Statistical model

1 Introduction

Predicting the current facial appearance of wanted or missing people who have been missing for several years is an important task in criminal investigation. Previously, sketches or photomontages have been used for this purpose. However, it is not easy to predict and depict one's realistic facial appearance only from photographs that are several years old and some interviews, because the quality of the resulting sketch or montage highly depends on the skill of the forensic artists. Thus, an automatic facial aging simulator, which can synthesize a photorealistic facial appearance based on statistics from the actual aging process, is required.

A. Maejima (✉) · A. Mizokawa · D. Kuwahara
Waseda University, 3-4-1, Okubo, Shinjuku-ku, Tokyo 169-8555, Japan
e-mail: akinobu@mlab.phys.waseda.ac.jp

S. Morishima
Waseda Research Institute for Science and Engineering, Waseda University, 3-4-1,
Okubo, Shinjuku-ku, Tokyo 169-8555, Japan
e-mail: shigeo@waseda.jp

Automatic facial aging simulation methods have been proposed by many researchers [2, 3, 5, 6, 9, 10, 12]. Most of the conventional methods adopts a linear combination model such as active appearance model [2, 3, 5, 6, 9] or 3D morphable model [10] to parameterize both facial geometry and appearance. However, in the case of linear combination, which relies on bases, the resulting image is somewhat blurred because the alignment between features, which appear for different individuals or ages is impossible. Burt and Perrett reported that perceived age of transformed faces by incorporating the difference between composite faces from different age groups onto real face images. Moreover, they described that the averaging operation involved in creating composite faces was observed to smoothen facial creases and result in composite faces appearing younger than those from their own age groups. Therefore, it is important to represent fine age-related features such as facial wrinkles, pores, pigments which are strongly affected by age perception [1].

To solve this problem, Tazoe et al. have proposed patch-based facial aging texture synthesis using human face images of the target age [12]. This method is based on the assumption that if we accurately reconstruct an input face using image patches of the target age, the resulting face would be the same individual's aged-face. Actually, their method can generate a photorealistic face image with fine age-related features. However, the resulting face sometimes does not look like the original person because the original face has been completely replaced by the reconstructed face. Moreover, age-related features such as wrinkles may be incorrectly expressed depending on the pixel intensity pattern of a patch in the input image.

In this chapter, we propose an automatic facial aging simulation method, which can overcome the above mentioned problems. Our method is based on patch-based texture synthesis using facial images of the target age. Using the statistical wrinkle aging pattern model, we predict the resulting facial wrinkle appearance (shape, depth, and length) of the target age. Moreover, we introduce a modified Poisson solver to seamlessly merge between image patches, and to keep the original facial appearance in the region which influences the individual perception and skin tone.

The principal contribution of this paper is to provide a simulator, which can synthesize photorealistic aged-face images with detailed textures such as spots and pigments of facial skin, as well as age-related facial wrinkles, while maintaining the identity of the original face.

2 Overview

Our facial aging simulation pipeline is depicted in Fig. 1. The facial aging simulation begins with the pre-processing to create patch libraries for age groups and to train the Statistical Wrinkle Aging Pattern Model. Then given an input face image, the actual simulation is done by the runtime processing.

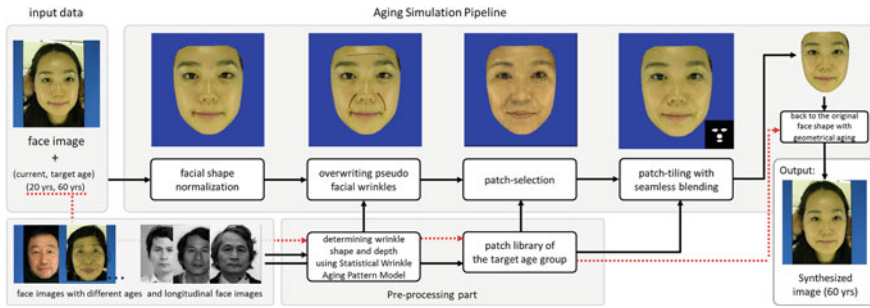


Fig. 1 Our facial aging simulation pipeline. Note that, any images of the target person are not contained in our database for creating path libraries and training the Statistical Wrinkle Aging Pattern Model

Preprocessing:

We collect face images corresponding to various ages, and create an average shaped face model for all faces. Then, shape normalized face image for each face image is generated by the same procedure as the facial shape normalization phase in the runtime processing. Then, we divide shape normalized face images into small square patches with some overlap between neighbors (in this paper, above and to the left). In this paper, the size of each patch is 42 by 42 pixels for 300 by 300 pixel face image and Each patch overlaps each other by 10 pixels. At the same time, an index representing the original position before cutting out are associated with each patch. Then, we construct patch library of age groups for which patches are grouped according to the labeled index. Also, the statistical wrinkle aging pattern model is trained by using longitudinal facial images (details are described in Sect. 4).

Runtime processing:

Shape normalization To remove geometric difference between individuals and to concentrate texture variations, we need to normalize an input facial shape to the average shape. First, 89 facial feature points including eyes, eye brows, nose, lips, and facial contour are detected from the input image using Zhang’s technique [13]. Second, the average face model is deformed using Radial Basis Functions [4] so that the vertices of the face model are matched to corresponding detected feature points. Finally, the deformed face model with the input image is rendered into the image. As a result, we can obtain a face image that the shape is normalized to that of the average face.

Overwriting pseudo facial wrinkles Pseudo facial wrinkles generated by the statistical wrinkle aging pattern model (details are described in Sect.4) consisted of curves representing wrinkles from individual face images over ages are overwritten onto the shape normalized image in order to simulate individual aging, especially with facial wrinkle, which significantly influence human perception of age.

Patch selection and tiling The shape normalized face image is reconstructed by patch-based texture synthesis using patches of the target age (details are described in Sect. 3).

Back to the original shape and geometric aging The shape of the resulting face is deformed toward that of the original face by performing an inverse operation of the facial shape normalization phase. At this time, to represent geometric aging effect, the facial shape is deformed using Tazoe’s technique [12]. The complete aged-face image is generated by embedding the resulting face into the input face image.

3 Patch-Based Texture Synthesis Using Age-Specific Patches

The shape normalized face image is reconstructed by patch-based texture synthesis using patches of the target age. The patch-based texture synthesis can be done by exhaustive search for the patch library at each labeled index (i, j) and tiling selected patches with minimum energy in the raster scan order. The energy function E is defined as weighted sum of the fitness term E_g and the regularization term E_l as shown in Eq. (1).

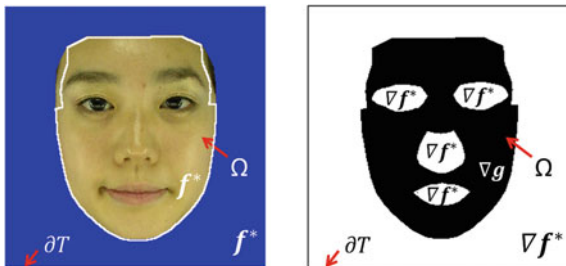
$$E(i, j, n) = \alpha * E_g(i, j, n) + (1 - \alpha) * E_l(i, j, n) \quad (1)$$

where n represents an unknown patch id in the library, which we would like to find at index (i, j) , E_g is the fitness term, which is defined as sum of squared difference between two feature vectors of a shape normalized face image and patches at index (i, j) . $\Omega_{i,j}$ represents the region of an indexed patch at (i, j) . \mathbf{I} means a feature vector of a normalized face image and \mathbf{P} is that of the patch of the target age. In this paper, a feature vector consists of the vector representation of the RGB pixel intensity $(r, g, b)^T$ and Laplacian of Gaussian filtered response l at pixel (x, y) of a patch, for example, $\mathbf{I} = (r, g, b, l)^t \in \mathfrak{R}^4$ (t is transpose).

$$E_g(i, j, n) = \sum_{x,y \in \Omega_{i,j}} \|\mathbf{I}(x, y) - \mathbf{P}_{i,j}^n(x, y)\|^2 \quad (2)$$

E_l is a regularization term, which preserves the spatial coherency between the selected patch and its neighboring patches. In other words, this term preserve visual consistency of the resulting face. Obviously, E_l is calculated by sum of squared difference between feature vectors that both overlap regions between the selected patch and its neighboring patches. In this paper, neighboring patches are on its above $(i, j - 1)$ and left $(i - 1, j)$ locations for (i, j) . In our implementation, the squared difference between all possible combinations of patches in each overlap region are calculated and stored at the pre-processing stage. Optimal patches with minimum energy are selected by exhaustive search for the patch library at each index (i, j) in the raster scan order. Of course, overall patch selection result is affected by the processing order depending on the weight for regularization term. Therefore, we need to adjust the weight α so that patches from different persons can be chosen

Fig. 2 Our modified Poisson settings



as possible, while preserving visual consistency. We set α to 0.8 empirically (All images in this paper were created by using same α).

Finally, the shape normalized face image is reconstructed by tiling selected patches from top-left to the bottom-right. In this patch-tiling process, the modified Poisson solver [11] seamlessly merges selected patches. Unlike seamless cloning algorithm with naive Poisson [7], the modified Poisson solver can preserve the color of a source image in the composite image by controlling the color preserving parameter. Assuming that \mathbf{f} is the vector representation of the desired pixel intensity, \mathbf{r} is also the same representation of the composed image intensity, and \mathbf{v} is the *guidance* vector field. The modified Poisson problem can be represented by the following equation.

$$\min_f \left(\int_T \|\text{div } \mathbf{v} - \Delta \mathbf{f}\|_2^2 dt + \varepsilon \int_T \|\mathbf{r} - \mathbf{f}\|_2^2 dt \right) \tag{3}$$

where, T is the region for the whole image, and ε is the color preserving parameter, which preserves the color of a source image in the resulting composite image. If we set a small ε , the resulting image would be affected by the composite image. In our case, we would like to preserve the original skin tone of the input face as much as possible rather than preserving the color of source image. Therefore, we modify the original equation proposed by Tanaka et al. [11] by replacing \mathbf{r} with \mathbf{f}^* which represents the corresponding intensity of the shape normalized face image (Our setting is depicted as shown in Fig. 2):

$$\min_f \left(\int_T \|\text{div } \mathbf{v} - \Delta \mathbf{f}\|_2^2 dt + \varepsilon \int_T \|\mathbf{f}^* - \mathbf{f}\|_2^2 dt \right) \tag{4}$$

By discretizing Eq. (4) and vanishing the derivative for a pixel value f_p at a pixel p , we can obtain the following equation.

$$(\varepsilon + |N_p|)f_p - |N_p| \sum_{q \in N_p} f_q = \varepsilon f_p^* - |N_p| \sum_{q \in N_p} v_{pq} \tag{5}$$

where f_p is the pixel intensity at a pixel p and ε is the weight, which decides the effect from an input image. If we set a small ε , the resulting image would be affected

by the color of the reconstructed face image. N_p is a set of neighboring pixels at pixel p , and $|N_p|$ is the number of neighbors. Also, v_{pq} is defined by the following equations.

$$v_{pq} = \begin{cases} g_p - g_q & \text{if } p \in \Omega \\ f_p^* - f_q^* & \text{otherwise} \end{cases} \quad (6)$$

where, g_p and g_q represent the pixel intensity at p and q of the reconstructed face image, Ω also represents the region in which the gradients can be transferred (in other words, age-related characteristics can be reflected). We set Ω containing the eyes, nose, and lips to retain the identity of the original face. We consider Eq. (5) for all pixels in the entire image and solve a sparse linear system. Finally, we can obtain the intensity f for each pixel.

4 Statistical Wrinkle Aging Pattern Model

In the texture synthesis phase, occurrence and properties such as shape, depth and length, and the number of wrinkles, pores and pigments of facial skin depend entirely on the pixel intensity pattern in the focusing patch. As mentioned before, it is important to represent fine age-related features such as facial wrinkles, pores, pigments which are strongly affected by age perception [1]. In this paper, we introduce a Statistical Wrinkle Aging Pattern Model (SWAPM) that can implicitly predict the wrinkle appearance, including shape, depth, and length. The SWAPM provides a cue where the appropriate patch can be selected in the patch selection process. The SWAPM can be constructed by carrying out the following procedures using a longitudinal facial image database such as the MORPH database [8]:

- (a) Manually marking on left-right laugh lines, left-right crow's feet, and facial wrinkles on left-right orbits and forehead for each image in the database.
- (b) Approximating each marked line using a parametric curve and parameterizing each wrinkle by acquired parameters and wrinkle density of each facial wrinkle. More specifically, we use a Ferguson curve for this approximation. The position and velocity at the start/end points of the approximated curve and depth that are decided based on the intensity distribution around the approximated curve are stacked into a vector form for all approximated curves. We refer to the resulting vector as the wrinkle vector for an individual.
- (c) Collecting all wrinkle vectors at existing ages in the database and train the aging pattern model in the same manner as the work by Park et al. [5] using principal component analysis. We refer to the resulting linear combination model as the SWAPM.

By changing the coefficient of the SWAPM, we can change the shape, depth, and length of the wrinkles of the model as shown in Fig. 3. The patch selection result can be modulated by adding or subtracting the pixel intensity from resulting wrinkles

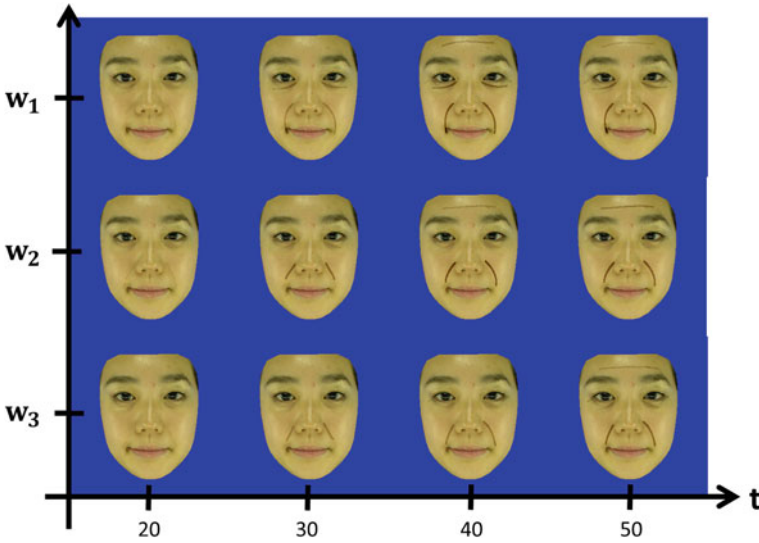


Fig. 3 Changing coefficients of SWAPM. The *vertical* axis corresponds different types of coefficient of SWAPM and the *horizontal* axis means age groups. The *brown lines* represent pseudo facial wrinkles generated by SWAPM. We found that shape, depth, and length of facial wrinkles vary depending on coefficients across the ages

to a shape normalized image. We demonstrate variations of aging simulation for an individual corresponding to Fig. 3 in Fig. 4. We found that our model can describe the variation of aging process for facial wrinkles implicitly.

5 Result and Discussion

Aging simulation results from 10 years old to 60 years for each individual old are shown in Fig. 5. Each row represents aging process for each individual and Each red outlined image represents the original image for each individual. As shown in Fig. 5, our method can represent fine-scale spots and pigments of facial skin that are difficult to represent using previous methods, as well as age-related facial wrinkles. In addition, unlike previous methods, our method can utilize different image databases for texture synthesis and aging pattern learning. As for texture synthesis, we use facial images of peoples of different ages taken under controlled environment. As for learning SWAPM, we utilize longitudinal facial images for each individual taken under different conditions for facial expression, posture, illumination, and resolution. Thus, we can preserve the quality of resulting texture. In general, it is usually hard or even impractical to collect a large database of large amount of individuals who can provide a series of individual images in different ages under the same shooting condition. Therefore, this point is also advantage of our method.

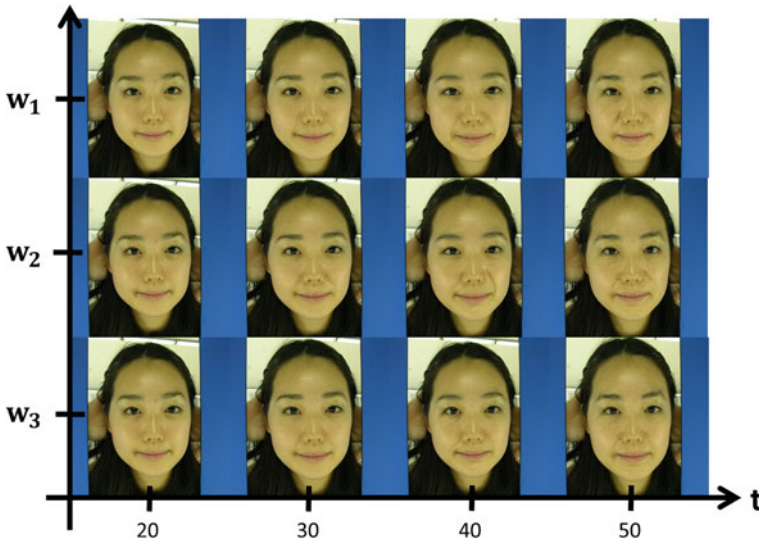


Fig. 4 Result of patch-based texture synthesis for Fig. 3. Similarly, the *vertical* axis corresponds different types of coefficient of SWAPM and the *horizontal* axis means age groups. We observed that variations of wrinkle aging pattern can be represented by SWAPM implicitly

Performance For every synthesized image in Fig. 5, shape normalization and determining wrinkle shape and depth and overwriting pseudo facial wrinkles 0.81 s, patch-based texture synthesis including patch selection and tiling 0.97 s, and geometric deformations 1 s. Entire process takes around 2.78 s in our experiments. Timings are executed on a 2.7 Ghz Intel Core i7 with 8 GB RAM (2011 VAIO Z).

Limitations In patch-based texture synthesis, original features such as moles are sometimes missing or placed in other locations. This is because the gradients of the original image have been completely replaced by those of patches, which consist from others except for the eyes, nose, and mouth region. However, this problem is easy to solve if some user interaction is allowed. More specifically, we manually specify the region Ω in which we would like to retain the identity, in Eq. (6). In addition, we need to consider for optimal size and shape of a patch to improve the quality of synthesized texture (Fig. 5).

Future works The capability to represent other aging effects including increasing/decreasing weight and changes in one's hair is one of the major concerns to be addressed in future work. Moreover, we plan to improve our facial texture synthesis for large pose and illumination changes between an input image and the patch library, and to evaluate the performance of our method on public facial databases like MORPH [8] database by performing a comparison with conventional methods.

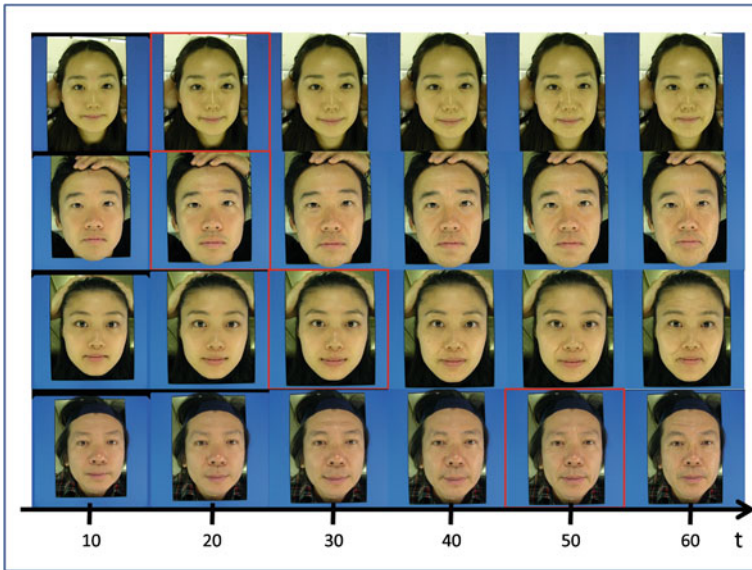


Fig. 5 Aging simulation result from 10 years old to 60 years old. Each *red* outlined image represents the original image for each individual. We observed that our method can represent fine-scale features like spots and pigments of facial skin that it are difficult to represent using previous methods

Acknowledgments This work was supported by the “R&D Program for Implementation of Anti-Crime and Anti-Terrorism Technologies for a Safe and Secure Society”, funds for integrated promotion of social system reform and research and development of the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

References

1. Burt M, Perrett DI (1995) Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour information. *J Roy Soc* 259:137–143
2. Geng X, Zhou Z-H, Smith-Miles KF (2007) Automatic age estimation based on facial aging patterns. *IEEE Trans PAMI* 29(12):2234–2240
3. Lanitis A, Taylor CJ, Cootes TF (2002) Toward automatic simulation of aging effects on face images. *IEEE Trans PAMI* 24(4):442–455
4. Noh J-Y, Fidaleo D, Neumann U (2000) Animated deformations with radial basis functions. In: *Proceedings of ACM Symposium on VRST’00*, pp 166–174
5. Park U, Tong Y, Jain AK (2010) Age invariant face recognition. *IEEE Trans PAMI* 32(5): 947–954
6. Patterson E, Ricanek K (2006) Automatic representation of adult aging in facial images. In: *Proceedings of the 6th IASTED international conference on VIIP 2006*, pp 171–176
7. Pérez P, Gangnet M, Blake A (2003) Poisson image editing. *ACM Trans Graph* 22(3):313–318
8. Ricanek K Jr, Tesafaye T (2006) MORPH: a longitudinal image database of normal adult age-progression. In: *IEEE 7th international conference on FG 2006*, pp 341–345

9. Suo J, Chen X, Shan S, Gao W, Dai Q (2012) A concatenational graph evolution aging model. *IEEE Trans PAMI* 34(11):2083–2093
10. Scherbaum K, Sunkel M, Seidel H-P, Blanz V (2007) Prediction of individual non-linear aging trajectories of faces. *Comput Graph Forum* 26(3):85–294
11. Tanaka M, Kamio R, Okutomi M (2012) Seamless image cloning by a closed form solution of a modified poisson problem. In: *ACM SIGGRAPH ASIA 2012, Posters*, 15
12. Tazoe Y, Gohara H, Maejima A, Morishima S (2012) Facial aging simulator considering geometry and patch-tiled texture. In: *ACM SIGGRAPH 2012, Posters*, 46
13. Zhang L, Tsukiji S, Ai H, Lao S (2005) A fast and robust automatic face alignment system. In: *IEEE international conference on ICCV 2005 (Demo Program)*