A Simple Face Aging/Reverse-Aging Synthesis Method Using Log-Gabor Wavelet

Ching-Tang Hsieh*, Chia-Shing Hu and Yu-Cheng Lee

Department of Electrical Engineering, Tamkang University,
Tamsui, Taiwan 251, R.O.C.

Abstract

A face aging/reverse-aging synthesis method based on face detection and Log-Gabor wavelet is proposed. Source images are first captured by the AdaBoost face detection algorithm and then individually normalized. A reference image from the target age group that closely resembles the test subject's face is then selected. Because the Log-Gabor wavelet is characterized by a broader bandwidth and takes less time to search a wide range of spectral information then the Gabor wavelet, the Log-Gabor wavelet method is used to determine the aged skin surface topography and the decomposition map is obtained. By adjusting the number of details to be extracted from the decomposition map, we can effectively synthesize facial images for different age groups. Experimental results are verified with wrinkle density estimation.

Key Words: Face Detection, Log-Gabor Wavelet, AdaBoost Algorithm, Aging Synthesis

1. Introduction

Recently, the applications of facial image synthesis have been extensively employed in many areas, including visual entertainment, motion picture production, the gaming industry, cosmetic examinations, missing persons cases, etc. Among these applications, facial image synthesis for locating missing persons through the use of facial recognition with an aging effect has become a popular research topic. When suffering memory loss, some elderly people go missing upon becoming lost as they are unable to recall their route home. By applying the reverse-aging technology proposed, the chances for them to be recognized by their relatives or acquaintances are significantly increased. In this study, a simple face aging/reverse-aging synthesis based on face detection and the Log-Gabor wavelet is proposed.

Generally, there are three types of methods for capturing the characteristics for skin surface topography. The first one focuses on skull bone structure. Based on

---

*Corresponding author. E-mail: hsieh@ee.tku.edu.tw

skull models, face shapes from different age groups are selected and then used as parameters to define the structure of the skull [1]. The second method selectively focuses on local areas of the face to capture the skin surface topography [2], which includes the pixel value distributions for wrinkles, age spots, etc. Age synthesis for a facial image is achieved by altering these pixel values. Since this technology is solely based on the distribution data for pixel values, it is difficult for younger subjects, as their faces often do not contain obvious skin surface topography, such as wrinkles or age spots. The third method is based on statistical analysis, for example, obtaining average values for facial images [3,4], or principal component analysis (PCA) [5] to capture the distribution of the aged skin surface topography. Statistical analysis can be performed on colors and shapes [3] to manipulate facial images of different age groups or genders. Parameterized statistical modeling [4], PCA and 3D face modeling [5] can also be used to manipulate aging effects.

In this study, a statistical approach is preferred, and the Log-Gabor wavelet is used to effectively evaluate the

topographic features that are representative of different age groups. The Log-Gabor wavelet offers the ability to perform multi-channel and multi-resolution analysis. In addition, when compared to the Gabor wavelet [6–10], the Log-Gabor wavelet is characterized by a broader bandwidth and takes less time to search a wide range of spectral information [10]. The topography data obtained are then used to process age synthesis on the source images. Topographic features including wrinkles are first obtained from older subjects; these features are then applied to younger subjects to perform age synthesis. The manipulated target images are evaluated [11] to ensure satisfactory results.

2. Log-Gabor Wavelet

In face detection studies, usually the size of the face, skin surface topography and facial features are used as the baselines. The Gabor Wavelet is mainly used in face detections to capture information relating to image frequencies and these facial features [6,7]. The Gabor Wavelet employs both multi-channel and multi-layer methods to capture the facial features while effectively identifying skin surface topographies in different frequencies on human faces. It can also be used in classifying facial expressions [8,9]. The Gabor Wavelet is defined as:

$$g(f,\theta,x,y) = s(f,\theta)w(x,y) \tag{1}$$

where $s(f, \theta)$ represents a complex sinusoidal wave, and $w(x, y)$ represents a second-order Gaussian-shaped function.

The application of the Gabor Wavelet has two areas of limitations. Its bandwidth is limited to approximately one octave, and it takes an excessive amount of time to search a wide range of spectral information if the image is too large. Field [10] proposed the Log-Gabor wavelet method to improve these limitations. The advantages of the Log-Gabor wavelet are that the DC components are removed and the frequency can be extended to a wider range.

In the frequency domain, a second-order Gaussian-shaped function can be regarded as a logarithmic frequency value. A second order Log-Gabor wavelet can be

represented in polar coordinates [8,9] as:

$$H(f,\theta) = H_f \times H_\theta \tag{2}$$

where H_f determines the radial components in each spectral band and H_θ determines the angular components for filters in different spatial directions. Equation (2) can also be represented as:

$$H(f,\theta) = \exp\left\{\frac{-\left[\ln\left(\frac{f}{f_0}\right)\right]^2}{2\left[\ln\left(\frac{\sigma_f}{f_0}\right)\right]^2}\right\} \exp\left(\frac{-(\theta-\theta_0)}{2\sigma_\theta^2}\right) \tag{3}$$

where f_0 is the centre frequency of the filter, defined as $\frac{1}{\lambda}$ for wavelength λ. The direction of the filter is given as θ_0, σ_f is the standard deviation for the radial components and σ_θ is the standard deviation for the angular components. The bandwidth shape of the filter is represented by $\frac{\sigma_f}{f_0}$ and should be held constant as f_0 varies. The number of octaves for the filter bandwidth varies with different values of $\frac{\sigma_f}{f_0}$, as shown in Table 1 below.

In this study, the experiment uses three octaves in the setting.

The bandwidth B — which represents the radial properties of the Log-Gabor wavelet — is defined as:

$$B = 2\sqrt{2/\ln 2} \times \left|\ln\left(\frac{\sigma_f}{f_0}\right)\right| \tag{4}$$

Finally, $\Delta\Omega$ — which represents the angular properties of the Log-Gabor wavelet — is defined as:

Table 1. Values of $\dfrac{\sigma_f}{f_0}$ and the number of octaves

Values of $\dfrac{\sigma_f}{f_0}$	Number of octaves
0.74	1
0.55	2
0.41	3

$$\Delta\Omega = 2\sigma_\phi 2\sqrt{2\ln 2} \tag{5}$$

From experiments, optimal effects are observed when the value of f_0 for the minimum wavelength is set to 3, where as any value above 6 would cause the image to be blur.

In the frequency domain, a Log-Gabor wavelet map is constructed based on eight different angles, namely 0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°, and five different frequencies, namely 1/3, 1/6, 1/12, 1/24 and 1/48 of the central frequency. Through the use of these filters, topographic features can be successfully detected in different directions with different levels of detail. The Log-Gabor wavelet map can be represented in the time domain as well, with both real and imaginary components. It is evident that the map is a combination of filters of different sizes and different angles, and that the Log-Gabor wavelet map in the time domain can also be used to analyze topography.

To analyze topographic features in facial images, we simply perform convolution on the sample facial image in the time domain with the Log-Gabor wavelet map, or transfer the sample face image into the frequency domain, convolve with the Log-Gabor wavelet map in the frequency domain, and then transfer them back into the time domain. At the end of this process, the topographic distribution for the sample facial image can be obtained in different sizes and different directions. This is known as the decomposition map.

3. System Structure

A number of technologies are employed to perform age synthesis, including the use of edge detection [11–13], median filters [11–13], face detection [14–16] and the Log-Gabor wavelet [8–10,17]. In addition, an age determination algorithm may be developed to verify the accuracy of the synthesized images.

Gradient vectors are defined as changes in the x and the y directions from the original point (x, y), calculated by partial derivatives.

$$\nabla f(x, y) = \left(\frac{\partial_f}{\partial_x}, \frac{\partial_f}{\partial_y} \right) \tag{6}$$

The magnitude is the greatest rate of change, defined as

$$M = \sqrt{\left(\frac{\partial_f}{\partial_x} \right)^2 + \left(\frac{\partial_f}{\partial_y} \right)^2} \tag{7}$$

And the direction θ is defined as

$$\theta = \tan^{-1}\left(\frac{\partial_f}{\partial_x} \bigg/ \frac{\partial_f}{\partial_y} \right) \tag{8}$$

Popular edge detectors include the Laplaceian edge detector, Prewitt edge detector and Sobel edge detector [12,13]. The Sobel edge detector combines differential calculation with low pass filtering, and it offers noise cancellation effects. It is the most frequently used method in edge detection.

Low pass filters effectively reduce noise but important information in the original image can be lost, leading to the blurring of boundaries. Median filters, on the other hand, minimize the edge-softening effects but maintain the ability to remove and isolate noise. In addition, they can be applied to an image multiple times and therefore are the preferred method for noise reduction.

3.1 Face Detection

Adaptive Boosting, also known as AdaBoost, was proposed in 1996 by Freund and Schapire [15] as an ensemble classification method to improve the calculation accuracy for classifiers with learning algorithms. To illustrate the algorithm, suppose we need to classify two types of data. The ideal situation would be as shown in Figure 1(a) where two sets of data are clearly divided by the decision line. However, in reality we often encounter a more complex situation where the data distributions are non-linear, as shown in Figure 1(b). Through the use of the AdaBoost algorithm, we can classify these two sets of data using four linear classifiers, as shown in Figure 1(c). These two sets of data are then classified into four entities, as shown in Figure 1(d). It can be observed that multiple linear classifiers allow us to effectively classify non-linear data while improving algorithm accuracy.

In 2001, Viola and Jones proposed real-time object detection by the AdaBoost algorithm for face detection [14] based on the concept of Integral Image, which searches for the lowest feature classification errors. Characteristic matrices are used to calculate the difference in images, including parameters such as color, shape, loca-

tion etc.; the Adaboost algorithm is then used to improve the identifying accuracy on the objects. The characteristic matrices are used as the baseline for classification. Figure 2 below demonstrates the characteristic matrices used in this system. They travel around the image, behaving like filters: the differences between the values in the matrices are calculated and integrated. The size of both the white and black matrices can be configured with different sizing parameters, but they must all have the same size. Based on the concept of Integral Image, we can easily calculate the total size the matrices in various locations.

The AdaBoost algorithm is a cascade classification method. From large selections of weak classifiers, the one with the most significance is identified [15]. In the specified learning samples, from all the possible weak classifiers, the Adaboost algorithm selects the one with the least amount of errors. The sampling weight is then configured accordingly, with the incorrectly classified

samples receiving further attention. The above process is repeated and a weak classifier is identified in each process. The linear combinations of these weak classifiers then form a strong classifier. It has been proven that, for the learning samples, as the number of weak classifiers increases, the classification errors declines exponentially for the strong classifiers. The cascade classifiers are formed by many strong classifiers. As shown in Figure 1, each level represents a strong classifier obtained by the AdaBoost algorithm. After adjusting the threshold values in each level, almost all of the sample images with a human face are passed, and sample images with non-human faces are eliminated, as shown in Figure 3.

The face detection algorithm used in this study is obtained by the built-in AdaBoost algorithm in OpenCV. Samples of human faces were preloaded into the system. Using the AdaBoost algorithm, the face tiles in the sample human facial images are selected. The database is generated after normalizing these samples into $120 \times 120 \times 3$ pixels. The AdaBoost algorithm automatically captures the positions for facial features, including eyes, nose and mouth, and each facial feature is then normalized for individual subjects. Figures 4(a) to 4(f) illustrate

Figure 1. Concept of AdaBoost algorithm.
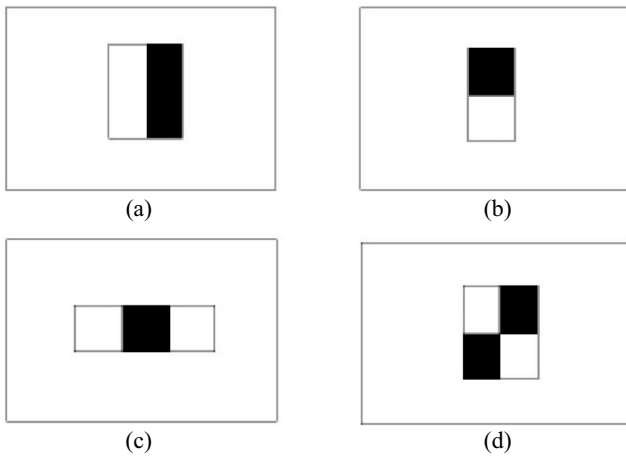
Figure 2. Characteristic matrices.
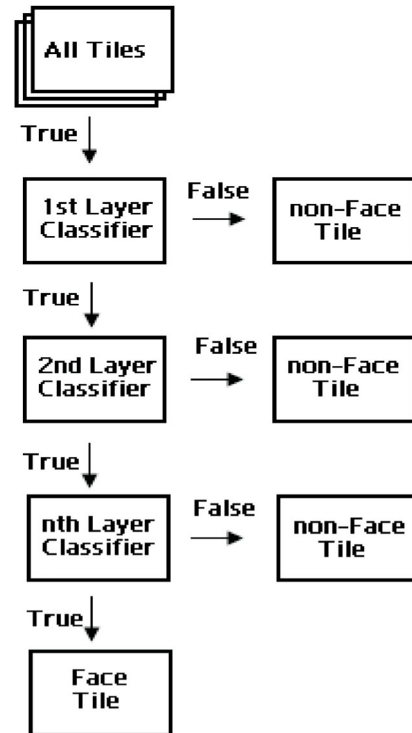
Figure 3. Cascade classification of AdaBoost algorithm.

the extraction process before normalization, and the normalized images are shown in Figure 4(g).

The original image is shown in Figure 4(a) and it is entered into our system to be processed by the AdaBoost algorithm, as shown in Figure 4(b). The landmark locations for the subject's facial features are detected in Figure 4(c) and the corresponding ROI image is obtained as shown in Figure 4(d). Cropping out the non-facial portion of the image, we obtain Figure 4(e). After zooming and resizing the image into $120 \times 120 \times 3$ pixels, the final image is obtained in Figure 4(f) and ready to be normalized and entered into the database, as shown in Figure
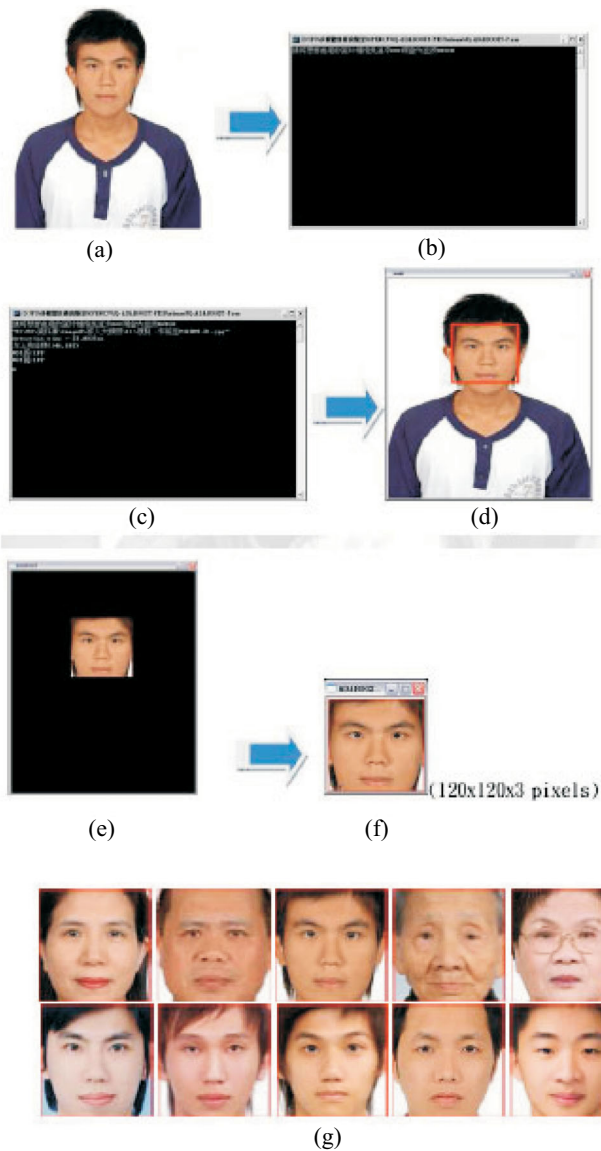
Figure 4. Example of face detection by AdaBoost.

4(g). Facial topographic analysis can then be performed to these normalized images.

3.2 Topographic Analysis

The Log-Gabor wavelet is used to analyze the skin surface topography of the normalized images. As people age, different types of skin surface topographies for different age groups appear on the faces, which include wrinkles, age spots, etc. It is important to effectively analyze the skin surface topographies in order to correctly determine the actual ages of the source image. The Log-Gabor wavelet extends to a wide frequency range, and topographies in different sizes as well as different angles can be analyzed at this stage. Table 2 below shows the parameters which are used in the experiments.

After convolving the face image with each Log-Gabor wavelet map in the time domain, images containing skin surface topographies are obtained. These images are referred to as the decomposition map.

It can be seen that in the decomposition map, as shown in Figure 5, images in the top two rows contain high frequency data, mainly used to define age texture or skin surface topology details, including wrinkles and age spots, whereas the rest of the images in the map contain low frequency data, mainly used for locating the position of major facial features, such as eyes, nose, and mouth.

Lastly, for age synthesis, the source image is matched with an image from the database that contains age texture information for the target age, as shown in Figure 6, and these two images are combined through the use of facial image synthesis to produce the final synthesized image.

As people age, more wrinkles appear on the face and hence the proportion of high frequency data would also increase. If the high frequency portion of the data is ex-

Table 2. Experimental parameters for Log-Gabor wavelet transform

Parameters	Value
Facial image size	$120 \times 120 \times 3$ pixels
Log-Gabor wavelet image size	240×240 pixels
Log-Gabor wavelet directions	20 sizes
Log-Gabor wavelet angles	4 angles
Bandwidth shape	3 Octaves
Minimum wavelength for f_c	3

(a)

(b)
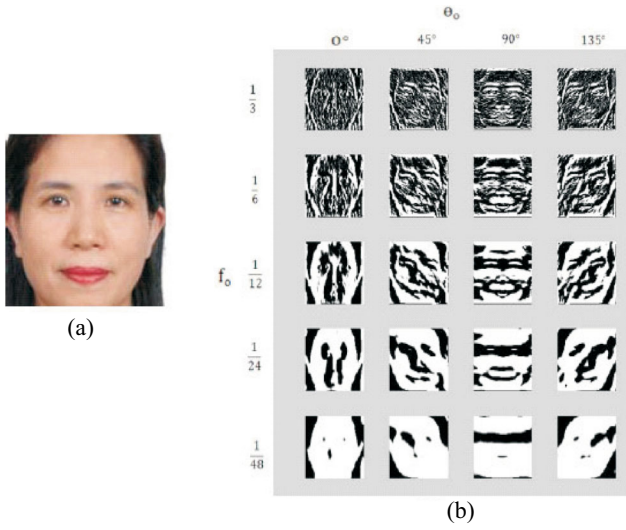
Figure 5. Decomposition map.

Figure 6. Age synthesis process.

tracted from the target image (facial image of the older person), and used to replace the high frequency portion of the source image (face of the younger person), age synthesis can then be successfully achieved.

Figure 7 demonstrates how the addition of high frequency data affects the age texture shown on the faces. The source image is a 48-year-old woman and the target image is an 88-year-old woman. High frequency data for the older woman is extracted and replaced for the corresponding mapping locations for the younger woman. During the age synthesis process, the image at the top left corner shows minimal effects of aging, where as the image at the lower right corner shows the maximum effects of aging.

To simulate reverse aging, the process described above is repeated with the face of the 88-year-old woman as the source image and the face of the 48-year-old woman as the target image, as illustrated in Figure 8. After extracting and replacing the appropriate high frequency information, age synthesis is again successfully achieved.

Figure 9 demonstrates how the addition of high frequency data removes age texture such as wrinkles on the faces when high frequency data for the younger woman is extracted and replaced for the corresponding mapping locations for the older woman. During the reverse age synthesis process, the image at the top left corner shows minimal effects of reverse aging, whereas the image at the lower right corner shows the maximum effects of reverse aging.

4. Results and Verification

Age determination calculations are performed on the age-synthesized images as well as reverse-age-synthesized images.

The Sobel edge detection method is used to detect wrinkles on the face, as wrinkles usually have darker colors than other areas of the skin. The pixel values as a result of the Sobel edge detection method can be used as a reference to represent the age. The threshold is con-
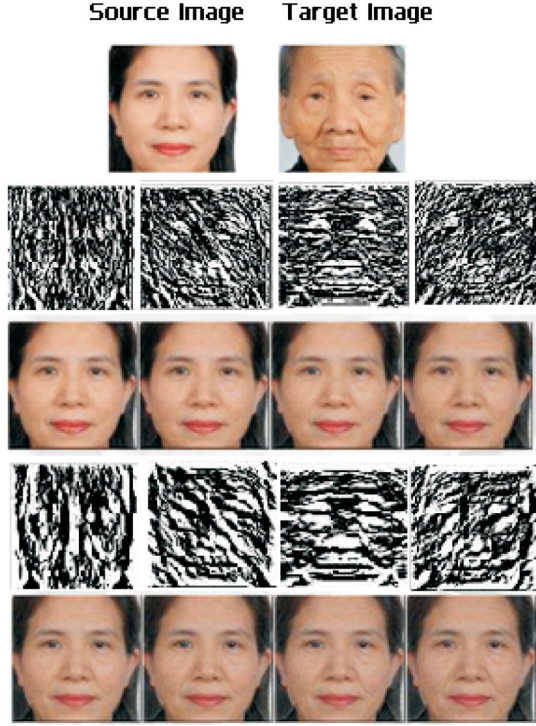
Figure 7. Age-synthesized images.

figured as 40, so that pixel values greater than 40 are set as dark skin color (binary value of 0, wrinkle) and pixel values less than 40 are set as light skin color (binary value of 1, non-wrinkle).

Median filters are used to remove unwanted noise. Calculation for wrinkle density [6] is then performed using the following equation:

$$W_{density} = \frac{|w|}{|p|} \tag{9}$$

Where p is the sum of the pixels in the region of interest (ROI), and w is the sum of dark pixels (wrinkle) in the ROI. $W_{density}$ varies between 0 and 1, with 1 represents a face full of wrinkles and 0 being a face with no wrinkles.

The wrinkle densities for the images in the database are calculated and summarized in Table 3 and Figure 10. Images in the database are grouped into young adults, adults and seniors. Young adults are from 20 to 39 years old, and adults are further divided into two age groups, 40 to 49 years old, and 50 to 59 years old. There are three age groups for seniors, 60 to 69 years old, 70 to 79 years old and 80 to 89 years old. The average values for wrinkle density W_p for each age group are calculated and
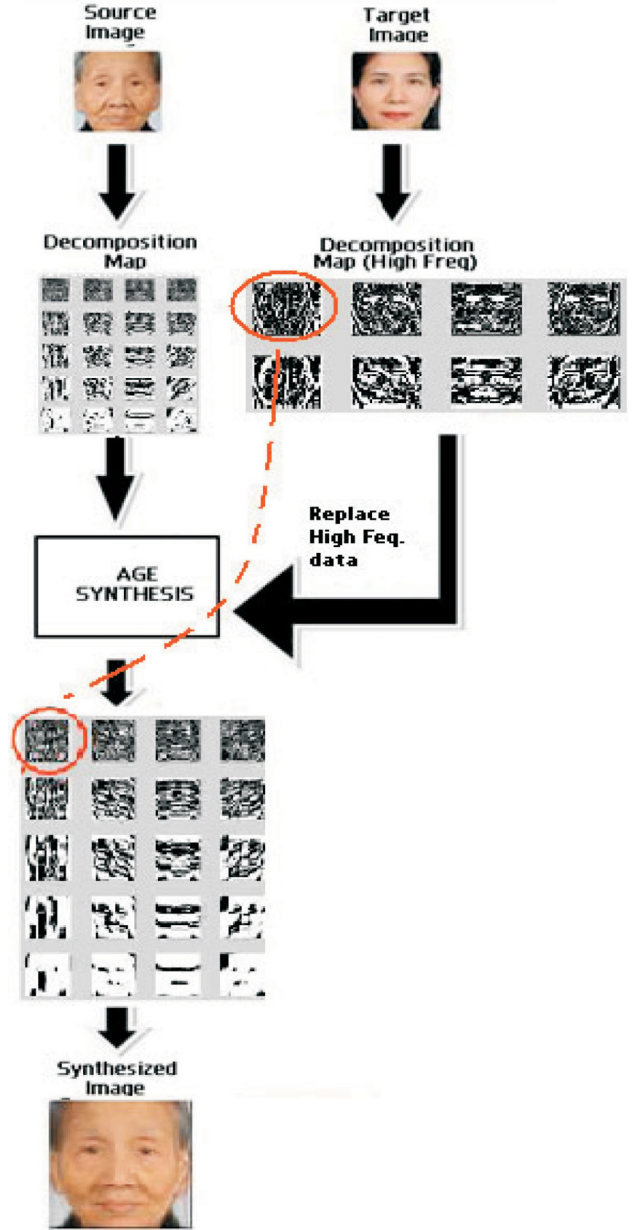
Figure 8. Reverse age synthesis process.

listed in Table 3. It can be observed that younger subjects have smaller values for W_p than older subjects. These wrinkle density values are used as references to determine how successful our age-synthesized images are. As a demonstration, we perform age synthesis to a test image and configure the target age to be approximately 80–89 years old with W_p of 0.2879. The resultant image is then analyzed to obtain its wrinkle density. If the W_p value is close to 0.2879, the age synthesis process is confirmed to be successful.

Source Image   Target Image
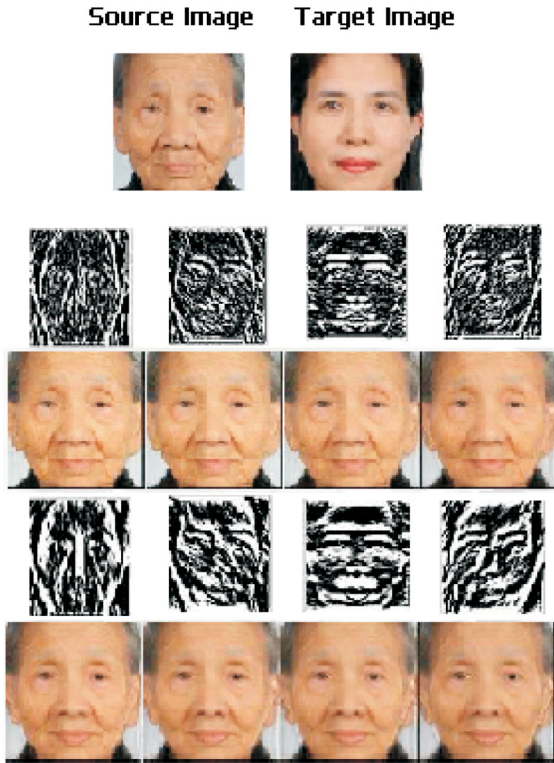
Figure 9. Reverse-age-synthesized images.

Table 3. Wrinkle densities of different age groups

Age group	W_p	Average W_p
20yr−39yr	0.0295	0.0689 ± 0.0219
	0.0766	
	0.0588	
	0.0757	
	0.0906	
40yr−49yr	0.1430	0.1156 ± 0.0154
	0.1354	
	0.1150	
	0.1313	
	0.1077	
50yr−59yr	0.1533	0.1517 ± 0.0204
	0.1517	
	0.1545	
	0.1432	
	0.1706	
60yr−69yr	0.2140	0.2207 ± 0.01
	0.2188	
	0.2124	
	0.2165	
	0.2142	
70yr−79yr	0.2397	0.246 ± 0.0169
	0.2528	
	0.2399	
	0.2410	
	0.2032	
80yr−88yr	0.2879	0.2842 ± 0.0208
	0.2710	
	0.2763	
	0.2618	
	0.2705	

As shown in Figure 11, the face of the 48-year-old woman has successfully transformed into faces with older ages. After performing age synthesis with different degrees of aging effects, three images are obtained and analyzed. A wrinkle density of 0.1430 corresponds to the age group of 50 to 59 years old, and a wrinkle density of 0.2221 corresponds to the age group of 60 to 69 years old. In the last image, where the source image is heavily processed by age synthesis, the wrinkle density is 0.2728 and this 48-year-old woman has been transformed into an age group of 80 to 89. Positive results obtained by the age determination calculation confirm the accuracy of the age synthesis process.

Similarly, in Figure 12, the wrinkles on the face of the 88-year-old woman have been successfully reduced. After performing age synthesis with different degrees of reverse aging effect, three images are obtained and analyzed. A wrinkle density of 0.2879 indicates the typical face of a woman in the age group of 80 to 89 years old, and an image with wrinkle density of 0.2137 corresponds to the age group of 60 to 69 years old. In the last image, the face of the 88-year-old woman is extensively

processed by reverse age synthesis, and the wrinkle density is reduced to 0.1722. This 88-year-old woman has been transformed into the age group of 50 to 59. Positive results obtained from the age determination calculation further confirm the accuracy for the reverse age synthesis process. However, our current system still experiences issues including edge distortion and hair discoloration. The Log-Gabor wavelet is used to perform detailed facial topographic analysis on human faces, but it cannot alter the shape or location of the facial features. When people age, some facial features may change their locations, as facial muscles are weakened and cannot hold the tissues in their original places; hence the corner of the mouth may sag, and skin around the eyelids may fold. In order for the resultant images to be more realistic — in other words, more convincing to the eye — additional

Figure 10. Example of wrinkle densities for images from different age groups.

technology is required to allows alteration of the facial features, for example, adjusting for the shape of the nose and lifting the location of the eyes. The development of such technology could be our research topic in the future.

5. Conclusion

A method based on the Log-Gabor wavelet is proposed to analyze skin surface topography on human faces. By replacing the high frequency data in the source image with the appropriate high frequency information from the corresponding sections in the composition map of the target image, both age synthesis and reverse age synthesis can be achieved. It is observed that as long as sufficient amount of information on skin surface topography is available for different age groups, both age synthesis and reverse age synthesis can be performed to transform a test image into any age group of interest. Edge detection technology is employed to calculate wrinkle density values and a wrinkle density index is used to verify the final images. Positive results further confirmed the accuracy of the system. This age synthesis

Figure 11. Age determination for age-synthesized images.

Figure 12. Age determination for reverse-age-synthesized images.

method based on face detection and the Log-Gabor wavelet can be employed with confidence to aid public services, for example in locating missing older persons.

References

[1] Nakagawa, M., Munetsugu, T., Kado, Y., Maehara, F. and Chihara, K., "The Facial Aging Simulation Based on the Skeletal Model," *IEICE(A)*, Japan, Vol. J80-A, pp. 1312−1315 (1997).

[2] Mukaida, S. and Ando, H., "Extraction and Manipulation of Wrinkles and Spots for Facial Image Synthesis," Proceedings. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 749−754 (2004). doi: 10.1109/AFGR.2004.1301624

[3] Rowland, D. A. and Perrett, D. I., "Manipulating Facial Appearance through Shape and Color," *Computer Graphics and Applications, IEEE*, Vol. 15, No. 5, pp. 70−76 (1995). doi: 10.1109/38.403830

[4] Lanitis, A., Taylor, C. J. and Cootes, T. F., "Modeling the Process of Ageing in Face Images," The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 1, pp. 131−136 (1999). doi: 10.1109/ICCV.1999.791208

[5] Choi, C. S., "Age Change for Predicting Future Faces," 1999 IEEE International Fuzzy Systems Conference (FUZZ-IEEE '99), Vol. 3, pp. 1603−1608 (1999). doi: 10.1109/FUZZY.1999.790144

[6] Pang, Y. W., Yuan, Y. and Li, X. L., "Gabor-Based Region Covariance Matrices for Face Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 7, pp. 989−993 (2008). doi: 10.1109/FUZZY.1999.790144

[7] Ho, J. T., "PSO-Based Modified Multilayer Neural Networks for Face Recognition," Department of Computer Science and Information Engineering, Chaoyang University of Technology (2006).

[8] Lajevardi, S. M. and Lech, M., "Facial Expression Recognition Using Neural Networks and Log-Gabor Filters," *Computing: Techniques and Applications, 2008. DICTA '08. Digital Image*, pp. 77−83 (2008). doi: 10.1109/DICTA.2008.13

[9] Rose, N., "Facial Expression Classification Using Gabor and Log-Gabor Filters," 7th International Conference on Automatic Face and Gesture Recognition (FGR 2006), pp. 346−350 (2006). doi: 10.1109/FGR.2006.49

[10] Field, D., "Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells," *Journal of Optical Society of American A*, Vol. 4, pp. 2379−2393 (1987). doi: 10.1364/JOSAA.4.002379

[11] Lee, C. P., "A Study on Classification of Age Groups Based on Facial Features," Department of Computer Science and Information Engineering, Tamkang University (2001).

[12] Chung, K. L., *Image Processing and Computer Vision*, 4th Edition, Dong-Hua Pub., Taipei, Taiwan (2002).

[13] Gonzalez, R. C. and Woods, R. E., *Digital Image Processing*, 2nd ed., Prentice-Hall (2002).

[14] Viola, P. and Jones, M., "Robust Real Time Object Detection," Proceedings Eighth IEEE International Conference on Computer Vision ICCV 2001 ICCV-01, Vol. 2, p. 747 (2001) doi: 10.1109/ICCV.2001.937709

[15] Freund, Y. and Schapire, R. E., "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, Vol. 55, No. 1, pp. 119−139 (1997). doi: 10.1006/jcss.1997.1504

[16] Chou, J. F., "Combine Color with Texture Feature of Face Detection and Recognition Technology," Department of Electronic Engineering, Ching Yun University (2008).

[17] Tu, C. H., "Enhanced Contour Detection Using Phase Preserving De-Noising Correction in Ultrasound Images," Department of Computer Science and Information Engineering, National Dong Hwa University (2006).