A Simple Automatic Facial Aging/Rejuvenating Synthesis Method

Ching-Tang Hsieh
Department of Electrical Engineering
Tamkang University
Taipei County, Taiwan, R.O.C
hsieh@ee.tku.edu.tw

Chia-Shing Hu
Department of Electrical Engineering
Tamkang University
Taipei County, Taiwan, R.O.C.

Chun-Wei Pan
Department of Electrical Engineering
Tamkang University
Taipei County, Taiwan, R.O.C.

***Abstract*—At present, the synthesizing faces of different ages does not emphasize on feature alignment and rectification of twisted images. If these situations do happen, they might cause failure and inaccuracy on synthesizing images. In this paper, we propose a reversible human facial aging/rejuvenating synthesis system which is implemented by Active Shape Model (ASM) integrated with Log-Gabor Wavelet, which can be used to search for the dementia elderly. First, we use AdaBoost and ASM algorithm to extract the feature set of human face, and rectify them by the concept of facial geometric invariance. The invariant concepts are the distance between inner corners of both eyes and the distance between the nose and chin. Then, we find manually one target image which is similar to the test image from the database, and analyze age texture of this human image by Log-Gabor wavelet in order to retrieve decomposition maps. Finally, we can effectively simulate human facial images of people of different ages by controlling the number of decomposition map of images and objectively judge the results via the density of wrinkles.**

Index Terms- aging synthesis; ASM algorithm; log-gabor wavelet;

I. INTRODUCTION

In recent years, as technologies have evolved and computational power has increased, the applications of facial image synthesis have been extensively employed in many areas, including visual entertainment, motion picture production, the gaming industry, cosmetic promotions, and missing person cases [1][2][3][4][5]. Among these applications, facial image synthesis for locating missing persons, through the use of facial recognition with aging effect has become a popular research topic.

Generally, there are three methods used in capturing skin surface topography. The first method [1] focuses on the bone structure of the skull. Based on skull models, various face shapes from different age groups can be selected and then used as the parameters to define the structural properties of the skulls. The second method [2] selectively focuses on local areas of the face and extracts the skin surface topography, or more specifically the distribution of pixel values for wrinkles, age spots, etc, in each of these areas. Age synthesis for facial images can then be achieved by altering the values of these pixels.

Since this technology is solely based on the distribution of pixel values, it is difficult to apply it to younger subjects whose faces very seldom exhibit obvious skin surface topographical features, such as wrinkles or age spots. The third method is based on statistical analysis, for example, obtaining average values for face images [3] [4], or principal component analysis (PCA) [6] to capture the distribution of the aged skin surface topographies. Statistical analysis can be performed on colors and shapes [3] to manipulate facial images from different age groups and different genders. Parameterized statistical modeling [4], PCA and 3D face modeling [6], and AdaBoost and Log-Gabor wavelet [5] can also be used to manipulate aging effects.

The three methods mentioned above do not emphasize facial feature alignment or correction for image distortion. In this study, an age synthesis method based on an integrated Active Shape Model (ASM) algorithm and Log-Gabor wavelet for face detection is proposed. The locations for eye brows, eyes, ears, nose, and lips are aligned, using the ASM algorithm to capture these facial features. A feature set containing 75 landmarks is generated and used to zoom and resize the images based on the concept of facial geometric invariance, which uses the distance between inner corners of both eyes and the distance between the nose and chin as references. The Log-Gabor wavelet, which offers the ability for multi-channel and multi-resolution analysis [5] [7], is used to effectively evaluate the topographical features which are representative of different age groups. The topography data obtained are used to process age synthesis on the test images. Results are compared for the conventional ASM algorithm and proposed integrated ASM algorithm. The manipulated target images are evaluated against reference [5] to ensure satisfactory results. The auto alignment process developed in the proposed system saves time by eliminating the need to perform manual alignments, and it also reduces learning time required for personnel training.

II. METHODOLOGY

A. ASM Algorithm

The Active Shape Model (ASM) algorithm was proposed by Cootes and Taylor [8] in 2000 to cut and divide images, including facial images. Its common applications include pattern recognition systems and object tracking systems which analyze rigid objects or objects with similar features, such as human faces and palms. The ASM algorithm is often used to extract the outlines and patterns of target objects for the purpose of evaluation and comparison during image post processing.

To generate the ASM model in this system, sets of features must be defined which are then used to describe the shape of the target object. Through the use of statistical analysis, the ASM model can be constructed to analyze the shape described by the feature sets. There are certain steps required to build the ASM model, including sample selection, landmark configuration, alignment training and model calculation. During the sample selection process, which in this case is for the purpose of face recognition, it is required that the images should only contain human faces and that any non-facial features be removed in order for the system to maintain accuracy. The ASM algorithm does not directly take test images as sample images; rather, it sets landmarks or important features on the test images. To minimize errors during age synthesis, alignments are performed during the training set where images are normalized before executing final evaluations of the skin topographies. In the model, the mean value, covariance of the training set, and the eigenvector and corresponding eigenvalue of the covariance are calculated.

For the ASM model to recognize the reference image, the shape of the face in the image should first be identified. For example, in the reference image, k grayscale pixels are selected from inside of the recognized shape, and k grayscale pixels are selected from outside of the recognized shape. In the end, $2k$+1 pixels form a grayscale vector. To minimize the influence of light settings and other environmental effects on the images, each grayscale vectors is normalized using the following equation:

$$v_1 = \frac{1}{\sum_i |v_{ij}|} v_i \tag{1}$$

where i is the i^{th} sample, and j is j^{th} entry in the grayscale vector v_{ij}. Then the mean value $\bar{v}$ and covariance S_v are both calculated. As the test image is input into the system, the ASM model extracts $2k$+1 pixels from the image and calculates the Mahalanobis distance between the reference image and the test image, as defined by the following equation:

$$f(v_s) = (v_s - \bar{v})^{\mathrm{T}} S_v^{-1} (v_s - \bar{v}) \tag{2}$$

where T is the transpose matrix. This step should be performed using iterative calculation until the smallest value of $f(v_s)$ is found. The final result defines the shape of the test image.

B. Log-Gabor Wavelet

In face recognition studies, usually the size of the face, skin surface topography and facial features are used as baselines. The Gabor wavelet is mainly used in face recognition systems to capture frequency information from facial features [5][7][9][10][11]. The Gabor wavelet employs multi-channel and multi-layer methods to capture facial features while simultaneously effectively identifying skin surface topographies of different frequencies on human faces. It can also be used in classifying facial expressions [9][10]. Gabor Wavelet is defined as:

$$g(f, \Theta, x, y) = s(f, \Theta) w(x, y) \tag{3}$$

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

The Gabor wavelet has two main limitations, namely that its bandwidth is limited to approximately one octave, and that search time for larger images is extensive. Field [7] proposed the Log-Gabor wavelet method to overcome these limitations. The advantage of the Log-Gabor wavelet is that the DC components are removed, and the frequency can be extended into 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 [9][10] as:

$$H(f,\theta) = H_f \times H_\theta \tag{4}$$

where H_f determines the radial components in different bandwidth, and H_θ determines the angular components for filters in different spatial directions. Equation (4) can also be represented as:

$$H(f,\theta)=\exp\left\{\frac{-\left[\ln\left(\frac{f}{f_0}\right)\right]^2}{2\left[\ln\left(\frac{\sigma_f}{f_0}\right)\right]^2}\right\}\exp\left(\frac{-(\theta-\theta_0)}{2\sigma_\theta^2}\right) \tag{5}$$

where f_0 is the centre frequency of the filter, defined as $^1/_\lambda$ for wavelength λ. The direction of the filter is given as θ_0, and σ_f and σ_θ are the standard deviations for the radial and angular components respectively. The bandwidth shape of the filter is represented by $\frac{\sigma_f}{f_0}$ and should be kept as a constant for different values of f_0. The number of octaves for the filter bandwidth varies with different values of $\frac{\sigma_f}{f_0}$. During the experiment performed in this study, 3 octaves are used with the value of $\frac{\sigma_f}{f_0}$ being 0.41.

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

$$B = 2\sqrt{2/\ \ln 2}\times\left|\ln\left(\frac{\sigma_f}{f_0}\right)\right| \tag{6}$$

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

$$\Delta\Omega = 2\sigma_\phi 2\sqrt{2\ln 2} \tag{7}$$

Experiment results show that optimal effects occur when the value of f_0 for the minimum wavelength is set to 3, whereas any value above 6 would cause the image to blur / result in a blurred image.

In the frequency domain, a Log-Gabor wavelet map is constructed based on 8 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, topographical 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 observed 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 skin topography.

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

III. System Structure

To execute age synthesis, the following three steps are implemented: image normalization (face alignment), topography evaluation (landmark analysis), and age synthesis, as shown in Figure 1.

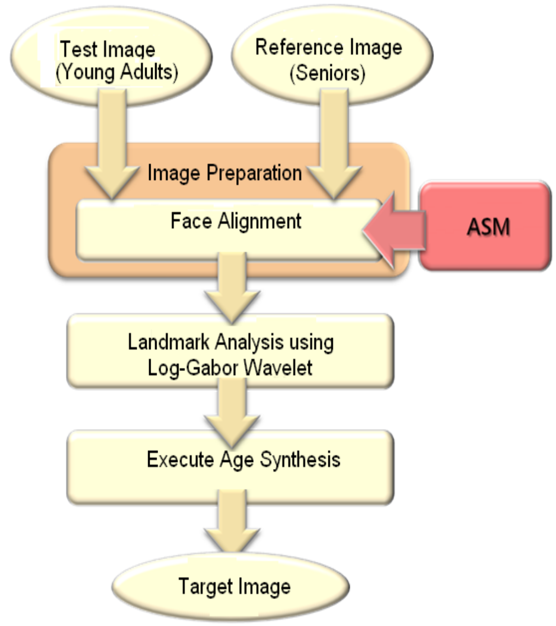

Fig. 1. System structure

A. Image Normalization

To ensure the realism of the target out image, it is essential to fix the locations for important facial features, or landmarks, including eye brows, eyes, ears, nose and lips. The ASM algorithm is used to effectively extract location information on these landmarks. Based on the concept of geometric invariance, i.e. the distance between the inner corners of both eyes, and the distance between the nose and the chin, these landmarks are normalized and the resultant images are resized to 120x160 pixels, as shown in Figure 2.

As mentioned above, the ASM algorithm is first employed to locate and mark sets of features for both the test and reference images, where each set contains 75 landmarks which represent facial features and outlines of the facial images. These feature sets, labelled P_{A75} and P_{B75}, are shown in Figure 3. The red dots visible in the figure are the common landmarks used in this study.

The images are then rotated according to the idea that for a person looking straight into the camera without tilting the head, the angle between the inner corners of the eyes should be horizontal (Figure 4).

The width and height of the face area are calculated based on the locations of the landmarks. The face area is then cropped out, as shown in Figure 5.

Finally, the images may be horizontally zoomed and resized, without distorting the faces, using the concept of geometric invariance, namely the distance between the inner corners of both eyes, as shown in Figure 6.

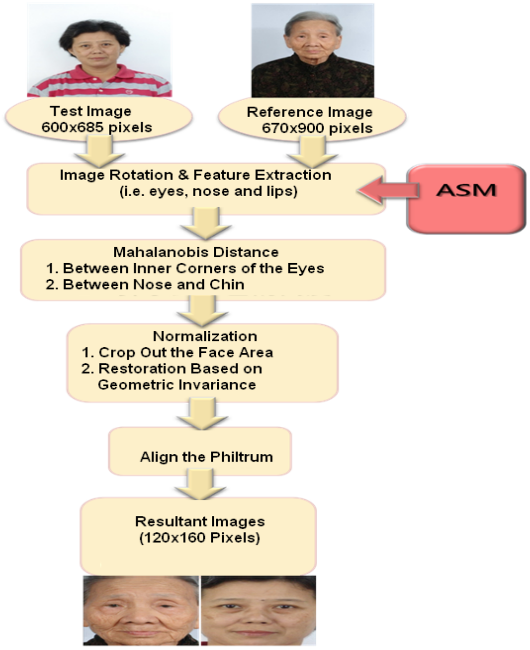

Fig. 2. Image normalization

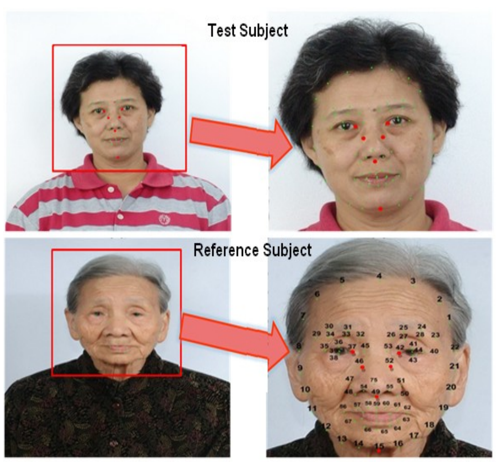

Fig. 3. Extract feature sets

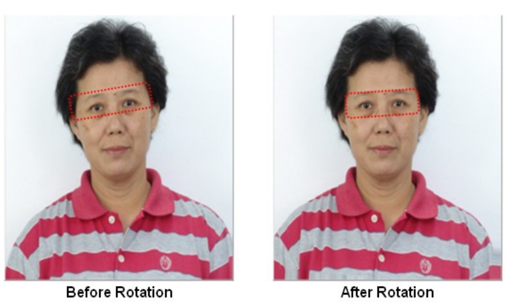

Fig. 4. Image rotation

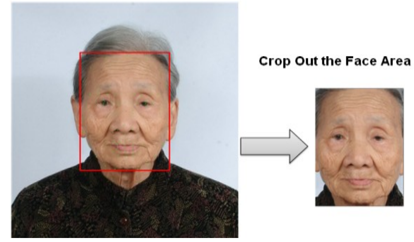

Fig. 5 Image cropping

Fig. 6 Image resizing

In Figure 6, the purple line defines the length between two inner corners of the eyes for the testing subject, and the green line is the length between two inner corners of the eyes for the reference subject. The image of the reference subject is zoomed and resized so the length of the green line is reduced to the same length as that of the purple line. The red line in the resultant image has the same length as the purple line and the adjustment is successfully achieved.

Mathematically, the locations of the two inner eye corners of the test and reference images are respectively represented by landmarks $P_{A75}[42]$ and $P_{A75}[37]$, and landmarks $P_{B75}[42]$ and $P_{B75}[37]$, where the 42^{nd} landmark is the left inner eye corner and the 37^{th} landmark is the right inner eye corner. The distances between these points are calculated as D_{x1} and D_{x2} and the resizing ratio $Scale_{WH}$ is obtained as follows:

$$Scale_{WH} = \frac{D_{x1}}{D_{x2}} \tag{8}$$
$$D_{x1} = P_{A75}[42] - P_{A75}[37]$$
$$D_{x2} = P_{B75}[42] - P_{B75}[37]$$

Once again following the concept of geometric invariance, the vertical distance between the nose and chin allows vertical zooming and image resizing. Suppose the 46^{th} landmark is the right side of the nose, the 52^{nd} landmark is the left side of the nose, and the 15^{th} landmark is the bottom of the chin. The distances between the noses and the chins are calculated as H_{y1} and H_{y2} and the resizing ratio $Scale_H$ is obtained as follows:

$$Scale_{WH} = \frac{H_{y1}}{H_{y2}}$$
$$H_{y1} = P_{A75}[8] - \frac{P_{A75}[46] + P_{A75}[52]}{2} \tag{9}$$
$$H_{y2} = P_{B75}[8] - \frac{P_{B75}[46] + P_{B75}[52]}{2}$$

After adjusting the width of the image using the distance between the inner corners of the eyes, the image is cut horizontally between the eyes and nose. The height of the lower half is then adjusted according to the distance between the nose and the chin. The resultant image combines the top portion of the face with the bottom portion, as illustrated in Figure 7.

During the final stage of image normalization, the philtrum of the test subject is aligned with that of the reference subject, as shown in Figure 8. The philtrum is determined by the ASM algorithm as the 49th landmark, marked in green in Figure 8. By this stage, both the test image and the reference image have been normalized and are ready to be used for age synthesis.

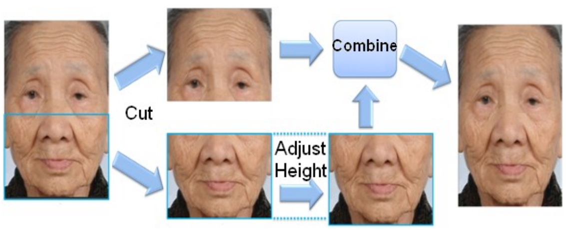

Fig. 7 Image adjustment.

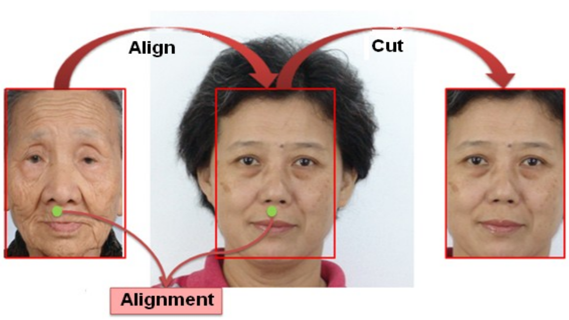

Fig. 8 Image alignment

B. *Topography Analyze*

The Log-Gabor wavelet is used to analysis the skin surface topography of the normalized images. As people age, different types of skin surface topographies for different age groups appear on their faces, which include wrinkles, age spots, etc. It is important to effectively analyze the skin surface topographies in order to correctly determine the actual age 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. Parameters for this study are listed in Table I below.

TABLE I. IMAGE ADJUSTMENT

Parameters	Value
Face Image	120x120x3 pixels
Log-Gabor wavelet image	320x320 pixels
Log-Gabor wavelet directions	20 sizes
Log-Gabor wavelet angles	4 angles
Bandwidth shape	3 Octaves

After convoluting the face image with each of the Log-Gabor wavelet maps in the time domain, images containing skin surface topographies are obtained. These images are collectively referred to as a decomposition map, as illustrated in Figure 9.

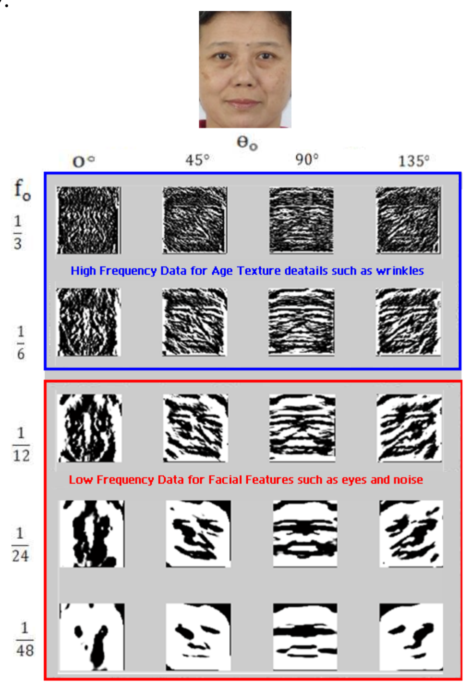

Fig. 9 Decomposition map

The images in the top two rows of the decomposition map contain high frequency data, mainly used to define age textures, 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.

C. *Age Synthesis*

As people age, more wrinkles appear on the face and hence the portions for high frequency data would also increase. If the high frequency portion of the data is extracted from the reference image (face of an older person), and used to replace the high frequency portion in the test image (face of an younger person), age synthesis can be successfully achieved, as shown in Figure 10 below.

Figure 11 demonstrates how the addition of high frequency data affects the age texture shown on the face. The test image is a younger woman and the reference image is an older woman. The high frequency features of the older woman are extracted and superimposed in the corresponding mapping locations for the younger woman. The image at the top left corner shows minimal effects of aging, and the image at the lower right corner shows the maximum effects of aging.

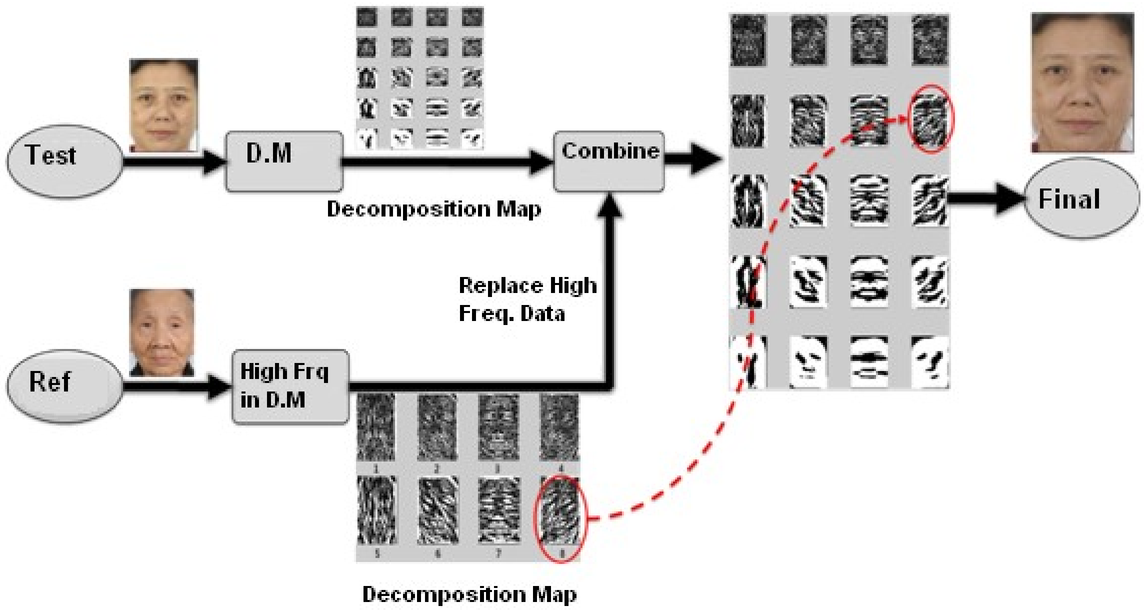

Fig. 10 Age synthesis using high frequency data

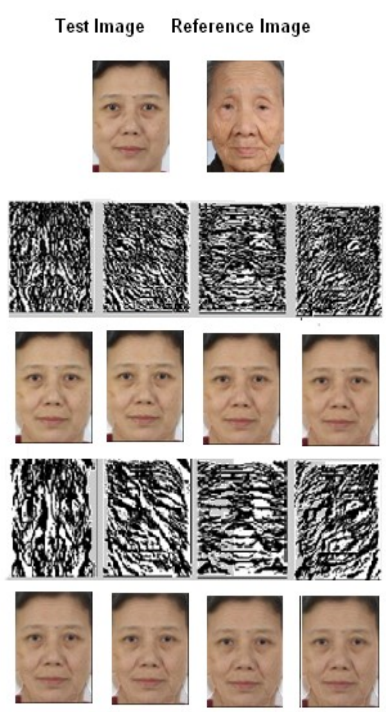

Fig. 11 Effects of high frequency data on face texture

IV. Results and Verification

During normalization of the images, the ASM algorithm is employed to identify facial areas and to calculate the locations of important facial features. In the proposed system with integrated ASM algorithm and Log-Gabor wavelet, the concept of geometric invariance is used and images are further adjusted to correct the alignments of those important facial features for the test and reference images. Figure 12 compares the resultant images after normalization, where image set 12(a) uses conventional ASM algorithm and image set 12(b) incorporates the concept of geometric invariance. The improved feature alignment (synthesized image quality) of the proposed system is clearly visible in the latter image set. This improved accuracy is also evident in Table 2, which gives comparative success rates for facial feature alignments using both the conventional ASM algorithm and the proposed system. The conventional ASM algorithm shows a maximum feature alignment success rate of 62.8% (for eye alignment) and a minimum of just 2.5% (for mouth alignment), while the proposed system achieved 100% success rates for all cases.

Normalization is a very important step in the age synthesis process, as accurate results can only be obtained when the important facial features of both the test and reference images are properly aligned.

TABLE II. Margin specifications

Success Rate	Original ASM (%)	Proposed System (%)
Eye alignment	62.8%	100%
Nose alignment	16.5%	100%
Mouth alignment	2.5%	100%
Three features all aligned	7.5%	100%
No features aligned	28.1%	0%

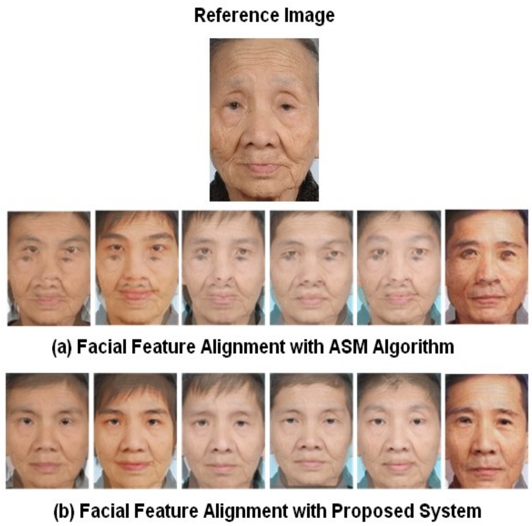

Fig. 12 Facial feature alignments

Fig. 13 Before and after age synthesis

The test images before and after age synthesis are shown in Figure 13. Wrinkle density is calculated using the method proposed in reference [5] to examine the final output image after age synthesis. The value for wrinkle density varies between 0 and 1, where 1 represents a face full of wrinkles and 0 represents a face completely without wrinkles. Generally, the wrinkle density of younger subjects would be smaller than that of older subjects. The wrinkle density value for the final output images after age synthesis is determined to be approximately 0.2207, which, according to reference [5], would correspond to an age group of 60 to 69 years old, i.e. age synthesis has been successfully achieved.

V. Conclusion

A method based on the ASM algorithm and Log-Gabor wavelet is proposed to perform age synthesis on human faces. Facial features, or landmarks, in full color images are first captured and analyzed by the ASM algorithm. They contain information to calculate the distance between the inner corners of both eyes and the distance between the nose and the chin. These distances are used for zooming and resizing of the images, during the image normalization process. After normalization, the Log-Gabor wavelet is used to evaluate the skin surface topography for both the test subject and the reference subject and Composition Maps are generated for both images. By replacing the high frequency data in the test image with that of the reference image, according to the appropriate sections in the Composition Map, age synthesis can be achieved. The final image is verified by wrinkle density index. The authors are confident that this age synthesis method based on integrated ASM algorithm and Log-Gabor wavelet will be beneficial to society, especially in locating missing persons.

References

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

[2] Mukaida, S.; Ando, H.; "Extraction and manipulation of wrinkles and spots for facial image synthesis", Proceedings. Sixth IEEE International Conference on Automatic Face and Gesture Recognition, p. 749 - 754, 2004.

[3] Rowland, D.A., Perrett D.I., "Manipulating facial appearance through shape and color", Computer Graphics and Applications, IEEE, volume: 15, Issue: 5, p.70-76, 1995.

[4] Lanitis, A., Taylor, C.J., Cootes, T.F., "Modeling the process of ageing in face images", The Proceedings of the Seventh IEEE International Conference on Computer Vision, volume: 1, p. 131 – 136, 1999.

[5] Ching-Tang Hsieh; Yu-Cheng Lee; Min-Nan Hsiao; "Facial Image Aging Synthesis Based on the Log-Gabor Wavelet", 18th International Conference on Software Engineering and Data Engineering (SEDE-2009), June 22 – 24, 2009

[6] Changseok Choi; "Age change for predicting future faces", 1999 IEEE International Fuzzy Systems Conference (FUZZ-IEEE '99), volume: 3, p.1603 – 1608, 1999.

[7] Field D., "Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells", Journal of Optical Society of American A, Volume 4, p. 2379-2393, 1987.

[8] T.F. Cootes and C.J. Taylor, "Statistical Model of Appearance for Computer Vision", Wolfson Image Analusis Unit, Image Science and Biomedical Engineering, University of Manchester, July 10, 2000.

[9] Lajevardi, S.M.; Lech, M.; "Facial Expression Recognition Using Neural Networks and Log-Gabor Filters", Computing: Techniques and Applications, 2008. DICTA '08.Digital Image, p. 77 – 83, 2008.

[10] Rose, N., "Facial Expression Classification using Gabor and Log-Gabor Filters", 7th International Conference on Automatic Face and Gesture Recognition(FGR 2006), p. 346 - 350, 2006.

[11] Cin-Hong Tu, "Enhanced Contour Detection Using Phase Preserving De-noising Correction in Ultrasound Images" Department of Computer Science and Information Engineering, National Dong Hwa University, 2006.