Pattern Recognition 44 (2011) 1262–1281

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Age estimation using a hierarchical classifier based on global and local facial features

Sung Eun Choi [a], Youn Joo Lee [a], Sung Joo Lee [a], Kang Ryoung Park [b], Jaihie Kim [a,*]

[a] School of Electrical and Electronic Engineering, Yonsei University, Biometrics Engineering Research Center, Republic of Korea
[b] Division of Electronics and Electrical Engineering, Dongguk University, Biometrics Engineering Research Center, Republic of Korea

ARTICLE INFO

Article history:
Received 17 June 2010
Received in revised form
20 October 2010
Accepted 7 December 2010
Available online 13 December 2010

Keywords:
Age estimation
Hybrid features
Hierarchical classifier
Gabor filter
Local binary pattern (LBP)

ABSTRACT

The research related to age estimation using face images has become increasingly important, due to the fact it has a variety of potentially useful applications. An age estimation system is generally composed of aging feature extraction and feature classification; both of which are important in order to improve the performance. For the aging feature extraction, the hybrid features, which are a combination of global and local features, have received a great deal of attention, because this method can compensate for defects found in individual global and local features. As for feature classification, the hierarchical classifier, which is composed of an age group classification (e.g. the class of less than 20 years old, the class of 20–39 years old, etc.) and a detailed age estimation (e.g. 17, 23 years old, etc.), provide a much better performance than other methods. However, both the hybrid features and hierarchical classifier methods have only been studied independently and no research combining them has yet been conducted in the previous works. Consequently, we propose a new age estimation method using a hierarchical classifier method based on both global and local facial features. Our research is novel in the following three ways, compared to the previous works. Firstly, age estimation accuracy is greatly improved through a combination of the proposed hybrid features and the hierarchical classifier. Secondly, new local feature extraction methods are proposed in order to improve the performance of the hybrid features. The wrinkle feature is extracted using a set of region specific Gabor filters, each of which is designed based on the regional direction of the wrinkles, and the skin feature is extracted using a local binary pattern (LBP), capable of extracting the detailed textures of skin. Thirdly, the improved hierarchical classifier is based on a support vector machine (SVM) and a support vector regression (SVR). To reduce the error propagation of the hierarchical classifier, each age group classifier is designed so that the age range to be estimated is overlapped by consideration of false acceptance error (FAE) and false rejection error (FRE) of each classifier. The experimental results showed that the performance of the proposed method was superior to that of the previous methods when using the BERC, PAL and FG-Net aging databases.

1. Introduction

The human face contains a great deal of information related to personal characteristics, including identification, emotion, age, gender and race. This information has been used extensively in the face-based human–computer interaction (HCI) systems capable of interpreting the facial information found in human-to-human communication [1]. Currently, the research related on age estimation using face images is more important than ever, because it has many applications, such as an internet access control, underage cigarette-vending machine use [1,2], age-based retrieval of face images [2], age prediction systems for finding lost children and face recognition robust to age progression. In addition, the estimated age of consumers who look at billboards is used in age specific target advertising as consumer preferences differ greatly by age.

Age estimation systems are generally designed to use two steps: an aging feature extraction and a feature classification. Feature extraction is very important in age estimation, since the extracted features greatly affect the classification performance. For this reason, a great deal of effort has been directed towards the extraction of discriminative aging features. These features can be categorized into local and global features, and hybrid features, which are a combination of the global and local features. Local features consist of the amount and depth of wrinkles, skin aging using freckles and age spots, hair color and the geometry of facial components. The local features have been commonly used to classify people into age groups (e.g. babies, young adults and senior adults) as they possess unique characteristics that distinguish

* Corresponding author.
E-mail addresses: choisungeun@yonsei.ac.kr (S.E. Choi), younjoo@yonsei.ac.kr (Y.J. Lee), sungjoo@yonsei.ac.kr (S.J. Lee), parkgr@dongguk.edu (K.R. Park), jhkim@yonsei.ac.kr (J. Kim).

0031-3203/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2010.12.005

specific age groups. For example, wrinkles are found in adulthood rather than in childhood, and geometric features, such as the distance ratios between features such as the eyes, the end of the nose, and the corners of the mouth, are noticeably changed in childhood rather than in adulthood. Consequently, these features are better suited to applications requiring an age group classification (e.g. the class of less than 20 years old, the class of 20–39 years old, etc.) rather than a detailed age estimation (e.g. 17, 23 years old, etc.). Conversely, unlike the local features, the global features are generally used to estimate a detailed age and contain not only aging characteristics, but other individual characteristics such as identity, expression, gender, ethnicity, among others. Therefore, individual characteristics are reflected better in global features than aging characteristics. For instance, active appearance models (AAM) [24] are mainly used to estimate an age as global features, because they offer greater amounts of information concerning the appearance and shape of a face than local features. However, AAM features do not include detailed wrinkle and skin informations, due to the dimensional reduction made by the principle component analysis (PCA). This defect in the global features can be resolved through combination with local features. In order to solve this problem, an age estimation method based on the hybrid features is proposed in [3], and the defects found in each feature can then be compensated for by combining the global and local features. Therefore, hybrid features are desirable for the accurate estimation of age.

When given a feature, feature classification steps are needed for age estimation. The feature classification can be divided into three approaches: the age group classification [6,16,17,23], the single-level age estimation [2,3,12,13,20] and the hierarchical age estimation [2,4,5]. Age group classification is an approach that roughly predicts an age group, whereas single-level and hierarchical age estimations are focused on detailed age prediction. The single-level age estimation is to find the age label in the total data set. On the other hand, the hierarchical age estimation is a coarse-to-fine method used to find the age label in a pre-classified group's small data set. Of these methods, the hierarchical age estimation provides the most improved performance [2,4,5]. As facial aging is perceived differently in different age groups, age estimation in a specific age group provides a more accurate result. Moreover age estimation on a smaller age group simplifies the computational load [4].

As mentioned above, the hybrid features and the hierarchical classifier provide a good performance for the estimation of age. However, they have been studied independently; research of combining them has not been conducted in the previous works. In this paper, we propose a new age estimation method that combines the hybrid features and the hierarchical classifier in anticipation of the fact that age estimation performance can be enhanced by a combination of good performance methods. In addition, new extraction methods for local features and a hierarchical classifier are proposed to improve them. For hybrid features, the wrinkle and skin features are used as local features, and the AAM features are used as the global features. Because wrinkles appear as edge or high frequency components in the image, the wrinkle features have generally been extracted by texture edge analysis using a Sobel filter [6–8], a Hough transform [9] and an active contour [10]. However, the edge components in a face image not only include wrinkles, they also include noise, such as hair, shadows, or a mustache. In order to reduce the influence of this noise, the predominant orientation of a wrinkle can be used to extract the feature. In this paper, a set of Gabor filters capable of determining the orientation of a wrinkle is used for the wrinkle feature extraction. By designing a wrinkle specific Gabor filter, a noise robust wrinkle feature can be extracted. Generally, the skin features have been extracted by analyzing high frequency components in the texture. Unlike wrinkles, however, facial skin texture slightly and randomly changes with age. Consequently, we were able to extract the skin features using the local binary pattern (LBP) method, which can extract detailed features of the skin. By using the proposed methods, more effective and detailed local features are extracted, thereby ensuring a good performance in age estimation.

In addition, improved hierarchical classifiers are proposed for an accurate age estimation. Typically, the hierarchical classifiers composed of age group classifications and detailed age estimations have a problem, in that an error in the age group classification is propagated to the age estimation. For example, if an image of a 15 year old is misclassified to the 20–39 years old class during the age group classification step, it is quite possible that the resulting estimation error is greater in instances, where the hierarchical age estimation using an age group estimator with a hard boundary that does not include the overlapped class, compared to a single-level age estimation. Because the aging function made in a specific age group can reflect specific age characteristics better than whole age range characteristics, the classification error during the age group classification step will affect the specific age estimation step. To reduce such classification error, each age group classifier is designed to overlap, taking into consideration the false acceptance error (FAE) and the false rejection error (FRE) of each classifier. By compensating for possible classification errors in this way, the performance of age estimation is greatly improved.

The publicly available FG-Net aging database is commonly used in many previous works for age estimation in order to evaluate performance [1,3,11–14,21]. However, it has many variations in expression, light, image resolution and noise caused by scanning. In addition, because it does not possess uniform age distribution, it becomes difficult to fairly analyze the performance of age estimation. In order to solve these problems, a new BERC database was collected for this research, in which the facial images have no variations in light, background, resolution, or expression. And, the number of images is uniformly distributed by age and gender. In this paper, age estimation performance is evaluated using our BERC database, the publically available PAL and the FG-Net aging database. Experimental results showed that the performance of the proposed method on all three databases was superior to the previous methods.

The remainder of the paper is organized as follows: in Section 2, we introduce an overview of the related works based on feature extraction and feature classification for age estimation. In Section 3, we explain the proposed methods for wrinkle and skin feature extraction, as well hierarchical age classifier design. In Section 4, the experimental results are reported and analyzed. Finally, in Section 5, our conclusions and future works are outlined.

2. Related works

There has been a great number of researches about how to accurately estimate ages from facial images. These age estimation methods were classified based on feature extraction and feature classification. In this section, previous works are described from these points of view.

2.1. Feature extraction

The facial features used in the previous works are divided into three categories: local features, global features and hybrid features. The local features such as wrinkles, skin, hair and geometric features have been commonly used to classify the age groups in many earlier studies. Kwon and Lobo [10] classified facial images into three age groups: babies, young adults and senior adults using the distance ratio of facial components and the wrinkle features.

They extracted the distance ratios based on the anthropometry of the face [15] as geometric features, which were then used to distinguish babies from adults. Wrinkle features were extracted by snakelets in the predesignated facial regions, and were then used to classify young and senior adults. Horng et al. [6] used wrinkle and geometric features to classify four age groups, using the Sobel filter to judge the degree of wrinkles. To determine the wrinkle features, the density and depth of the wrinkle and the average variance of the skin were extracted, using the Sobel edge magnitude. Hayashi et al. [9] used the Digital Template Hough Transform (DTHT) to extract the wrinkle features. After the skin regions were extracted by skin color, the wrinkles were modeled using the DTHT. Takimoto et al. [7] used both the Sobel filter and the Gabor jet in order to distinguish a deep wrinkle from a fine wrinkle. The skin features were extracted with a difference image between the original image and an averaged image. They also used skin color features in the HSV color space. Jun-Da Xia and Chung-Lin Huang [8] proposed an age classification method using wrinkle features extracted by the Sobel filter along with the hair color features. The regions for feature extraction were selected based on facial landmarks detected by the AAM. In these previous methods, local features were used mostly for age group classification, because local features indicate characteristics of an age group rather than a specific age.

The global features that represent overall facial characteristics have been largely used to estimate detailed age. Methods used to determine global features are Active Appearance Models (AAM), Gabor wavelet transform (GWT) [16,17], subspace features using the image intensity [13,18], and image frequency [19]. In the literature reviewed for this paper, many previous works to determine facial features based on the AAM model have been proposed [1,2,11,12,20]. The AAM is a well-known method that represents faces with statistical appearance and shape models using PCA. The first age estimation algorithm using AAM features and regression methods was proposed by Lanitis et al. [20]. In their work, the relationship between age and features was defined as quadratic aging functions, and the facial age was then estimated using aging features. After that, Lanitis et al. [2] used the AAM features to compare different classifiers for age estimation. Geng et al. [11] proposed the aging pattern subspace (AGES) method, using the subspace representation of a sequence of individual aging face images. In their work, an AAM was also used as a feature extractor. Yan et al. [12] designed a regressor based on training samples with uncertain nonnegative labels using AAM features. Though the AAM has been frequently used in the previous works, its features do not include information about wrinkles and skin, because of the dimensional reduction done by the PCA.

In order to solve this problem with the AAM features, hybrid features, which are a combination of global and local features, were proposed by Suo et al. [3] who designed sparse features consisting of the AAM, wrinkles, skin, hair and the configuration of the facial components features using the hierarchical face model. In each component, four types of features were extracted: topology, geometry, photometry and configuration features. However, elaborate work is required in order to extract the features in [3]. For example, the geometric features of wrinkles were described by a set of landmarks. However, extracting the landmarks of the wrinkles is more unstable than analyzing their texture. The landmarks of the wrinkles represent the positions on the wrinkle edges, and because facial wrinkles are composed of both fine and deep wrinkles, they are not always clearly portrayed in facial images. Furthermore, in order to detect the positions on the wrinkle edges, robust operations of edge representation and edge line tracing are required. Consequently, extracting the landmarks of wrinkles becomes a very challenging problem, and a texture analysis of the wrinkle area is a more stable method of analyzing various wrinkles. In addition, hair features were extracted using a graph cut and generative hair model [21]. However, the problems associated with this approach when used for actual age estimation system remained unsolved, due to the inherent difficulties associated with hair modeling [22]. In spite of having used this complicated method, the performance of the age estimation was not improved compared to other research conducted using only the AAM features [12–14].

In order to improve performance, we propose a new extraction method for wrinkle and skin features. The wrinkle features are effectively extracted by a Gabor filter set based on the direction of wrinkles on the face. The skin features are extracted by an LBP for the skin aging analysis. The pros and cons of both the previous methods and the proposed method based on the facial features are summarized in Table 1.

Table 1
Comparisons between the previous works and the proposed method for feature extraction.

Category	Methods	Strength	Weakness
Local features	Wrinkle, skin, geometric features and hair [6–10]	• Ability to extract aging features without an influence of personal characteristics	• Better suited to age group classification than detailed age estimation
Global features	AAM [1,2,11,12,20] GWT [16,17], image intensity[13,18], image frequency [19]	• Provides sufficient information for detailed age estimation compared to local features	• Does not include enough information on wrinkles and skin aging
Hybrid features (global+local features)	Sparse features [3,21]	• Combining global and local features to solve the problem of global features • Use of hair information	• The extraction of local features is unstable (e.g. hair line, wrinkle line) • The performance of age estimation is not improved • The performance is dependent on AAM fitting result
	Proposed method • Extracting wrinkle and skin features considering each biological trait	• Improvement of age estimation performance of local features	• The performance is dependent on AAM fitting result
	1) Wrinkle: Gabor filter set considering regional direction of wrinkles 2) Skin: local binary pattern method	1) Wrinkle feature: robust to noise such as hair, shadows and mustache 2) Skin: microstructural analysis of skin	

2.2. Feature classification

From the point of view of feature classification, the previous methods are divided into three categories: age group classification, single-level age estimation and hierarchical age estimation. The age group classification is used to classify multiple age groups. In order to determine the classifier, SVM [16], LDA [17], distance measure [23] and neural networks [6] have been used in the previous works. To estimate the detailed age, single-level or hierarchical age estimation was used. Numerous learning methods, such as quadratic regression [20], SVR [13], MLP [2,3], or RUN [12] were used for single-level age estimation. Although the single-level age estimation has the virtue of simplicity. However, because the aging features can differ depending on age group, the single-level age classifier has been trained in all age groups and has limits in estimating an accurate age. Therefore, the hierarchical classifier trained in a small age group will provide a better performance [2,4,5]. Lanitis et al. [2] demonstrated that the hierarchical classifiers performed better than the single-level classifier through experimental results gained using hierarchical classifiers based on appearance or age or both. It was found that the appearance and age specific (AAS) method performed the best. Guo et al. [4] proposed a hierarchical age estimation method based on smaller gender and age groups rather than on all ages. Luu et al. [5] proposed hierarchical age estimation based on a support vector machine (SVM) for age group estimation and support vector regression (SVR) for specific age estimation.

Taking this into consideration, it is known that hierarchical classifiers provide better performance than single-level classifiers in terms of age estimation, as seen in the previous works. However, the hierarchical classifiers using age group estimator with a hard boundary that does not include overlapping has a problem in that errors occurring in the age group classification step are propagated into the sub-class estimation step (i.e. the specific age estimation step). In other words, if an input face image is misclassified during the age group classification, it is quite possible that the age estimation errors will increase. To reduce such classification errors, each age group classifier is designed to have an overlapping estimated age range, which considers the false acceptance error (FAE) and false rejection error (FRE) of each classifier. By compensating for classification errors using overlapping classes, the total age estimation performance is improved. The pros and cons of both the previous methods and the proposed method based on the age estimation methods are summarized in Table 2.

3. Proposed method for age estimation

3.1. Overview of the proposed method

The proposed age estimation method consists of a global feature extraction, a local feature extraction, a feature fusion and a hierarchical age estimation, including an age group classification and a specific age estimation as shown in Fig. 1. In the global feature extraction step, the appearance and shape parameters of the AAM are extracted as global features. In the local feature extraction step, the wrinkle and skin features are extracted by our proposed methods. In the feature fusion step, the global and local features are combined into a feature vector by performing a feature level fusion. In the age group estimation step, the input face images are classified into different age groups. Finally, the detailed age is estimated within the selected age group in the specific age estimation step. A detailed description of each step is given in the following subsections.

3.2. Active appearance models (AAM)

The active appearance model (AAM) is used to estimate age as facial global features in this paper. The AAM is a generative parametric model that contains both the shape and appearance of a human face, which it models using the principal component analysis (PCA), and is able to generate various instances using only a small number of parameters [24]. Therefore, an AAM has been widely used for face modeling and facial feature point extraction.

In order to construct the shape model, v points were selected in the facial images to compose a facial shape $s=(x_1, y_1, x_2, y_2, \ldots, x_v, y_v)^T$. For facial shape s, 68 landmark points on each face image were utilized in this paper, as shown in Fig. 2, therefore $v=68$. For the test data, the 68 landmark points can be manually obtained by human or automatically obtained by an AAM fitting. In order to concentrate on problems of extracting facial features and age estimation without the effects of an AAM fitting error, manually obtained points were used from Sections 4.3.1 to 4.3.3. Also, in order to compare performance using manually and automatically obtained facial points, automatically obtained points by an AAM fitting were used in Section 4.3.4. The facial shape s is expressed by the linear combination of the mean shape s_0 and the n

Table 2
Comparison between the previous works and the proposed method for feature classification.

Category	Methods	Strength	Weakness
Age group classification	SVM [16], Neural network [7,6] Fuzzy LDA [17]	• Suitable for classification into coarse age group	• Unsuitable to estimate detailed age
Single-level age estimation	Quadratic regression [20], SVR [13], MLP [2,3], RUN [12]	• Ability to estimate age, simply using a single estimator	• Does not consider the differences of age feature values according to age groups
Hierarchical age estimation	Hierarchical age estimation [2,4,5]	• The performance is improved by considering the differences of age feature values according to age groups	• The errors in the age group estimation step are propagated to the specific age estimation step when using an age group estimator with a hard boundary not including overlapped class
	Proposed method • Hierarchical age estimation based on SVM and SVR • In age classification step using SVM, each class is designed to have an overlapped age range to be estimated	• The errors of age group classification are reduced by using the overlapped age range of two classifiers	• Although the errors in the age group estimation are smaller than those in the hierarchical age estimation, the errors can be propagated to the specific age estimation step

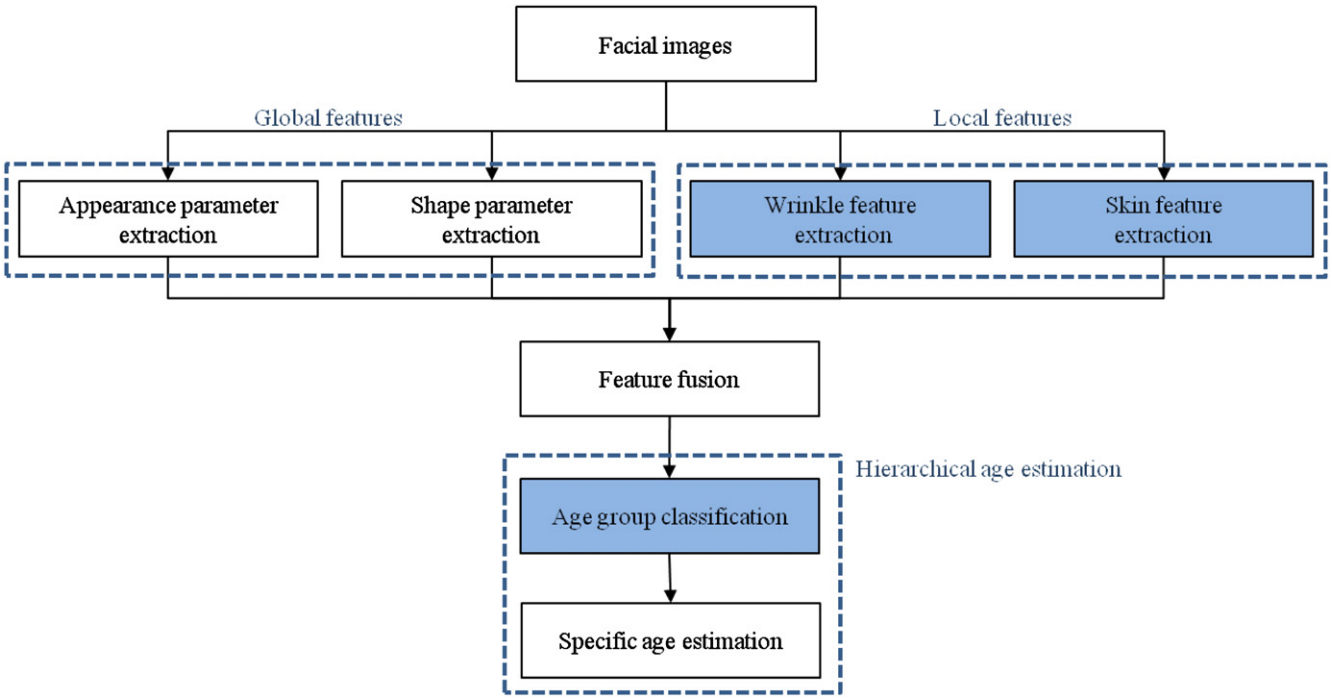

Fig. 1. Flow chart of the proposed age estimation method.

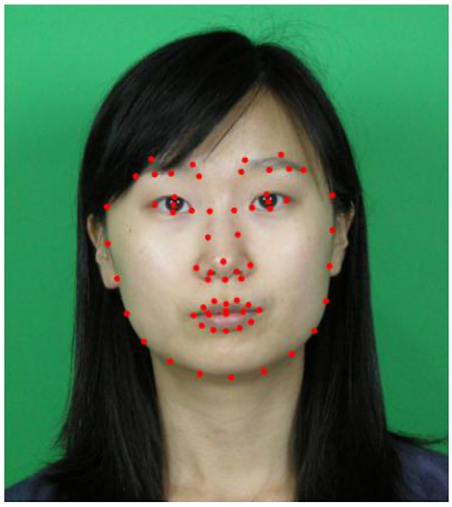

Fig. 2. The 68 facial feature points.

shape vectors s_i for $i=1, \ldots, n$

$$s = s_0 + \sum_{i=1}^{n} p_i s_i \tag{1}$$

where P_i, for $i=1, \ldots, n$, are the shape parameters. In order to obtain the appearance model, the shape of an image is normalized by warping it to the mean shape s_0. The facial appearance, $A(x)$ is expressed by a linear combination of the mean appearance A_0 and the m appearance images $A_i(x)$

$$A(x) = A_0 + \sum_{i=1}^{m} \lambda_i A_i(x) \quad \forall x \in s_0 \tag{2}$$

where λ_i is the appearance parameter [25]. The shape parameter P_i and the appearance parameter λ_i are utilized as facial global features for age estimation. The dimension of each parameter is decided in order to preserve approximately 95% of the variability in the training set [11].

When no unlabeled face image is provided, the AAM fitting process needs to automatically match the model to the new image data. The AAM fitting methods aim to find the shape and appearance parameters that minimize errors between the reconstructed image and the warped original input image. The objective function for an AAM fitting is as follows:

$$\sum_{x \in s_0} \left[A_0(x) + \sum_{i=1}^{m} \lambda_i A_i(x) - I(W(x;p)) \right]^2 \tag{3}$$

where W is the warping function needed to change the image position from position x in the input image to mean shape s_0 and $I(W(x;p))$ is the intensity of the warped original input image. To minimize Eq. (3) simultaneously with respect to the shape parameter p and the appearance parameter λ, the gradient descent algorithm is used. A detailed description can be found in [25].

Because the AAM is designed for face modeling, individual characteristics such as the variation of identity are reflected stronger than the aging characteristics in the AAM features. In addition, some crucial aging features such as wrinkles and skin textures are removed from the AAM features in the dimension reduction process performed by the PCA. This problem is shown by the simple experiments shown in Figs. 3 and 4. Fig. 3 shows the comparison of the original image (O) and the reconstructed image by the AAM appearance parameter (R). The original image (O) is normalized by warping it to the mean shape of the training data, and the difference image (R−O) is defined by the absolute difference between the reconstructed image (R) and the original image (O). As shown in Fig. 3, the reconstructed image does not sufficiently express the facial aging features such as wrinkles or skin. For example, the wrinkles around the eyes and mouth are not represented in the reconstructed image, as shown in the dashed circles of Fig. 3(b), and the total pixel errors between the reconstructed and the original images increase with age, as shown in Fig. 4. Based on Figs. 3 and 4, it can be seen that facial aging is not sufficiently represented by the appearance parameter of the AAM, so additional local features for age estimation are needed.

3.3. Wrinkle feature extraction

The facial wrinkles provide extremely important information regarding the estimation of age, and so have been given a great deal of consideration by researchers [6–10]. We propose a new method to extract wrinkle features based on the biological characteristics of the wrinkles. Many studies have been conducted on facial wrinkles in cosmetic surgery and biology [26–28]. Facial wrinkles are created by repeated facial muscular movements and expressions, and are therefore formed perpendicularly to the direction of the underlying facial muscles. For instance, forehead wrinkles have a horizontal direction, and nasolabial folds (e.g. wrinkles on each side of the nose) are in a diagonal line. The various types of facial wrinkles are shown in Fig. 5 [27]. Consequently, facial wrinkles have unique characteristics depending upon the direction of the muscles, the frequency of use and the effect of gravity. These characteristics of facial wrinkles can be used to extract wrinkle features for age estimation.

In order to analyze facial wrinkles possessing directional characteristics, a Gabor filter with different parameters according to the direction and frequency of the selected facial area was used. By using the Gabor filter utilizing the dominant direction of the wrinkles, regional wrinkle features are extracted effectively. The facial wrinkles, which have diverse attributes, such as deep and fine wrinkles, are extracted by the Gabor filter using various frequencies.

The Gabor filter is commonly used for texture classification and segmentation, image recognition, registration, and motion tracking [29]. The two dimensional Gabor filter in the spatial domain is defined by

$$g(x,y)=\left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right] \quad (4)$$

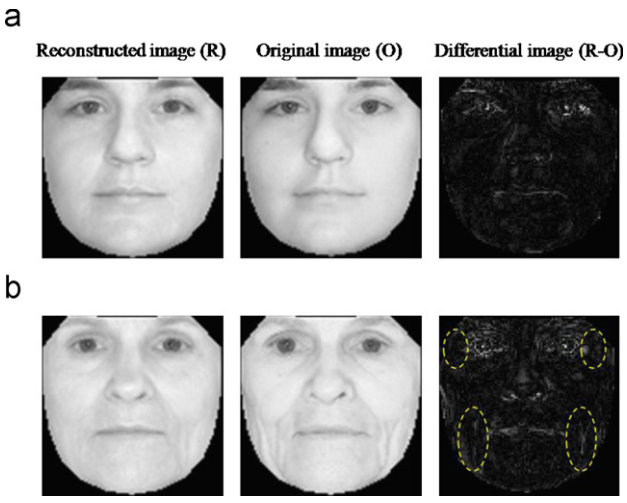

Fig. 3. Comparison of the original image and the reconstructed image from the appearance parameter of the AAM. (a) 18 years old and (b) 79 years old.

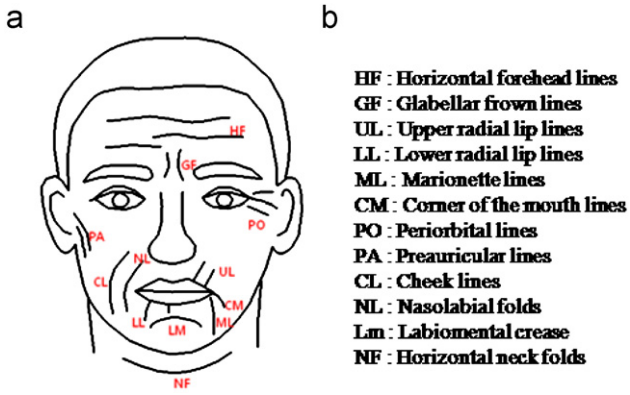

Fig. 5. Types of facial wrinkles [27]: (a) wrinkle position and (b) wrinkle name.

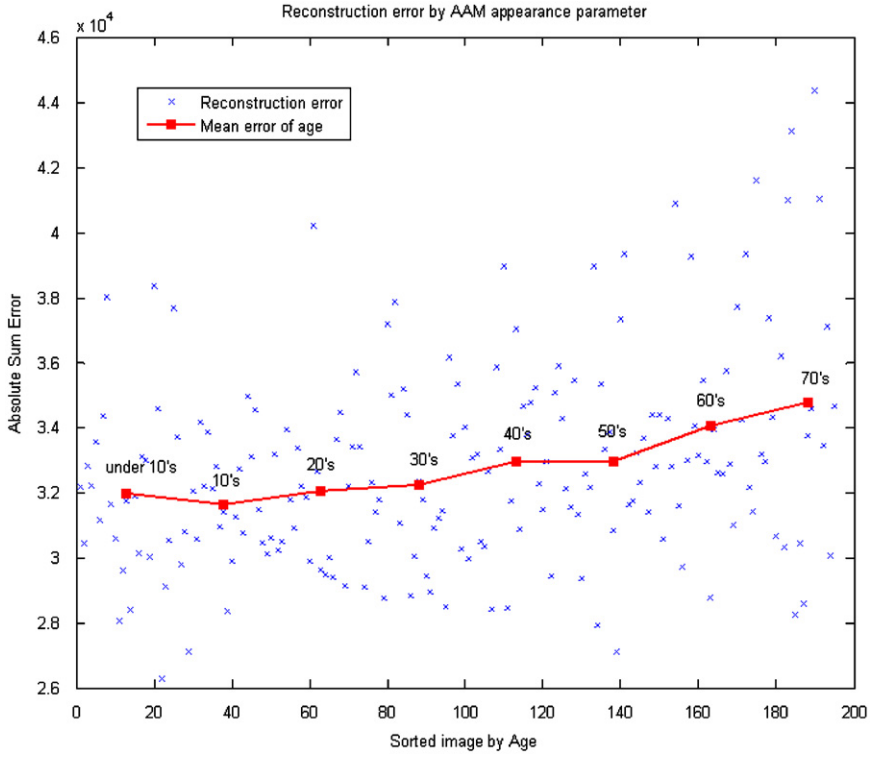

Fig. 4. Reconstruction error from the appearance parameter of the AAM according to age.

where σ_x and σ_y are the standard deviations of the x- and y-axes, respectively, and W is the radial frequency of the sinusoid. The Fourier transform of the Gabor function found in Eq. (4) is expressed by

$$G(u, v) = \exp\left[-\frac{1}{2}\left(\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right] \quad (5)$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$.

To make the Gabor wavelet, $g(x,y)$ of Eq. (4) is regarded as the mother Gabor wavelet. The Gabor wavelet is made by the dilations and rotations of $g(x,y)$ and is expressed by

$$g_s(x, y) = a^{-m} g(\tilde{x}, \tilde{y}) \quad (6)$$

$$\begin{cases} \tilde{x} = x\cos\theta + y\sin\theta \\ \tilde{y} = -x\sin\theta + y\cos\theta \end{cases}$$

where θ is the filter orientation expressed by $\theta = n\pi/K$, K is the number of the filter's orientation, a^{-m} is the filter scale, $m = 0, \ldots, S$, S is the number of scales. To prevent redundancy in the frequency domain, the Gabor wavelet is designed by

$$a = \left(\frac{U_h}{U_l}\right)^{1/(S-1)}, \quad \sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}},$$

$$\sigma_v = \tan\left(\frac{\pi}{2K}\right)\left[U_h - 2\ln\left(\frac{\sigma_u^2}{U_h}\right)\right]\left[2\ln 2 - \frac{(2\ln 2)^2\sigma_u^2}{U_h^2}\right]^{-(1/2)} \quad (7)$$

where U_l and U_h denote the lower and upper average frequencies, respectively, and $W = U_h$. For more detail refer to [29,30]. Fig. 6 shows the 24 Gabor filters that have the orientation number of $K = 6$, the scale number of $S = 4$, the lower average frequency of $U_l = 0.05$ and the upper average frequency of $U_h = 0.4$ with considering references [29,30].

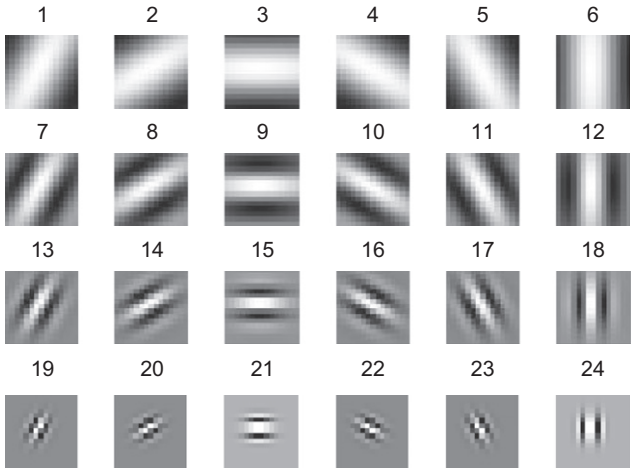

Fig. 6. Gabor filters with 4 scales and 6 orientations.

To extract the wrinkle features, we set the wrinkle analysis areas in the facial images by referring to the types of wrinkles found in Fig. 5. The wrinkle analysis areas are selected based on the 68 facial feature points seen in Fig. 2. For rotation-invariant processing, rotation of all images is normalized using two eye points before selecting the wrinkle analysis areas as shown in Fig. 7. Based on the facial feature points, nine wrinkle areas are selected, as shown in Fig. 8. In each wrinkle area, the predominant orientation of wrinkles can be estimated, and the Gabor filter sets in each wrinkle area are selected using the predominant orientation of the wrinkles. For example, diagonal and vertical filters are assigned to part 4 of Fig. 8 containing the nasolabial folds, as shown in Fig. 9(d). All filters are allocated for parts 5 and 6 of Fig. 8, because the sides of the eyes having wrinkles going in all directions, as shown in Fig. 9(e). Fig. 9 shows the allocated filter set for each wrinkle area. To determine the wrinkle features, we use the mean and variance of the magnitude response of the Gabor filter in each wrinkle area, because the mean and the variance of the magnitude represent both the strength and quantity of wrinkles.

3.4. Skin feature extraction

As a person is getting older, facial blemishes such as freckles, lentigines (age spots) and fine wrinkles increase on the skin. Inside aging skin, melanocytes, which are melanin-producing cells in the bottom layer of the skin's epidermis, are damaged by exposure to

Fig. 8. Wrinkle areas for the extraction of the features.

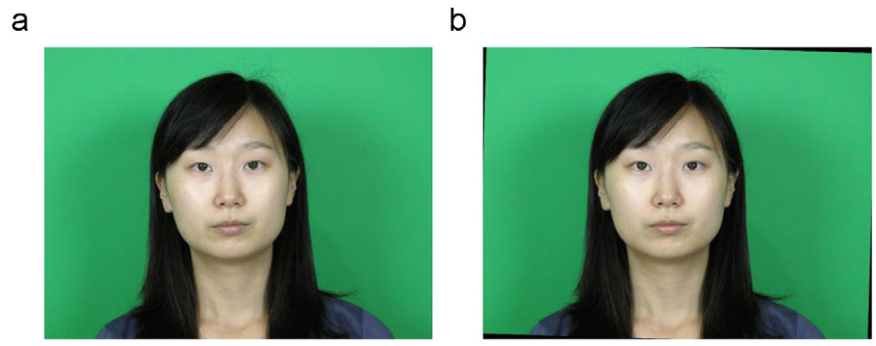

Fig. 7. Rotation normalization with two eye points: (a) original image and (b) rotated image.

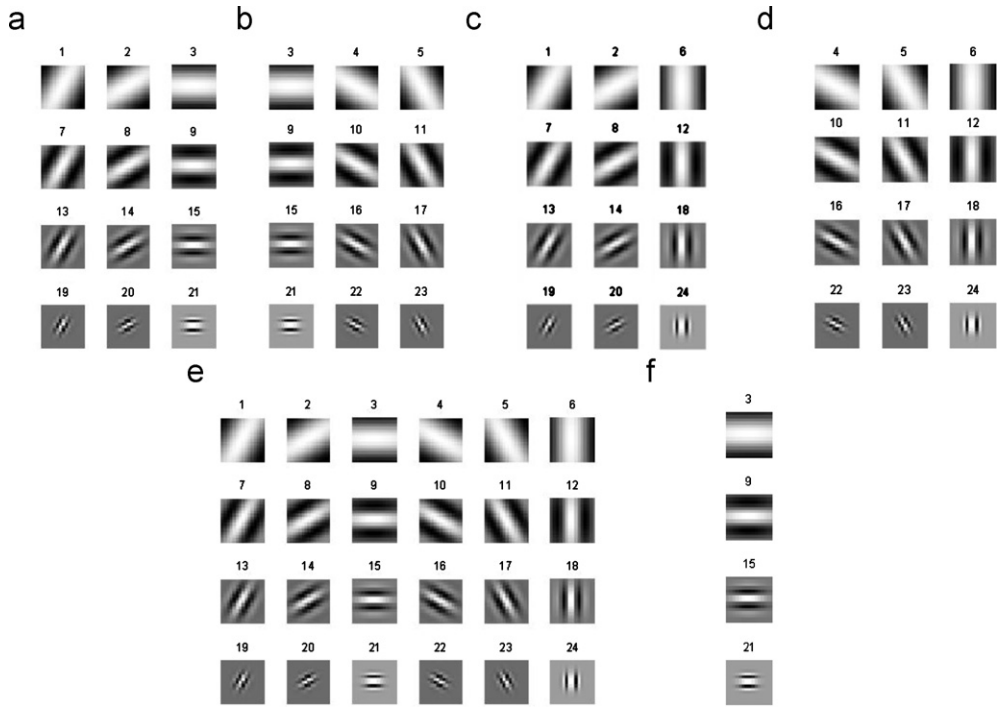

Fig. 9. Gabor filter sets according to the direction of facial wrinkles. (a) part 1 of Fig. 8, (b) part 2 of Fig. 8, (c) parts 3 and 8 of Fig. 8, (d) parts 4 and 9 of Fig. 8, (e) parts 5 and 6 of Fig. 8 and (f) part 7 of Fig. 8.

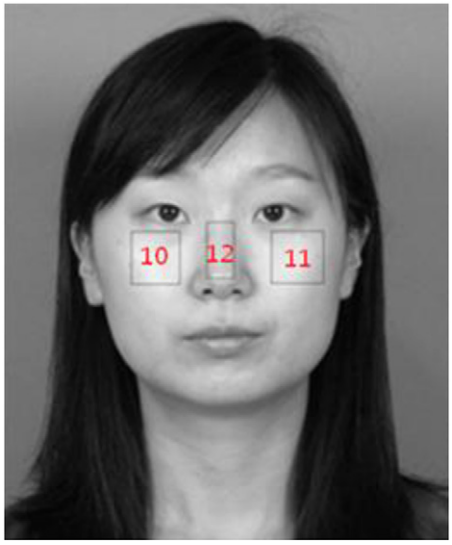

Fig. 10. Skin areas used for extracting the features.

ultra violet rays from the sun. This damage to the melanocytes results in the overproduction of melanin in the skin, resulting in freckles and lentigines. In addition, the collagen that plays a role in reflecting light decreases and is distributed non-uniformly. Therefore, the overall tone of the facial skin becomes non-uniform [28]. Thus, the aging of facial skin is mainly affected by sunlight exposure. For instance, facial parts affected by a lot of sunlight exposure are the upper cheek, the bridge of the nose, etc. Consequently, we selected parts of the upper cheek and the bridge of the nose to analyze facial skin, as shown in Fig. 10. These skin areas were selected based on the 68 facial feature points seen in Fig. 2. Unlike wrinkles, skin aging appears randomly and non-uniformly for each facial part. In addition, aging skin has slight variations. Therefore, a method capable of analyzing very fine details of skin texture is required. In this paper, the skin features are extracted using the LBP method, which detects microstructures such as edges, lines, spots and flat areas [31,34]. In addition, the LBP has the advantage of computational simplicity.

The LBP operator is used for texture classification, segmentation, face detection, face recognition, gender and age classification [23]. The basic concept of the LBP is to assign a code to each pixel comparing it to its neighboring pixels. The creation of the LBP code is expressed by

$$LBP_{P,R}=\sum_{p=0}^{P-1}s(g_p-g_c)2^p,\quad \text{where } s(x)=\begin{cases}1, x\geq 0\\ 0, x<0\end{cases} \tag{8}$$

where P is the number of neighboring pixels, R is the distance from the center to the neighboring pixels, as shown in Fig. 11(b). g_c Corresponds to the gray value of the center pixel, $g_p(p=1,\ldots,P-1)$ corresponds to the gray values of the p equally spaced pixels on the circle of radius $R\,(R>0)$ that forms a circularly symmetric neighbor set, and s is the threshold function of x. The LBP example is shown in Fig. 11(a). The gray value of center pixel (g_c) of the sub-window is assigned an LBP code using Eq. (8). The process for an LBP code generation, for the first sub-window in Fig. 11(a), is as follows:

$$\begin{aligned}LBP_{P,R}&=\sum_{p=0}^{P-1}s(g_p-g_c)2^p=s(g_0-g_c)2^0+s(g_1-g_c)2^1+s(g_2-g_c)2^2\\&\quad+s(g_3-g_c)2^3+s(g_4-g_c)2^4+s(g_5-g_c)2^5\\&\quad+s(g_6-g_c)2^6+s(g_7-g_c)2^7\\&=s(46-35)2^0+s(12-35)2^1+s(30-35)2^2\\&\quad+s(24-35)2^3+s(5-35)2^4+s(37-35)2^5\\&\quad+s(80-35)2^6+s(70-35)2^7\\&=1\times 2^0+0\times 2^1+0\times 2^2+0\times 2^3+0\times 2^4+1\times 2^5\\&\quad+1\times 2^6+1\times 2^7\\&=1+32+64+128=225\end{aligned}$$

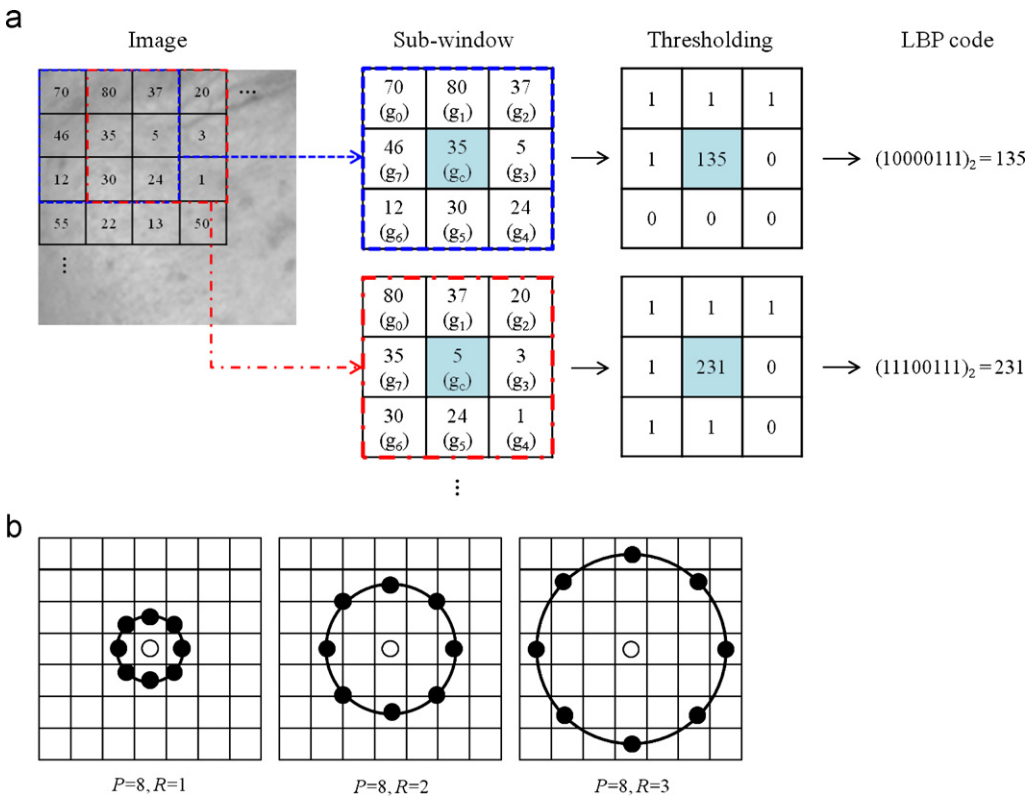

Fig. 11. An example of the LBP and the position of the neighboring pixels according to the P and R: (a) an example of the LBP and (b) positions of the neighboring pixels according to the P and R.

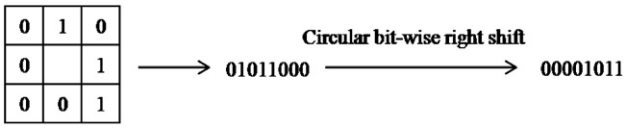

Fig. 12. An example of the rotation invariant code.

All of the pixels in the texture are assigned to the LBP code in this way, and the pixel position according to the P and R value is described in Fig. 11(b). The multi-resolution texture analysis is applied using the various P and R values.

In order to extract features robust to rotation, the rotation invariant code of the LBP is generated by circularly rotating the original code until its minimum value is reached [32]. A simple example of the rotation invariant is shown in Fig. 12 [33].

The LBP codes were divided into uniform and non-uniform patterns, as was done in [31]. A uniform pattern is defined as one that includes at most two bitwise transitions from 0 to 1 (or 1–0). For example, 00000000, 00011000 and 11101111 are uniform patterns. On the other hand, 01010000, 11001100 and 01010101 are non-uniform patterns. The uniform patterns operate as a template for microstructures such as a bright spot, a flat area or a dark spot and the edges of the varying positive and negative curvatures, as shown in Fig. 14 [31]. However, the non-uniform patterns do not have enough information to describe a texture as the number of non-uniform patterns in a texture is relatively small. Consequently, all non-uniform patterns are assigned arbitrarily to the same code, which is not included in the uniform pattern codes. For instance, the rotation invariant uniform patterns are assigned to codes 0–8 and non-uniform patterns are assigned to code 9, which is not included in the uniform pattern code, when $P=8$ and $R=1$. Some example patterns and codes meeting the requirements of the rotation invariant and uniform patterns when $P=8$ and $R=1$ are shown in Fig. 13. This results in the texture features of the LBP being composed of 10 codes, the uniform pattern codes from 0 to 8 and the non-uniform code 9. The characteristics of the texture, which are represented by LBP codes when $P=8$ and $R=1$, are described in Fig. 14. For instance, code 0 represents a spot and code 4 represents an edge in the textures [32]. The final feature of the skin texture is a histogram of the LBP codes accumulated over the texture. To reduce the effect of the difference in an image size, the histogram is normalized by an image size. Because the histogram of the LBP codes contains detailed information of the texture, such as edges, lines and flat areas, it is suitable for analysis of the skin features.

In this paper, the LBP is used to extract the skin features for age estimation. Fig. 15 shows a comparison of the skin features of an adult and a child. As people get older, freckles, lentigines and fine wrinkles increase on the skin, as shown in Fig. 15(a) and (b). Fig. 15(c) shows the normalized histogram of the LBP codes of the skin texture of a 10 and a 76 years old when $P=8$ and $R=1$ of Fig. 11(b). Compared to the skin of the younger subject, the aged skin has more 2, 3 (corner), 4 (edge) and 5 codes. On the other hand, the younger skin has more 1, 7, 8 (spot/flat) and 9 (non-uniform pattern) codes. Fig. 15(d) and (e) shows the tendency of the normalized histogram of codes 4 and 8 in the BERC database. Code 4, which represents edges, tends to increase with age as shown in Fig. 15(d). On the other hand, code 8, which represents spots or flat areas, tends to decrease with age as shown in Fig. 15(e). This is explained by the fact that aging skin texture has many edge and corner components. On the contrary, young skin has more flat and non-uniform pattern components than an aging skin. Thus, the LBP histogram reflects the characteristics of skin aging and can be used for the skin features.

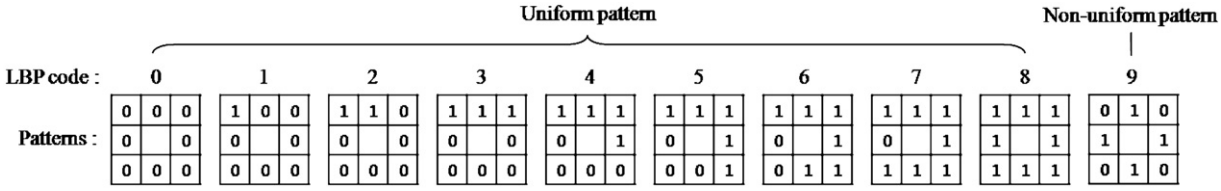

Fig. 13. Examples of the LBP codes for rotation invariant and uniform patterns when $P=8$ and $R=1$ from Fig. 11(b).

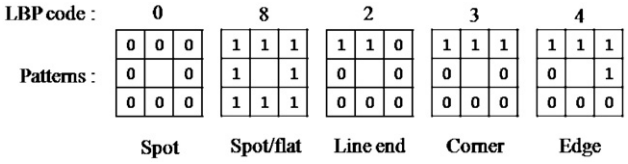

Fig. 14. Texture characteristics of the LBP code when $P=8$ and $R=1$ from Fig. 11(b).

3.5. Feature fusion method

To combine multiple features, a feature level, a score level, or a decision level fusion is used in biometrics [35]. The feature set, which is made by the feature level fusion, contains richer information about the raw data [46]. Also, the feature level fusion is able to eliminate the redundant information resulting from the correlation of fused features [45]. Therefore, we adopt the feature level fusion to combine global and local features. Our feature level fusion is summarized in Fig. 16. Given a face image, the appearance ($f_1=[a_1, \ldots, a_k]$) and the shape parameter ($f_2=[s_1, \ldots, s_l]$) of the AAM, the wrinkle features ($f_3=[w_1, \ldots, w_m]$) and the skin features ($f_4=[t_1, \ldots, t_n]$) are extracted in the feature extraction step (k, l, m and n are dimensions of each feature). In the feature normalization step, each feature is normalized by the z-score normalization as

$$f_i' = \frac{f_i - \mu_i}{\sigma_i}, \quad i = 1, 2, 3, 4 \tag{9}$$

where f_i is the ith feature vector, μ_i and σ_i are the mean and the standard deviation of the ith feature, respectively, f_i' is the normalized ith feature vector and i is a feature index. The fused feature is created by concatenating the normalized features as

$$f_{new} = [f_1' f_2' f_3' f_4'] = [a_1, \ldots, a_k, s_1, \ldots, s_l, w_1, \ldots, w_m, t_1, \ldots, t_n] \tag{10}$$

where f_{new} is the combined features. Concatenation results in an increase in the dimension of the feature, and leads to severe requirements for computation time and storage [47]. Therefore, the feature dimension is then reduced by the PCA as

$$f_{PCA} = [f_1, \ldots, f_p], \quad p < k+l+m+n \tag{11}$$

where p is the reduced dimension of the concatenated feature. Finally, f_{PCA} is used to estimate the age of the facial images.

The dimensions of f_{PCA} (e.g. p) in Eq. (11) were decided experimentally based on the resulting performance of the age estimation for each database. The performance of the age estimation is measured by increasing the PCA dimension, and the dimension with the best performance is then selected. The selected dimensions for the f_{PCA} are shown in Table 3. In the case of the FG-Net aging database, the noise caused by scanning the pictures and the various resolutions of the images are an obstacle to the analysis of the skin features. Consequently, the skin features are excluded in the FG-Net aging database. Because there are numerous irrelevant features and a great deal of noise in facial images, small dimensions are obtained by the PCA in the FG-Net database, where much noise is caused scanning the pictures with large variations in lighting, expression, background, pose, resolution, etc.

3.6. Hierarchical age estimation

The distribution of age features differs according to the age group. For example, facial wrinkles are usually created in adulthood, while geometric features change while growing up. If facial features are trained in the whole age range by a single estimator, this characteristic is not taken into account when estimating an accurate age. For this reason, a hierarchical age estimation method is proposed in the previous works to improve its performance in comparison to the single-level age estimation [2,4,5]. However, the hierarchical age estimation method using an age group estimator with a hard boundary that does not include overlapping classes has a problem, in that an error in the age group classification step is propagated to the specific age estimation step. For instance, if the face image of 15 year old is misclassified into the age 20–39 class, during the age group classification step, using an age group estimator with a hard boundary that does not include overlapping classes, the estimation error will be potentially larger than in a single-level age estimation. It is quite probable that the boundary data of a class be misclassified. That is, the face data of 19 year old is more likely to be misclassified into the age 20–39 class than that of 15 year old. To reduce such a classification error, a new hierarchical age estimation method, using overlapped classes between two adjoining classes, is proposed in this paper. It is composed of an age group classification step using a support vector machine (SVM) and an age estimation step, using support vector regression (SVR). The process of the proposed method is shown in Fig. 17. The combined features of the global and local features, explained in the previous section, are used in the hierarchical age estimation. The input data is first classified into an age group class by the SVM. The detailed age is then estimated by the SVR in the selected age group class. In the age group classification step, each classifier is determined by which age range is to be estimated and is overlapped taking into consideration the false acceptance error (FAE) and the false rejection error (FRE) of each classifier. The overlapped age class is set around the boundary of the class in order to reduce the boundary classification error. By compensating for the classification error, the total performance of the age estimation is improved.

The overlapped class has a similar concept to soften boundaries as a fuzzy classifier. To determine the overlap for an age class, we use the decision value from the SVM. The SVM is defined by [36]

$$f(X) = \text{sign}\left(\sum_{i=1}^{k} \alpha_i y_i K(s_i, x) + b\right) \tag{12}$$

where k is the data number, $y_i \in \{-1,1\}$ is the class labels of the training data, α_i is the Lagrange multiplier and b is a bias. $K(s_i,x)$ is a kernel function which is used for the non-linear classification, s_i is the support vector and $\sum_{i=1}^{k} \alpha_i y_i K(s_i,x)+b$ is the SVM decision value. If $f(x)$ is positive, the input feature vector (x) is classified as class 1, whereas if $f(x)$ is negative, the input feature vector (x) is classified as class 2. A conventional SVM uses just the decision value sign to classify, but the decision value can be used for additional classifiers. For instance, a fuzzy SVM using the decision value in a prediction process is proposed by Xie et al. [37]. An overlapped age class is decided based on two output distributions of the SVM classifier. For example, by using the training data of

classes 1 (ω_1) and 2 (ω_2), the two output distributions of the SVM classifier are obtained as shown in Fig. 18. Inevitably, the two output distributions have an overlapped area, as shown in Fig. 18(a). In the case of a classification without an overlapped class, the classification error occurring in the FRE of class 1 (or the FAE of class 2) and the FAE of class 1 (or the FRE of class 2) is represented by a diagonal pattern in Fig. 18(a). In order to reduce this classification error, the overlapped class of classes 1 and 2 is set based on the threshold position (e.g. th^+ and th^-) in the overlapped region of the two output distributions of the SVM classifier, as shown in Fig. 18(b). If the threshold position (e.g. th^+ and th^-) of the SVM decision value is decided in training, the data, most likely to be misclassified, can be predicted using the threshold position in the test. By processing the misclassified data through the overlapped class as distinguished from other data, the classification error is reduced as represented by the diagonal pattern in Fig. 18(b). The threshold value (e.g. th^+ and th^-) used to decide the overlapped class is defined as $\mu \pm \alpha$, the μ is the mean of the SVM decision value of the misclassified data in class 1 (e.g. FAE class 1) and class 2 (e.g. FRE class 1), α is an additional value. To find the optimal threshold value, the additional value α was selected in a range -0.1 to 0.1 through experimentation.

The experimental results of the classification using training data for the FG-Net database are shown in Fig. 19. The training data are classified into two classes based on a 15 year old, using the SVM. The SVM decision value of each data point is shown in Fig. 19(a), and the data distribution according to the SVM value is shown in Fig. 19(b). In Fig. 19(a), the misclassified data is represented by a filled marker. From the misclassified data, the overlapped range of the data distribution appears in Fig. 19(b) to be similar to Fig. 18(a). Therefore, the classification error of the practical data can be reduced by an overlapped age class such as that found in Fig. 18(b). In this paper, we use seven classes, including four age group classes and three overlapped classes, as shown in Fig. 17. The age criterion for the classification is selected experimentally at each database, because it is affected by age distribution. In each age class (e.g. classes 1–7) and each overlapped age class (e.g. the overlapped age region of classes 1 and 2, 2 and 3 and 3 and 4), the age estimator based on the SVR is applied and the specific age is estimated.

4. Experiments

4.1. Databases

Three databases are used to evaluate the performance of the proposed method: the BERC database, the PAL aging database and the FG-Net aging database. The BERC database was collected by Biometrics Engineering Research Center (BERC) [38] in order to estimate the ages of Koreans. The database contains 390 people's images in the age range 3–83 years old. The facial images were obtained using a digital camera (NIKON coolpix p5000) with a very

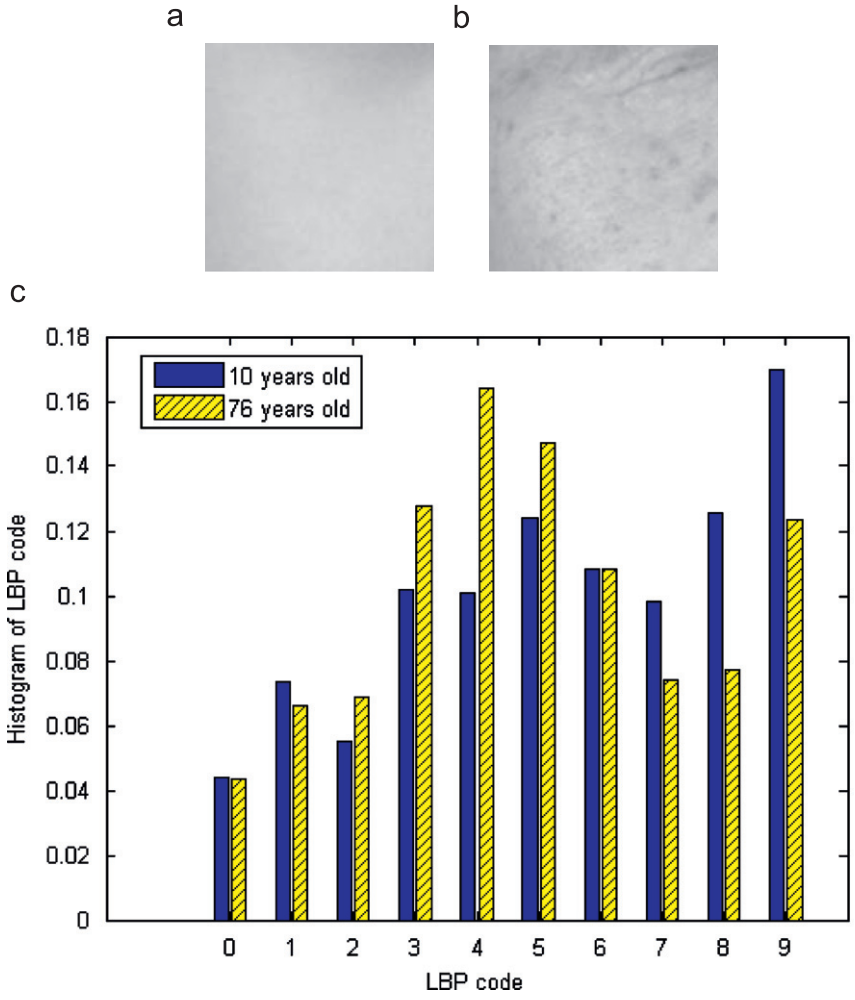

Fig. 15. Analysis of skin texture using LBP histogram: (a) skin texture of 10 year old, (b) skin texture of 76 year old, (c) normalized histogram of LBP codes of skin texture of 10 and 76 years old, where $P=8$, $R=1$, (d) normalized histogram of LBP code 4 according to age, where $P=8$, $R=1$ and (e) normalized histogram of LBP code 8 according to age, where $P=8$, $R=1$.

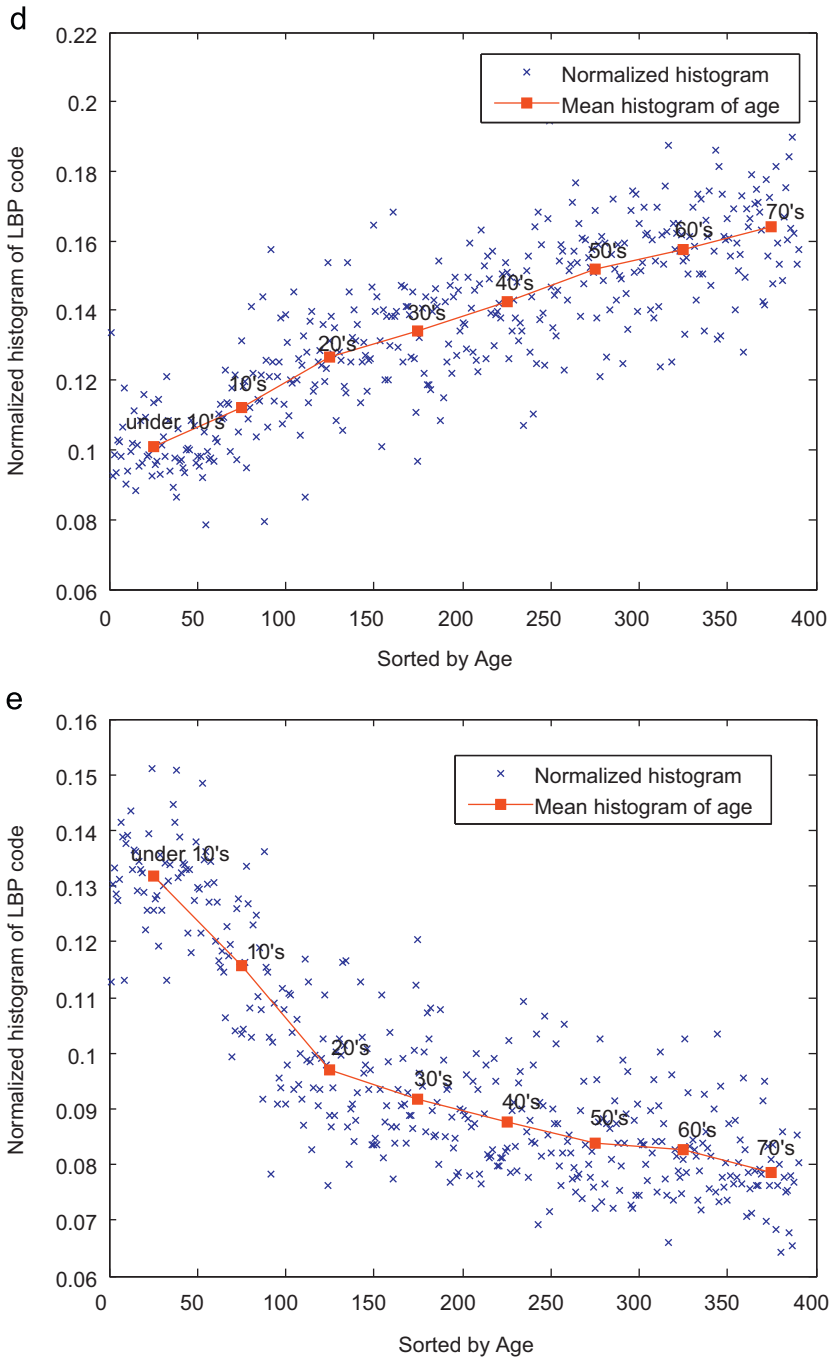

Fig. 15. (*Continued*)

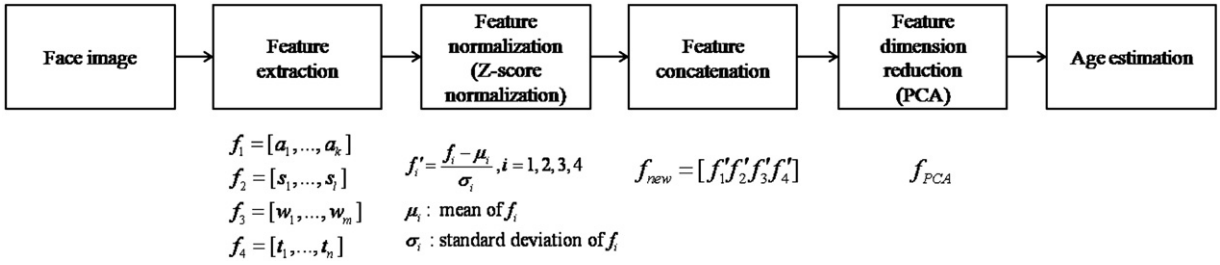

Fig. 16. Block diagram of our feature level fusion.

Table 3
Features dimension for each database.

Database	f_1 (appearance)	f_2 (shape)	f_3 (wrinkle)	f_4 (skin)	f_{new} (global+local)	f_{PCA} (global+local with PCA)
BERC DB	166	32	232	60	490	483
PAL DB	133	37	232	60	462	319
FG-Net DB	89	28	232	0	349	260

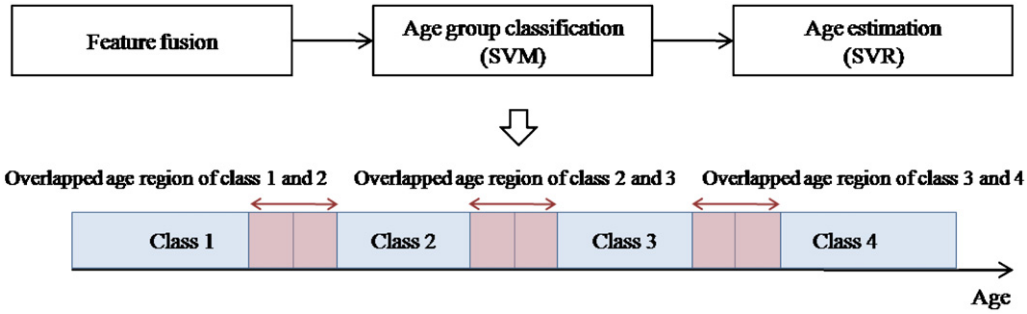

Fig. 17. The proposed hierarchical age estimation method.

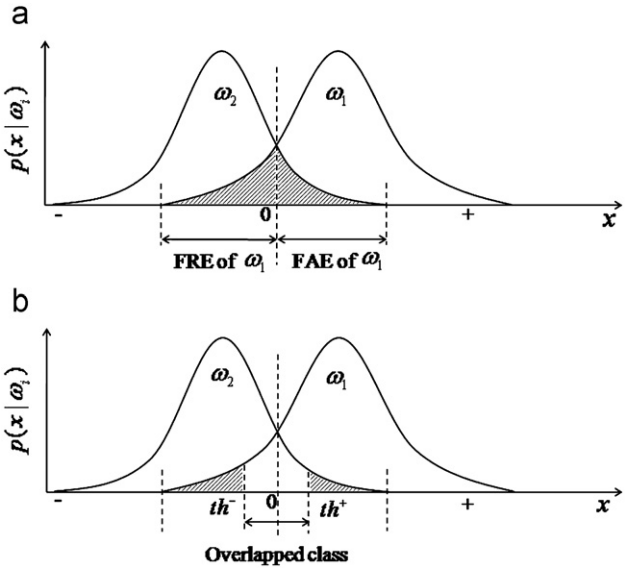

Fig. 18. A comparison of the classification errors with and without an overlapped class: (a) data distributions according to the SVM decision value (x) for both class 1 (ω_1) and class 2 (ω_2), without having an overlapped class and (b) data distributions according to the SVM decision value (x) for both class 1 (ω_1) and class 2 (ω_2) with an overlapped class.

high resolution of 3648×2736 pixels. In our experiments, we scaled down the images to 1600×1200 pixels in order to reduce the processing time. All of the images have no variance in light and expression. The subjects, with respect to gender and age, are uniformly distributed as shown in Fig. 23(a). Therefore, the database is believed to be better suited for analysis of the estimation performance than the others. Sample images from the BERC database are shown in Fig. 20.

The PAL aging database contains 430 Caucasians with age range 18–93 years old [39]. The images in the database were captured using a digital camera with fixed light and position conditions. The resolution of the images is 640×480 pixels. The database includes various expressions such as smiling, sadness, anger, or neutral faces. In the experiments, we used only neutral faces in order to exclude the facial expression effect. Sample images used in our experiments are shown in Fig. 21.

The FG-Net aging database is the most frequently used database for estimating age in the previous works, as it is a publicly available database. The database has 1002 images composed of 82 Europeans in the age range 0–69 years old. Individuals in the database have one or more images included at different ages. Images were obtained by scanning, unlike the other databases. Therefore, there are extreme variations in lighting, expression, background, pose, resolution and noise from scanning. Sample images of the FG-Net aging database are shown in Fig. 22. Each facial image has 68 labeled facial feature points used for the shape features [40].

The data distributions of each database according to age are shown in Fig. 23.

4.2. Experimental setup

With the BERC and PAL aging databases, five-folds cross validations are performed to evaluate the performance of age estimation, as was done in [3]. The age and gender are evenly distributed for each fold. With the FG-Net aging database, Leave-One-Person-Out (LOPO) is performed as it contains a number of images of the same person. The LOPO is used mainly to prevent the same person from being included in both the training and the test data set [11–14]. Finally, 82-folds are used on the FG-Net aging database. For facial feature extraction, manually obtained facial feature points were used. Manually obtained points were used, because we concentrated on problems of extracting facial feature and age estimation without the effects of AAM fitting errors. Also, the FG-Net database, the most frequently used public database for age estimation, provides 68 facial feature points with images. Therefore, manually obtained facial points are used for age estimation in many previous works [11–14]. Additionally, experiments for comparison of using manually and automatically obtained facial points were shown in Section 4.3.4.

The optimal parameters of the SVR, SVM and threshold value of the overlapped age range were selected experimentally from the training set. In the hierarchical age estimation, the SVM is used as the classifier in the age group classification step. A conventional SVM is designed for two-class problems by selecting the optimal linear decision hyper-plane [41]. To extend this to non-linear

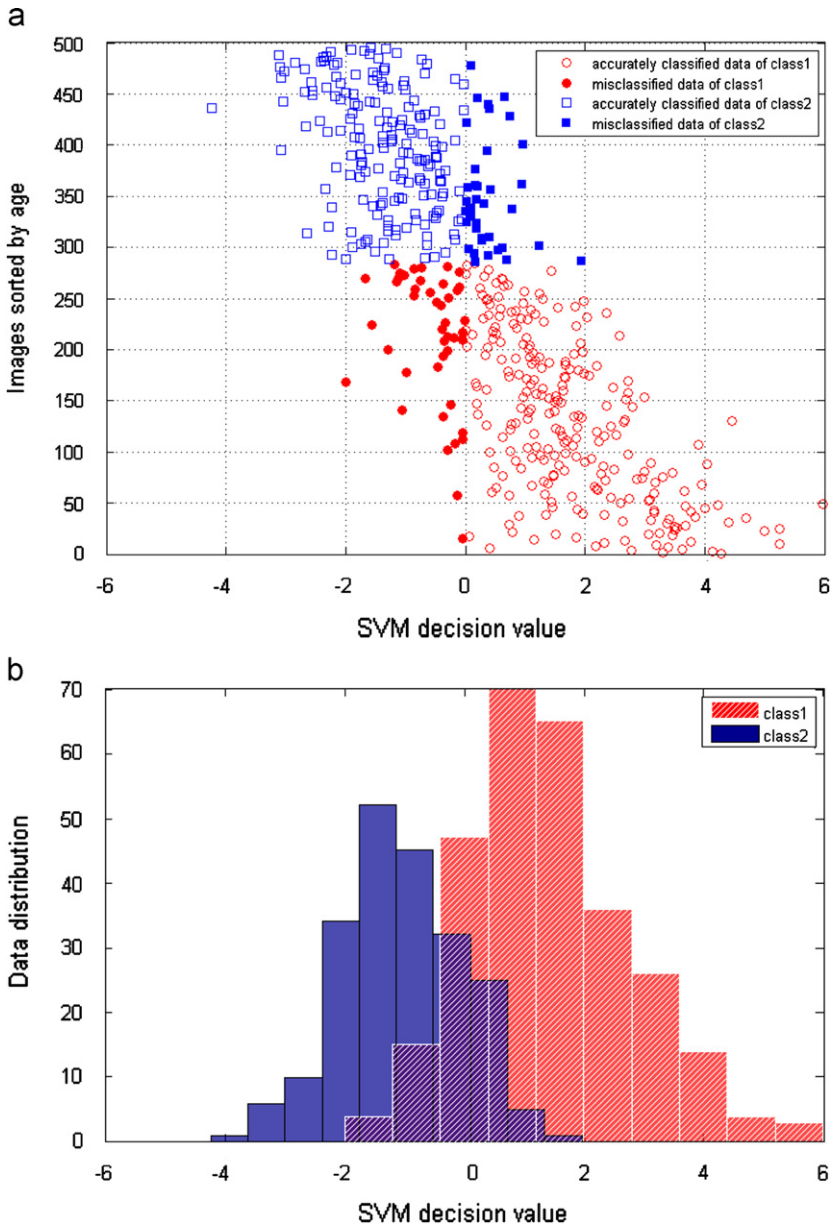

Fig. 19. The distribution of the training data according to the SVM decision value from the FG-Net database: (a) the SVM decision value of the images sorted by age and (b) the data distribution according to the SVM decision value.

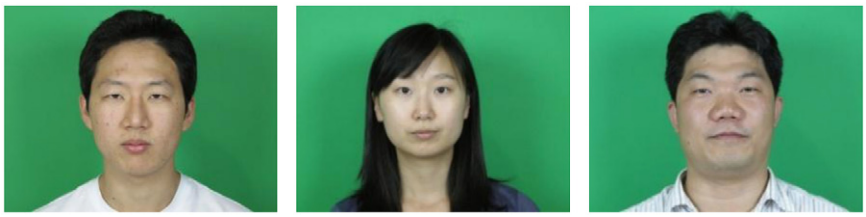

Fig. 20. Sample images from the BERC database.

surfaces, kernel function $K(s_i,x)$ is used. In this paper, the radial basis function (RBF) kernel was used, as shown in Eq. (13) [36]

$$K(x, y) = e^{-||x-y||^2/2\sigma^2} \tag{13}$$

To estimate an accurate age, the regression method is used. The relationship between an age and a facial feature vector (b) is defined by an aging function as

$$age = f(b) \tag{14}$$

where b is the facial feature vector, f is the aging function and age is the estimated age. In this research, we used the Support Vector Regression (SVR) [42]. The SVR is a version of an SVM for regression problems used to estimate real-valued functions [36]. The main

Fig. 21. Sample images from the PAL aging database.

Fig. 22. Sample images from the FG-Net aging database.

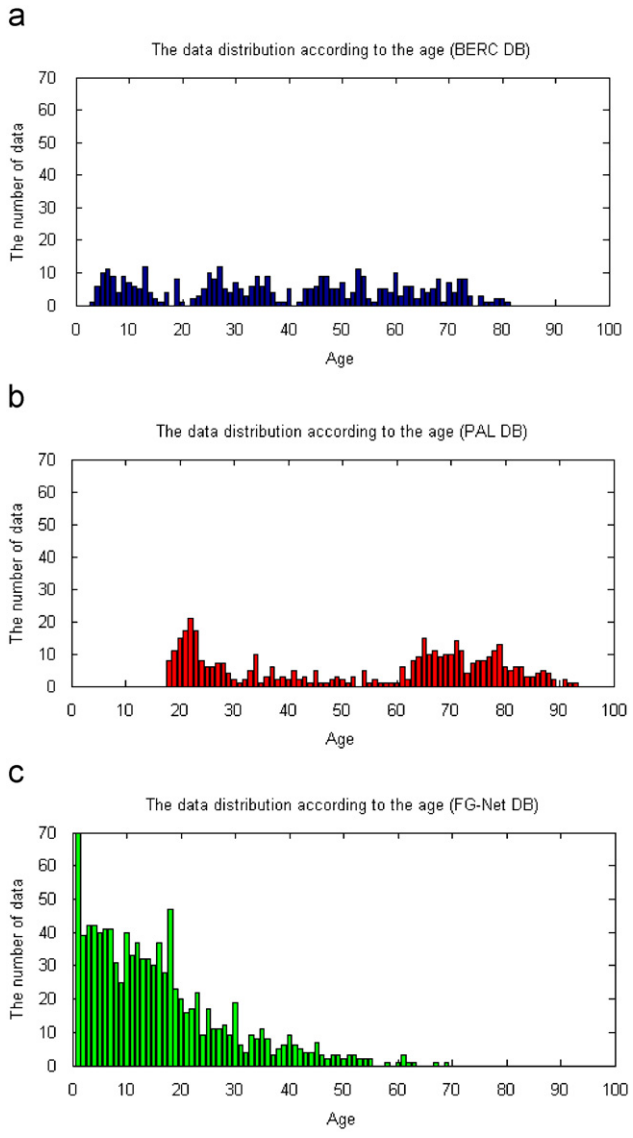

Fig. 23. The data distribution of databases according to age: (a) BERC database, (b) PAL database and (c) FG-Net database.

goal of the SVR is to find a function $f(b)$ that has at most ε deviation from the actually obtained targets label (age) for all training data, and is as flat as possible simultaneously [44]. As in the SVM, the RBF kernel function was used in the SVR for the nonlinear regression. In the training of the SVM and the SVR, the regularization parameter C and the kernel parameter σ are determined with a training set, experimentally. The mean absolute error (MAE) and the cumulative score (CS) are used to evaluate performance, which have been commonly used in the previous works [11–14]. The MAE is defined as the mean of the absolute difference between the estimated age and the ground-truth age as

$$MAE = \sum_{k=1}^{N} |\hat{a}_k - a_k| / N \tag{15}$$

where N is the number of the test data, $\hat{a}_k$ is the estimated age of the test image k and a_k is the ground-truth age of the test image k. The cumulative score (CS) is defined as the ratio of the number of data whose errors are lower than a threshold value to the total number of data. It is expressed by

$$CS(th) = N_{e \le th} / N\, 100(\%) \tag{16}$$

where $N_{e \le th}$ is the number of data whose estimation error (e) is lower than threshold (th), N is the number of total data.

4.3. Experimental results

In this subsection, we show the results of the performance evaluation on the proposed age estimation method in terms of the MAE and CS using the BERC database, the PAL aging database and the FG-Net aging database.

4.3.1. Performance evaluation when using the proposed local features

As was mentioned in Section 3.3, the wrinkle features are extracted by Gabor filter sets selected using the predominant orientation of the wrinkles. In order to analyze the effects of the number of the Gabor filters, the performances of the wrinkle features extracted by different Gabor filter sets were compared. Two Gabor filter sets with 4 scales, 6 orientations and 5 scales, 8 orientations were used in this experiment as shown in Fig. 24. Table 4 shows the performance of the wrinkle feature according to different scales and orientations of Gabor filter sets. As a result, the performances of wrinkle features of two cases of Gabor filters were

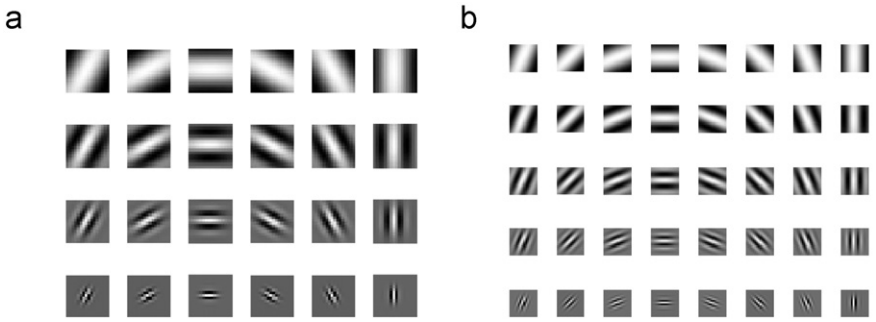

Fig. 24. Gabor filter with different scales and orientations: (a) Gabor filter with 4 scales, 6 orientations and (b) Gabor filter with 5 scales, 8 orientations.

Table 4
The performance of wrinkle feature according to scale and orientation of Gabor filter in terms of the MAE/standard deviation of error. (unit: years old.)

Database	Gabor filter with 4 scales, 6 orientations	Gabor filter with 5 scales, 8 orientations
BERC DB	6.1769/4.8076	6.0307/5.0088
PAL DB	7.7126/6.7062	8.1612/6.5747
FG-Net DB	8.1996/7.5774	8.0140/7.7892

Table 5
The performance of skin feature according to P and R of an LBP in the BERC database in terms of the MAE/standard deviation of error (unit: years old).

(P, R)	Skin feature
(8, 1)	6.9948/6.0051
(8, 2)	8.2769/6.7846
(8, 3)	8.8846/6.9441
(8, 1) (8, 2)	7.1333/6.1253
(8, 2) (8, 3)	7.1948/6.1613
(8, 1) (8, 3)	6.7307/5.9160
(8, 1) (8, 2) (8, 3)	6.8846/6.0400

similar. Therefore, we used 24 Gabor filters with 4 scales and 6 orientations to reduce the computational time.

The skin features are extracted by an LBP as mentioned in Section 3.4. The multi-resolution analysis is possible using various values of the number of the neighboring pixels (P) and the distance from the center to the neighboring pixels (R) in an LBP as shown in Fig. 11(b). The increase of the number of neighboring pixels causes the increase of the dimension of features, and the calculation of Eq. (8) is more required. As a result, the complexity of the process increases. However, the overall performance is not necessarily significantly affected by the increase of the number of neighboring pixels [43]. Therefore, we used eight neighboring pixels (P) with various distances from the center to the neighboring pixels (R).

In this paper, the optimal P and R are obtained by experiments such as that in [42]. In order to select the optimal values of P and R, the combinations of $P=8, R=1$, $P=8, R=2$ and $P=8, R=3$ were used to extract skin features. Table 5 shows the performance of skin feature with various combinations of P and R of an LBP for multi-resolution in the BERC database. As a result, we used the varying values of $P=8, R=1$ and $P=8, R=3$ which gives better performance than other combinations.

To evaluate the performance on the proposed wrinkle and skin features, experiments were performed to compare the proposed method with a previous method. In the previous works, the wrinkle features were usually extracted by analysis of the edges in the images. In this experiment, the Sobel filter is selected as the wrinkle features extractor for the previous methods and is usually used for analysis of facial wrinkles [6–9]. The mean and variance of the Sobel filtered images are used for the wrinkle features, as found in [7]. In both the proposed method and previous methods, nine wrinkle areas are used, as shown in Fig. 8. The age is estimated with the SVR using only the wrinkle feature extracted by each method. The experimental results are shown in the first and second columns of Table 6. It shows that our proposed method for extracting wrinkle features offers a better performance than previous works using the Sobel filter on all three databases. The MAE of the proposed method improved by about 23.35% on the BERC database, 30.84% on the PAL aging database and 4.83% on the FG-Net aging database as compared to those using the Sobel filter [7]. The Sobel filter analyzes only the strength of the wrinkles, but the proposed method analyzes both the strength and orientation of the wrinkles. By using the orientation of the wrinkles, the proposed method gives a better performance in age estimation.

The age estimation performance on the proposed skin features was compared with the previous works. In the previous works, the skin features were usually extracted by analyzing high frequency components in the facial images. One of them, the statistics using the difference of the average filtered image was used in [7]. The method to extract skin features in [7] is summarized as follows. The original image is smoothed by a 5×5 pixel average filter, and a difference image between the original and the smoothed images is created. In the difference image, the blotches on the facial skin remained. The mean and variance of the pixel values in the difference image containing blotches are used for the skin features. In this experiment, we compared the proposed LBP method with the statistics of blotches in the difference images as the skin feature extractor. The experimental results are shown in the third and fourth columns of Table 6. In the case of the FG-Net aging database, the noise caused by scanning the pictures and the various resolutions of the images are an obstacle to the analysis of the skin features, therefore resulting in its having been excluded.

The results show that the proposed LBP method is superior to the previous method [7] on the BERC and PAL aging databases. The MAE of the proposed method is improved by about 52.03% on the BERC database and 22.12% on the PAL aging database as compared to those using the difference image. The skin features extracted by the LBP include detailed texture features, such as the edge, line end and flat areas. Therefore, they are more discriminative features to represent texture than the statistics of the difference image between the smoothed and the original images of [7].

Additionally, the performances of skin feature are compared in various image resolutions for analysis of the influence of resolution. For experiments, we used the BERC database which has the highest resolution, because it is able to resize to various resolutions. As a result of these experiments, the skin feature extracted by an LBP is influenced by image resolution, as shown in Fig. 25. Nevertheless,

Table 6
Comparison of the proposed method for extracting the wrinkle and skin features with a previous method in terms of the MAE/standard deviation of error (unit : years old).

Method	Wrinkle feature		Skin feature	
	Sobel filter [7]	**Region specific Gabor filter (proposed method)**	Difference image [7]	**LBP (proposed method)**
BERC DB	8.0590/6.6448	6.1769/4.8076	14.0333/10.1293	6.7308/5.9160
PAL DB	11.1519/8.2920	7.7126/6.7062	12.0093/10.0474	9.3528/7.7211
FG-Net DB	8.6158/8.1433	8.1996/7.5774	–	

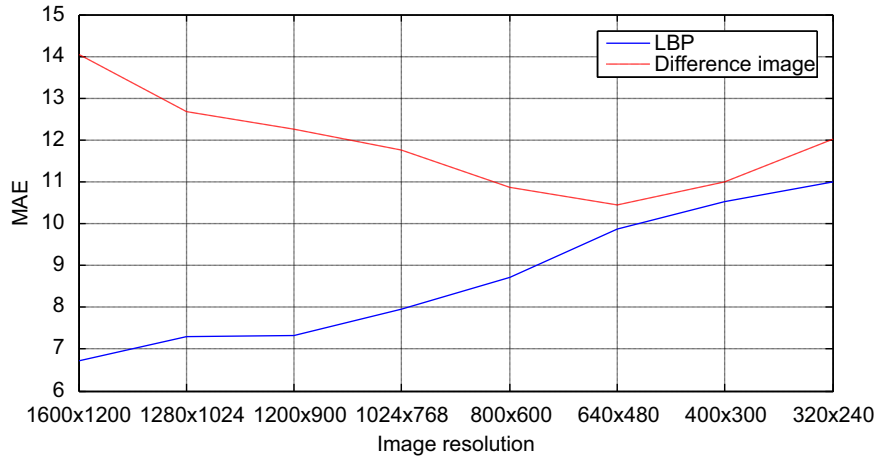

Fig. 25. The performance of skin feature according to an image resolution on the BERC database.

Table 7
Comparison of the proposed hierarchical age estimation method and the other estimation methods in terms of the MAE/standard deviation of error (unit: years old).

Database	Single-level age estimation [3]	Hierarchical age estimation without the overlapped age class [2,4,5]	Hierarchical age estimation with the overlapped age class (proposed method)
BERC DB	4.8436/3.8956	4.7846/4.2335	4.6821/4.0519
PAL DB	5.3551/4.1669	4.5164/4.4713	4.3271/4.1664
FG-Net DB	5.3194/5.7841	5.0050/6.2242	4.6577/5.7274

the skin feature extracted by an LBP is superior to the previous methods using a difference image [7] in all the cases of image resolutions.

4.3.2. Performance evaluation when using the proposed hierarchical age estimation

In this research, the new hierarchical age estimation method which considers the overlapped age class, shown in Fig. 17, was proposed. To measure the performance of the proposed method, three methods for age estimation were compared in this experiment. These are the single-level age estimation [3], the hierarchical age estimation without the overlapped age class [2,4,5], and the hierarchical age estimation with the overlapped age class. The combined features of the AAM, wrinkle and skin features were used in this experiment. The results of the experiments are shown in Table 7. When the hierarchical age estimation without the overlapped age class was used, the performance was improved as compared to that of the single-level age estimation. The reason for the improvement is that the effective features for age estimation are different in each age range. By the proposed method using hierarchical age estimation with overlapped age classes, the performance was even more improved due to the compensation for the classification error of the boundary data by the overlapped age classes. The MAE is improved by about 2.14% on the BERC database, 4.19% in the PAL aging database and 6.93% on the FG-Net aging database as compared to those of the hierarchical age estimation without the overlapped age classes.

The cumulative scores for the comparison of the proposed hierarchical age estimation method (with the overlapped age classes) and the previous hierarchical age estimation (without the overlapped age classes) on each database at the error level from 0 to 15 (years old) are shown in Fig. 26. The cumulative score moved up an average of 0.3846 on the BERC database, 1.0514 on the PAL aging database and 1.5032 on the FG-Net aging database using the proposed hierarchical age estimation compared to the previous method. In the FG-Net aging database, many data is concentrated under the age of 20, with over 70% of the data as shown in Fig. 23(c). The proposed hierarchical age estimation is more helpful for non-uniform distributed databases than uniform databases. As shown in Eq. (16), the higher the cumulative score becomes, the accuracy of the age estimation also becomes higher. From Table 7 and Fig. 26, the age estimation performance is improved by the proposed hierarchical age estimation in both the MAE and the CS sense.

4.3.3. Comparison of using only the global features to the combined global and local features

As previously described, the AAM features do not include detailed information about wrinkles or skin. In order to

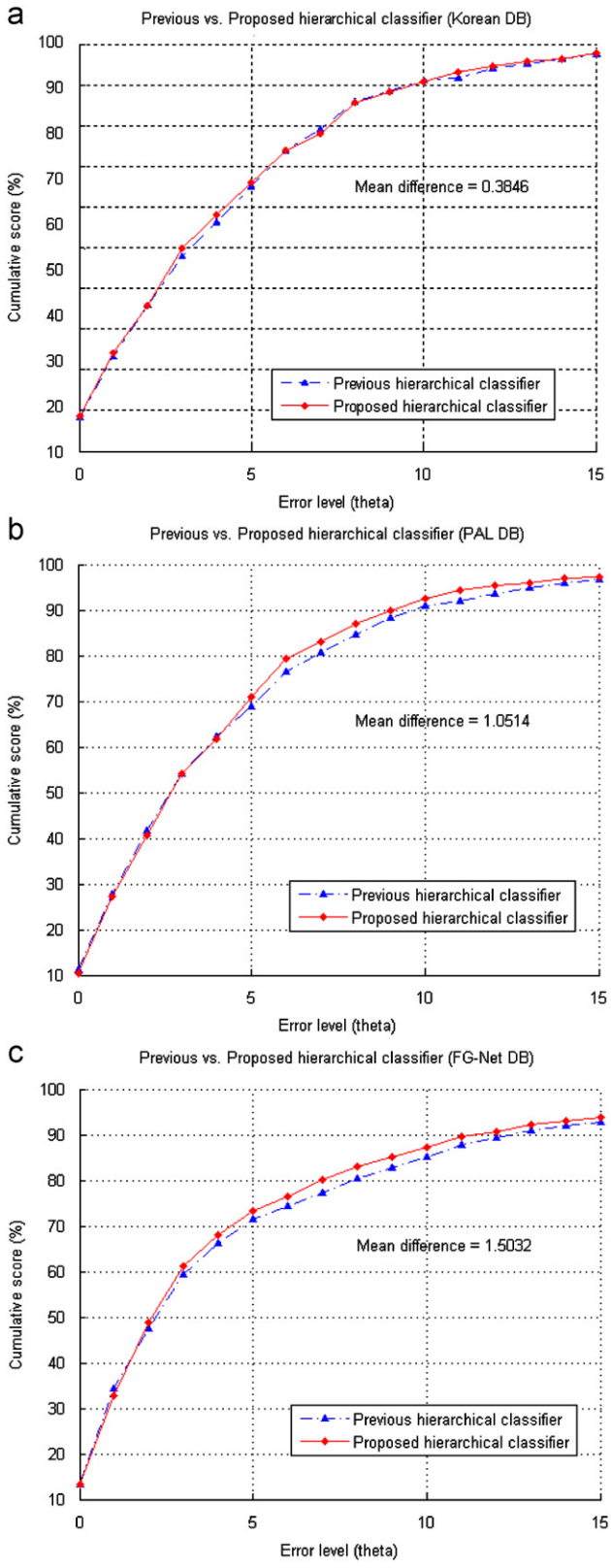

Fig. 26. Cumulative scores comparing the proposed hierarchical age estimation method (with the overlapped age class) and the previous hierarchical age estimation (without the overlapped age class): (a) BERC database, (b) PAL aging database and (c) FG-Net aging database.

complement the AAM features, the wrinkle and skin local features can be combined with the AAM parameters [3,21]. In this experiment, the performance using only the global features was compared to the combined global and local features, using the proposed hierarchical age estimation methods. For the global features, the appearance and shape parameters of the AAM were used. Table 8 shows the experimental results. The MAE is improved by about 14.31% on the BERC database, 4.68% on the PAL aging database and 4.67% on the FG-Net aging database, as compared to those using only the global features. Because the resolution of the images is very important for the representation of the local facial features, the performance of age estimation is most improved on the BERC database, which uses high resolution images, when combining the global and local features.

Table 8
Comparison of using only the global features to the combined global and local features in terms of the MAE/standard deviation of error (unit: years old).

Database	Global features	Global+local features
BERC DB	5.4641/5.4208	4.6821/4.0519
PAL DB	4.5397/4.9286	4.3271/4.1664
FG-Net DB	4.8862/6.0528	4.6577/5.7274

Also, the cumulative score moved up on an average of 3.3814 on the BERC database, 0.5403 on the PAL aging database and 0.8982 on the FG-Net aging database by combining with the proposed local features, as shown in Fig. 27. As a result, the combined global and local features improve the performance of the age estimation.

4.3.4. *Comparison of using manually and automatically obtained facial points*

In order to properly analyze the effects of an AAM fitting error, the performance was compared using manually and automatically obtained facial points. The hybrid features and proposed hierarchical age estimation methods were used for this experiment, and the experimental results shown in Table 9. As this table shows, the age estimation errors when facial feature points are automatically detected by an AAM increase marginally compared to those with the manually detected points, allowing confirmation that the proposed method can work with an automatic feature extraction by an AAM.

4.3.5. *Comparisons of the proposed method to the previous works*

In this experiment, the performance of the proposed method is compared with the previous works, in which the performances are reported on the FG-Net aging database, using the AAM features and the LOPO method. As with the previous works, facial feature points provided in the FG-Net aging database were used to extract aging features. The Weighted Appearance Specific (WAS) is a method using the aging function learned by the weighted sum of the known aging functions [20]. The aging pattern subspace method (AGES) is a subspace representation of a sequence of individual aging face images [11]. The regression with uncertain nonnegative labels (RUN) is a regression method using an age uncertainty [12]. The local adjustment of the regression results (LARR) and the probabilistic fusion approach (PFA) are complementary methods for the regression error proposed in [13,14]. The method of design sparse features, which combines global and local features [3], is compared to evaluate the performance of combining the global and local features.

The MAEs of the previous methods and the proposed method are tabulated in Table 10. Our method has an MAE of 4.6577, which is lower than all the previous methods listed in Table 10. Our method is superior to the methods, using only AAM features (WAS [11], AGES [11], RUN [12], LARR [13], PFA [14]) or uses a hybrid feature

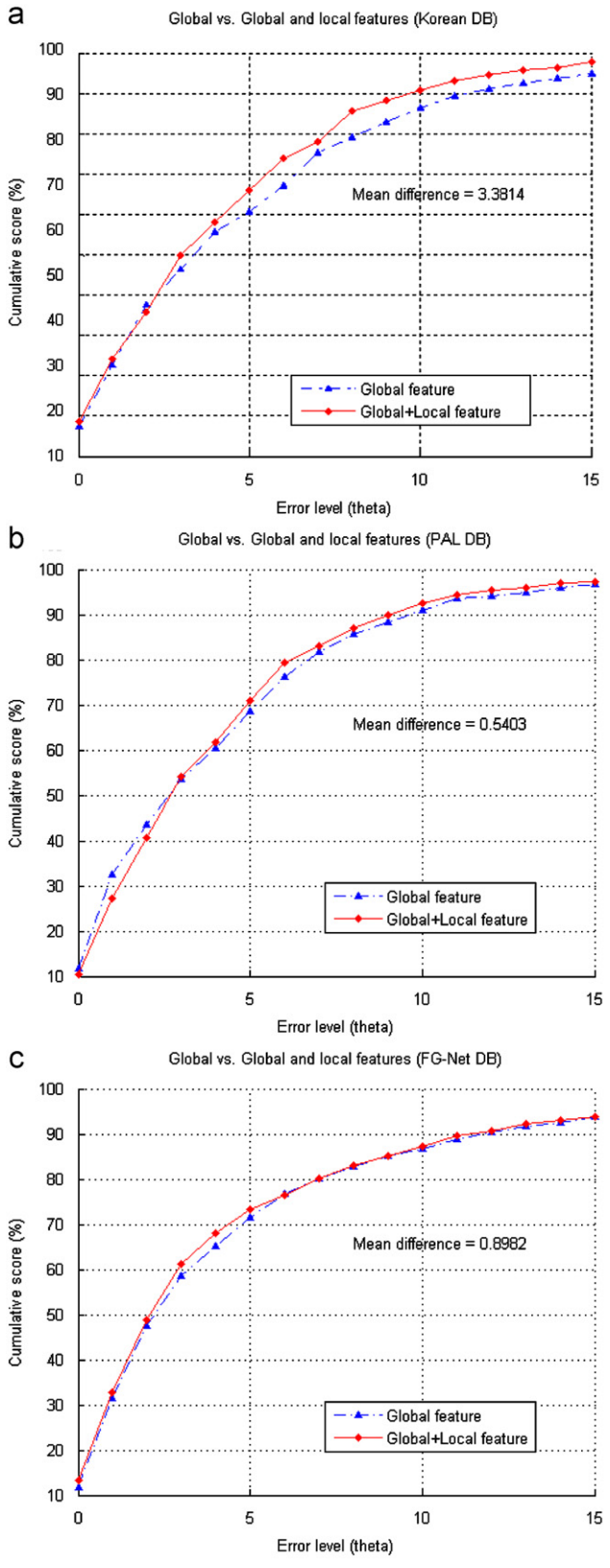

Fig. 27. Cumulative scores comparing the use of only the global features to the combined global and local features: (a) BERC database, (b) PAL aging database and (c) FG-Net aging database.

(design sparse features [3]). This result shows that the proposed method improves the performance of the age estimation better than the previous works.

Table 9
Comparison of using manually and automatically obtained facial points in terms of the MAE/standard deviation of error (unit: years old).

Database	Manually obtained facial points	Automatically obtained facial points
BERC DB	4.6821/4.0519	5.3846/4.8601
PAL DB	4.3271/4.1664	6.0234/5.5578
FG-Net DB	4.6577/5.7274	5.5938/7.0863

Table 10
Comparison of the proposed methods to the previous works.

Methods of age estimation	MAE
WAS [11]	8.06
AGES [11]	6.22
RUN [12]	5.78
LARR [13]	5.07
PFA [14]	4.97
Design sparse features [3]	5.974
Proposed methods	4.6577

5. Conclusions and future works

In this paper, we proposed an age estimation method combining hybrid features and a hierarchical classifier. In addition, feature extraction methods for wrinkles and skin, as well as a new hierarchical age estimation method, are proposed. The wrinkle features are extracted by a region specific Gabor filter set and the skin features are extracted, using the LBP method. The hierarchical age estimation is designed to overlap the age classes in order to reduce classification errors of the boundary data. The experimental results showed that the local features extracted by the proposed methods provide a better performance than the previous methods. We also improved the performance by using the proposed hierarchical method.

Also, we were able to ascertain that the facial local features such as wrinkles and skin are affected more by image conditions as opposed to the global features. And, as the skin texture in images from the FG-Net aging database does not have enough aging information, we only used the wrinkle feature in the FG-Net aging database. In the future, we will conduct research into extracting the local features under various conditions, such as expression, light, gender, etc.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) through the Biometrics Engineering Research Center (BERC) at the Yonsei University (R112002105070030(2010)).

References

[1] Xin Geng, Zhi-Hua Zhou, Yu Zhang, Gang Li, Honghua Dai, Learning from facial aging patterns for automatic age estimation, in: Proceedings of the 14th annual ACM international conference on Multimedia (2006) 307–316.
[2] Andreas Lanitis, Chrisina Draganova, Chris Christodoulou, Comparing different classifiers for automatic age estimation, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 34 (1) (2004) 621–628.
[3] Jinli Suo, Tianfu Wu, Songchun Zhu, Shiguang Shan, Xilin Chen, Wen Gao, Design sparse features for age estimation using hierarchical face model, in: Proceedings of the Eighth IEEE International Conference on Automatic Face & Gesture Recognition, FG '08. (2008) 17–19.
[4] G.-D. Guo, G. Mu, Y. Fu, C.R. Dyer, T.S. Huang, A study on automatic age estimation using a large database, in: Proceedings of the IEEE Conference on Computer Vision (ICCV'09) (2009).
[5] Khoa Luu, Karl Ricanek Jr, Tien D. Bui, Ching Y. Suen, Age estimation using active appearance models and support vector machine regression, in: Proceedings of the IEEE Third International Conference on Biometrics: Theory, Applications, and Systems (2009) 1–5.

[6] Wen Bing Horng, Cheng Ping Lee, Chun Wen Chen, Classification of age groups based on facial features, Tamkang Journal of Science and Engineering 4 (3) (2001) 183–192.
[7] Hironori Takimoto, Yasue Mitsukura, Minoru Fukumi, Norio Akamatsu, Robust gender and age estimation under varying facial pose, Electronics and Communications in Japan 91 (7) (2008) 32–40.
[8] Jun-Da Txia, Chung-Lin Huang, Age estimation using AAM and local facial features, in: International conference on Intelligent Information Hiding and Multimedia Signal Processing (2009) 885–888.
[9] J. Hayashi, M. Yasumoto, H. Ito, H. Koshimizu, A method for estimating and modeling age and gender using facial image processing, in: Proceedings of the Seventh International Conference on Virtual Systems and Multimedia (VSMM'01) (2001) 439.
[10] Young H. Kwon, Niels da Bitoria Lobo, Age classification from facial images, Computer Vision and Pattern Recognition 74 (1) (1999) 1–21.
[11] Xin Geng, Zhi-Hua Zhou, K. Smith-Miles, Automatic age estimation based on facial aging patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (12) (2007) 2234–2240.
[12] Shuicheng Yan, Huan Wang, Xiaoou Tang, T.S. Huang, Learning auto-structured regressor from uncertain nonnegative labels, in: IEEE 11th International Conference on Computer Vision, ICCV (2007) 1–8.
[13] Guodong Guo, Yun Fu, C.R. Dyer, T.S. Huang, Image-based human age estimation by manifold learning and locally adjusted robust regression, IEEE Transactions on Image Processing 17 (7) (2008) 1178–1188.
[14] Guodong Guo, Yun Fu, C.R. Dyer, T.S. Huang, A probabilistic fusion approach to human age prediction, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (2008) 1–6.
[15] G. Leslie, M.D. Farkas, Anthropometry of the head and face, second ed., Raven Press, New York, 1994.
[16] H.C. Lian, B.L. Lu, Age estimation using a min–max modular support vector machine, in: 12th International Conference on Neural Information Processing(ICONIP 2005) (2005) 83–88.
[17] Feng Gao, Haizhou Ai, Face age classification on consumer images with gabor feature and fuzzy LDA method, Lecture Notes In Computer Science, in: Proceedings of the Third International Conference on Advances in Biometrics 5558 (2009) 132–141.
[18] Yun Fu, T.S. Huang, Human age estimation with regression on discriminative aging manifold, IEEE Transactions on Multimedia 10 (4) (2008) 578–584.
[19] H. Fukai, H. Takimoto, Y. Mitsukura, M. Fukumi, An apparent age estimation system using the evolutionary algorithm, in: International Conference on Control, Automation and Systems, ICCAS ′07. (2007) 2146–2149.
[20] Andreas Lanitis, Christopher J. Taylor, Timothy F. Cootes, Toward automatic simulation of aging effects on face images, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (4) (2002) 442–455.
[21] Jinli Suo, Song-Chun Zhu, Shiguang Shan, Xilin Chen, A compositional and dynamic model for face aging, IEEE Transactions Pattern Analysis and Machine Intelligence 32 (3) (2009) 385–401.
[22] Hong Chen, Song-Chun Zhu, A generative sketch model for human hair analysis and synthesis, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (7) (2006) 1025–1040.
[23] A. Gunay, V.V. Nabiyev, Automatic age classification with LBP, in: 23rd International Symposium on Computer and Information Sciences, ISCIS ′08 (2008) 1–4.
[24] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (6) (2001) 681–685.
[25] Iain Matthews, Simon Baker, Active appearance models revisited, International Journal of Computer Vision 60 (2) (2004) 135–164.
[26] Alexandre Cruz Berg, Silvana Cardoso Justo, Aging of orbicularis muscle in virtual human faces, in: Seventh International Conference on Information Visualization, Proceedings (2003) 164–168.
[27] G. Lemperle, R.E. Holmes, S.R. Cohen, S.M. Lemperle, A classification of facial wrinkles, Plastic and Reconstructive Surgery 108 (6) (2001) 1735–1750.
[28] P&G beauty & grooming, Defining issues. Available: 〈http://www.pgbeautyscience.com/defining-issues-skintone.html〉 (accessed 18.03.10).
[29] B.S. Manjunathi, W.Y. Ma, Texture features for browsing, retrieval of image data, IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (8) (1996) 837–842.
[30] Ho Gi Jung, Jaihie Kim, Constructing a pedestrian recognition system with a public open database, without the necessity of re-training: an experimental study, Pattern Analysis and Applications 13 (2) (2010) 223–233.
[31] Timo Ojala, Matti Pietikäinen, Topi Mäenpää, Multiresolution Gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) (2002) 971–987.
[32] T. Mäenpää, M. Pietikäinen, Texture analysis with local binary patterns, Handbook of Pattern Recognition and Computer Vision, third ed., World Scientific, pp.197–216.
[33] M. Pietikäinen, T. Ojala, Z. Xu, Rotation-invariant texture classification using feature distributions, Pattern Recognition 33 (1) (2000) 43–52.
[34] Timo Ahonen, Abdenour Hadid, Matti Pietikainen, Face recognition with local binary pattern, Lecture Notes in Computer Science 3021 (2004) 469–481.
[35] Arun Ross, Anil Jain, Information fusion in biometrics, Pattern Recognition Letters 24 (13) (2003) 2115–2125.
[36] J.C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery 2 (1998) 121–167.
[37] Zongxia Xie, Qinghua Hu, Daren Yu, Fuzzy output support vector machines for classification, in: The Proceedings of the International Conference on Advances in Natural Computation (2005) 1190–1197.
[38] Biometric Engineering Research Center Yonsei University. Available: 〈http://berc.yonsei.ac.kr〉 (accessed 30.03.10).
[39] M. Minear, Denise C Park, A lifespan database of adult facial stimuli, Behavior Research Methods, Instruments and Computers 36 (4) (2004) 630–633.
[40] The FG-NET Aging Database. Available: 〈http://www.fgnet.rsunit.com〉 (accessed 05.12.09).
[41] V. Vapnik, Support Vector Estimation of Functions, in Statistical Learning Theory, Wiley, 1998, pp. 375–570.
[42] Chih C. Chang, Chih J. Lin, LIBSVM: a library for support vector machines. Available: 〈http://www.csie.ntu.edu.tw/~cjlin/libsvm〉 (accessed 18.03.10).
[43] Timo Ahonen, Abdenour Hadid, Matti Pietikäinen, Face Description with Local Binary Patterns: Application to Face Recognition 28 (12) .
[44] J. smola Alex, Bernhard Schölkopf, a tutorial on support vector regression, Statistics and Computing 14 (3) (2006) 199–222.
[45] Jian Yang, Jing-yu Yang, David Zhang, Jian-feng Lu, Feature fusion: parallel strategy vs. serial strategy, Pattern Recognition 36 (2003) 1369–1381.
[46] Arun Ross, Rohin Govindarajan, Feature level fusion in biometric systems, in: Proceedings of the Biometric Consortium Conference (BCC), September 2004.
[47] Richard O. Duda, Peter E. Hart, David G. Stork, Pattern Classification, Wiley, 2001 pp.161–214.

Sung Eun Choi received the B.S. degree in Information Electronics Engineering at Ewha Womans University, Seoul, Korea, in 2004, and the M.S. degree in Biometrics Engineering at Yonsei University, Seoul, Korea, in 2010, where she is currently working towards her Ph.D. degree in Electrical and Electronic Engineering. Her current research interests include face age estimation, biometrics, pattern recognition and computer vision.

Youn Joo Lee received her B.S. degree in Electronic Engineering from Konkuk University, Seoul, Republic of Korea, in 2003 and her M.S. degree in Electrical and Electronic Engineering from Yonsei University, Seoul, Republic of Korea, in 2006, where she is currently working towards her Ph.D. degree in Electrical and Electronic Engineering. Her current research interests include 3D face reconstruction, biometrics, pattern recognition, computer vision and image processing.

Sung Joo Lee received his B.S. degree in Electrical and Electronic Engineering and his M.S. degree in Biometric Engineering in 2004 and 2006, respectively, from Yonsei University, Seoul, Korea, where he is currently working towards his Ph.D. degree in Electrical and Electronic Engineering. His current research interests include 3D face reconstruction, biometrics, pattern recognition and computer vision.

Kang Ryoung Park received B.S. and M.S. degrees in Electronic Engineering from Yonsei University, Seoul, Korea, in 1994 and 1996, respectively. He also received the Ph.D. degree in computer vision at the Department of Electrical and Computer Engineering in Yonsei University in 2000. He was an assistant professor in the division of digital media technology at Sangmyung University from March 2003 to February 2008. He has been an associate professor in the Division of Electronics and Electrical Engineering at the Dongguk University from March 2008. He has been also a research member of Biometrics Engineering Research Center (BERC). His research interests include computer vision, image processing and biometrics.

Jaihie Kim received the B.S. degree in Electronic Engineering at Yonsei University, Seoul, Korea, in 1979, and the M.S. degree in data structures and the Ph.D. degree in artificial intelligence at Case Western Reserve University, Cleveland, OH, in 1982 and 1984, respectively. Since 1984, he has been a professor in the School of Electrical and Electronic Engineering, Yonsei University. He is currently as the Director of the Biometric Engineering Research Center in Korea. His research areas include biometrics, computer vision and pattern recognition. Prof. Kim is currently the Chairman of Korean Biometric Association.