Classification of Age Groups Based on Facial Features

Wen-Bing Horng, Cheng-Ping Lee and Chun-Wen Chen

Department of Computer Science and Information Engineering
Tamkang University
Tamsui, Taipei, Taiwan 251, R. O. C.
E-mail: horng@mail.tku.edu.tw

Abstract

An age group classification system for gray-scale facial images is proposed in this paper. Four age groups, including babies, young adults, middle-aged adults, and old adults, are used in the classification system. The process of the system is divided into three phases: location, feature extraction, and age classification. Based on the symmetry of human faces and the variation of gray levels, the positions of eyes, noses, and mouths could be located by applying the Sobel edge operator and region labeling. Two geometric features and three wrinkle features from a facial image are then obtained. Finally, two back-propagation neural networks are constructed for classification. The first one employs the geometric features to distinguish whether the facial image is a baby. If it is not, then the second network uses the wrinkle features to classify the image into one of three adult groups. The proposed system is experimented with 230 facial images on a Pentium II 350 processor with 128 MB RAM. One half of the images are used for training and the other half for test. It takes 0.235 second to classify an image on an average. The identification rate achieves 90.52% for the training images and 81.58% for the test images, which is roughly close to human's subjective justification.

***Key Words*:** Age Classification, Facial Feature Extraction, Neural Network

1. Introduction

Human facial image processing has been an active and interesting research issue for years. Since human faces provide a lot of information, many topics have drawn lots of attentions and thus have been studied intensively. The most of these is face recognition [1]. Other research topics include predicting feature faces [2], reconstructing faces from some prescribed features [11], classifying gender, races, and expressions from facial images [5], and so on. However, very few studies have been done on age classification.

Kwon and Lobo [8] first worked on the age classification problem. They referred to cranio-facial research, theatrical makeup, plastic surgery, and perception to find out the features that change with age increasing. They classified gray-scale facial images into three age groups: babies, young adults, and senior adults. First, they applied deformable templates [13] and snakes [7] to locate primary features (such as eyes, noses, mouth, etc.) from a facial image, and judged if it is an infant by the ratios of the distances between primary features. Then, they used snakes to locate wrinkles on specific areas of a face to analyze the facial image being young or old.

Kwon and Lobo declared that their result was promising. However, their data set includes only 47 images, and the infant identification rate is below 68%. Besides, since the methods they used for location, such as deformable templates and snakes, are computationally expensive, the system might not be suitable for real time processing.

In this paper, a fast and robust system is

proposed for age group classification. The Sobel edge operator is employed to obtain facial features and two back-propagation neural nets are constructed to classify facial images into babies, young adults, middle-aged adults, and old adults. With 230 experimental mug shots, the identification rate of the system archives 81.58% for test images, which is roughly close to human's performance.

2. Preliminaries

In this section, some notations, definitions, and techniques of digital image processing used in this paper are briefly reviewed. For more information, readers may refer to conventional image processing books [3,4].

Let N be the set of natural numbers, (x, y) be the spatial coordinates of a digitized image, and $G = \{0, 1, \ldots, l - 1\}$ be a set of positive integers representing gray levels. Then, an *image* function can be defined as the mapping

$$f : N \times N \rightarrow G$$

The *brightness* (i.e., *gray level*) of a pixel with coordinates (x, y) is denoted as $f(x, y)$. The origin is at the upper-left corner of an image with the x-axis horizontal and the y-axis vertical.

Let $t \in G$ be a threshold, $B = \{b_0, b_1\}$ be a pair of binary gray levels, and $b_0, b_1 \in G$. The result of *thresholding* an image function $f(x, y)$ at gray level t is a *binary image* function

$$f_t : N \times N \rightarrow B$$

such that $f_t(x, y) = b_0$ if $f(x, y) < t$, and b_1 otherwise.

The gray level *histogram* of an image function $f(x, y)$ with gray levels in G is a discrete function

$$h_g : G \rightarrow N$$

such that $h_g(k) = n_k$, where $k \in G$, and n_k is the number of pixels in the image with gray level k.

Let $C = \{c_0, c_1, \ldots, c_m\}$ be a subset of G, in which $m < l$, $c_i \in G$, for $i = 0, 1, \ldots, m$, and $c_{j+1} - c_j = 1$, for $j = 0, 1, \ldots, m - 1$. That is, $[c_0, c_m]$ is a subrange of $[0, l - 1]$. A *histogram stretch*, or *range normalization*, of an image function $f(x, y)$ on C is a mapping

$$g : N \times N \rightarrow C$$

such that

$$g(x, y) = \left(\frac{f(x, y) - f_{min}}{f_{max} - f_{min}} \right) \times (c_m - c_0) + c_0$$

where f_{max} and f_{min} stands for the maximum and minimum gray levels of image function $f(x, y)$.

The *horizontal projection* of a binary image function is a discrete function

$$p : N \rightarrow N$$

such that $p(y) = n_y$, where y is an y-axis coordinate, and n_y is the number of pixels at y-axis coordinate y in the binary image with gray level b_0.

An *smoothing* of equally weighted moving average of a horizontal projection $p(y)$ is a transformation

$$q : N \rightarrow N$$

such that

$$q(y) = \frac{1}{2R + 1} \sum_{i=-R}^{R} p(y + i)$$

where $R > 0$ is a smoothing range.

The *gradient* of an image function $f(x, y)$ at coordinates (x, y) is defined as the vector

$$\nabla f(x, y) = \left(\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right)$$

The *magnitude* of this vector is

$$\text{mag}(\nabla f(x, y)) = \sqrt{\left(\frac{\partial f(x, y)}{\partial x} \right)^2 + \left(\frac{\partial f(x, y)}{\partial y} \right)^2}$$

The *Sobel operators* (or *edge detection masks*) are usually used for approximating the magnitude of the gradient. These masks look for edges in both the horizontal and vertical directions. The masks, as defined as follows, are each convoluted with an image function $f(x, y)$:

–1	–2	–1
0	0	0
1	2	1

–1	0	1
–2	0	2
–1	0	1

At each pixel with coordinates (x, y), two numbers are obtained: s_x corresponds the result from the right (column) mask, and s_y from the left (row) mask. The *edge magnitude*, which may be used to approximate to the gradient magnitude, of pixel (x, y) is defined as

$$M(f(x, y)) = \sqrt{s_x^2 + s_y^2}$$

or, in practice, a simpler approximation

$$M(f(x, y)) \approx |s_x| + |s_y|$$

The four horizontal and vertical neighbors of a pixel p are called the 4-*neighbors* of p, denoted by $N_4(p)$. These points, together with the four diagonal neighbors of p, are called the 8-*neighbors* of p, denoted by $N_8(p)$. For a binary image $f(x, y)$, two pixels p and q with the same gray level

are 4-*connected* (8-*connected*) if q is in the set $N_4(p)$ ($N_8(p)$). A pixel p is 4-*adjacent* (8-*adjacent*) to a pixel q if they are 4-connected (8-connected). A *region*, or *connected component*, of a binary image can be loosely defined as a collection of adjacent pixels that have the same gray level b_0. A region is said to be 4-connected (8-connected) if for every pair of pixels in the region there is some finite chain of pixels that connects them such that each consecutive pair of pixels is 4-adjacent (8-adjacent). An algorithm for finding all regions and labeling them are called a *region labeling* algorithm, which can be found in [3,4]. The region labeling algorithm used in this paper is 8-connected.

3. The Proposed System

The proposed age group classification system is briefly outlined in this section. The process of the system is mainly composed of three phases-location, feature extraction, and age classification, as illustrated in Figure 1. In the location phase, the symmetry of human faces helps find vertical central lines of faces. Since eyes, noses, and mouths have significant brightness changes, the Sobel edge operator and region labeling are applied to locate them.

Both geometric and wrinkle features are employed in the system for classification. In the feature extraction phase, two geometric features are evaluated as the ratios of the distances between eyes, noses, and mouths. Three different wrinkle features are defined to quantify the degrees of facial wrinkles.

In the age classification phase, two backpropagation neural networks [10] are constructed. The first one employs the geometric features to distinguish whether a facial image is a baby. If it is not, then the second network uses the wrinkle features to classify the image into one of three adult groups.

Notice that the dynamic range of each facial image is different. Thus, the preprocessing of histogram stretch operation is performed on all experimental images so that the ranges of gray-level of all images are mapped to the range [0, 255].

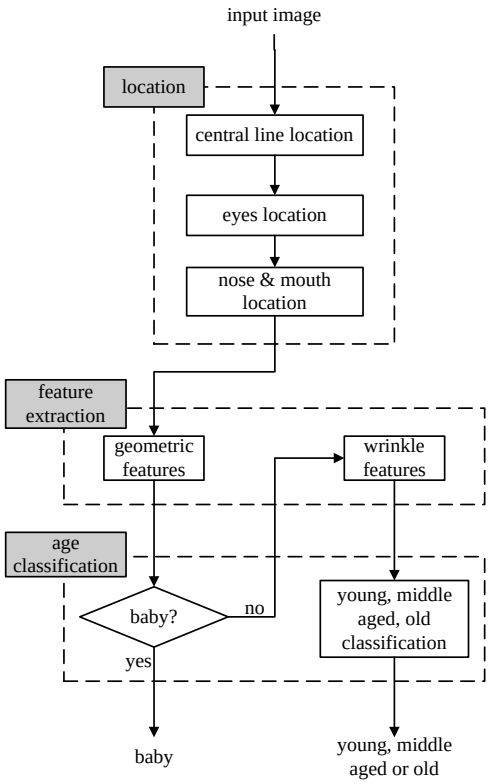


Figure 1. Process of the system

4. Location Phase

Suppose that the width and height of an image are W and H, respectively. Since the system is mainly designed to deal with mug shots, it is assumed that the width of a face is about $W/2$, one half width of the whole facial image. The face vertical central line could be easily located due to the symmetry of human faces [6]. Because most of the eyes in mug shots are in the middle height, $H/2$, of the images, the *eyes' searching region* could be defined as illustrated in Figure 2. The left and right boundaries are located from the vertical central line of the face by extending both directions $W/4$ width horizontally. The top and bottom boundaries are positioned, respectively, at $H/3$ and $3H/5$ of the image from the top of the image (properly positioning the bottom boundary of the searching region below $H/2$).

Once the eyes' searching region is found, the eyes are located as follows. First, the Sobel edge operator is applied to the original image to find edges. (In this research, the threshold is empirically defined to 100 according to the experiment.) Then, the horizontal projection is performed on the eyes' searching area. Since eyes have strong intensity variation, the major peak is formed at the position of the eyes, as shown in Figure 3. Finally, the areas of the eyes are located

by finding the nearest big enough regions (whose areas are greater than 130 pixels in this paper) to the left and right from the vertical central line via region labeling.

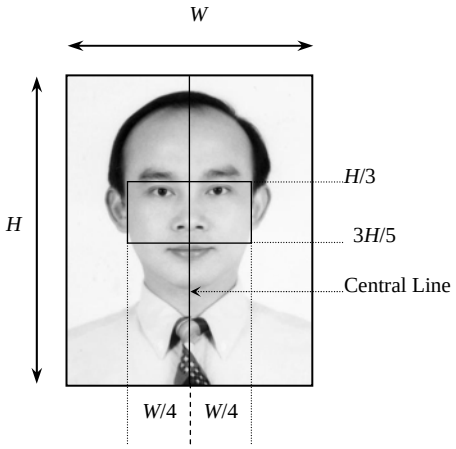


Figure 2. The eyes' searching region

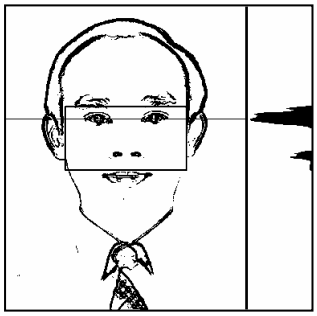


Figure 3. The position of eyes

The nose and mouth in an image could be located in a similar way. Let D_{eyes} be the distance between the outer margins of the eyes, as depicted in Figure 4. According to the facial model described in [9], the distance from eyes to mouth is around $2D_{eyes}/3$, and the left and right margins of the mouth are located about the centers of their corresponding eyes, respectively. The *nose-mouth's searching region* could be defined as illustrated in Figure 4. The top boundary is located at the lower bottom boundary of the two eyes. The bottom boundary is positioned at $3D_{eyes}/4$ downward from the top boundary (properly extending the height of the searching region larger than $2D_{eyes}/3$). The left and right boundaries of the searching region are located inwards 1/4 of the eyes' width from the outer margins of the eyes, respectively.

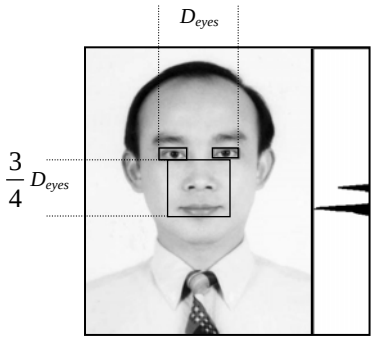


Figure 4. The nose-mouth's searching region

Similarly, the horizontal projection is first performed on the nose-mouth's searching region, and the projection is then smoothed with smoothing range $R = 2$, as shown in Figure 4. At the positions of the nose and the mouth, two obvious peaks are revealed.

By calculus, the first order derivative is zero or does not exist at the position where the extreme takes place. Because the horizontal projection is discrete, it may not be able to find positions where the first order derivatives are zeros. However, it is much easier and more precise to find zero-crossing positions, as illustrated in Figure 5. On peaks the first order derivative becomes negative from positive, while on valleys the first order derivative becomes positive from negative. The difference of two adjacent positions is applied to approximate the first order derivative, and all of the peaks and valleys could be found. The highest two peaks, therefore, are easily discovered. The one close to eyes is the position of the nose and the other is that of the mouth.

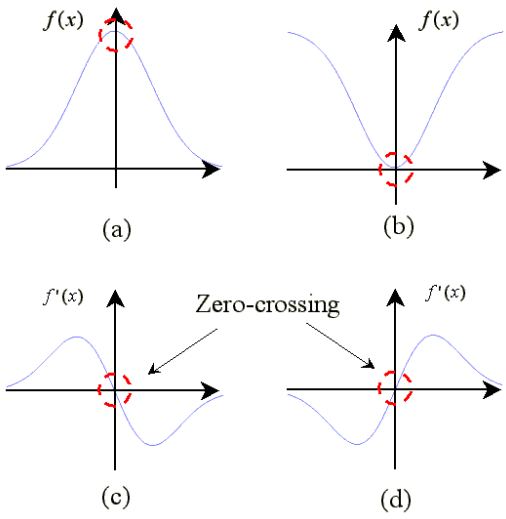
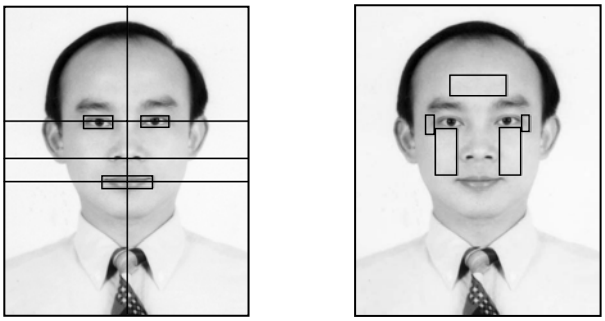


Figure 5. (a) Peak (b) Valley (c) Derivative of the peak (d) Derivative of the valley

After the mouth is located, region labeling is applied in the nose-mouth's region to find the area of the mouth. Figure 6 (a) shows the result of the entire location phase.



(a) (b)

Figure 6. (a) Result of the location phase (b) Five regions for wrinkle features extraction

5. Feature Extraction Phase

The key issue of any classification systems is how to find a set of reliable features as the basis for categorization. According to the studies of facial sketch [9] and dramatic makeup [12], many facial features change with age increasing. These features could be divided into two categories: *wrinkle features* and *geometric features*.

5.1 Wrinkle Features

With age increasing, wrinkles on faces become more and more clear. Aged people often have clear wrinkles on the following areas of a face [12]:

- The forehead has horizontal furrows.
- The eye corners have crows' feet.
- The cheeks have obvious cheekbones, crescent shaped pouches, and deep lines between the cheeks and the upper lips.

Therefore, as illustrated in Figure 6 (b), five areas—the forehead, two eye corners, and two cheeks—are located for feature extraction.

The Sobel edge magnitudes, approximating gradient magnitudes, are used in this paper to judge the degree of wrinkles, since wrinkles have obvious changes in intensity and some even form clear lines. If a pixel belongs to wrinkles, then its Sobel edge magnitude is larger. This is because the variance of gray levels is obvious. From this point of view, a pixel is labeled as a *wrinkle pixel* if its Sobel edge magnitude is larger than some threshold (40, in this research). Figures 7 (a) and (c) show a young adult and an old adult, respectively. Figures 7 (b) and (d) illustrates the results after the thresholded Sobel operator. It is easy to see that the wrinkles on the old adult are more and clearer than those on the young adult. In order to quantify the degree of the skin creases, three wrinkle features are defined as follows.

5.1.1 Wrinkle Density

The density of wrinkles in area A is defined as:

$$D_{1,A} = \frac{|W_A|}{|P_A|}$$

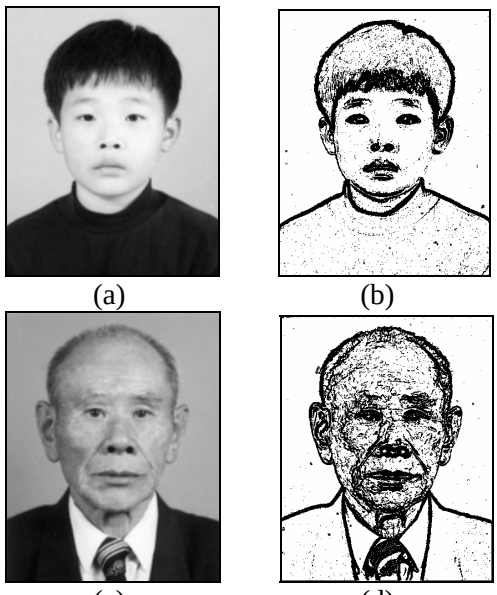
where W_A stands for the set of all wrinkles pixels in area A, P_A is the set of all pixels in area A, and $|\cdot|$ represents the number of the elements in the set. The value of $D_{1,A}$ is between zero and one. The more the wrinkles are, the closer to one the value of $D_{1,A}$ is.

5.1.2 Wrinkle Depth

The *depth of wrinkles* in area A is defined as:

$$D_{2,A} = \frac{1}{\alpha |W_A|} \sum_{(x,y)\in W_A} M(f(x,y))$$

where $M(f(x,y))$ stands for the Sobel edge magnitude (the simpler absolute values are used in this implementation) of wrinkle pixel with coordinates (x, y) in W_A. In order to avoid the saturation problem during neural network training [10], the depth of wrinkles is divided by a constant α (255, in this study). The deeper the wrinkles are, the larger the value of $D_{2,A}$ is.



(a) (b)

(c) (d)

Figure 7. (a) (c) Original images (b) (d) Results after the Sobel operator

5.1.3 Average Skin Variance

Besides wrinkles, the old adults also have the characteristics of outward eyebrows, sinking temples, and obvious cheekbones. The variances of these traits are smooth and global in intensity.

They are different from wrinkles, which change drastically and locally. In order to obtain these traits, the *average skin variance* in area A is defined as:

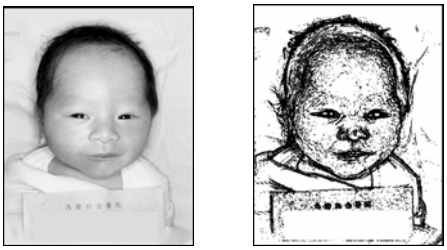
$$V_A = \frac{1}{\alpha |P_A|} \sum_{(x,y)\in P_A} M(f(x,y))$$

Similarly, to avoid saturation during neural net training, the same constant α is used as above. The less the average skin variation in area A is, the smaller the value V_A is. It means the area is smoother.

It is noticed that, generally, the light source may make the brightness of a facial image uneven and may create shadows (mostly affected by the nose). Therefore, only the features on the brighter side of a face are used, which contain one eye corner, one cheek, and the forehead.

5.2 Geometric Features

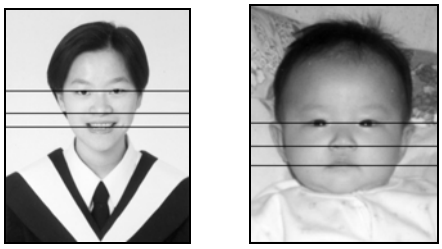
Newborn babies often have lots of wrinkles on their faces, as illustrated in Figure 8. The structure of head bones in infanthood is not full-grown and the ratios between primary features are different from those in other life periods. Therefore, using the geometric relations of primary features is more reliable than wrinkles when an image is judged to be a baby or not.



(a) (b)

Figure 8. (a) A baby (b) Result after the Sobel operator

In babyhood, the head is near a circle. The distance from eyes to mouths is close to that between two eyes. With the growth of the head bone, the head becomes oval, and the distance from the eyes to the mouth becomes longer. Besides, the ratio of the distance between babies' eyes and noses and that between noses and mouths is close to one, while that of adults' is larger than one, as illustrated in Figures 9 (a) and (b). Therefore, two geometric features are defined below for recognizing babies.



(a) (b)

Figure 9. (a) An adult's face (b) A baby's face

5.2.1 First Geometric Feature

The *first geometric feature* is the ratio defined as:

$$R_{em} = \frac{D_{em}}{D_{ee}}$$

where D_{em} is the distance between the eyes and the mouth, and D_{ee} is the distance between two eyes' centers. Since a baby's face is near a circle, R_{em} is small. On the other hand, an adult's facial shape becomes longer, so R_{em} becomes larger.

5.2.2 Second Geometric Feature

The *second geometric feature* is the ratio defined as:

$$R_{enm} = \frac{D_{en}}{D_{nm}}$$

where D_{en} is the distance between the eyes and the nose, and D_{nm} is the distance between the nose and the mouth. The distance between the eyes and the nose of a baby is close to that between the nose and the mouth. Thus, the value of R_{enm} is close to one, and may be even less than one. The distance between the eyes and the nose of an adult is longer than that between the nose and the mouth, so the value of R_{enm} becomes larger.

6. Age Classification Phase

Two back-propagation neural networks are constructed in the age classification phase. The first one is used to classify whether a facial image is a baby. The network has three layers as shown in Figure 10. The input vector consists of R_{em} and R_{enm}. The output is a value between zero and one. If the output is larger than 0.5, then the image is classified as a baby.

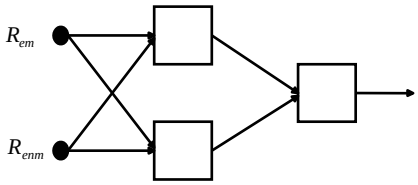


Figure 10. Network 1 for classifying babies

If an image is not classified as a baby, the second neural network is activated to determine whether it is a young adult, a middle-aged adult, or

an old adult. The network has three layers as shown in Figure 11. The input vector of the second network consists of nine wrinkle features obtained from three areas, each with three wrinkle features. There are three neurons in the output layer—the old adults, the middle-aged adults, and the young adults. The outputs are all between 0 and 1. The highest output value of the neurons decides to which age group the image belongs.

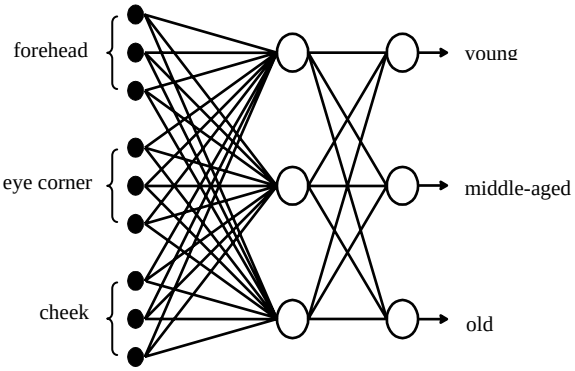


Figure 11. Network 2 for classifying young, middle-aged, and old adults

7. Experimental Results

In this research, there are 230 gray-scale with 256 gray levels facial images used for experiment. Each image size is 480×380. Thirty people are invited to classify, subjectively, these images into twelve age intervals, as depicted in Table 1. The *weighted average age* of these subjective judgements are then computed. For instance, image f005, as shown in Figure 12, is classified between 20 and 29 years old by 27 people, while classified between 30 and 39 years old by 4 people. Thus, its weighted average age is (25×27 + 35×3) / (27+3) = 26. In this system, the images with the weighted average age below three are marked as babies; those between 3 and 39 are marked as young adults; those between 40 and 59 are marked as middle-aged adults; and those over 60 are marked as old adults. Henceforth, image f005 is classified as a young adult.

Figure 12. Image f005

According to the above age group marking rule, 42 facial images are marked as babies, 89 images as young adults, 64 images as middle-aged adults, and 35 images as old adults. Among the 230 experimental images, one half of the images of each group are used as training data, and the other half are used as test images, as shown in Tables 2 to 5. The results of the age groups marking are applied to neural network training, and finally, are used for evaluating the performance of the classification system on the test images.

Table 1. Age intervals and age groups

Age Group	Age Intervals
Baby	0~2
Young Adults	3~12
	13~19
	20~29
	30~39
Middle-aged Adults	40~49
	50~59
Old Adults	60~69
	70~79
	80~89
	90~99
	100~

In the training phase, all the 21 baby images are correctly classified. Thus, the identification rate (or correct rate, used interchangeably) for babies is 21/21 = 100%. As to the adults, the correct rates for young, middle-aged, and old adults are 41/45 = 91.11%, 29/32 = 90.36%, and 14/18 = 77.78%, respectively. Thus, the total correct rate for adults is 84/95 = 88.42%, as illustrated in Table 3. Therefore, the overall identification rate for the training images is 105/116 = 90.52%.

In the test phase, the identification rate for babies is 20/21 = 99.12%. The correct rate for three adult groups are 37/44 = 84.41%, 25/32 = 78.13%, and 11/17 = 64.71%, respectively. Thus, the total correct rate for adults is 73/93 = 78.49%, as shown in Table 5. Therefore, the overall identification rate for test images is 93/114 = 81.58%. The average recognition time of each test image is around 0.235 second on a Pentium II 350 processor with 128MB RAM.

Table 2. Training phase for network 1

Age group	Sample size	Correctly labeled	Correct rate	Total correct rate
Babies	21	21	100%	100%
Adults	95	95	100%	(116/116)

Table 3. Training phase for network 2

Age group	Sample size	Correctly labeled	Correct rate	Total correct rate
Young	45	41	91.11%	88.42% (84/95)
Middle	32	29	90.36%	
Old	18	14	77.78%	

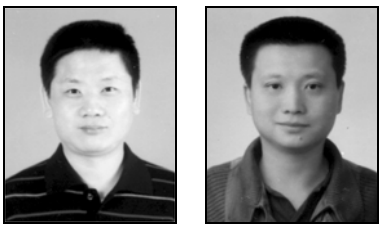
Table 4. Test phase for network 1

Age group	Sample size	Correctly labeled	Correct rate	Total correct rate
Babies	21	20	95.24%	99.12%
Adults	93	93	100%	(113/114)

Table 5. Test phase for network 2

Age group	Sample size	Correctly labeled	Correct rate	Total correct rate
Young	44	37	84.41%	78.49% (73/93)
Middle	32	25	78.13%	
Old	17	11	64.71%	

(a) f029 (b) f050 (c) f053



(d) f209 (e) f211

Figure 13. Images of reasonable classification

Some of the classification errors are due to the reason that there are no clear cuts for human beings to classify each age group, while the system adopts crisp boundaries for each group. For example, as shown in Figure 13 (b), fifteen people marked it to be between 40 and 59 years old, and the other fifteen people marked it as between 60 to 79 years old. The computed average age is 60, which is marked as old since it is just at the boundary. However, it may be reasonable to accept that if the system classified it as a middle-aged adult. There are five images, as illustrated in Figures 13 (a)~(e), each with its weighted average age close to the boundary within one year. The detailed information is depicted in Table 6. If they are considered to be acceptably correct, then the total correct rate for test adults promotes to 78/93 = 83.87%, and the overall identification rate for test images achieves 98/114 = 85.96%.

8. Conclusion and Discussion

In this paper, a fast and efficient age classification system based on facial features is proposed to classify a facial image into one of four age groups: babies, young adults, middle aged adults, and old adults. The process of the system is composed of three phases—location, feature extraction, and age classification.

The positions of the eyes, nose, and mouth of a facial image are located first. Then, two geometric features and three wrinkle features are extracted. Finally, two back-propagation neural networks are constructed for classification. The first neural network is used for distinguishing babies based on geometric features. The second network is applied for classifying adults based on wrinkle features.

There are 230 gray-scale facial images of each size 480×380 used for experiment. Thirty people are invited to classify these images, subjectively, as the base of the classification. One half of the 230 images are used to train the neural nets. The rest images are used to evaluate the trained system's performance. For the test images, the correct rate for distinguishing babies is 99.12%, and the total correct rate for three adults groups is 78.49%. Thus, the overall identification rate of the test images is 81.58%.

After investigating the factors that may affect the classification system, some important ones are summarized as follows. (a) In order for the mug shots looking better, facial wrinkles are usually smoothed or removed by photographers. (b) For some images, the light sources are too strong so

Table 6. Reasonable classification

Image no.	Average age	Human class	System class	Young		Middle		Old	
				20~29	30~39	40~49	50~59	60~69	70~79
f029	41	Middle	Young	0	13	16	1	0	0
f050	60	Old	Middle	0	0	1	14	14	1
f053	59.3	Middle	Old	0	0	1	15	14	0
f209	40.3	Middle	Young	3	11	13	3	0	0
f211	39	Young	Middle	2	15	12	1	0	0

that some detailed facial features are lost. (c) It is difficult for human beings to correctly classify a facial image, since there are no clear boundaries for each age group. Besides, each person's viewpoint for each age group may be different.

As mentioned in the previous section, the age group labeling is based on the subjective opinions of thirty people. It is interesting to see how consistent these people's viewpoints are. For babies the correct rate is 100%, however, for adults the total correct rate is 85.52% (since the correct rates for young, middle-aged, and old adults are 95.76%, 76.88%, and 74.90%, respectively). Thus, the overall identification rate for all the 230 experimental images is 88.19% by humans.

It can be seen that, in the training phase the system outperforms humans' judgment. In the test phase, including the reasonable classified images, the overall identification rate of the system is 85.96%, which is a little bit lower than humans' classification. As illustrated from the above discussion, it could be concluded that the system's performance is rather close to human's performance for age group classification.

References

[1] Chellappa, R., Wilson, C. L. and Sirohey, S., "Human and machine recognition of faces: A Survey," *Proc. of the IEEE*, Vol. 83, pp. 705-740 (1995).

[2] Choi, C., "Age change for predicting future faces," *Proc. IEEE Int. Conf. on Fuzzy Systems*, Vol. 3, pp. 1603-1608 (1999).

[3] Gonzales, R. C. and Woods, R. E., *Digital Image Processing*, Addison-Wesley, Reading, MA, U. S. A. (1992).

[4] Gose, E., Johnsonbaugh, R. and Jost, S., *Pattern Recognition and Image Analysis*, Prentice Hall, Upper Saddle River, New Jersey, U. S. A. (1996).

[5] Gutta, S. and Wecheler, H., "Gender and ethnic classification of human faces using hybrid classifiers," *Proc. Int. Joint Conference on Neural Networks*, Vol. 6, pp. 4084-4089 (1999).

[6] Huang, J. C., "A study on gray scale face recognition," Master Thesis, National Chiao Tung University, Hsinchu, Taiwan, R. O. C. (1996).

[7] Kass, M., Witkin, A. and Terzopoulos, D., "Snake: active contour models," *Proc. First Int. Conf. on Computer Vision*, London, England, pp. 259-268 (1987).

[8] Kwon, Y. H. and da Vitoria Lobo, N., "Age classification from facial images," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Seattle, Washington, U. S. A., pp. 762-767 (1994).

[9] Loomis, A., *Drawing the Head and Hands*, ninth printing, Viking Press, New York , U. S. A. (1974).

[10] Looney, C. G., *Pattern Recognition Using Neural Networks*, Oxford University Press, New York, U. S. A. (1997).

[11] Shepherd, J. W., "An interactive computer system for retrieving faces," *Aspects of Face Processing*, Ellis, H. D. et al. Eds, Martinus Nijhoff International, Dordrecht, The Netherlands, pp. 398-409 (1986).

[12] Thomas, C., *Make-Up: The Dramatic Student's Approach*, Theatre Arts Books, New York, U. S. A. (1968).

[13] Yuille, A. L., Choen, D. S. and Hallinan,

P. W., “Feature extraction from faces using deformable templates,” *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, San Diego, California, U. S. A., pp. 104-109 (1989).

Manuscript Received: Jul. 19, 2001
And Accepted: Aug. 10, 2001