Computer Vision and Image Understanding
Vol. 74, No. 1, April, pp. 1–21, 1999
Article ID cviu.1997.0549, available online at http://www.idealibrary.com on IDEAL®

Age Classification from Facial Images

Young H. Kwon* and Niels da Vitoria Lobo†

School of Computer Science, University of Central Florida, Orlando, Florida 32816

Received May 26, 1994; accepted September 6, 1996

This paper presents a theory and practical computations for visual age classification from facial images. Currently, the theory has only been implemented to classify input images into one of three age-groups: babies, young adults, and senior adults. The computations are based on cranio-facial development theory and skin wrinkle analysis. In the implementation, primary features of the face are found first, followed by secondary feature analysis. The primary features are the eyes, nose, mouth, chin, virtual-top of the head and the sides of the face. From these features, ratios that distinguish babies from young adults and seniors are computed. In secondary feature analysis, a wrinkle geography map is used to guide the detection and measurement of wrinkles. The wrinkle index computed is sufficient to distinguish seniors from young adults and babies. A combination rule for the ratios and the wrinkle index thus permits categorization of a face into one of three classes. Results using real images are presented. This is the first work involving age classification, and the first work that successfully extracts and uses natural wrinkles. It is also a successful demonstration that facial features are sufficient for a classification task, a finding that is important to the debate about what are appropriate representations for facial analysis.

1. INTRODUCTION

As humans, we are easily able to categorize a person's age group from an image of the person's face and are often able to be quite precise in this estimation. This ability has not been pursued in the computer vision community. In order to begin researching the issues involved in this process, this research addresses the limited task of age classification of a mugshot facial image into a baby, young adult, and senior adult.

Any progress in the research community's understanding of the remarkable ability that human's have with regard to facial-image analysis will go a long way toward the broader goals of face-recognition and facial-expression recognition. In the long run, besides leading to a theory for automatic precise age identification which would assist robots in numerous ways, analysis of facial features such as aging-wrinkles will assist in wrinkle analysis for facial-expression recognition. However, in the shorter term too, an improvement of our understanding of how humans may classify age from visual images can be used in the domain of indexing into a face database by the person's age, in the area of newspaper-story understanding [10, 11], and in the application areas such as gathering population age-statistics visually (for example, getting the ages of patrons at entertainment and amusement parks or in television network viewer-rating studies.)

To gain an understanding for the aging process of the face, we consulted studies in cranio-facial research [1], art and theatrical makeup [2, 7], plastic surgery, and perception [5]. The main theory in the area of cranio-facial research is that the appropriate mathematical model to describe the growth of a person's head from infancy to adulthood is the revised cardioidal strain transformation, written in polar form as: $\theta' = \theta$, $R' = R(1 + k(1 - \cos\theta))$, where θ is the angle formed from the Y-axis, R is the radius of the circle, k is a parameter that increases over time, and (R', θ')'s are the successive growths of the circle over time [1]. The revised cardioidal strain transformation describing head growth can be visualized as a series of ever growing circles all attached at a common tangent "base" point, in this case the top of the head. With this transformation the growth of lower parts of the face is more pronounced than that of the upper part. Thus, for example, within the top and bottom margins of the head, the eyes occupy a higher position in an adult than in an infant (note that this is due not to eye migration, but instead to an outgrowing and dropping of the chin and jaw).

Another consequence of this development into adulthood is that, relative to the margins formed by the eyes and the mouth, the position of the nostrils (nose) drops. Hence, to distinguish babyfaces from the two older groups, this research has evaluated a set of ratios. These ratios only require the automatic localization of primary features, namely the eyes, nose, mouth, chin, and virtual top of the head (see Fig. 9). Ratio 1 is the T-ratio formed by two segments: the segment T_1 joining the two eyes and the segment T_2 between the midpoint of T_1 and the nose. Ratio 2 is the T-ratio formed by two segments: the segment T_1 as above, and the segment T_3 between the midpoint of T_1 and the mouth. Ratio 3 is the T-ratio formed by two segments: the segment T_1 as above, and the segment T_4 between the midpoint of T_1 and the chin. Ratio 4 is the ratio of the segment representing the difference in height between nose and eye-midpoint, and the segment representing the difference in height between mouth and

---

*Also at RSI Baseband Technologies, Orlando, FL 32822. E-mail: kwon@cs.ucf.edu.

† E-mail: niels@bohr.cs.ucf.edu.

1077-3142/99 $30.00

eye-midpoint. Ratio 5 is the ratio of the segment representing the difference in height between mouth and eye-midpoint, and the segment representing the difference in height between chin and eye-midpoint. Ratio 6 is the height of the eyes within the top and bottom head-margins.

After distinguishing the face as belonging to the union of the two older classes, our classification goals require us to determine whether the person is a young adult or a senior. Bone structural changes do not occur after the person is fully grown (i.e., the geometric relationships of primary features do not vary); hence secondary features need to be identified and exploited. From studying the aging process of adult humans, one can observe that the facial skin of an older person is not as taut as in a younger adult.

Wrinkles are a good indication of the loosening skin; thus we have selected wrinkles as the next important feature (although, in general, these aging-wrinkles must not be confused with creases formed from facial expressions).

2. PREVIOUS WORK

No previous work has been reported on any aspects of age information in images of human faces. However, it is appropriate to review research on facial image analysis, as many of the issues encountered in our problem are similar to those encountered in related problems. Previous computational work on face-images has been carried out in two distinct paradigms. In the first paradigm researchers first extract features such as the eyes, nose, etc., then they relate these features geometrically, and finally they use the geometric relationships to aid in analysis and recognition. The current research has adopted this paradigm of locating features and analyzing them for age classification. The second paradigm treats the complete face image as an input vector and bases analysis and recognition on algebraic transformations of the input space.

Attempts at performing automated analysis of facial images using geometrical features date back to Bledsoe [3]. A subsequent attempt at this task was undertaken by Kanade [13]. He started by detecting 16 features. Then he analyzed inter- and intraclass variances, finding that some of the parameters were ineffective. The remaining 13 parameters were used for the recognition task. However, the features were not always accurately located.

More recently, Yuille, Hallinan, and Cohen [29] used deformable templates to detect features of faces, in particular, eyes and lips. External energy functions of valley, peak, edge, and image intensity are defined. These external images are computed from the original intensity image. The internal energy consisting of the template parameters interact dynamically with the external energy. The best fit occurs when the total energy equation, consisting of internal and external energy, is minimized.

Pursuing the approach of Yuille, Hallinan, and Cohen [29], Shackleton and Welsh [26] performed feature-locating steps for detecting the eye, having first added template parameters for regions of enhanced whites (of the eyes). Once they detected the eye, they performed geometric normalization of the eye in preparation for a recognition step. This recognition step uses the eigen analysis approach (see below) of Turk and Pentland [27] and Kirby and Sirovich [15]. The contribution of Shackleton and Welsh's method is that most of the background is eliminated and only the feature of interest, in this case the eye, is considered for recognition.

Yang and Huang [28] use pyramids to find bright regions (at the coarsest level) and then confirm the location of a face by finding supporting features.

Craw, Tock, and Bennett [8] designed a system to locate the features in a face using statistical knowledge to create an ideal template. This system is capable of labeling the hair, the eyes, the mouth, and the nose in a mugshot image.

Reisfeld and Yeshurun [25] use generalized symmetry to localize the eyes and mouth in images of human faces. Their system is able to cope with a wide variety of sizes, rotations, and orientations.

Nagamine, Uemerra, and Masuda [20] have developed methods to match features in range images of faces. These images might be produced by, for instance, a binocular system.

Using color images of light-skinned faces, Novak [21] uses skin tones to find the face, lip-pinks to find the lips, and blue-greens to find the eyes.

In an attempt at recognizing facial expressions, Matsuno, Lee, and Tsuji [19] use potential nets, which undergo structural deformations at features such as the eyebrows, nose, and mouth. Based on the pattern of deformations, classification is achieved.

Working in the other paradigm, Turk and Pentland [27] convert an $N \times N$ image of a face into a single vector of size N^2 by concatenating scan lines. Then they compute the eigenvectors of the covariance matrix of the set of face images. Only a few of the eigenvalues are significant, thus characterizing the low-dimensional "face-space." A face can be represented in this new space by a few weights. Recognition is considered successful when an image's weights fall within some neighborhood of a set of weights already stored in a database. This method is sensitive to scale, viewing angle, lighting changes, and background noise. Similar work in this paradigm has been reported by Kirby and Sirovich [15].

As another embodiment of this approach, O'Toole, Abdi, Deffenbacher, and Bartlett [23] used autoassociative memory techniques, and this proved useful in classifying faces by gender and race, and in recognition. This method sets up an autoassociative memory of J completely interconnected units, where J is the number of pixels in an image. The connection strengths were stored in a $J \times J$ matrix. Eigenvectors were extracted and the first seven eigenvectors were used to differentiate gender and race. For the training sets they used, it was discovered that the difference in coefficients for the eigenvectors is useful for female/male classification and for Caucasian/Japanese classification. For other work that examined the ability to classify gender using neural networks, see [9].

Recently, Brunelli and Poggio [4] have compared the utility of the two paradigms described above, in the task of face recognition. They concluded that the paradigm of using the complete image patches and a correlation technique (which is related to eigen-analysis in the sense that eigenvectors are linear combinations of the image patches in the database) outperforms the feature-based geometric approach (note that they called their correlation approach "template-matching," which truly must be distinguished from the templates used in the feature-based geometric approaches described earlier). However, in their experiments, their features were not located with precision (examine, for example, the location of the left eye in their Fig. 6), and hence, their conclusion about the inferiority of the approach, may be incorrect. In this research, we determine that facial features, when they can be found, are robust for performing age classification. In addition, since we need to detect wrinkles for the age classification task and since we expect to be able to transfer insight from its performance here to the area of facial expression analysis in the long run, we have chosen to adopt explicit feature extraction and analysis. It should be noted that for a task such as ours, in which there are only a finite number (and possibly only a handful) of categories, a few types of geometric relationships of explicit feature analysis may suffice. On the other hand, recognition of a specific individual amongst a large set of possibilities (possibly hundreds) requires a much larger parameter space. It may be that the few geometric relationships measured and tested by other researchers (e.g., Kanade [13]), carry insufficient information on which to base recognition.

3. OVERVIEW OF THE APPROACH IN THIS RESEARCH

A high level description of the steps in the approach taken in this research is presented next.

A. Find facial features
  1. Find initial rough oval
  2. Find chin; adjust oval
  3. Find sides of face
  4. Compute virtual top of head
  5. Find eyes
  6. Find mouth
  7. Find nose
B. Compute facial feature ratios
C. Compute wrinkle analysis
D. Combine B and C to conclude age category.

Our approach to finding the initial oval, and the eyes follows that of Yuille, Hallinan, and Cohen [29] and Shackleton and Welsh [26]. An energy equation is defined according to $E_{\text{total}} = E_{\text{ext}} + E_{\text{int}}$, where $E_{\text{ext}} = E_{\text{eye}} + E_{\text{edge}} + E_{\text{valley}} + E_{\text{intensity}}$ and E_{int} contains the geometric template for the face. These energy terms are related to potential fields that guide the template fitting process. The potential fields are formed from the image operations. Corresponding to the image forces used in previous work (which used morphological operators such as open, close, erode, and dilate) we employ appropriate image transformations to localize the features.

A series of stages (similar to epochs in previous work) are set up where at each stage an intermediate goal is accomplished. The minimization is performed using gradient descent. Once a stabilized energy sum is reached, the intermediate goal of that stage has been realized and processing continues to the next stage.

Our approach to finding the chin and the sides of the face involves dropping a population of *snakelets* [14] around the boundaries of the initial oval, in three rectangular regions, and performing a Hough transform to find a parabolic curve in each of the three regions. A *snakelet* is a small snake segment designed to find a small individual curve segment in the image. The nose and the mouth are found by convolution with dark-bar detectors. From these features, geometric ratios are computed.

Next, snakelets are used to find wrinkles. A wrinkle geography map drops multiple snakelets in polygonal regions, where wrinkles may typically be found. The main focus of this stage of analysis is concentrated on the wrinkles on the forehead, wrinkles next to the eyes, and the wrinkles near the cheek bones. The presence of wrinkles in a region is concluded positively if there are several curves in the region. The different locales of evidence for the presence of wrinkles are then weighted appropriately to infer the age group within adults.

4. FACIAL FEATURE DETECTION AND LOCALIZATION

The localization of the facial features is performed in stages. At each stage, a particular facial feature parameter is found. The center position of the head in the image is initialized manually, with an allowance for a large margin of error. In separate work, we show how the center of a face can be located automatically with no knowledge of the scene [18] (another approach is discussed in [28]). The initial oval-finding stage finds an oval that best fits the face/head, and consequently the center position of the head is automatically updated. The chin-finding stage finds the best chin in the rectangular area specified by the oval parameters. The face sides-finding stage finds the left and right sides of the face in the area specified by the chin and oval parameters. The virtual top of the head is then acquired from the oval generated from the chin and the two face sides. The chin parameter, if it is found robustly, is then used to refine the initial oval. Otherwise, the initial oval is used for the remaining stages. The iris-attracting stage places both the left and the right iris centers of the template near the respective iris centers in the image specified by the oval parameter. The iris-fitting stage tries to fit more accurately the iris contour by estimating the iris radius and simultaneously improving estimates of both iris center positions. The mouth-finding stage finds the position of the center of the mouth and finally the nose-finding stage finds the position of the bottom of the nose.

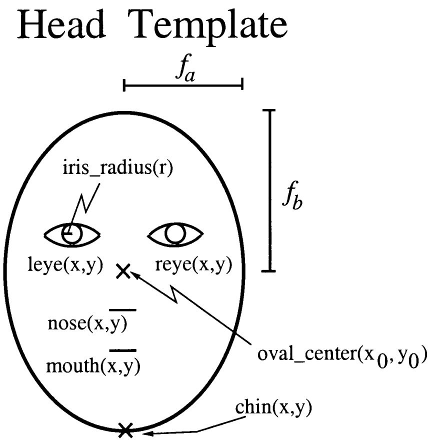


FIG. 1. This shows a face template, including the parameters used in oval-fitting and eye-fitting.

Figure 1 explains some of the parameters used. For the gradient descent process, potential images are created. The potential image for the edge, Φ_{edge}, was generated by first applying a morphological operator of *dilate*(5) − *erode*(5) and then blurring it by using an exponential filter $e^{-\alpha(x^2+y^2)^{1/2}}$ with an α value of 0.8. The *dilate* operator causes expansion while the *erode* operator causes shrinkage. Further details of morphological operators can be found in Haralick and Shapiro [12]. The potential image for the valley, Φ_{valley}, was generated by first applying a morphological operator of *close*(21) and then blurring it by using an exponential filter with an α value of 0.4. A similar operator was used by Shackleton and Welsh [26]. The *close* operator smooths contours and fills gaps [12]. The eye potential image Φ_{eye} was generated with a first-order difference operator of size 3, in both X and Y directions.

4.1. Initial Oval-Finding Stage

The initial oval-finding stage uses the edge-image potential Φ_{edge} to find the best oval that fits the outline of the face:

$$\begin{aligned}
E_{\text{oval\_total}} &= E_{\text{edge}} + E_{\text{internal}}, \\
E_{\text{edge}} &= \frac{c_2}{\text{length}} \int_{\text{face\_edge}} \Phi_{\text{edge}}(\mathbf{x})\, ds, \\
E_{\text{internal}} &= k_1(X_e - X_c) + k_1(Y_e - Y_c - (0.1) f_b) \\
&\quad + k_2(f_a - 0.7 f_b) + k_3 e^{-\beta_1 f_a} + k_3 e^{-\beta_3 r}.
\end{aligned}$$

The coefficients used are: $c_2 = 200$, $c_3 = 100$, $k_1 = 100$, $k_2 = 100$, $k_3 = 100000$, $\beta_1 = 0.04$, $\beta_2 = 0.028$. These numeric values were chosen by a combination of trial/error and contemplation about their roles in the computation.

The oval axes half-lengths (f_a, f_b) are initialized and the oval is positioned inside the face perimeter. The k_1-term force draws the oval center (X_c, Y_c) downward because the oval center found in stage 1 is biased towards the top due to the inclusion of hair in stage 1. The k_1-term also maintains the horizontal component of the center close to where it was initially dropped at (X_e, Y_e). The k_3-terms are outward-forcing for each half-length to prevent the oval from becoming too small. The k_3-term also helps to overcome the possibility of the lower edge of the oval stabilizing around the mouth. The k_2-term tries to maintain a ratio that is reasonable for a face. The fitting of the oval is performed by iteratively updating the oval center position and the oval axes half-lengths (f_a, f_b). The update formula is obtained by gradient descent with respect to these four parameters. The stage is complete when the total energy stabilizes around a minimum value. Figure 2 illustrates the outcome of this stage.

4.2. Chin-Finding Stage

The chin-finding stage uses the edge-image potential to find the bottom of the chin. First, multiple snakelets are dropped in random orientations in a grid, one per grid point, in a rectangular region of width $\frac{3}{2} f_a$ centered at the vertical axis of the oval center, and of height $\frac{1}{4} f_b$ to the top and $\frac{1}{8} f_b$ to the bottom, positioned from the horizontal axis tangent to the bottom of the oval. After all the snakelets have stabilized, they are used to obtain a Hough transform for the strongest parabola in that region. The Hough transform is implemented in two stages. First, each snakelet votes for a coarse candidate position for the base of the parabola. Nonmaximal suppression is performed on these votes, and the best five are chosen. Amongst these five, the one that is located in the highest position on the face (to eliminate parabolas corresponding to necklines of clothing, or of the shadow formed by the chin) is chosen. Then, the snakelets that voted for this winning position of the base are identified, and a finer search of the base is performed in a similar manner. This leads to robust identification of the chin, when it is delineated in the raw image. If the image quality is poor, or illumination conditions are bad, the correct position of the chin is very difficult to find.

Determination of whether a chin is found is currently done by human visual inspection. If it has been found, it is used in three places: to refine the initial oval; to find the virtual top of the head using also the sides of the face (see upcoming sections on "face side-finding stage" and "finding the virtual top of the head"); and to be used in ratios 3a and 5a. If it is determined that it is not found, the bottom of the initial oval is used in the ratios, thus requiring ratios 3b and 5b. In these ratios, the chin is referred as "ochin" (for oval-chin). Further research should explore how it can be automatically determined whether the chin has been found robustly. Figure 3 shows the stages involved in locating the chin.

4.2.1. Parabola-fitting algorithm.

The following is a numerical description of the parabola-fitting process.

With two known points on the parabola,

$$\begin{aligned}
y &= k(x-h)^2 + c \qquad \text{(U-shaped parabola equation)} \\
y_1 &= k(x_1-h)^2 + c \qquad (c, h, k \text{ are the same for both equations}) \\
y_2 &= k(x_2-h)^2 + c.
\end{aligned}$$

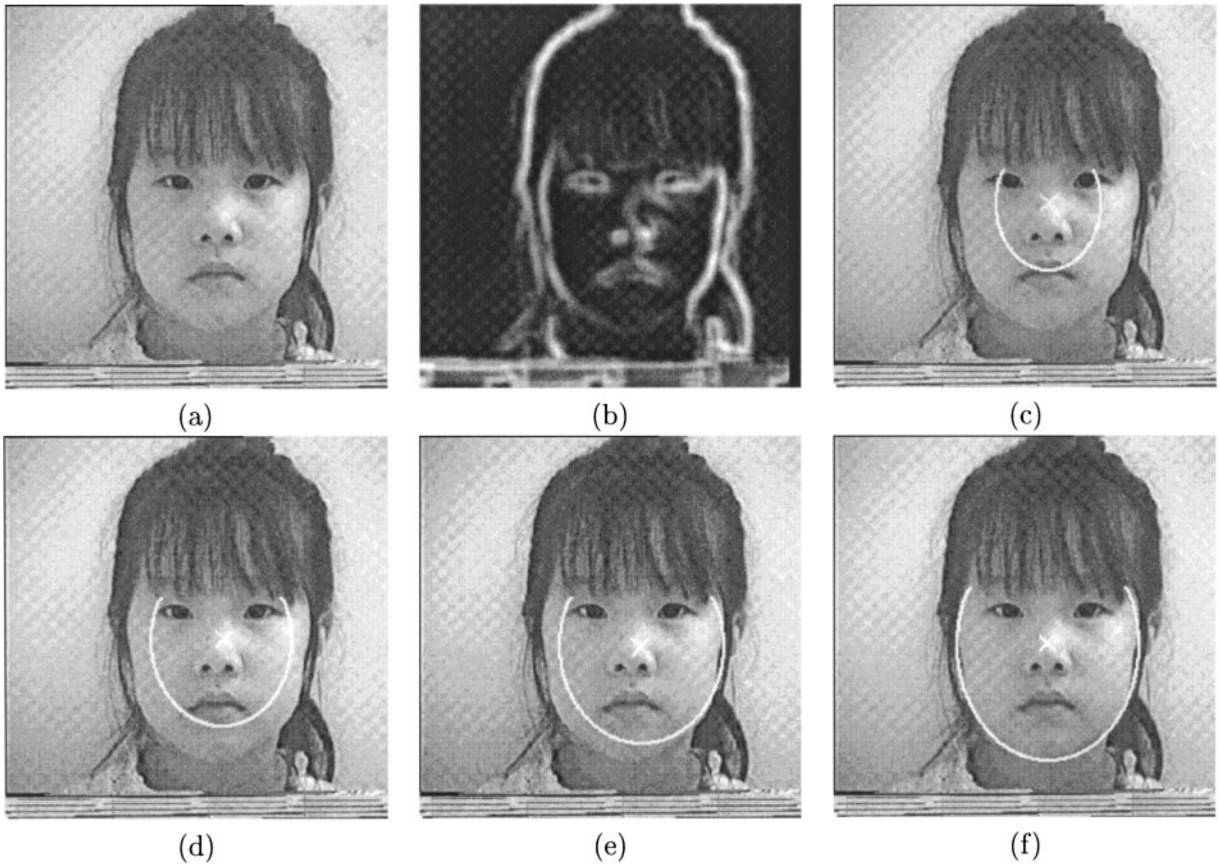


FIG. 2. Oval-fitting: (a) original image; (b) edge potential; (c)–(f) various stages before the final fit.

Algebriac manipulations yield

$$c = \frac{(x_2 - h)^2 y_1 - (x_1 - h)^2 y_2}{(x_2 - h)^2 - (x_1 - h)^2}.$$

Since (x_1, y_1) and (x_2, y_2) are known, solve for c and h by using a Hough transform.

4.3. *Face Side-Finding Stage*

The face side-finding stage uses the edge-image potential to find the left and right sides of the face. This stage uses the parabola-fitting algorithm, except that the parabola equations are modified to suit the left and right open-ended parabolas. Figure 4 shows the right and left open-ended parabolas.

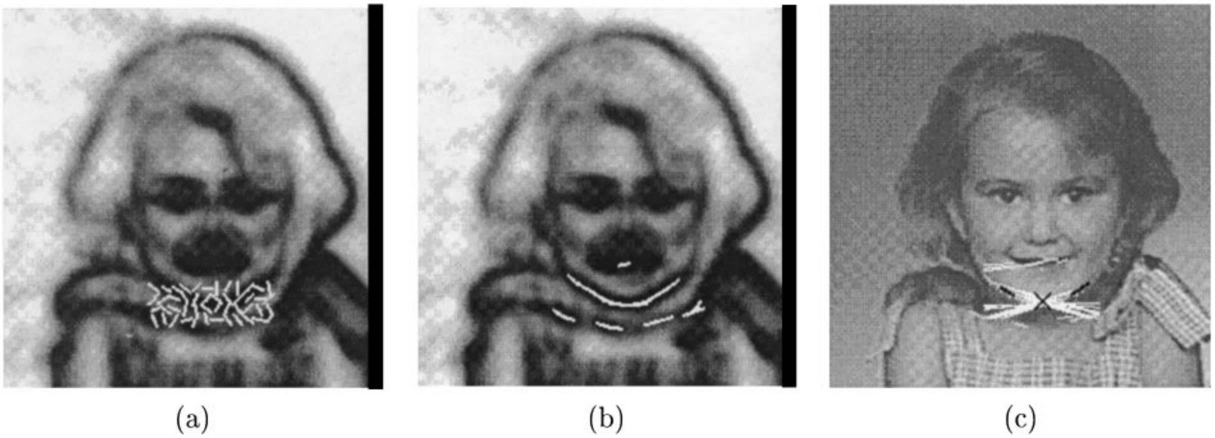
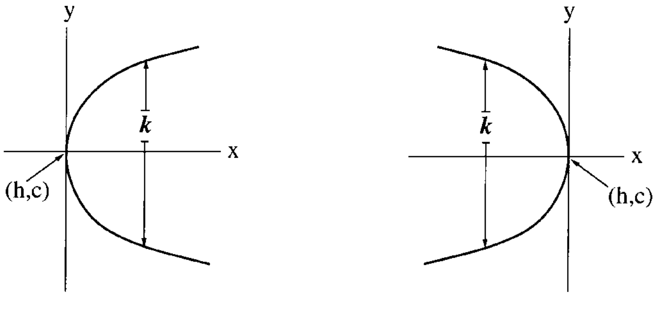


FIG. 3. Chin-finding: (a) initial snakelets dropped on the negative of edge potential image; (b) stabilized snakelets; (c) the winner in Hough space is marked with an “X,” which is the chin.



(a) (b)

FIG. 4. (a) Right open-ended parabola to fit the left side of the face. (b) Left open-ended parabola to fit the right side of the face.

The following are the parabola equations for the right open-ended parabola and the left open-ended parabola, respectively:

$$x = k(y - c)^2 + h, \quad x = -k(y - c)^2 + h.$$

Figure 5 shows the stages in locating the left and right sides of the face. This process is also very robust, except when the ear is too close to the side of the face, or when there is disheveled hair within the lower half of the face. Evaluation of the goodness of fit for the side-finding stage is performed by human visual inspection.

4.4. Finding the Virtual Top of the Head

The top of the skull is very difficult to estimate when the person has hair. Hence, an alternative strategy is used here. It is described next.

At this stage, the three pools of snakelets that voted for winning parabolas are pooled, and a single oval is fit to all the snakelets' points. The goal here is experimental and tentative in its purpose. The oval is used to find the virtual top of the head, as confined by the snakelets that support the parabolas of the chin and sides of the face. The purpose of this is to have an estimate of the ratio of the height of the eyes from the chin, relative to the virtual top of the head. This process is obviously affected by the errors in finding the sides of the face, which in turn are affected by disheveled hair in the lower half of the face, and a ear that is too close to the side of the face. Figure 5 shows the outcome for one of the successful cases.

The following is a numerical description of the process of estimating the oval that yields the virtual top of the head. The parametric form of the oval equation is

$$\begin{aligned} x &= f_a * \cos\theta + x_0 \\ y &= f_b * \sin\theta + y_0, \end{aligned}$$

where

$$\begin{aligned} (x_0, y_0) &= \text{oval\_center}, \\ (x, y) &= \text{point on the oval\_edge}, \\ \theta &= \text{angle from the } x\text{-axis}, \\ (f_a, f_b) &= \text{the respective } x \text{ and } y \text{ distances} \\ &\quad\ \text{from the oval\_center}. \end{aligned}$$

Given (x, y), (x_0, y_0), and f_b, we can solve for f_a:

$$\begin{aligned} \theta &= \sin^{-1}[(y - y_0)/f_b] \\ f_a &= (x - x_0)/\cos\theta. \end{aligned}$$

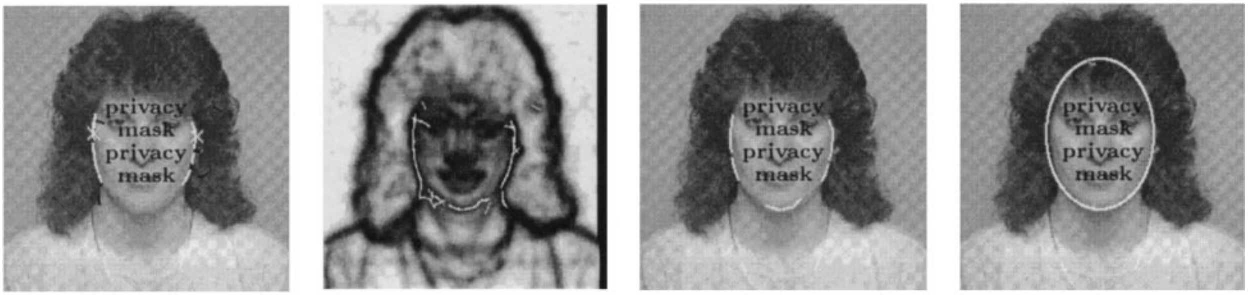
Figure 6 shows the resulting snakelets of the parabola-fitting for the three groups. These snakelets are then used as an input to the virtual top of the head finding algorithm.

4.5. Eye-Finding Stage

This consists of an iris-attracting stage and an iris-fitting stage.

4.5.1. Iris-attracting stage. The iris-attracting stage places both the left and the right iris centers of the template near the respective iris centers in the image specified by the computed oval.

Using the parameters from the refined oval (or the initial oval, if the refinement cannot be computed), both the left and the right iris centers of the template are placed near the respective irises



(a) (b) (c) (d)

FIG. 5. Face side-finding: (a) the "X" on each side marks the base point of the parabola for that side of the face. Virtual head finder: (b) snakelets after settling down; (c) snakelets that support the three parabolas; (d) the fit of the face-oval has been refined. The virtual top of the head is the top of this oval.

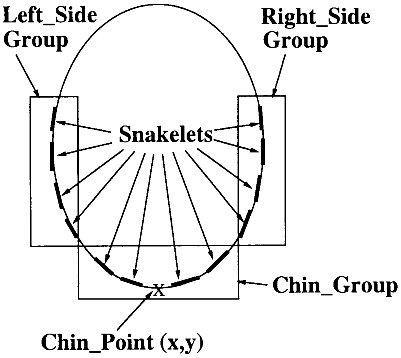


FIG. 6. This shows the snakelets and the chin-point used in the virtual top of the head finding algorithm.

in the image. The left iris center is placed $\frac{1}{10} f_b$ above the oval center for y and $\frac{1}{2} f_a$ left of the oval center for x. Similarly, the right iris center is placed $\frac{1}{10} f_b$ above oval center for y and $\frac{1}{2} f_a$ right of the oval center for x. The iris-attracting stage uses the valley image potential Φ_{valley} and the intensity image potential $\Phi_{\text{intensity}}$ to attract both left and right iris centers to their true positions. The following expression is minimized:

$$\begin{aligned}
E_{\text{eye\_total}} &= E_{\text{intensity}} + E_{\text{valley}} + E_{\text{eye}}\\
E_{\text{intensity}} &= -\frac{c_5}{\text{area}} \iint_{\text{left\_iris\_area}} \Phi_{\text{intensity}}(\mathbf{x})\, dA\\
&\quad - \frac{c_5}{\text{area}} \iint_{\text{right\_iris\_area}} \Phi_{\text{intensity}}(\mathbf{x})\, dA\\
E_{\text{valley}} &= \frac{c_5}{\text{area}} \iint_{\text{left\_iris\_area}} \Phi_{\text{valley}}(\mathbf{x})\, dA\\
&\quad + \frac{c_5}{\text{area}} \iint_{\text{right\_iris\_area}} \Phi_{\text{valley}}(\mathbf{x})\, dA\\
E_{\text{eye}} &= \frac{c_4}{\text{length}} \int_{\text{left\_iris\_edge}} \Phi_{\text{eye}}(\mathbf{x})\, ds\\
&\quad + \frac{c_4}{\text{length}} \int_{\text{right\_iris\_edge}} \Phi_{\text{eye}}(\mathbf{x})\, ds.
\end{aligned}$$

The coefficients used are $c_4 = 100$, $c_5 = 300$, and their numeric values were selected after deliberation about their roles relative to each other.

The iris centers are initially positioned, guided by the oval parameters. The valley potential is strong around the actual iris area and it draws the iris/eye template over a substantial distance. The update is by gradient descent. The oval parameters are not allowed to change in this stage. Thus, only a change in the iris parameters will update the total energy.

4.5.2. Iris-fitting stage. For the iris-fitting stage, the equations from the iris-attracting stage and additional equations for each of the energy potentials are used. An internal energy term is also introduced to prevent the iris radius from collapsing to a point. The relevant expression that is minimized is

$$\begin{aligned}
E_{\text{iris\_total}} &= E_{\text{intensity\_total}} + E_{\text{valley\_total}} + E_{\text{eye\_total}} + E_{\text{internal}}\\
E_{\text{intensity\_total}} &= E_{\text{intensity\_eye}} + E_{\text{intensity\_iris}}\\
E_{\text{valley\_total}} &= E_{\text{valley\_eye}} + E_{\text{valley\_iris}}\\
E_{\text{eye\_total}} &= E_{\text{eye\_eye}} + E_{\text{eye\_iris}}
\end{aligned}$$

$E_{\text{intensity\_eye}}$ is the same as in iris-attracting stage

$$\begin{aligned}
E_{\text{intensity\_iris}} &= -\frac{2 \times c_9}{r \times \text{length}} \int_{\text{left\_iris\_edge}} \Phi_{\text{intensity}}(\mathbf{x})\, ds\\
&\quad + \frac{2 \times c_9}{r \times \text{area}} \iint_{\text{left\_iris\_area}} \Phi_{\text{intensity}}(\mathbf{x})\, dA
\end{aligned}$$

$E_{\text{valley\_eye}}$ is the same as in iris-attracting stage

$$\begin{aligned}
E_{\text{valley\_iris}} &= \frac{2 \times c_8}{r \times \text{length}} \int_{\text{left\_iris\_edge}} \Phi_{\text{valley}}(\mathbf{x})\, ds\\
&\quad - \frac{2 \times c_8}{r \times \text{area}} \iint_{\text{left\_iris\_area}} \Phi_{\text{valley}}(\mathbf{x})\, dA
\end{aligned}$$

$E_{\text{eye\_eye}}$ is the same as in iris-attracting stage

$$\begin{aligned}
E_{\text{eye\_iris}} &= \frac{c_7}{\text{length}} \int_{\text{left\_iris\_edge}} \Phi_{\text{eye}}(\mathbf{x})\, ds\\
E_{\text{internal}} &= k_4 e^{-\beta_3 r}.
\end{aligned}$$

The coefficients used are $c_4 = 100$, $c_5 = 300$, $c_7 = 200$, $c_8 = 100$, $c_9 = 150$, $k_4 = 80000$, $\beta_3 = 0.6$, and these were picked by thoughtful observation about their relative importance.

The iris-fitting stage uses the intensity image potential $\Phi_{\text{intensity}}$, the valley image potential Φ_{valley}, and the eye image potential Φ_{eye} to fit the iris template to the actual image. The k_4-term is an outward force that prevents the radius of the iris from becoming too small. The darkness of the iris area from the intensity image and the brightness of the iris area from the valley image potential, along with the iris edge information from the eye image potential, guide the fitting of the iris template to the image. The eye parameters a, b, c (which are constants controlling the boundary of the eye) are scaled with respect to the radius of the iris and these parameters are used to clip the iris disc to a partial disc. The only parameters allowed to change in this stage are both the iris centers and the scale of the iris template. The parameters are updated by gradient descent. Figure 7 shows the outcome of the iris-attracting and -fitting stages. Evaluation of the goodness of fit for the eye-finding stage is conducted by human visual inspection.

4.6. Finding the Mouth

After finding both iris centers, a vertical axis half way between the two irises is computed. The search area for the mouth is determined by a rectangular region of width $\frac{4}{3} f_a$ centered at

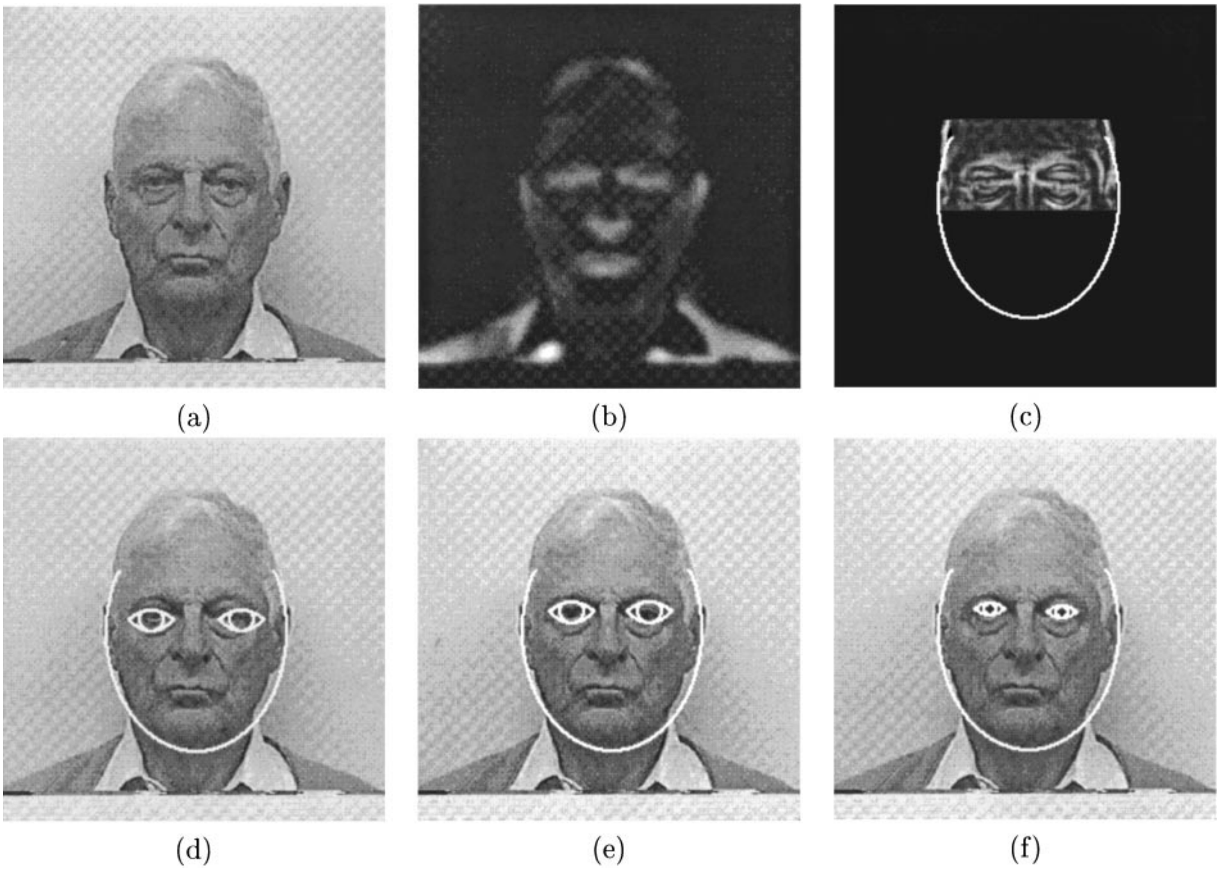


FIG. 7. Iris-attracting and iris-fitting: (a) original image; (b) valley potential; (c) edge potential shown just for eyes; (d)–(e) iris-attracting stages; (f) final iris fit.

the vertical axis, and of height $0.5\,f_b$ centered at $0.5\,f_b$ from the bottom of the oval. This area is large enough so that the mouth is guaranteed to be within this region. An averaging filter (simple arithmetic average) with mask size of 31×5 is convolved with the mouth area of the intensity image to produce the mouth image potential Φ_{mouth}. From the mouth image potential, each horizontal line is summed and the one with the maximum sum is selected as the position of the center of the mouth.

4.7. *Finding the Nose*

After the mouth is found, an area for the nose is determined, guided by the vertical axis, the mouth, and the eye. The nose is searched for in a rectangular region bounded vertically by the middle 60% of the distance between the eyes and the mouth, and bounded horizontally by the 2 eye-centers. Similar to the mouth stage, an averaging filter with mask size (31×5) is convolved with the nose area to produce a nose image potential Φ_{nose}. From the nose image potential, each horizontal line is summed and the one with the maximum is selected as the position of the bottom of the nose. The size of this mask varies with the size of the oval. Figure 8 shows the outcome of locating the center of the mouth, and the bottom of the nose. The correctness of the outcome of both processes can be evaluated by human visual inspection.

5. EVALUATING DIFFERENT FACIAL FEATURE RATIOS

After the primary facial features have been located, they can be used to compute the ratios for age classification. Six ratios are evaluated and their usefulness is discussed. A facial database of 47 faces, comprising babies, seniors, and young/middle-aged adults, is used for this study. Figure 9 graphically explains these ratios.

Tables 1 and 2 show the calculated ratio results. Ratio 6 is not included because it was difficult to obtain robustly. The discussion of Ratio 6 is presented in the upcoming section titled "Ratio 6." Ratios 3 and 5 use the chin parameter and thus have corresponding ratios 3b and 5b for when the chin is obtained from the initial oval. In these tables, "ochin" refers to oval-chin.

For the five ratios, the individual column data from Tables 1 and 2 are used to obtain thresholds for classification. The first five ratios were recomputed after dropping the data evaluated as unfavorable due to facial expression or rotation of the head. The bimodal threshold for each ratio is calculated according to Otsu's method [22].

Table 3 tabulates the statistics of each ratio. The results using Ratio 6 were not included as the virtual top of the head could only be computed for a small set of faces.

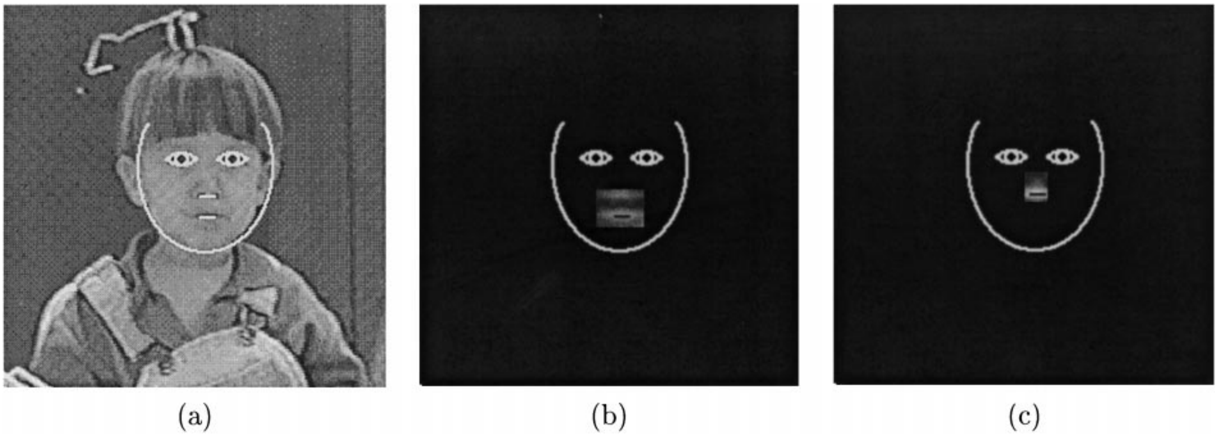


FIG. 8. Mouth- and nose-fitting: (a) final mouth and nose fit shown over original image; (b) mouth potential (from image); (c) nose potential (from image).

5.1. Ratio 6—Eye–Chin : Top_Head–Chin

Ratio 6 is the height of the eyes within the top and bottom head-margins. For this ratio, it is not practical to obtain the true top of the head; hence, the virtual top of the head is used. This ratio makes use of features found in steps 1 through 5 of the overview. It is the most reliable of the ratios, if the top of the head could be found. It has low intraclass variance, and high interclass variance. However, our method of measurement of the virtual top of the head is not robust to how the hair lies around the ears, whether or not the person is bald, and to variation in the actual shape of the lower part of the head. This measurement could be improved with better methods for estimating hair configurations.

5.2. Discussion of Ratios

Several ratios have been examined above. The most promising is Ratio 1. This ratio uses features which are not affected by any facial expressions or facial motions. However, it too is subject to

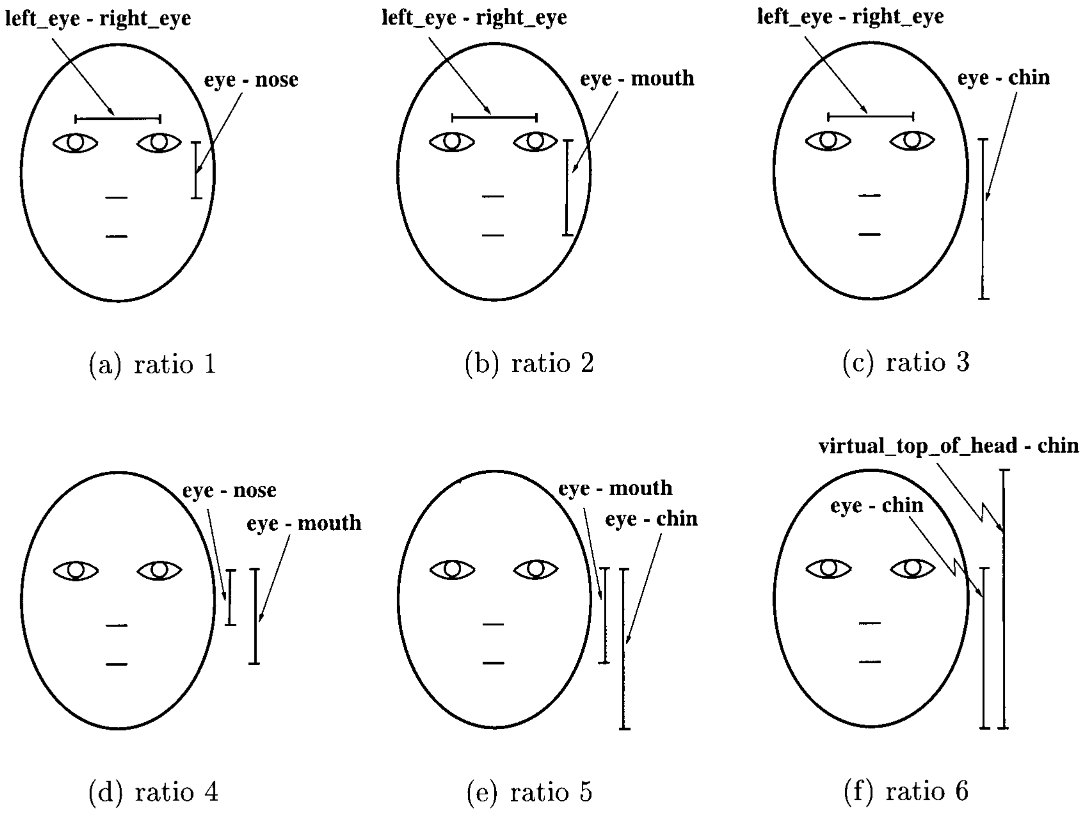


FIG. 9. The six ratios used.

TABLE 1
Results of the Ratio Computation with Real Images for Baby Class

Subject	Ratio1 nose-T	Ratio2 mth-T	Ratio3a chin-T	Ratio3b ochin-T	Ratio4 e-n-m	Ratio5a e-m-c	Ratio5b e-m-ochin
baby01	1.4857	0.9286	0.5843	0.5909	0.6250	0.6292	0.6364
baby02	1.5385	0.9836	0.6186	0.6250	0.6393	0.6289	0.6354
baby03	1.4167	0.8947	0.5763	0.5313	0.6316	0.6441	0.5938
baby04	1.5000	0.9333	0.5753	0.5833	0.6222	0.6164	0.6250
baby05	1.2941	0.9362	0.5789	0.5714	0.7234	0.6184	0.6104
baby08	1.3500	0.8710	0.5455	0.5510	0.6452	0.6263	0.6327
baby10	1.5556	0.9825	—	—	0.6316	—	—
baby12	1.5625	0.9434	—	0.5495	0.6038	—	0.5824
baby13	1.5862	0.9583	—	0.6216	0.6042	—	0.6486
baby14	1.4483	0.9130	0.6000	0.5833	0.6304	0.6571	0.6389
baby15	1.5152	0.9434	—	—	0.6226	—	—
baby16	1.5172	0.9778	—	—	0.6444	—	—
baby17	1.5429	0.9153	—	0.6000	0.5932	—	0.6556
baby18	1.3684	0.8387	0.5200	0.5361	0.6129	0.6200	0.6392
baby19	2.0769	1.0189	—	—	0.4906	—	—
baby20	1.6296	0.9565	0.6667	0.6377	0.5870	0.6970	0.6667
baby21	1.7333	0.9811	—	—	0.5660	—	—
baby22	1.4286	0.9302	—	0.5263	0.6512	—	0.5658
baby24	1.7333	1.1304	0.6047	0.6047	0.6522	0.5349	0.5349
baby25	1.8000	0.9818	—	—	0.5455	—	—
b02	1.5484	0.8727	—	—	0.5636	—	—
b06	1.3939	0.9020	0.5169	—	0.6471	0.5730	—
b07	1.4545	0.9057	—	0.5581	0.6226	—	0.6163
b09	1.7419	1.0588	—	—	0.6078	—	—
b18	1.7391	1.0000	—	—	0.5750	—	—
Sum	38.9609	23.7579	6.3870	8.6702	15.3384	6.8453	9.2819
Num elements	25	25	11	15	25	11	15
Average	1.55844	0.95032	0.58064	0.57801	0.61354	0.62230	0.61879
Std. deviation	0.16851	0.06086	0.04138	0.03434	0.04352	0.03988	0.03469
Variance	0.02840	0.00370	0.00171	0.00118	0.00189	0.00159	0.00120

imprecise localization. If it can be made robust to shading, shadowing, and occlusion effects, it should serve as a good classifier. Ratio 2 appears to be the ratio that can be measured reliably and also shows promise in providing reliable classification. Ratios 3, 4, and 5 are not as promising. In theory, Ratio 6 is the most reliable, but, in practice, suffers from errors in estimating the virtual top of the head.

Ratios 1, 2, and 3 will suffer if the face is rotated in depth, and, as such, some measure needs to be adopted to compensate for this rotation, before the ratios are computed. Ratios 4, 5, and 6 are more robust to this occurrence.

Further research should explore enhancing this approach by combining several ratios to make the final ratio classification. Such combination could be based on statistical analysis.

6. WRINKLES

6.1. *Finding Wrinkles*

Once the primary features have been found for the face, the wrinkle geography map is used to determine where snakelets should be dropped to search for wrinkles (see Fig. 10).

Since the resolution of a 256 × 256 image does not capture any wrinkle information, it is necessary to zoom in to the areas depicted by the wrinkle geography to capture further detail. For now, to prove our concept, the zooming-in process is

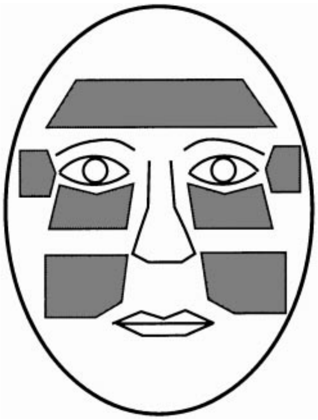


FIG. 10. Wrinkle geography. This shows the regions that are searched for facial wrinkles, after the eyes, nose, mouth, chin, and sides of the face have been located.

TABLE 2

Results of the Ratio Computation with Real Images for Adult and Senior Classes

Subject	Ratio1 nose-T	Ratio2 mth-T	Ratio3a chin-T	Ratio3b ochin-T	Ratio4 e-n-m	Ratio5a e-m-c	Ratio5b e-m-ochin
snr01	1.1923	0.9118	—	0.6019	0.7647	—	0.6602
snr02	1.7333	—	—	0.5778	—	—	—
snr03	—	—	—	—	—	—	—
snr05	1.4286	0.9434	—	—	0.6604	—	—
snr06	1.6800	0.9767	—	0.6176	0.5814	—	0.6324
snr07	—	—	—	—	—	—	—
snr10	—	—	—	—	—	—	—
snr11	—	1.0000	—	0.5455	—	—	0.5455
snr15	1.3621	0.7119	—	0.4468	—	—	0.6277
snr18	1.1481	0.8158	—	0.5439	0.7105	—	0.6667
s01	1.4737	0.9333	0.5957	0.5657	0.6333	0.6383	0.6061
s10	1.3500	0.8710	0.5934	0.5934	0.6452	0.6813	0.6813
s11	1.4359	0.9492	—	—	0.6610	—	—
s12	1.5263	0.9355	0.6042	0.5979	0.6129	0.6458	0.6392
adult01	1.4167	0.8718	0.5313	0.5440	0.6154	0.6094	0.6240
adult04	1.2778	0.7931	0.5055	0.5111	0.6207	0.6374	0.6444
a01	1.1765	0.8333	0.5195	—	0.7083	0.6234	—
a02	1.2941	0.9362	0.5714	0.5789	0.7234	0.6104	0.6184
a04	1.2121	0.9302	0.5405	0.5479	0.7674	0.5811	0.5890
a06	1.1111	0.8163	0.5263	0.5479	0.7347	0.6447	0.6712
a12	1.2000	0.8571	—	—	0.7143	—	—
a19	1.6296	0.9778	0.5867	0.5867	0.6000	0.6000	0.6000
Sum	23.2862	16.0644	5.5745	8.4071	10.7537	6.2718	8.8060
Num elements	18	18	10	15	16	10	14
Average	1.36977	0.89247	0.55745	0.56047	0.67210	0.62718	0.62900
Std. deviation	0.15074	0.05675	0.03475	0.02681	0.05445	0.02700	0.02971
Variance	0.02272	0.00322	0.00121	0.00072	0.00296	0.00073	0.00088

TABLE 3

Results of the Ratio Computation with Real Images (47 Faces), Indicating That It Is Possible to Computationally Distinguish between Baby and Nonbaby Images

Ratio	Threshold		Correctly labeled	Sample size	% Correct
1	1.48	Baby	14	21	67%
		Adult	9	13	69%
2	0.91	Baby	16	21	76%
		Adult	7	13	54%
3a	0.57	Baby	7	9	78%
		Adult	4	9	44%
3b	0.53	Baby	12	13	92%
		Adult	2	13	15%
4	0.62	Baby	8	21	38%
		Adult	8	12	67%
5a	0.64	Baby	6	9	67%
		Adult	3	9	33%
5b	0.63	Baby	6	13	46%
		Adult	7	12	58%

accomplished manually. Figure 11 shows how the area around an eye could be zoomed into to obtain a new 256 × 256 image. With an actively controlled zoom lens, the zooming-in task could be made automatic. Another possibility is to take higher resolution images at the outset and search for wrinkles in the areas depicted by the wrinkle geography. Recently, commercially

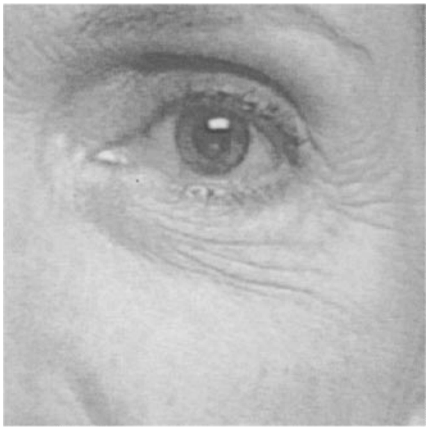


FIG. 11. Zooming in to increase resolution. While this could be automated with a camera that has software-controlled zoom capability, it is accomplished manually at this stage of the research.

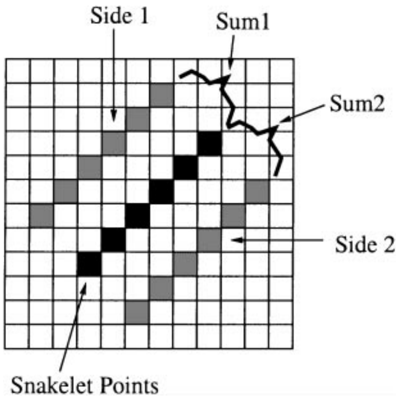


FIG. 12. This shows the directional derivative orthogonal to the snakelet curve for each side of the snakelet.

available high resolution cameras (1280 × 1024) have become available from manufacturers such as Kodak. Pursuing this approach with such high resolution images is a topic for further research.

6.2. *Possible Strategies for Finding and Analyzing Wrinkles*

Once an image (zoomed-in) in which the presence or absence of wrinkles is to be determined has been obtained, the choices outlined at the start of this paper for face analysis can be reapplied. That is, one option is to treat the whole image as an input vector and conduct an algebraic analysis of the vector space to achieve classification (for example, Shackleton and Welsh [26] used this strategy for recognition based on eye images) into wrinkled and nonwrinkled areas. The other option is to detect individual features (in this case, individual wrinkle-candidates) and conduct geometric analysis of them to confirm whether these features are wrinkles or not. As an embodiment of the principles of geometric feature analysis, one could also examine an individual wrinkle-candidate for its curvature and how deeply it is embedded in the skin. Another embodiment would be to perform a match between a deformable template of a wrinkle and the image. The elementary embodiment we use here is one where we simply confirm that the candidates for pieces of wrinkle are not all lying on just one underlying curve. As such, if all the candidates for wrinkles-pieces lie on one curve (with very few outliers), the candidates are not labeled as wrinkles. In this manner, curves arising from shading marks, or noisy specks of skin, are not mistakenly labeled as wrinkles.

6.3. *Steps for Detection and Classification of Wrinkle-Candidates*

First, snakelets [14] are dropped in random orientations along the nodes of a grid using the raw intensity image as the potential for the snakelets. The bounding region for the grid is chosen by the areas shown in the wrinkle geography map.

6.3.1. Steps for the detection of wrinkle-candidates. When these snakelets have stabilized, those snakelets that have found shallow valleys are eliminated. A shallow valley is detected according to the following: For each point in a snakelet, the directional derivative (of the raw image intensity) taken orthogonal to the snakelet curve is computed.

Figure 12 shows the directional derivative orthogonal to the snakelet curve for each side of the snakelet. These are summed separately for each side of the snakelet and normalized for

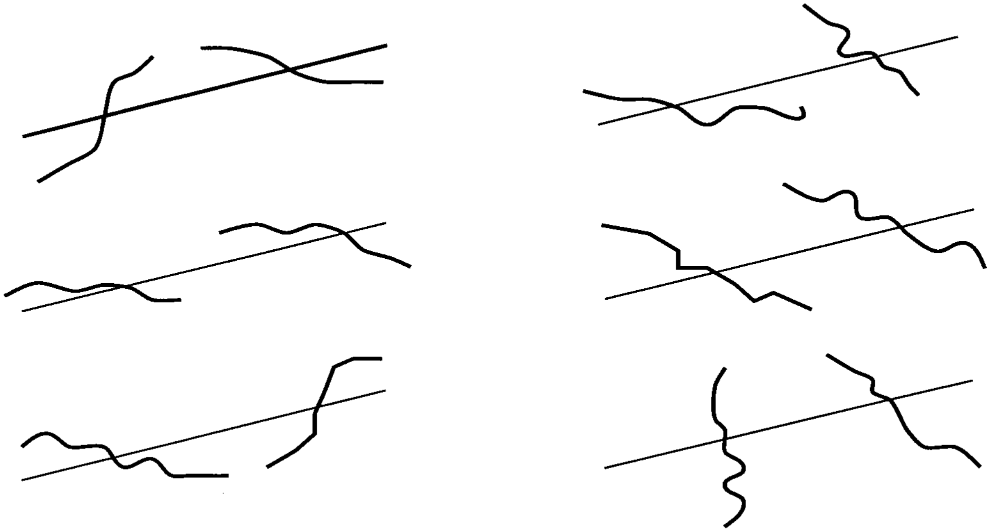


FIG. 13. Wrinkle pattern test. This shows how pairs of snakelets are judged as to whether they are part of a wrinkle. The pairs on the left are classed as not being reliable wrinkle patterns, as they both may belong to a single curve. The pairs on the right are taken as appropriate evidence for a wrinkle pattern.

FIG. 14. Wrinkle-finding process for a region beneath the right eye. Top row shows data for an older adult, middle row for a young adult, and bottom row for a baby. (a) Initial snakelets dropped in the image shown in 11. First column, except for (a), shows the original image. Second column (b, e, h) shows stabilized snakelets. Third column (c, f, i) shows results of snakelets that survive the steepness test. It is obvious that only the older adult data will pass the Wrinkle Pattern Test.

the number of points in the snakelet, to obtain two sums that indicate the steepness of the valley the snakelet is occupying. If these steepness sums do not exceed some preselected threshold, the snakelet is eliminated. In this manner, only those snakelets that lie in a deep enough intensity valley survive. The deep intensity valleys correspond to narrow and deep wrinkles. Note that the relationship between the depth of a wrinkle and the depth of the intensity valley is fortuitous; shading effects cause deeper wrinkles to appear darker.

6.3.2. Classification of wrinkle-candidates: Wrinkle pattern test. Finally, the snakelets that survive the steepness test, are analyzed for their group pattern, to ensure that there are enough of them and that they do not all lie on only one curve. First,

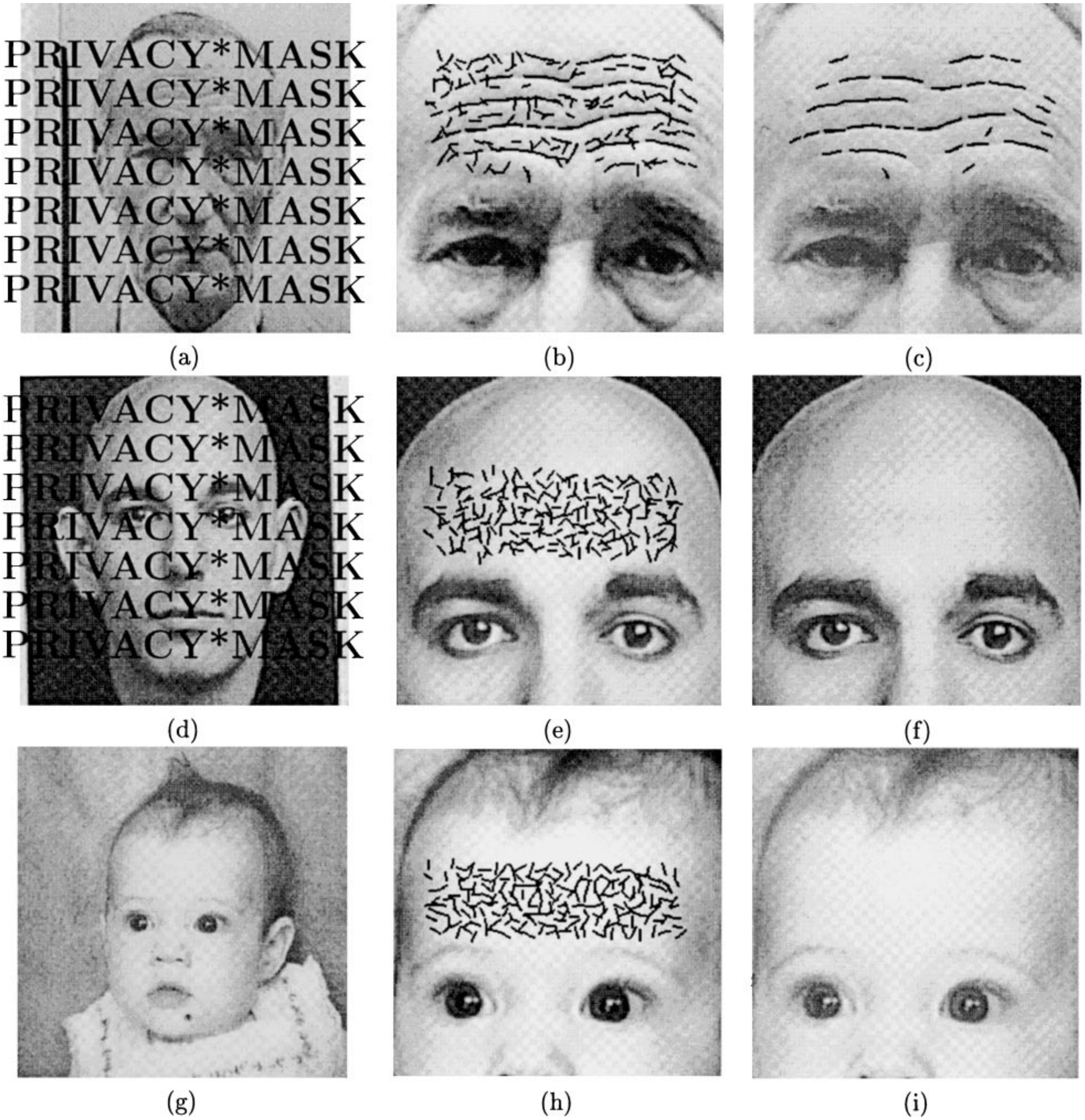


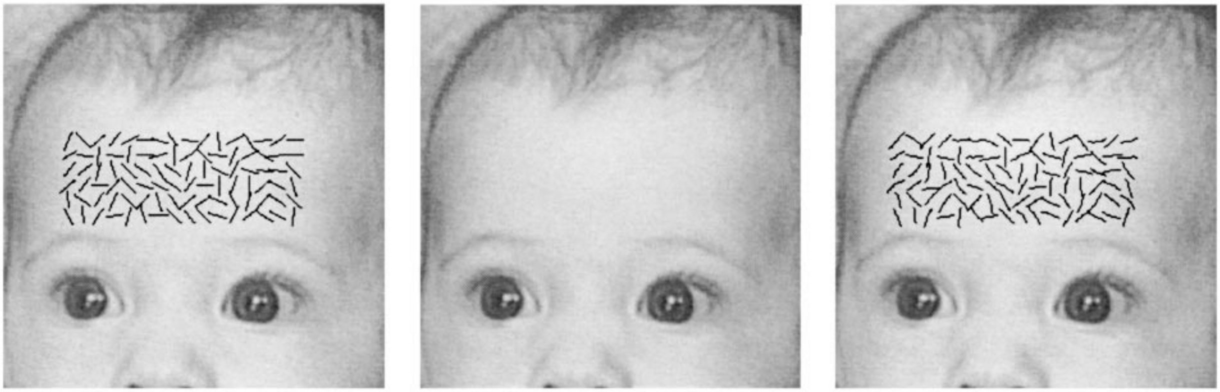
FIG. 15. Wrinkle-finding process for the forehead region. Top row shows data for an older adult, middle row for a young adult, and bottom row for a baby. First column (a, d, g) shows the original image. Second column (b, e, h) shows stabilized snakelets. Third column (c, f, i) shows results of snakelets that survive the steepness test. It is obvious that only the older adult data will pass the Wrinkle Pattern Test.

it is verified that there are more than five snakelets. Next, for each pair of snakelets (there are a square number of pairs), the following is done. The line segment ($\overline{P_i P_j}$) joining the snakelets is computed, and the average orientation for each of the two snakelets with respect to this line is computed. If either of the two orientations is sufficiently orthogonal to the line, or if they satisfy directional constraints indicating they could not belong to the same shallow curve, the snakes are taken to be multiple curves. The region is classified as wrinkled if there is substantial evidence for multiple curves. The evidence is computed from

$$\sum_{P_i, P_j} e^{-(|\overline{P_i P_j}|^2/\sigma^2)} \left(\overline{P_i P_j}^{\perp} \cdot \widehat{\tau_{P_i}}\right) \left(\overline{P_i P_j}^{\perp} \cdot \widehat{\tau_{P_j}}\right),$$

where P_i and P_j are centers of two snakelets, $\widehat{\tau_{P_j}}$ is the unit tangent to the snakelet centered at P_j, $\overline{P_i P_j}^{\perp}$ is the unit

(baby24fw - 44 snakes) (baby24fw - 1 snakes) (baby24fw - 43 snakes)



(baby25fw - 96 snakes) (baby25fw - 0 snakes) (baby25fw - 96 snakes)

(baby14ew - 50 snakes) (baby14ew - 0 snakes) (baby14ew - 50 snakes)

FIG. 16. Wrinkle analysis process. First column: initial snakelets dropped into an image. Second column: wrinkle snakelets (including quantity). Third column: nonwrinkle snakelets.

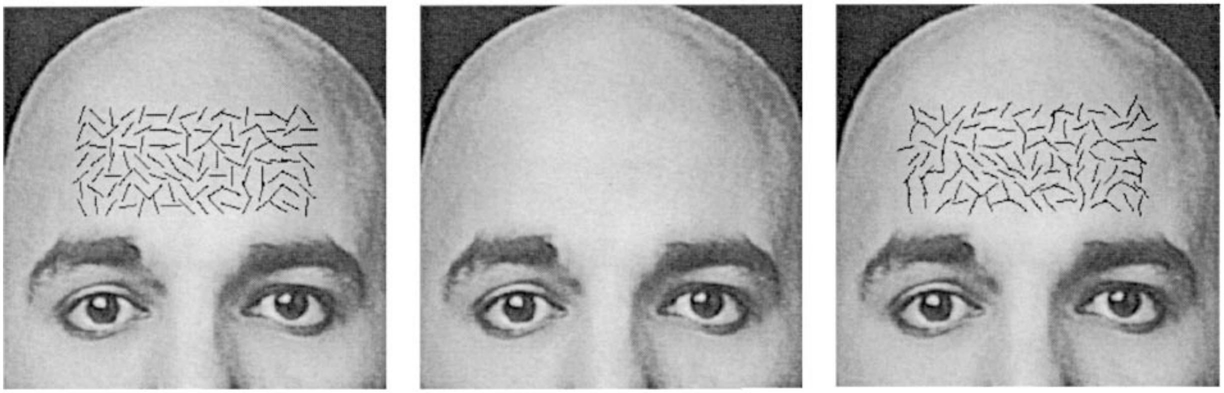
perpendicular to the line joining P_i and P_j, and $e^{-(|\overline{P_i P_j}|^2/\sigma^2)}$ is a decaying function of the distance between the two snakes. To be classified as a region with multiple curves, this computed evidence must exceed a preset threshold. (See Fig. 13 for relevant examples of wrinkle patterns.)

Figure 14 shows the outcomes of the process applied to a senior, a young adult, and a baby in the area beneath the image's right eye. Figure 15 shows the outcomes of the process applied to a senior, a young adult, and a baby in the area around the forehead.

In addition, Figs. 16 to 21 show how snakelets are detected, counted, and used in confirming the presence of wrinkled skin. In these figures, the first column shows the initial snakelets dropped, the second column shows the snakelets that are found to be candidates (the numbers in the columns are the quantities

(baby17ew - 72 snakes) (baby17ew - 0 snakes) (baby17ew - 72 snakes)

(baby22ew - 98 snakes) (baby22ew - 3 snakes) (baby22ew - 98 snakes)



(adult01fw - 112 snakes) (adult01fw - 0 snakes) (adult01fw - 112 snakes)

FIG. 17. Wrinkle analysis process. First column: initial snakelets dropped into an image. Second column: wrinkle snakelets (including quantity). Third column: nonwrinkle snakelets.

of snakelets), and the third column shows those snakelets that are not candidates.

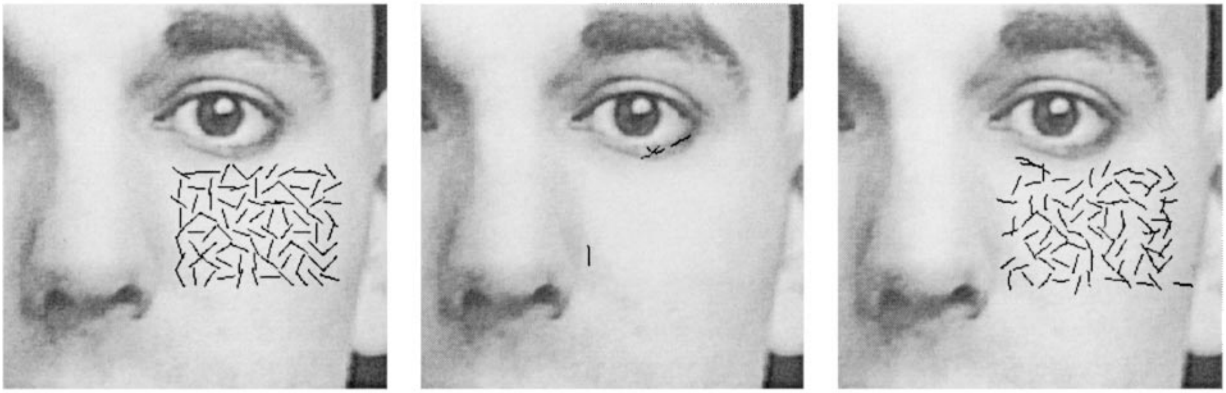
6.3.3. Combining different locales of evidence for wrinkles. At this stage in the research, the different locales which are inspected for wrinkles (e.g., forehead and area beneath each eye) are weighted equally in determining whether the face has wrinkles.

7. COMBINING RATIOS AND WRINKLE INFORMATION TO JUDGE AGE CATEGORY

With only three age groups to be categorized, the combination rule is simple. If the ratios are baby-like and no wrinkles are found, the face is labeled a baby. If wrinkles are found, and the ratios are adult-like, the face is labeled a senior adult. For all other cases, the face is labeled a mid-aged adult. To decide if

(adult04fw - 108 snakes) (adult04fw - 0 snakes) (adult04fw - 108 snakes)

(adult04rew - 144 snakes) (adult04rew - 5 snakes) (adult04rew - 139 snakes)



(adult05ew - 88 snakes) (adult05ew - 4 snakes) (adult05ew - 84 snakes)

FIG. 18. Wrinkle analysis process. First column: initial snakelets dropped in an image. Second column: wrinkle snakelets (including quantity). Third column: nonwrinkle snakelets.

the ratios are baby-like, we use Ratios 1 and 2, as they yielded the greatest correct percentages in Table 3. They are also the two ratios with the widest spread between their averages for the groups. Thus, a face is considered baby-like if it had no wrinkles, and if either Ratio 1 or 2 indicates it is a baby.

Our complete database contains 47 faces comprising babies, seniors, and young/mid-aged adults. Wrinkle detection was not performed on all faces because the zoomed-in images were only obtained for 15 faces. For these 15 faces, the complete classification algorithm was run, and the classifications were 100% correct. The 15 faces contained five babies, five mid-age adults, and five seniors.

Table 4 summarizes the measurements and computed inferences for the 15 faces. In it, column 2 shows Ratio 1, along

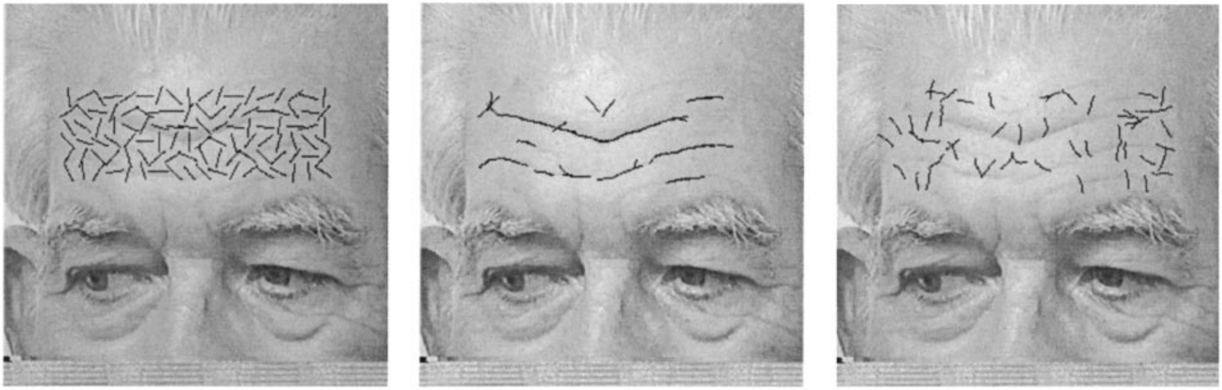
(snr01ew - 128 snakes) (snr01ew - 56 snakes) (snr01ew - 72 snakes)

(snr02ew - 180 snakes) (snr02ew - 126 snakes) (snr02ew - 54 snakes)

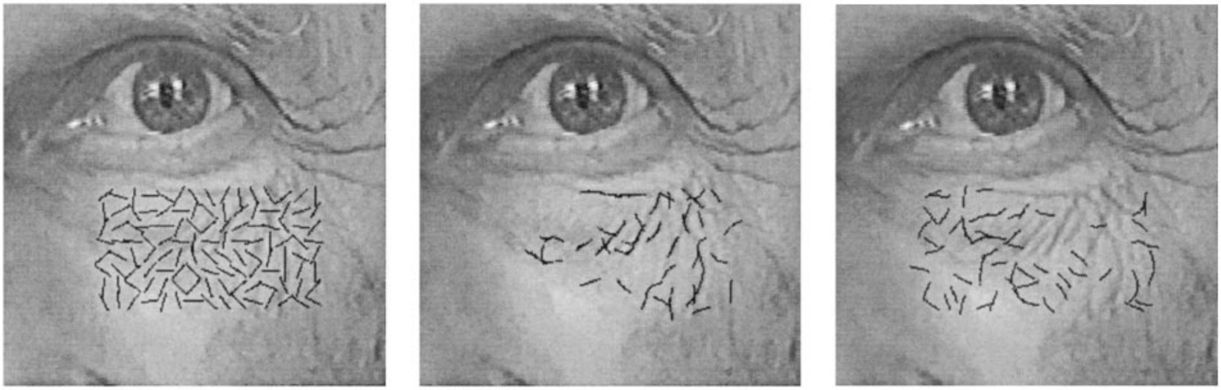
(snr09ew - 120 snakes) (snr09ew - 77 snakes) (snr09ew - 43 snakes)

FIG. 19. Wrinkle analysis process. First column: initial snakelets dropped in an image. Second column: wrinkle snakelets (including quantity). Third column: nonwrinkle snakelets.

(snr13ew - 135 snakes) (snr13ew - 102 snakes) (snr13ew - 33 snakes)



(snr14fw - 108 snakes) (snr14fw - 57 snakes) (snr14fw - 51 snakes)



(snr15ew - 120 snakes) (snr15ew - 59 snakes) (snr15ew - 61 snakes)

FIG. 20. Wrinkle analysis process. First column: initial snakelets dropped in an image. Second column: wrinkle snakelets (including quantity). Third column: nonwrinkle snakelets.

with the threshold used to decide if the ratio is baby-like or not. Column 3 similarly shows Ratio 2. Column 4 shows the number of dropped snakelets that passed the steepness test and are thus wrinkle snakelets. Column 5 shows the result of applying the wrinkle pattern test to these wrinkle snakelets. Column 6 shows the final inference that combines the information from columns 2, 3, and 5, in the manner described in this section.

8. DISCUSSION AND CONCLUSION

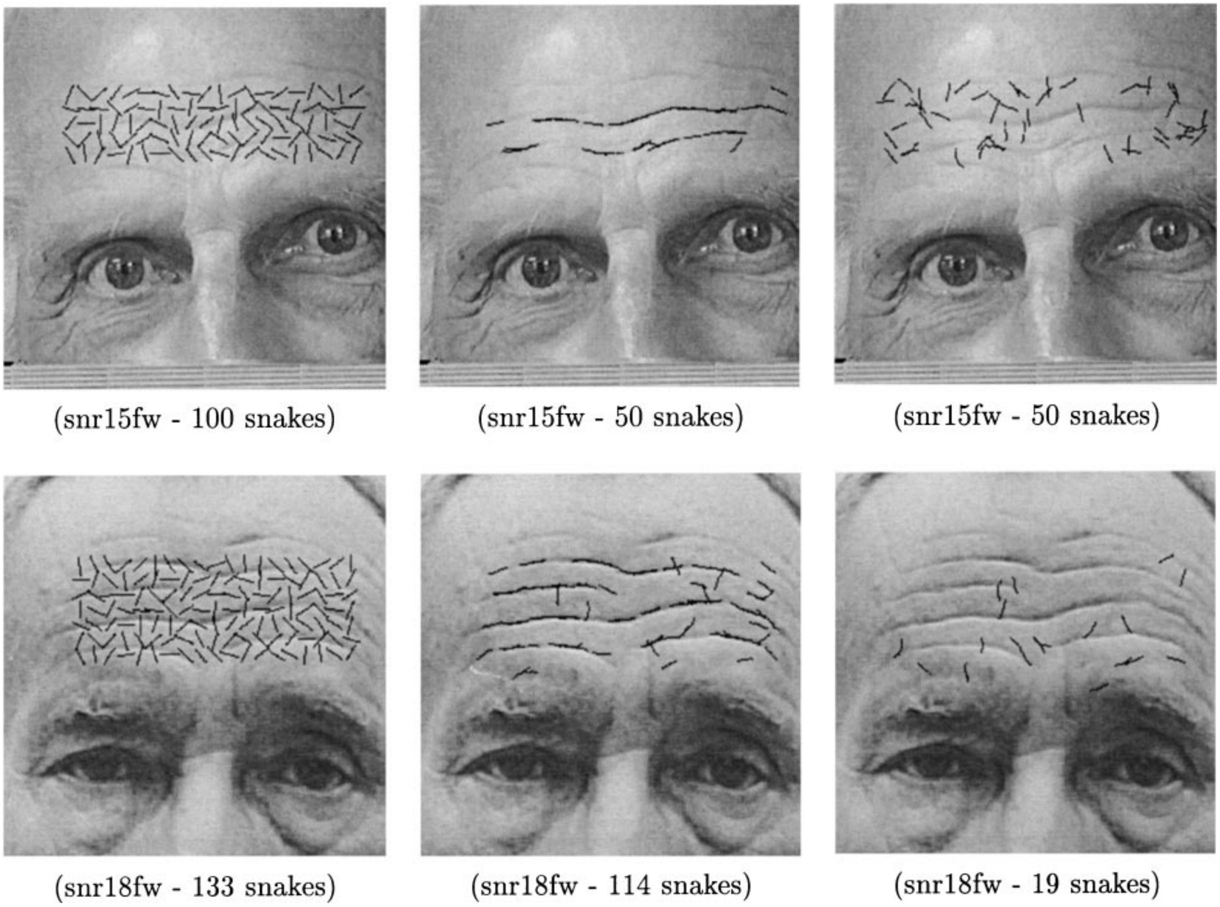
We have outlined a computational theory for visual age classification from facial images. For now, only three age-groups were considered: babies, young adults, and senior adults. First, primary features of the face, namely the eyes, nose, mouth, chin, and virtual top of the head, are found. From these, ratios are

TABLE 4

Results of the Complete Classification Scheme Applied to 15 Faces: 5 Babies, 5 Mid-age Adults, and 5 Seniors

Subject	Ratio 1 threshold = 1.48	Ratio 2 threshold = 0.912	Wrinkle snakelets found	Decide wrinkled?	Computed label
Baby14	1.45	0.913	0	No	Baby
Baby17	1.54	0.915	0	No	Baby
Baby22	1.43	0.930	3	No	Baby
Baby24	1.73	1.13	1	No	Baby
Baby25	1.80	0.982	0	No	Baby
Adult01	1.41	0.872	0	No	Adult
Adult04	1.28	0.793	5	No	Adult
Adult05	1.30	0.731	4	No	Adult
A02	1.29	0.936	0	No	Adult
A12	1.20	0.857	0	No	Adult
Snr01	1.19	0.911	56	Yes	Senior
Snr09	1.39	0.78	77	Yes	Senior
Snr13	1.25	0.81	102	Yes	Senior
Snr15	1.36	0.71	59	Yes	Senior
Snr18	1.14	0.81	114	Yes	Senior

Note. Column 2 shows Ratio 1, along with the threshold used to decide if the ratio is baby-like or not. Column 3 similarly shows Ratio 2. Column 4 shows the number of dropped snakelets that passed the steepness test and are thus wrinkle snakelets. Column 5 shows the result of applying the wrinkle pattern test to these wrinkle snakelets. Column 6 shows the final inference that combines the information from columns 2, 3, and 5, in the manner described in this section.



(snr15fw - 100 snakes) (snr15fw - 50 snakes) (snr15fw - 50 snakes)

(snr18fw - 133 snakes) (snr18fw - 114 snakes) (snr18fw - 19 snakes)

FIG. 21. Wrinkle analysis process. First column: initial snakelets dropped in an image. Second column: wrinkle snakelets (including quantity). Third column: nonwrinkle snakelets.

computed that permit the distinguishing of babies from others. Next, secondary features, namely wrinkles, are detected and analyzed. This step permits the distinguishing of seniors from those in the two younger categories.

This work has shown that computing ratios and detecting the presence of wrinkles can yield age categorization. These criteria were suggested by cranio-facial research and the observation that aging skin develops wrinkles. There are several directions that need to be further explored. The problem of varying orientation of the face needs to addressed. The work thus far has assumed mugshot viewpoints, and this makes the ratio computations easy. For faces that do not satisfy this, analysis along the lines of that using symmetry [25] will need to be employed to compute the facial orientation, and to compensate for a nonfrontoparallel orientation.

Next, skin-color and eye-color restrictions will need to be loosened. Then, the computation of age in the presence of eye patches and dark glasses, and other occlusions and shadowing effects needs to be explored. Age computation must be made robust to varieties of moustaches, facial scars, and disheveled hair. Finally, an accurate estimation of the top of the skull has defied all approaches thus far. As we showed, this estimate enhances the ability to tell age and, thus, needs to be computed. In all of these endeavors, additional age-related information can be exploited: the growth of the nose and the nose-bridge, and the relative shrinking of the iris-size over time, and changes to the outline of the face. This work has but examined and documented some computational cues that exist for this task.

ACKNOWLEDGMENTS

The authors are grateful to M. Shackleton (and British Telecom) for providing computer code for finding eye features in facial images and for access to images of middle-aged faces; to D. Terzopoulos for providing code for snakes; to M. Turk (and the M.I.T. Media Lab) for access to images of middle-aged faces; to A. O'toole and A. Yuille for helpful advice; to the reviewers for valuable comments; and to the numerous babies and senior adults whose pictures were used.

REFERENCES

1. R. A. Alley, *Social and Applied Aspects of Perceiving Faces*, Erlbaum, Hillsdale, NJ, 1988.
2. D. Arnink, *Creative Theatrical Makeup*, Prentice–Hall, Englewood Cliffs, NJ, 1984.
3. W. W. Bledsoe, *Man–Machine Facial Recognition*, Tech. Rep. PRI:22, Panoramic Res. Inc., Palo Alto, CA, 1966.
4. R. Brunelli and T. Poggio, Face recognition: Features versus templates, *IEEE PAMI* **15**, No. 10 (1993), 1042–1052.
5. R. Bruyer, *The Neuropsychology of Face Perception and Facial Expression*, Erlbaum Assoc. Inc., Hillsdale, NJ, 1986.
6. A. Clark and M. Kokuer, Feature identification and model tracking for model-based coding, in *Proc. of ICPR, Hague, Netherlands, 1992*, Vol. 3, pp. 79–82.
7. R. Corson, *Stage Makeup*, Prentice–Hall, Englewood Cliffs, NJ, 1986.
8. I. Craw, D. Tock, and A. Bennett, Finding face features, in *Proc. ECCV, Europe, 1992*, pp. 92–96.
9. A. J. Golomb, D. T. Lawrence, and T. J. Sejnowski, Sexnet: A neural network identifies sex from human faces, *Advances in Neural Information Processing Systems* **3** (1991), 572–577.
10. V. Govindaraju, D. B. Sher, and R. Srihari, Locating human faces in newspaper photographs, in *Proc. CVPR, San Diego, 1989*, pp. 278–285.
11. V. Govindaraju, R. Srihari, and D. B. Sher, Caption-aided face location in newspaper photographs, in *Proc. ICPR, The Hague, 1992*, pp. 474–477.
12. R. Haralick and L. Shapiro, *Robot and Computer Vision, Vol. 1*, Addison–Wesley, Reading, MA, 1992.
13. T. Kanade, Picture processing by computer complex and recognition of human faces, Tech. Rep., Dept. Inform. Sci., Kyoto University, 1973.
14. M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active contour models, in *IJCV, 1988*, pp. 321–331.
15. M. Kirby and L. Sirovich, Application of the K-L procedure for the characterization of human faces, *IEEE PAMI* **12**, No. 1 (1990), 103–108.
16. Y. Kwon, *Age Classification from Facial Images*, M.Sc. thesis, Univ. Central Florida, Dec. 1993.
17. Y. Kwon and N. Da Vitoria Lobo, Age classification from facial images, in *Proc. CVPR, Seattle, WA, June 1994*, pp. 762–767.
18. Y. Kwon and N. Da Vitoria Lobo, Face detection using templates, in *Proc. ICPR, Israel, Oct. 1994.*
19. K. Matsuno, C.-W. Lee, and S. Tsuji, Recognition of human facial expression without feature extraction, in *Proc. of ECCV, Stockholm, Sweden, May 1994*, pp. 513–520.
20. T. Nagamine, T. Uemura and I. Masuda, 3D facial image analysis for human identification, in *Proc. ICPR, Hague, Netherlands, 1992*, Vol. 1, pp. 324–327.
21. C. Novak, Finding facial features in color images, in *CVPR Wkshp on Color Vision, New York, June 1993*.
22. N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Systems, Man, Cybernetics* **SMC-9** (1979), 62–66.
23. A. J. O'Toole, H. Abdi, K. A. Deffenbacher, and J. C. Bartlett, Classifying faces by race and sex using an autoassociative memory trained for recognition, in *Proc. of Annual Meeting of the Cog. Sci. Soc., Chicago, 1991*, pp. 847–851.
24. M. J. T. Reinders, B. Sankur, and J. C. A. van der Lubbe, Transformation of a general 3D facial model to an actual scene face, in *Proc., International Conference on Pattern Recognition, The Hague, Netherlands, 1992*, Vol. 3, pp. 75–78.
25. D. Reisfeld and Yeshurun, Robust detection of facial features by generalized symmetry, in *Proc. ICPR, The Hague*, 1992, pp. 117–120.
26. M. A. Shackleton and W. J. Welsh, Classification of facial features for recognition, in *Proc. CVPR, Hawaii, 1991*, pp. 573–579.
27. M. A. Turk and A. P. Pentland, Face recognition using eigenfaces, in *Proc. CVPR, Hawaii, 1991*, pp. 586–591.
28. G. Yang and T. Huang, Human face detection in a scene, in *Proc. CVPR, New York, 1993*, pp. 453–458.
29. A. L. Yuille, P. W. Hallinan, and D. S. Cohen, Feature extraction from faces using deformable templates *IJCV* **8**, No. 2 (1992), 99–111.
30. A. W. Young and H. D. Ellis (Eds.), *Handbook of Research on Face Processing*, North-Holland, Amsterdam, 1989.