Age Synthesis and Estimation via Faces: A Survey

Yun Fu, *Member*, *IEEE*, Guodong Guo, *Senior Member*, *IEEE*, and
Thomas S. Huang, *Fellow*, *IEEE*

Abstract—Human age, as an important personal trait, can be directly inferred by distinct patterns emerging from the facial appearance. Derived from rapid advances in computer graphics and machine vision, computer-based age synthesis and estimation via faces have become particularly prevalent topics recently because of their explosively emerging real-world applications, such as forensic art, electronic customer relationship management, security control and surveillance monitoring, biometrics, entertainment, and cosmetology. Age synthesis is defined to rerender a face image aesthetically with natural aging and rejuvenating effects on the individual face. Age estimation is defined to label a face image automatically with the exact age (year) or the age group (year range) of the individual face. Because of their particularity and complexity, both problems are attractive yet challenging to computer-based application system designers. Large efforts from both academia and industry have been devoted in the last a few decades. In this paper, we survey the complete state-of-the-art techniques in the face image-based age synthesis and estimation topics. Existing models, popular algorithms, system performances, technical difficulties, popular face aging databases, evaluation protocols, and promising future directions are also provided with systematic discussions.

Index Terms—Face aging, age estimation, age synthesis, age progression, survey.

✦

1 INTRODUCTION

As a "window to the soul" [1], the human face conveys important perceptible information related to individual traits. The human traits displayed by facial attributes, such as personal identity, facial expression, gender, age, ethnic origin, and pose, have attracted much attention in the last several decades from both industry and academia since face image processing techniques yield extensive applications in graphics and computer vision fields [2], [3], [4]. There are two fundamental problems inspiring the development of these techniques.

- *Face image synthesis*: Render face images with customized single or mixed facial attributes (identity, expression, gender, age, ethnicity, pose, etc.).
- *Face image analysis*: Interpret face images in terms of facial attributes (identity, expression, gender, age, ethnicity, pose, etc.).

Among them, face image-based age synthesis and estimation have become particularly interesting topics in recent years because of their emerging new applications. People

---

- *Y. Fu is with the Department of Computer Science and Engineering, University at Buffalo (SUNY), 241 Bell Hall, Buffalo, NY 14260-2000. E-mail: yunfu@buffalo.edu.*
- *G. Guo is with the Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506. E-mail: guodong.guo@mail.wvu.edu.*
- *T.S. Huang is with Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign (UIUC), 405 North Mathews Avenue, Urbana, IL 61801. E-mail: huang@ifp.uiuc.edu.*

have the ability, developed early in life, to determine age between 20 and 60 years and conceive aging appearance from the face with high accuracy, on average, with a group decision [5], [6]. For example, we can easily figure out the aging process on Albert Einstein's faces, as shown in Fig. 1. Especially, the forensic artist [7] can imagine and make realistic age progression pictures in terms of photos or semantic description of given faces [8]. Well-trained Swedish alcohol salespeople have professional skills for accurate age estimation with low bias [9]. Age of face has also been considered as an important semantic or contextual cue in social networks [2], [10]. Can a machine perform the same as a human? Technology advances in computer science and engineering have given a positive answer to this question. There are two basic tasks in this field, computer-based age synthesis and estimation, that are described as follows:

- *Age synthesis*: Rerender a face image aesthetically with natural aging and rejuvenating effects on the individual face.
- *Age estimation*: Label a face image automatically with the exact age (year) or the age group (year range) of the individual face.

To further understand the tasks, we want to differentiate four concepts about human age in the paper.

- *Actual age*: The real age (cumulated years after birth) of an individual.
- *Appearance age*: The age information shown on the visual appearance.
- *Perceived age*: The individual age gauged by human subjects from the visual appearance.
- *Estimated age*: The individual age recognized by machine from the visual appearance.
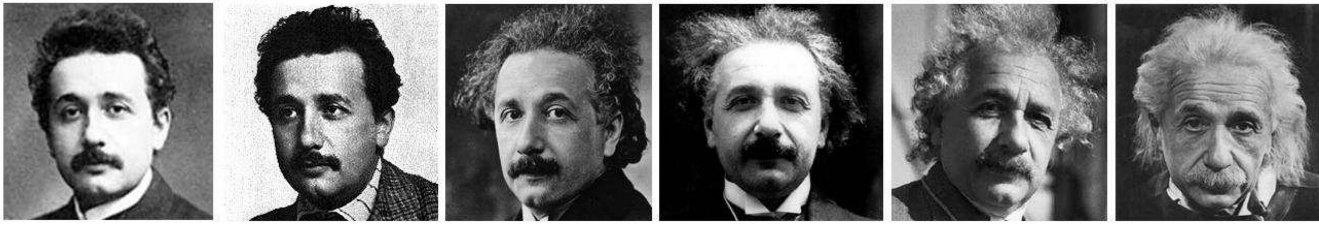
Fig. 1. Albert Einstein's face aging (collected by Internet image search).

The appearance age is typically consistent with the actual age. However, the variation is often inevitable due to the generic difference between different individuals and environmental/artificial factors. Both perceived age and estimated age are defined on the appearance age. The actual age is often defined as the ground truth.

1.1 Real-World Applications

There are many popular real-world applications related to age synthesis [4] and estimation. Computer-aided age synthesis significantly relieves the burden of tedious manual work while at the same time providing more photorealistic effects and high-quality pictures. Age estimation by machine is useful in applications where we don't need to specifically identify the individual, such as a government employee, but want to know his or her age.

1.1.1 Forensic Art

The forensic art involves interdisciplinary knowledge of anthropometry, psychology, postmortem reconstruction, human aging, perception, and computer graphics. As a principal artistic technique in forensic art, age progression is used to modify and enhance photographs by computer or manually (with professional hand drawing skills) for the purpose of suspect/victim and lost person identification with law enforcement [8], [7]. This technique has evolved when police investigative work and art united throughout history. When the photos of missing family members (especially children [11], [12], [13]) or wanted fugitives are outdated, forensic artists can predict the natural aging of the subject faces and produce updated face images, utilizing all available individual information, such as facial attributes, lifestyle, occupation, and genetics. Age synthesis by machine can significantly enhance the efficiency of the forensic artist while at the same time providing more photorealistic aging/rejuvenating effects that can satisfy the needs of aesthetics.

1.1.2 Electronic Customer Relationship Management (ECRM)

The ECRM [14] is a management strategy to use information technology and multimedia interaction tools for effectively managing differentiated relationships with all customers and communicating with them individually. Since different groups of customers have very different consuming habits, preferences, responsiveness, and expectation to marketing, companies can gain more profits by acknowledging this fact, responding directly to all customers' specific needs, and providing customized products or services. The most challenging part hereby is to obtain and analyze enough personal information from all customer groups, which needs companies to establish long-term customer relationships and sustain a large amount of cost input. For example, a fast food shop owner might want to know what percentage of each age group prefers and purchases what kind of sandwiches; the advertisers want to target specific audiences (potential customers) for specific advertisements in terms of age groups; a mobile phone company wants to know which age group is more interested in their new product models showing in a public kiosk; a store display might show a business suit as an adult walks by or jeans as a teenager walks by. Obviously, it is almost impossible to realize those due to privacy issues. However, with the help of a computer-based automatic age estimation system, a camera snapping photos of customers could collect demographic data by capturing customers' face images and automatically labeling age groups. All of these can be done without violating anyone's privacy.

1.1.3 Security Control and Surveillance Monitoring

Security control and surveillance monitoring issues are more and more crucial in our everyday life, especially when advanced technologies and explosive information become common to access and possess [15]. With the input of a monitoring camera, an age estimation system can warn or stop underage drinkers from entering bars or wine shops; prevent minors from purchasing tobacco products from vending machines; refuse the aged when he/she wants to try a roller coaster in an amusement park; and deny children access to adult Web sites or restricted movies [16], [17]. In Japan, police found that a particular age group is more apt to money transfer fraud on ATMs, in which age estimation from surveillance monitoring can play an important role. Age estimation software can also be used in health care systems, such as robotic nurse and intelligent intensive care unit, for customized services. For example, a personalized Avatar will be selected automatically from the custom-built Avatar database to interact with patients from different age groups with particular preferences.

1.1.4 Biometrics

Age estimation is a type of soft biometrics [18] that provides ancillary information of the users' identity information. It can be used to complement the primary biometric features, such as face, fingerprint, iris, and hand geometry, to improve the performance of a primary (hard) biometrics system. In real face recognition or identification applications, it is often the case that the system needs to recognize or identify faces after a gap of several years [19], [20], [21], [22], such as passport renewal and border security [23], which reveals the importance of age synthesis. Integrated

with a dynamic aging model, the face recognition or identification system can dynamically tune the model parameters by considering the shape or texture variations during the aging process. System robustness to time gap can be significantly improved [24].

1.1.5 Entertainment

Aging and rejuvenating are popular special visual effects in film making, especially for science fiction films such as "*The Curious Case of Benjamin Button*" (2008). Without any noticeable artifacts in many such movies, the actor's appearance can be transformed from young to old or reverse instantly or gradually with extremely realistic aging effects. Some of these mysterious visual effects are generated by age synthesis techniques to provide fantastic experiences to audiences. Image morphing [25] is often used to generate a seamless transition for animation purpose, such as Michael Jackson's music video "*Black or White*" (1991). Automatic album management of consumer photographs [26] is a possible application for age estimation [10]. Users can select photo groups from the automatically organized album according to the facial attributes labeling.

1.1.6 Cosmetology

People want to have their faces look younger. Rejuvenation, implemented by cosmetic surgery (shape) [27], [28] or skin care (texture), is a popular beauty treatment in cosmetology. Computer-aided age synthesis is obviously a useful technique to predict the rejuvenating results as references that can help beauticians or surgeons in the treatments and operations. With a face rejuvenating system available [29], more individual-oriented services can be provided to patients, who can directly see the computer synthetic photorealistic results and make changes with their preferences beforehand [30].

1.2 Problems and Motivations

Although, as aforementioned, the real-world applications are very rich and attractive, existing facts and attitudes from the perception field reveal the difficulties and challenges of automatic age synthesis and estimation by computer. Different people have different rates of the aging process [31], [16], which is determined by not only the person's genes but also many other factors, such as health condition, living style, working environment, and sociality [32]. The effects of ultraviolet radiation, usually through exposure to sunlight, may cause solar aging, which is another strong cause for advanced signs of face aging. In particular, Stone also stated in [33] that aging can be accelerated by smoking, genetic predisposition, emotional stress, disease processes, dramatic changes in weight, and exposure to extreme climates. As investigated by Berry and McArthur [34] using an ecological approach, age-related variations in craniofacial growth play an important role in social perception. Facial characteristics may affect impressions if they typically reveal psychological attributes. Boissieux et al. [35] also pointed out the challenge to accurately modeling aging skin because of the different aesthetic values and focuses of attention people have. Thus, face aging is uncontrollable and personalized [23], [36], [37], which makes the task of accurately capturing aging variations more difficult. Moreover, age ranking with ordinal

labels is often a hard task because of the uncertainty within the labels [38], [39].

From a technical point of view, males and females may have different face aging patterns displayed in images due to the different extent in using makeups and accessories [40], [31]. Many female face images may potentially show younger appearances. There are still two open problems remaining on how to 1) build personalized age synthesis models and 2) extract general discriminative features for age estimation, reducing the negative influence of individual differences. Moreover, the burden of acquiring large-scale databases which cover enough age range with chronological face aging images makes the estimation tasks more difficult to achieve. Although Web image mining can help the data collection [41], it is usually hard or even impractical to collect a large database of large amount of subjects who can provide a series of personal images in different ages.

How can a machine fulfill the job of age synthesis and estimation? What kinds of techniques can we use to deal with the problems and overcome difficulties? What has been done and what will be done? These questions have partially drawn attention at some recent IEEE conferences, such as BTAS'08[1] and FG'08.[2] Some industry products have also been developed by companies such as VideoMining[3] and NEC.[4] System demos and softwares are also extensively reported in [29], [42], [43], [44], [45], [46], [47], [48], [49], [50]. Following this trend, we answer those questions in the paper with a complete survey and general discussions that cover the state-of-the-art technique advances in computer-based age synthesis and estimation. A recently published survey [4] has partially addressed the topic of computational methods for modeling facial aging. From more broad views in this survey, we particularly focus on the approaches and solutions of both age synthesis and estimation emerging in recent years that significantly improve the traditional techniques by considering the particularity and complexity of the problems.

1.3 Organization of the Paper

The remainder of the paper is organized as follows: In Section 2, we summarize the natural aging process happening on human faces in two stages. The perception and aesthetics points of views are also introduced. Starting from the discussion in face modeling aspect, Section 3 surveys the age synthesis techniques in three aspects: shape, texture, and appearance. We introduce the existing age image representation models and state-of-the-art age estimation techniques in Section 4. The performance comparison of existing systems is also discussed. Section 5 lists some popular face aging databases with description details available. Evaluation protocols for age synthesis and estimation are summarized in Section 6. We conclude the paper in the last section and provide promising future directions.

2 HUMAN AGING ON FACES

Human face aging is generally a slow and irreversible process, even though some retinoids (e.g., tretinoin) may
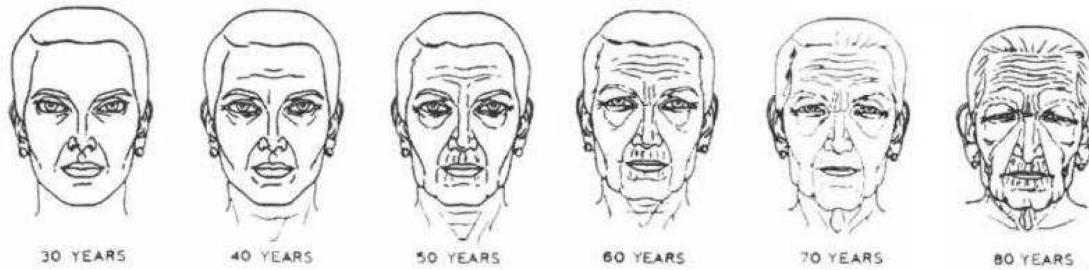

Fig. 2. Face aging sketches from 30 to 80 years with 10 years per sketch, originally shown in [55], also in [33].

slightly reverse minor photoaging effects.[5] Although people are aging differently and aging shows different forms in different ages, there are still some general changes and resemblances we can always describe [51], [1]. From the biological or anthropometric point of view, there are roughly two stages during the human life that are quite different in face growth, development, and aging forms [52].

During the early growth and development of the face, from birth to adulthood, the greatest change is the craniofacial growth (shape change) [5]. Overall, the face size is getting larger gradually during the craniofacial growth. Fig. 3d shows six prototype faces, originally shown in [53], with craniofacial growth from 2 to 18 years, 3 years per sketch. The shape synthesis is based on the anthropometric measurements from [54]. With the growth of the cranium, the forehead slopes back, shrinks, and releases spaces on the surface of the cranium, while the facial features, such as eyes, nose, ears, and mouth, expand their areas and tend to cover these interstitial spaces. Cheeks extend to larger areas and the chin becomes more protrusive. The facial skin relatively does not change too much compared with the craniofacial growth. But facial hairs, such as a mustache, may become dense and even bushy. Skin color may change a little bit. Pores may open and enlarge [5]. More details about craniofacial growth and aging can be retrieved from [51].

During adult aging, from adulthood to old age, the most perceptible change becomes skin aging (texture change). The shape change still continues, but less dramatically, mostly due to typical patterns in skin and tissue. Originally shown in [55] and also in [33], Fig. 2 shows six face aging sketches from 30 to 80 years, with 10 years per sketch. Biologically [33], [5], as the face matures and ages with loss of collagen beneath skin as well as gravity effects, the skin becomes thinner, darker, less elastic, and more leathery. Adynamic wrinkles and blemishes due to biologic aging gradually appear. Dynamic wrinkles and folds due to muscle motion become more distinct. In the areas of deeper attachment, such as cheeks, eyelids, chin, and nose, elasticity of muscles and soft tissues gets weak and fat continues depositing. In other areas, fat may atrophy or be absorbed. These changes lead to the downward descent or sagging of the skin, such as double chin, dropping cheek, and lower eyelid bags. Although the craniofacial growth is not dramatic during this aging period, the facial geometry change is still evident from 30 to 80 years, especially in the female faces. Faces change from a U-shaped or upside-down triangle shape to a trapezoid or rectangle [55], [33].

5. http://www.farrellmanual.com/topical/retina.htm.

The bony framework underneath the skin may also deteriorate to accelerate the development of skin aging, such as wrinkles, creases, and droops. In addition, face aging during this age period may cause the loss of flexible control of facial muscles so that the facial movements and behaviors may also change unintentionally [5]. More details about general ways the head and face age in adulthood can be obtained from [51].

From the perception or aesthetics point of view, face aging can be explained in a different way. Assuming that faces are points in a high-dimensional space [56], the distinctiveness indicates the distance between any two faces. The attractiveness is negatively correlated with distinctiveness because less distinctive and more typical is seen as more attractive, presented by Rhodes and Tremewan [57] in 1996. Following this theory and concepts, Deffenbacher et al. found that face aging is actually related to attractiveness and distinctiveness [58]. Their experiments are based on a 3D face model built on European-descent male and female young adults. When the facial caricature technique is performed, they found that young faces are more attractive but less distinctive than old faces. When the degree of facial caricature increases, the distinctiveness increases while the attractiveness decreases linearly and, at the same time, faces are looking older. Langlois and Roggman [59] support the "Averageness Hypothesis" of attractiveness, which indicates that an average face is the most attractive face. Braun et al. [60] believe that the merging of several faces can remove unpleasant asymmetries, irregularities, wrinkles, and pimples so that the skin looks perfectly smooth, clear, and younger. Moreover, babyfaceness makes faces look younger and can enhance attractiveness. Alley [61] also found that the morphological changes of face aging using shape transformations can induce the decrease of perceived facial cuteness. Based on the facial aging perception, O'Toole et al. [62], [63] applied a facial caricature algorithm to a 3D face model for face aging synthesis. They support the conclusion that an average face is younger and more attractive than an individual face. They suggest that the distinctiveness of an individual face, defined as its distance from the average face in a 3D face space, embodies the information highly related to face aging. So, in perception and aesthetics, facial aging is often explained and implemented by the concepts of attractiveness and distinctiveness with explicit caricatures or transformations from the empirical knowledge of biology anthropometry (face growth, development, and aging forms).

Fig. 3. Face aging by shape synthesis models. (a) Cardioidal strain transformation model, originally shown in [108]. (b) 3D shape model, originally shown in [62]. (c) Anthropometric measurements of human face, originally shown in [54]. (d) Craniofacial growth model for age progression, originally shown in [53], by considering the models in both (a) and (c).

3 AGE SYNTHESIS ON FACES

3.1 Face Modeling

Age synthesis, also called age progression, is often implemented by first building a generic face model. Face modeling has been prevalent for a long time in both the computer graphics and computer vision fields. The pioneering research of computer-generated face model can be traced back to Parke's work [64] in 1972. A 3D mesh model is built to generate cartoon faces. Facial expression animation is synthesized by analyzing a typical pair of real face photographs. Thereafter, a large number of face models—3D or 2D, photorealistic, or nonphotorealistic—have been developed and reported for different purposes of applications [65], [66], [67], [68], [69]. Typically, there are two common categories of face models: empirical knowledge-based (with subclasses of geometry based and image based) and statistical learning-based (appearance-based). We discuss the three subcategories as follows.

3.1.1 Geometry-Based Model

This kind of model generates automatic facial animations with generic geometric mesh [70], dynamic skin-muscle deformation [71], active contours [72], or anthropometric growth [53]. They are mainly designed for nonphotorealistic rendering. It digitizes facial mesh through geometric units representing face muscles, tissues, and skin in either 2D or 3D. The animation effect and motion smoothness are achieved by vertex displacements and interpolation, such as cosine and bilinear functions. A prevalent geometry-based face animation model is the caricature generator [73], [74] in the 1980s, which is formalized as the process of exaggerating differences (distance) from a given face to a mean face (the average face of a set of samples). Shape caricaturing and anticaricaturing can also be applied to improve the animation effectiveness [75]. As a widely used shape geometry analysis tool developed by Cootes et al., the Active Shape Model (ASM) [76] is a statistical measure of face shapes represented by a set of points in the training set. It iteratively deforms to match the model points to a face in a given image. In each iteration, a local search is made around the current position of each point to update with a neighbor point that better matches the model. Some more systematic developments related to this modeling technique evolve the popular Facial Action Coding System (FACS)

[77] and MPEG-4 Facial Animation Parameter (FAP) [78] that are used as standards. Since this modeling technique can generate facial animations by parameterization [79], many real-world avatar systems [80], [81], [82] were developed. More discussions can be retrieved from the survey [83] by Deng and Noh in 2007. Some prevalent applications of geometry-based model involve cartoon face and caricature generation by contour learning [72], [84]. In these cases, face shape and facial feature components are captured and processed to generate cartoon effects. Individual traits are exaggerated for caricatures.

3.1.2 Image-Based Model

Image-based models focus on generating photorealistic face images from other images rather than from geometric primitives. A heuristic technique is to generate texture details on the given face images to simulate human traits, e.g., face skin rerendering with creases and aging wrinkles [35], [85]. This technique is simple to implement but too empirical to be generalized for photorealistic rendering. Another way to improve the photorealistic effect is to capture and clone face attributes from a source image to a target image, e.g., cloning facial expressions [86] or aging skin details [87]. An interesting idea, so-called Merging Ratio Images (MRI) [88], is to extend the Ratio Image (RI) concept to multiple face attributes and merge expression RI [89], illumination RI (quotient image) [90], and aging RI [91] for photorealistic face rendering. In a reverse way, given a set of sample images with different facial attributes for the same individual, the natural (neutral) image of the individual's appearance can be synthesized by image-based rendering [92]. Image morphing [93] technique is often combined with the image-based models for smooth lighting morphing [94], view morphing [95], and expression morphing [96].

3.1.3 Appearance-Based Model

Appearance-based models consider both shape and texture rendering to achieve highly realistic results. The shape and texture are both vectorized for image representation [97]. Instead of heavily using empirical knowledge like the previous two models, this kind of model usually uses statistical learning to build the model. Usually, a large size face database is associated with a generic model considering linear object class [98], [99]. One of the most popular 3D

models is the Morphable Model (MM) [25] developed by Blanz and Vetter in 1999. Constructed by hundreds of 3D face samples, the morphable model can transform an individual face by caricaturing shape and texture vectors in terms of different attributes. Variations of face attributes are diversified. Some attributes that are invariant for each individual—gender, fullness, smile, frown, weight, and hooked nose—can be isolated and synthesized by the model. Because of the difficulties in 3D database collections and model correspondence, some alternative 2D models have also been reported. Cootes et al. presented the Active Appearance Model (AAM) [100] in 1998 that contains a statistical model of the shape and gray-level appearance of the face. Fu and Zheng presented the Merging Face (M-Face) [30] model in 2006, the integration of shape caricaturing and texture MRI-mapping, to generate emotional facial attributes in rotated views. M-Face is a 2D parameter-driven model in which 3D modeling is avoided without weakening photorealistic effects. Appearance-based face models can also be created by using a set of photographs corresponding to different views and considering both shape and texture [69]. By manually marking the key feature points on the face images, the personalized face model is finalized by interactively fitting a 3D generic model to the given set of images.

3.2 Age Synthesis Algorithms

Based on different face models, age synthesis algorithms can be applied to rerender a face image aesthetically with natural aging and rejuvenating effects. Three popular synthesis algorithms are discussed as follows.

3.2.1 Explicit Data-Driven Synthesis

Based on the particular face model, shape, texture, or appearance can be synthesized effectively. The explicit data-driven synthesis focuses on the shape analysis, which is more related to craniofacial growth in age progression [19]. As skin textures do not change too much for young faces, the distinct shape changes during craniofacial growth are more prone to be observed and modeled for the purposes of appearance prediction and face recognition/verification across age progression [101], [102]. Shape models are usually based on a set of predefined points, lines, or grids to represent landmarks and positions of key facial features. Model fitting is implemented by constraining and shifting the positions of the landmarks through scaling, rotation, and translation to best match a given face image. A more advanced shape model may be constructed by learning the common statistics among a group of 3D laser scan data, which need dense correspondences due to the high resolution.

D'Arcy Thompson was the known pioneer because of the theory of geometrical transformations for species development and evolution in his book—"*On Growth and Form*" [103]—published in 1917 [104]. He believed that the differences between the forms of related species can be represented geometrically. A representative example reproduced in Arthur's paper [104] shows how to use geometrical transformations to implement the morphogenesis among a human skull, a chimpanzee skull, and a baboon skull. The three outlines of the skull shape can be mutually turned from one into another by warping the original grid

to another corresponding grid. This phenomenon implies a possible explanation for how evolution works.

D'Arcy's pioneer studies inspired some existing work on craniofacial growth for face aging, especially since computer is well developed and has been widely used since the 1970s [105]. Pittenger and Shaw presented a theory of event perception [106], [107] in which craniofacial growth is modeled to be a viscal-elastic event. The perception of these events is analyzed by using the concepts of transformational and structural invariants. Craniofacial shear and strain transformations are applied to the face shape profiles for modeling the facial growth. They concluded that for relative age judgments, the craniofacial strain transformation is more reliable than the craniofacial shear transformation. Todd et al. [108] presented the "revised" cardioidal strain transformation for craniofacial growth and used a fluid-filled spherical object to remodel faces. Mark et al. [109] found that any changes during a specific craniofacial growth preserve the same geometric invariants. So, a growth-related transformation can be uniquely constrained by the geometric invariants. As an extension to the existing 2D models, he developed a 3D craniofacial growth model [110] to synthesize younger faces for perception purposes. Mark et al. [111] further found that the cardioidal strain for growth transformations is more effective when the structural contours become more curved and less angular. Bruce et al. [112] extended Mark's experiments to further find that cardioidal strain can be perceived in 3D craniofacial growth. Note that the cardioidal strain is effective to simulate the craniofacial growth of young faces. Such aging changes occur little during adulthood and thereafter [113]. The "revised" cardioidal strain transformation model described in [108] is shown in Fig. 3a.

O'Toole et al. [62], [63] applied a caricaturing algorithm to a 3D face model for 3D age synthesis, as shown in Fig. 3b. They calculated the average face of many sample faces captured by 3D laser scanning and dense correspondence. The distance from each sample face to the average face is computed by warping each sample face to the average face shape. Increasing the distance (caricature) may make the skull become more distinct so that the face appear older, while decreasing the distance (antiaricature) may make the face appear younger. Fig. 4a shows a 3D face aging result using the shape caricaturing model. Mukaida et al. [114] proposed using Principal Component Analysis (PCA) [115] for age-related shape feature extraction. They found that the principal component of PCA is highly related to face age variation, which can be used to synthesize novel age-manipulated face images. Ramanathan and Chellappa [53] proposed a craniofacial growth model, as shown in Fig. 3d, for age progression of young faces under 18 years of age. Their work was originally from Todd's "revised" cardioidal strain transformation for craniofacial growth [108]. But they also considered the anthropometric evidences—measurements over facial landmarks collected by Farkas [54], as shown in Fig. 3c—to estimate the local growth parameters for craniofacial changes across facial growth. This model provides more realistic shape synthesis results and reasonable age progression models in a psychophysical point of view. They also investigated the face similarity measurement across age
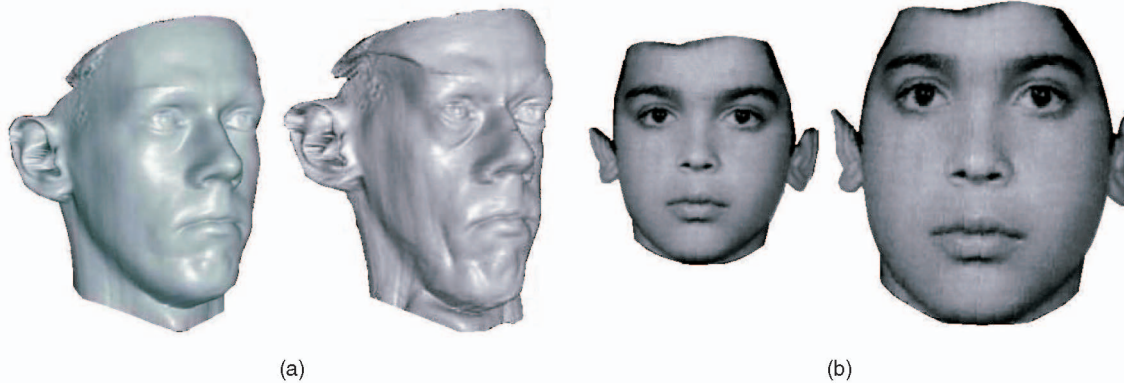
Fig. 4. Face aging results of shape models. (a) 3D shape aging results, originally shown in [62]. (b) Age progression for a young face shape, originally shown in [53].

progression and face verification across age progression by Bayesian age difference classifiers (intrapersonal and extrapersonal spaces built on age difference images from the passport images) [23]. Fig. 4b shows the age progression result of a young face using the craniofacial growth model.

3.2.2 Explicit Mechanical Synthesis

The explicit mechanical synthesis focuses on the texture analysis, which is more related to skin aging, the most distinct facial changes after adulthood. During skin aging, wrinkles emerge and become more pronounced due to the nature of skin and muscle contraction. This technique is usually developed using image-based rendering for the purpose of photorealistic appearance prediction across age progression. Given a face image for rendering, this technique either constructs a skin model for aging synthesis or captures aging patterns from example older faces and clones to a young face. Both ways need to align the faces or calculate the correspondence between image pairs.

Wu et al. [116], [117], [118] synthesized dynamic wrinkles using a 3D dynamic model for facial aging and animation. The facial skin is modeled according to its anatomy structure, which is composed of three layers—muscle, connective tissue, and skin. In the model shown in Fig. 6a, the connective tissue is simulated as springs that can connect and constrain the two layers. Muscle masks, designed as B-spline patches, are constructed to control the facial expression and movement. Wrinkles caused by muscle contraction can be simulated by an elastic process combined with plastic-visco units. This synthesis can be automatically adapted to different individual face shape and tissue structures. Boissieux et al. [35] extended Wu's work by considering skin as a volumetric substance, as shown in Fig. 6b. Their model still has multiple layers to model the skin structures, such as the epidermis, dermis, and hypodermis. For simplicity, only two layers to model the epidermis and the other underneath tissues are considered and approximated to be linear, isotropic, and elastic materials. The elastic deformation is computed by a finite element analysis. In addition, Boissieux et al. also presented an image-based model using eight generic aging masks, as shown in Fig. 6c. Extracted from real photos, the masks are person-specific and contain different types of wrinkles corresponding to gender, face shape, and facial expression. Aging simulation is performed by scaling the wrinkle intensities linearly in terms of age and mapping the corresponding wrinkle mask to the specific face. Shan et al. [87] presented an image-based method to transfer geometric details between two object surfaces. They found that geometric details can be captured without knowing the surface reflectance. After alignment, the geometric details can be cloned to render other surfaces. Both face aging and rejuvenating can be synthesized using this method. The first and third columns in Fig. 5a show the original images of two people with different ages. The second and fourth columns show the two rendering results for face aging and rejuvenating (exchange the age between the two faces). Fu and Zheng [88], [91] used global ratio image to capture the aging skin details. A series of aging patterns are generated by morphing a real ratio image starting from the average face of a set of young faces. Aging

Fig. 5. Face aging and rejuvenating results. (a) Image-based surface detail transfer results, originally shown in [87]. (b) M-Face results, originally shown in [30].
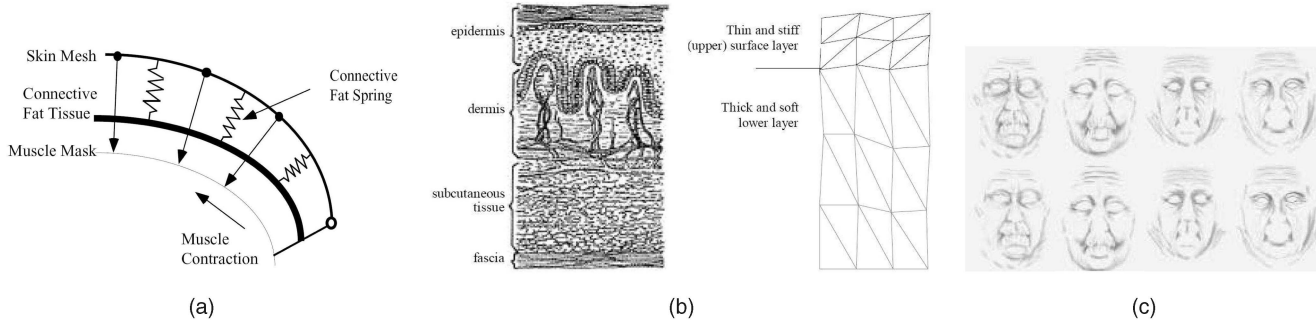
Fig. 6. Face aging by texture synthesis models. (a) Plastic-visco-elastic model, originally shown in [116]. (b) Layered structure of the skin model, originally shown in [35]. (c) Eight generic masks, originally shown in [35].

animation can be generated by multiplying a destination image and the sequential ratio images. By analyzing the properties of pixel distributions in local areas, Mukaida and Ando [119] presented extracting and separating facial wrinkles and spots for skin aging synthesis. Gandhi reported a real-world age synthesis system [120]. The associated aging database contains 800 face images collected from the Internet with age range from 15 to 99. The system starts by localizing facial features and normalizing facial attributes. After correspondence, training images are grouped in terms of human age to form signature faces showing significant aging differences. The age synthesis, mainly texture synthesis, is carried out by considering both signature images and regression-based age prediction. Jiang and Wang [121] presented synthesizing aging or rejuvenating faces from superresolution. The database is first categorized into different age groups. In each group, two tensor models with three modes (pixels, identities, and ages) are built on original images and downsampled images, respectively. Based on the assumption that the variances in the low-resolution images with different age can be negligible, the aging or rejuvenating synthesis is made by projecting the identity parameters from the low-resolution tensor space to the corresponding high-resolution tensor space. This technique mainly works on adult face images. Since there is no face correspondence implemented, the synthesis results still need improvements.

3.2.3 Implicit Statistical Synthesis

The implicit statistical synthesis focuses on the appearance analysis, which considers shape and texture synthesis simultaneously and often uses statistical methods. This needs to collect a database that contains a large number of face images with a broad range of ages. In this case, each face image is considered as a high-dimensional point in the age space. So, the age synthesis can be animated by tuning the distances between faces with different ages or the model parameters controlling different appearance variations. Person-specific aging can also be realized to synthesize different aging styles for different individuals.

Lanitis et al. [122] used AAM [100], [123] to build aging functions of young faces under 30 years of age, in which PCA is applied to extract the shape and texture variations from a set of training examples. The PCA coefficients for the linear reconstruction of training samples are considered as model parameters, which control different types of appearance variations. Parameter 6 in this work accounts

for the aging. This aging model can be used for age normalization and further improves the face recognition performance. Fig. 7 shows the AAM example and the aging appearance synthesis resulted by Lanitis's method. Scandrett et al. [124] used a similar statistical model with more theoretical extensions to build the aging space for human appearance. Person-specific aging trajectories are calculated in the model space to synthesize face aging effects according to historical and familial correlations. Patterson et al. [125] reported an initial comparison of these synthetic face aging results using the AAM model with age progression drawn by a forensic artist.

Since the RGB color of face images contains a variety of aging cues, such as skin pigmentation, wrinkles, and hair color, Burt and Perrett [113] presented an algorithm for European-descent face aging and rejuvenating by considering both shape and color information. The basic idea is to extract the aging patterns by capturing the appearance differences between the composite (average) faces from different age groups after intergroup correspondence. These differences are scaled and added to a new face image for shape and texture caricaturing with perceived aging or rejuvenating effect.

Fu and Zheng [30] presented the M-Face framework for appearance-based photorealistic facial modeling. The Aging Ratio Image (ARI) is captured from their aging database for texture rendering. The shape caricaturing is integrated for associated shape deformation. Derived from an average young face, the caricatured shape is warped by exaggerating individual distinctiveness, while the rerendered texture multiplies the ARI during the caricaturing. The proposed
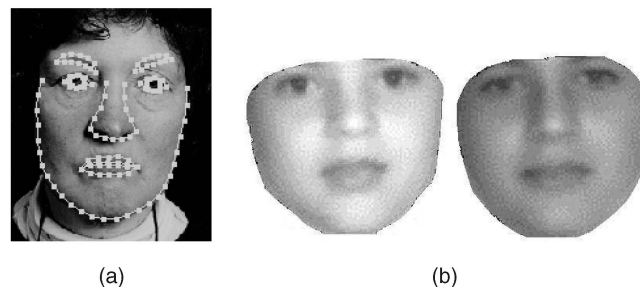

Fig. 7. Face aging using active appearance model and principal component analysis. (a) Active appearance model, originally shown in [100], [123]. (b) Aging appearance synthesis result, originally shown in [122].
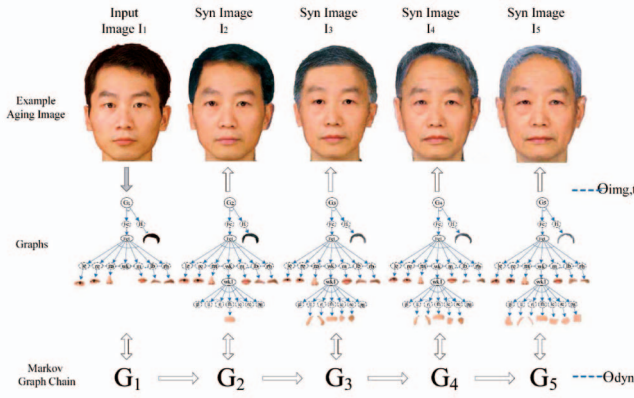
Fig. 8. Multilayer dynamic face aging model and results, originally shown in [42].

2D parameter-driven model can manipulate chronological aging or rejuvenating with satisfying photorealistic effects, avoiding tedious 3D face modeling. Fig. 5b shows a rendering example from [30] by M-Face.

Ramanathan and Chellappa [126] proposed an appearance-based adult aging framework. The shape aging is manipulated by a physically based parametric muscle model, while the texture aging is manipulated by an image-gradient-based wrinkle extraction function. This model can predict and simulate facial aging in two cases: weight gain and weight loss. The wrinkle synthesis module can generate different effects such as subtle, moderate, and strong.

Suo et al. [42] presented a dynamic face aging model with multiresolution and multilayer image representations. Associated with a layered And-Or graph, all 50,000 training face images are decomposed into different parts. The general aging effects are learned from global hair style and shape changes, facial components deformations, and facial wrinkles geography. A dynamic Markov process model is built on the graph structures of the training data. The graph structures over different age groups are finally sampled in terms of the dynamic model to synthesize new aging faces. Fig. 8 shows this model and some aging synthesis effects which exhibit highly photorealistic results.

Scherbaum et al. [13] presented synthesizing 3D aging faces via the 3D morphable model, which is applicable to images at any given pose and illumination. Using a nonlinear SVR function, they calculated the individual aging trajectories for given faces from a database of 3D scans of 238 teenagers and 200 adults. The teenager faces cover an age range from 8 to 16 years, with 125 males and 113 females. The application of the system is to help police find missing children.

3.2.4 Summary

If there is a large training database available which covers enough variations, the implicit statistical synthesis is more applicable and effective. When the training database is small, the explicit data-driven synthesis is more effective for craniofacial growth synthesis in age progression and the explicit mechanical synthesis is more effective for aging synthesis after adulthood. Fig. 9 shows the categorical diagram of the relationship between age synthesis algorithms.
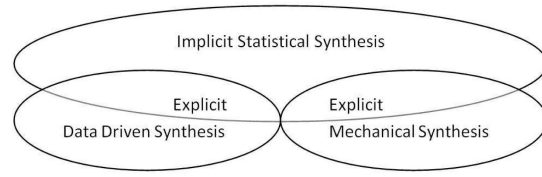

Fig. 9. Categorical diagram of age synthesis algorithms.

4 AGE ESTIMATION VIA FACES

Compared with the large body of research on aging synthesis presented above, there are relatively few publications specifically on age estimation. This situation is generally caused by three reasons: 1) Age estimation is not a standard classification problem. Derived from different application scenarios, it can be taken as either a multiclass classification problem or a regression problem. 2) A large aging database is often hard to collect, especially the chronometrical image series for an individual. 3) As mentioned in [37], [23], [36], the age progression displayed on faces is uncontrollable and personalized. Such special characteristics of aging variation cannot be captured accurately due to the prolific and diversified information conveyed by human faces. The existing age estimation systems using face images typically consist of two concatenated modules: age image representation and age estimation techniques.

4.1 Age Image Representation

4.1.1 Anthropometric Models

The earliest paper published in the area of age classification from facial images was the work by Kwon and Lobo [127], [128]. The main idea of this approach is to consult studies in cranio-facial development theory [32]. The theory of cranio-facial research uses the mathematical model—"revised" cardioidal strain transformation model [32]—to describe the growth of a person's head from infancy to adulthood: $\theta' = \theta, R' = R(1 + k(1 - \cos\theta))$, where θ is the angle formed from the vertical 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 [32]. Farkas [54] provided a comprehensive overview of face anthropometry. Face anthropometry is the science of measuring sizes and proportions on human faces. Farkas defined face anthropometry in terms of measurements taken from 57 landmarks or fiducial points on human faces over different ages, e.g., from infancy to adulthood. For age growth characterization, people usually utilize the distance ratios measured from facial landmarks, instead of directly using the mathematical models given in [32]. There are two reasons that people do not use the mathematical formulation for age estimation: 1) The mathematical model cannot characterize the head profile in a natural way, especially when the ages are close to adults [53], and 2) the head profile is difficult to measure from 2D face images.

In age image characterization, Kwon and Lobo [127], [128] computed six ratios of distances on frontal face images and used them to separate babies from adults. The experiment was performed on a small database of 47 faces. The authors did not report the overall performance on this

small database, but showed the results from using each ratio. Ramanathan and Chellappa [53] used eight ratios of distance measures to model age progression in young faces, e.g., 0 to 18 years. The purpose is to predict an individual's appearance across age and to perform face recognition across age progression. Their experiment was performed on a database of 233 images of 109 individuals, partially collected from the FG-NET aging database [129] and partially from their own collection. They reported a face recognition improvement by using the age progression transformation, for example, improved from 8 percent (without age prediction) to 15 percent (with age transformation) for the rank 1 face recognition. Gunay and Nabiyev [130] also proposed a variant of such a method to detect anthropometric features for age estimation.

Age estimation approaches based on the anthropometry model can only deal with young ages since the human head shape does not change too much in its adult period. Kwon and Lobo [127], [128] computed wrinkles from face images to separate young adults from senior adults. The wrinkles were computed in several regions, such as on the forehead, next to the eyes, and near the cheek bones. The presence of wrinkles in a region is based on the detection of curves in that region.

In sum, the anthropometry model might be useful for young ages, but is not appropriate for adults. In practice, only frontal faces can be used to measure facial geometries because the ratios of distances are computed from 2D face images which are sensitive to head pose. So far, there are no reported results on a large database for human age estimation using the anthropometry model. Finally, the anthropometric model-based representations only consider the facial geometry but not any texture information, and the anthropometric measurements from Farkas [54] are mainly for European-descent faces rather than for people from other origins.

4.1.2 Active Appearance Models

The active appearance model is a statistical face model proposed initially in [123] for coding face images. Given a set of training face images, a statistical shape model and an intensity model are learned separately, based on the principal component analysis. The AAMs were used successfully for face encoding. Lanitis et al. [122] extended the AAMs for face aging by proposing an aging function, $age = \mathbf{f}(\mathbf{b})$, to explain the variation in age. In the aging function, age is the actual age of an individual in a face image, $\mathbf{b}$ is a vector containing 50 raw model parameters learned from the AAMs, and $\mathbf{f}$ is the aging function. The aging function defines the relationship between the age of individuals and the parametric description of the face images.

The experiments were performed on a database of 500 face images of 60 subjects among which about 45 individuals have age progression (the same individual with different age images). The authors mainly focused on small age variations such as from 1 to 30 years old. They showed that the age simulation can improve face recognition performance from 63 to 71 percent or from 51 to 66 percent (when the training and test subsets were switched).

Lanitis et al. [17] also tried different classifiers for age estimation based on their age image representation, especially the quadratic aging function [122]. In comparison with the anthropometry model-based approaches [128], [127], [53], the AAMs can deal with any ages in general, rather than just young ages. In addition, the AAMs-based approaches to age image representation consider both shape and texture rather than just the facial geometry as in the anthropometric model-based methods. These approaches are applicable to the precise age estimation since each test image will be labeled with a specific age value chosen from a continuous range. The further improvements on these aging pattern representation methods are: 1) to provide evidence that the relation between face and age can be essentially represented by a quadratic function, 2) to deal with outliers in the age labeling space, and 3) to deal with high-dimensional parameters.

4.1.3 Aging Pattern Subspace

Instead of dealing with each aging face image separately, a sequence of an individual's aging face images might be used all together to model the aging process. This idea was explored by Geng et al. [131], [37], which is called AGing pattErn Subspace (AGES). An aging pattern is defined as a sequence of personal face images, coming from the same person, sorted in the temporal order [131]. If the face images of all ages are available for an individual, the corresponding aging pattern is called a complete aging pattern; otherwise, it is called an incomplete aging pattern. The AGES method presented in [131], [37] can synthesize the missing ages by using an EM-like iterative learning algorithm.

This AGES method works in two stages: the learning stage and the age estimation stage. In learning of the aging pattern subspace, the PCA technique is used to obtain a subspace representation. The difference from the standard PCA approach is that there are possibly some missing age images for each aging pattern. The proposed EM-like iterative learning method is used to minimize the reconstruction error characterized by the difference between the available age images and the reconstructed face images. The initial values for missing faces are set by the average of available face images. Then, the eigenvectors of the covariance matrix of all face images and the means can be computed. Given the eigenvectors and mean face, the reconstruction of the faces can be computed. The process iterates until the reconstruction error is small enough. In the age estimation stage, the test face image needs to find an aging pattern suitable for it and a proper age position in that aging pattern. The age position is returned as the estimate of the age of the test face image. To do this, the test face is verified at every possible position in the aging pattern and the one with the minimum reconstruction error is selected. It is observed that when the test face is placed at an improper position in the aging pattern, the reconstructed faces become ghost-like twisted faces [131], [37].

In the aging pattern subspace approach, each face image is first encoded by using the AAMs method [123], similar to that of Lanitis et al. [17]. But the emphasis of the AGES method is to use face images of the same individual at different ages altogether to represent the aging pattern. So, we categorize this approach separately from AAMs-based method such as [17]. In terms of utilizing the AAMs for encoding face images, Geng et al. [131], [37] used 200 AAMs

features to encode each face image, while only 50 features were used in [17].

The experiment for evaluating the AGES method [131], [37] was performed on the FG-NET aging database [129]. The Mean Absolute Error (MAE) was reported as 6.77 years.

In practical use of the AGES method, a problem is that in order to estimate the age of an input face image, the AGES method assumes that there exist face images of the same individual but at different ages or at least a similar aging pattern for that face image in the training database. This assumption might not be satisfied for some aging databases, such as the Yamaha Gender and Age (YGA) database [36]. And also, it is not easy to collect a large database containing face images of the same individual at many different ages with close imaging conditions. Another problem of the AGES method is that the AAM face representation might not encode facial wrinkles well for senior people because the AAM method only encodes the image intensities without using any spatial neighborhood to calculate texture patterns. Intensities of single pixels usually cannot characterize local texture information. In order to represent the facial wrinkles for elder adults, the texture patterns at local regions need to be considered.

4.1.4 Age Manifold

Instead of learning a specific aging pattern for each individual, such as that in [131], [37], a common aging trend or pattern can be learned from many individuals at different ages. For each age, many personal faces might be adopted to represent that age. For each individual, he or she might have face images at a single age or at a range of ages. This kind of aging pattern learning makes the task of face aging representation very flexible in comparison with the age pattern subspace approach [131], [37]. One possible way to learn the common aging pattern is the age manifold [40], which utilizes a manifold embedding [132] technique to learn the low-dimensional aging trend from many face images at each age. There is no requirement to present many different ages for each individual. In an extreme case, each individual may only have images at a single age in the database. This flexibility makes it easier to collect a large aging face database. Scherbaum et al. [13] also presented a method to perform statistical age estimation with manifold learning over the 3D morphable model. The manifolds are isosurfaces of the nonlinear SVR function, and aging is performed by finding a curved trajectory orthogonal to the isosurfaces.

Suppose the image space is sampled by a set of face images $\mathcal{X} = \{\mathbf{x}_i : \mathbf{x}_i \in \mathbb{R}^D\}_{i=1}^n$ of m subjects with data dimension D in the order of subject ages. A ground truth set $\mathcal{L} = \{l_i : l_i \in \mathbb{N}\}_{i=1}^n$ associated with the images provides the age labeling. The goal of age manifold learning is to find a low-dimensional representation in the embedded subspace, capturing the intrinsic data distribution and geometric structure, as well as its representation $\mathcal{Y} = \{\mathbf{y}_i : \mathbf{y}_i \in \mathbb{R}^d\}_{i=1}^n$ with $d \ll D$. Thus, the mapping from the image space to the manifold space can be formulated as either linear or nonlinear functions, such as $\mathcal{Y} = P(\mathcal{X}, \mathcal{L})$. As suggested in [40], [36], the manifold embedding can be calculated by learning an $D \times d$ matrix $\mathbf{P}$ satisfying $\mathbf{Y} = \mathbf{P}^T \mathbf{X}$ in a linear way [133], [134] or directly learning $\mathbf{Y}$ in a nonlinear way [135], [136], where $\mathbf{Y} = [\mathbf{y}_1 \, \mathbf{y}_2 \, \cdots \, \mathbf{y}_n]$,
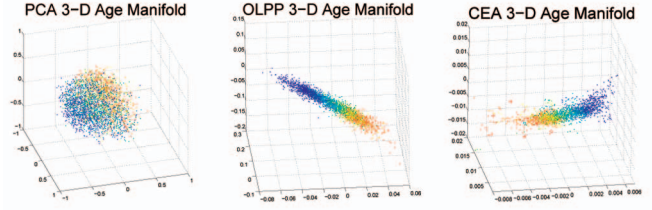

Fig. 10. Aging manifold using PCA, OLPP, and CEA, originally shown in [40], [36].

$\mathbf{X} = [\mathbf{x}_1 \, \mathbf{x}_2 \, \cdots \, \mathbf{x}_n]$, and $\mathbf{P} = [\mathbf{p}_1 \, \mathbf{p}_2 \, \cdots \, \mathbf{p}_d]$. An effective algorithm, called Orthogonal Locality Preserving Projections (OLPP) [137], is to connect the manifold learning with subspace learning. The OLPP basically produces orthogonal basis functions based on the Locality Preserving Projections (LPP) [138], [139] formulation, which searches the embedding to preserve essential manifold structures by measuring the local neighborhood distances.

Fig. 10 illustrates the aging manifold visualization results shown in [40], [36], which reveal that the age manifold can be learned using a large aging database, and demonstrated that the OLPP could represent the age manifold well, while the traditional PCA is not good for this purpose. Fu and Huang [36] also showed that Conformal Embedding Analysis (CEA) can be used to further improve the age estimation accuracy (highlighted by PhysOrg.com [3]). Guo et al. [31], [16] provided more evidence to show that the age manifold learned with the OLPP method can represent the age images well and the MAE was reported about 5.30 years on a large aging database when a so-called Locally Adjusted Robust Regression (LARR) method was used. The MAE could achieve 5.07 years on the FG-NET database. Yan et al. [140] proposed Synchronized Submanifold Embedding (SSE) for person-independent age image representation, which considers both age labels and subject identities to promote the generalization capability on age estimation. It achieves the MAE of 5.21 years on the FG-NET database.

Age manifold here is actually connected to the subspace analysis for aging patterns. It finds an embedded low-dimensional space when each age has many people's face images. A nice property of using the OLPP or CEA method for age manifold learning is that the age labels can be incorporated into the embedding process in a "supervised" manner. While the traditional PCA method can only work in an "unsupervised" manner without using any age information in learning the subspace representation. In comparison with the aging pattern subspace [131], [37], the age manifold does not require learning a specific aging pattern for each individual. All ages of different individuals could be used altogether. The only requirement for the age manifold representation is that the size of the training data should be large enough in order to learn the embedded manifold with statistical sufficiency.

4.1.5 Appearance Models

Aging-related facial feature extraction is more focused by the appearance model. Both global and local features were used in existing age estimation systems. Hayashi et al. [44], [45] considered both texture (wrinkle) and shape (geometry) features to characterize each face images. A semantic-level

description of the face was also used to characterize facial features. The system estimates human ages through a multiple-group classification scheme with five year intervals. It works on a small Japanese face database with 300 subjects covering 15-64 years. Gender estimation was also implemented to improve the age estimation since the aging patterns are different for the male and the female. An extension work was proposed by Fujiwara and Koshimizu [43] using the system. Fukai and Takimoto [49] presented using FFT to extract feature spectrum from facial appearance and using Genetic Algorithm (GA) for feature selection. This kind of feature can hardly capture the texture details.

The effective texture descriptor, Local Binary Patterns (LBP) [141], has been used for appearance feature extraction in an automatic age estimation system [142]. It achieves 80 percent accuracy on the FERET database [143] with nearest neighbor classification, and 80-90 percent on the FERET and PIE databases with AdaBoost [144]. The Gabor feature [145] has also been tried on the age estimation task [146], which has been demonstrated to be more effective than LBP.

Since local variations in the face appearance during aging are often evident, Yan et al. [147], [148] proposed using Spatially Flexible Patch (SFP) as the feature descriptor. Since SFP considers local patches and their position information, face images with small misalignment, occlusion, and head pose variations can still be handled effectively. Moreover, it can also enrich the discriminating characteristic of the feature set when insufficient samples are provided. Modeled by a Gaussian Mixture Model (GMM), their system can achieve the MAE of 4.95 years on the FG-NET aging database, and MAEs of 4.94 and 4.38 years for female and male on the YGA database, respectively.

Suo et al. [149] proposed designing graphical facial features—topology, geometry, photometry, and configuration—based on their early developed multiresolution hierarchical face model [42]. Basically, within the hierarchical model, they applied specific filters to different parameters at different levels for feature extraction. Their system can achieve the MAE of 5.974 years on the FG-NET aging database using Multilayer Perceptron (MLP), and MAE of 4.68 years (with error tolerance of 10 years estimation rate attains 91.6 percent) on their own database, respectively.

Guo et al. [150] proposed using the Biologically Inspired Features (BIFs) [151], [152] for age estimation via faces. The basic idea is derived from a feedforward model of the primate visual object recognition pathway—the "HMAX" model. It contains alternating layers called Simple (S) and Complex (C) cell units creating increasing complexity as the layers progress from the primary Visual cortex (V1) to Inferior Temporal (IT) cortex [151], [150]. The first layer S1 is created with a Gabor filtering on the input image. The second layer C1 is obtained with a "MAX" operation on S1. This kind of feature has the advantage of handling small translations, rotations, and scale changes, which, if effective, captures aging patterns. The BIF feature combined with Support Vector Machines (SVMs) [153], [154] can achieve the MAE of 4.77 years on the FG-NET aging database, and MAEs of 3.91 and 3.47 years for female and male on the YGA database,

respectively. Considering both age and gender estimation in an automatic framework [155], the BIF+Age Manifold feature combined with SVM can achieve MAEs of 2.61 and 2.58 years for female and male on the YGA database, respectively. These results demonstrate the superior performance of the BIF for the task of age estimation.

4.2   Age Estimation Algorithms

Given an aging feature representation, the next step is to estimate the age. Age estimation can be viewed as a pattern recognition task, but it is a "special" one. Each age label can be viewed as a class; thus, age estimation can be considered as a classification problem. On the other hand, age numbers are a set of sequential values, e.g., 0, 1, 2,...; thus, age estimation can also be considered as a regression problem. So, a question might be asked: Is it more effective to consider age estimation as a classification or a regression problem [31], [16]? We will first review different existing approaches on both sides and then answer the question in our opinions.

4.2.1   Classification

Lanitis et al. [17] evaluated the performance of different classifiers for age estimation, including the nearest neighbor classifier, the Artificial Neural Networks (ANN), and a quadratic function classifier. The face images are represented by the AAM method. The quadratic function is actually a regression function to relate the face representations to age labels; however, the authors called it a quadratic function classifier [17]. From experiments on a small database containing 400 images at ages ranging from 0 to 35 years, it was reported that the quadratic function classifier can reach 5.04 years of MAE, which is slightly lower than the nearest neighbor classifier, but higher than the ANN and an unsupervised approach—the Self-Organizing Map (SOM) method. The authors also suggested some extensions, for example, first do clustering, and then to refine the age estimation in a hierarchical fashion. The extended methods reduce the error rates a little bit, but the results were reported on a small database with a small age range. The authors also compared the age estimation by a computer with human observers. A random subset of 32 images was selected from the 400 images for testing the accuracy of humans in age estimation. With 20 observers in the test, an average error is achieved by 3.64 years.

Formulated as an 11-class classification problem for the WIT-DB database (separate male and female), Ueki et al. [50] built 11 Gaussian models in a low-dimensional 2DLDA+LDA feature space using the EM Algorithm. The age group classification is determined by fitting the test image to each Gaussian model and comparing the likelihoods. For the five-year range age group classification, their system achieves accuracies of about 50 percent for male and 43 percent for female. For 10-year range age group classification, it achieves accuracies of about 72 percent for male and 63 percent for female. For 15-year range age group classification, it achieves accuracies of about 82 percent for male and 74 percent for female. The SVM was applied to age estimation by Guo et al. [31], [16] on a large YGA database with 8,000 images. The MAEs are 5.55 and 5.52 years for females and males, respectively. The MAE

is 7.16 for the FG-NET aging database. Kanno et al. [156] presented using artificial ANN for the four-class age group classification, which achieves 80 percent accuracy on 110 young male faces. Supervised SOM was also reported in the age estimation system [49].

4.2.2 Regression

Considering age estimation to be a regression problem, Lanitis et al. [122] investigated three formulations for the aging function: linear, quadratic, and cubic, respectively, with 50 raw model parameters. The optimal model parameters are learned from training face images of different ages based on a genetic algorithm. Inspired by Lanitis's work, Fu et al. [40], [36] used a multiple linear regression [157] function to fit the CEA aging manifold, which achieves significant improvements over some existing techniques. The Support Vector Regression (SVR) method was applied by Guo et al. [31], [16] on the OLPP age manifold for age estimation. The MAEs are 7.00 and 7.47 years on the YGA database for females and males, respectively, and 5.16 for the FG-NET aging database. Yan et al. [39] used a Semidefinite Programming (SDP) formulation to solve the regression problem for age estimation, in which the regressor is learned from uncertain nonnegative labels. The MAEs are 9.79 and 10.36 years for the females and males, respectively, on the YGA database. The authors also showed that the SDP formulation for age regression gives much better results than the quadratic regression function and the multilayer perceptrons. The SDP is an effective tool but computationally very expensive [39]. When the size of the training set is large, the solution to SDP may be difficult to achieve. An Expectation-Maximization (EM) algorithm was used to solve the regression problem and speed up the optimization process [38]. The MAEs are reported as 6.95 years for both female and male on the YGA database and 5.33 for the FG-NET aging database. As mentioned above, when SFP representation is modeled by GMM, the age regression can be further improved with patch-kernel measurement [147], [148]. Zhou et al. [158] presented the generalized Image-Based Regression (IBR) aimed at multiple-output settings. A boosting scheme is used to select features from redundant Haar-like feature set. The proposed training algorithm can also significantly reduce the computational cost. The IBR can achieve 5.81 MAE of a fivefold cross-validation test on the FG-NET aging database. Suo et al. [149] compared Age-group-specific Linear Regression (ALR), MLP, SVR, and logistic regression (multiclass Adaboost [159]) on FG-NET and their own databases, and finally, achieved the best performance with MLP in their experiments. Ni et al. [41] proposed the Robust Multi-instance Regression (RMIR), which can deal with label outliers, when building a universal age estimator based on a large Web-collected database. This system can achieve reasonable age estimation accuracy with automatic online training. Metric learning has also been considered to discover the latent semantic similarity between age data, which can potentially improve the regression results [160].

4.2.3 A Hybrid Approach

From the above, the age estimation problem can be viewed as either a classification or regression problem. Which is better for age estimation: classification or regression? To get
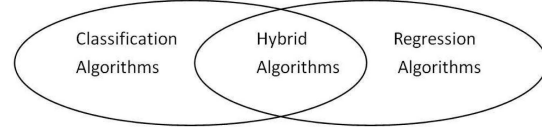

Fig. 11. Categorical diagram of age estimation algorithms.

an answer, one might choose some representative classifiers and regressors and compare them on the same databases to see the performance difference. Guo et al. [31], [16] chose the SVM as a representative classifier and the SVR as a representative regressor, and compared their performance using the same input data. From their experiments, the SVM performs much better than the SVR on the YGA database (5.55 versus 7.00 and 5.52 versus 7.47 years for females and males, respectively), while the SVM performs much worse than the SVR on the FG-NET database (7.16 versus 5.16 years). From the experimental results, we can see that the classification-based age estimation can be much better or much worse than the regression-based approach in different cases.

A better way for robustness is to combine the classification and regression methods to take advantage of the merits from both. Based on this motivation, Guo et al. [31], [16] proposed a method called LARR. They showed that a consistently better performance can be obtained by combining a classifier and a regressor. Using the combination scheme, the MAEs can reach 5.25 and 5.30 years for female and male on the YGA database, and 5.07 years on the FG-NET aging database, respectively. The idea of the LARR is to first do a regression using all available age data, and then use the regression results to constrain classifiers with a small local search range, which improves the age estimation performance significantly. However, the LARR method [31] cannot determine the range of local search for the classifier automatically. It has to heuristically try different ranges, such as 4, 8, 16, 32, and 64, and requires the user to empirically choose a best solution among those results. To determine the local search range parameters automatically in a data-driven manner, Guo et al. [161] proposed a probabilistic approach to combine regression and classification results. The idea is to transform or interpret the regression results into probabilities using a uniform distribution, and then the probabilities from the classification result are cut off by the uniform distribution. In addition to determining the parameters automatically in a data-driven fashion, the accuracies are also improved. The MAEs are 5.11 and 5.12 years for female and male on the YGA database, and 4.97 years on the FG-NET aging database, respectively.

4.2.4 Summary

Age estimation can be approached by either a classifier or regressor since different databases and systems may be too biased or unbalanced for evaluation. But it is not purely a classification or regression problem. A promising approach to age estimation is to combine regression and classification methods as demonstrated in [31], [16], [161]. Fig. 11 shows the categorical diagram of the relationship between age estimation algorithms. It is still interesting to develop more

advanced schemes for the combination of classifiers and regressors so that the accuracy of age estimation might be improved further.

5 AGING DATABASES

Data collection is crucial for the purposes of photorealistic age synthesis and precise age estimation. However, it is extremely hard, in practice, to collect large size aging databases, especially when one wants to collect the chronometric image series from an individual. We summarize as follows some existing benchmark aging databases.

5.1 FG-NET Aging Database [129]

The FG-NET aging database is publicly available. It contains 1,002 high-resolution color or gray-scale face images of 82 multiple-race subjects with large variation of lighting, pose, and expression. The age range is from 0 to 69 years with chronological aging images available for each subject (on average, 12 images per subject). The aging features can be either a shape model with 68 key points, the AAM with 200 model parameters, or the appearance [131], [39], [38].

5.2 MORPH Database [162], [163]

The publicly available MORPH face database was collected by the Face Aging Group at the University of North Carolina at Wilmington for the purpose of face biometrics applications. This longitudinal database records individuals' metadata, such as age, gender, ethnicity, height, weight, and ancestry, which is organized into two albums. Album 1 contains 1,724 face images of 515 subjects taken between 1962 and 1998. The ages range from the average of 27.3 to maximum 68 years. There are 294 images of females and 1,430 images of males. The age span is from 46 days to 29 years. Album 2 contains more than 20,000 face images obtained from more than 4,000 individuals.

5.3 YGA Database [36], [40]

The YGA database contains 8,000 high-resolution outdoor color images of 1,600 Asian subjects, 800 females and 800 males, with ages ranging from 0 (newborn) to 93 years. Each subject has about five near frontal images at the same age and a ground truth label of his or her approximate age as an integer. The photos contain large variations in illumination, facial expression, and makeup. The faces are cropped automatically by a face detector [159] and resized to 60×60 gray-level patches.

5.4 WIT-DB Database [50]

The Waseda human-computer Interaction Technology-DataBase (WIT-DB) contains about 5,500 different Japanese subjects with 2,500 females and 3,000 males. With 1-14 images per subject, there are 12,008 female face images and 14,214 male face images in total. The faces are unoccluded frontal views with neutral expression and wide variety of lighting conditions. Faces are cropped and resized to 32×32 gray-level patches. Age labels are from 3 to 85 years, which are divided into 11 nonoverlap age groups.

5.5 AI&R Asian Face Database [30]

The AI&R Asian face database, collected for photorealistic face synthesis and rendering, contains four subdatabases:

AI&R V1.0 (Expression), AI&R V2.0 (Aging), AI&R V3.0 (View), and AI&R V4.0 (Illumination). The images are 640×480 JPEG files with 24 bits color depth. The AI&R V2.0 contains 34 frontal images of 17 individuals with age between 22 and 61 years.

5.6 Burt's Caucasian Face Database [113]

The database used in [113] contains 147 European-descent male faces, with neutral expressions, aged between 20 and 62. The images were taken in the same conditions with 531×704 resolution and 24 bits color depth. All individuals were shaven and without makeup and glasses. The facial shape features are represented by 208 key points.

5.7 LHI Face Database [164], [42], [149]

The Lotus Hill Research Institute (LHI) face database includes 50,000 midresolution adult Asian face images of subjects at different age groups with little illumination and pose variations, which can be used for face age synthesis [42]. Part of the database is used for age estimation [149], which consists of 8,000 color face images of 120×160 resolution, half males and half females, one image per subject, with an age range from 9 to 89 and roughly 100 images for each age.

5.8 HOIP Face Database [165]

The Human and Object Interaction Processing (HOIP) face database contains in total 306,600 images of 300 subjects, half males and half females. The age label is distributed in 10 groups from 15 to 64 with five year interval. There are 15 males and 15 females in each group. The face images with neutral expression have 640×480 resolution and 24 bits color depth.

5.9 Iranian Face Database [166]

The database contains 616 subjects, 487 males and 129 females, with in total 3,600 color images. The age label is from 2 to 85. The images, with variations of pose, expression, without glasses, have 640×480 resolution and 24 bits color depth.

5.10 Gallagher's Web-Collected Database [10], [167]

The Web-collected database reported in [10] contains 28,231 faces (in 5,080 images) from the Flickr.com image search engine using three-group searches such as "wedding+ bride+ groom+portrait" "group shot" or "group photo" or "group portrait" "family portrait." The age labels are of seven categories: 0-2, 3-7, 8-12, 13-19, 20-36, 37-65, and 66+. The median face has only 18.5 pixels between the eye centers, and 25 percent of the faces have under 12.5 pixels [167].

5.11 Ni's Web-Collected Database [41]

The Web-collected database reported in [41] contains 219,892 faces (in 77,021 images) from Flickr.com and the Google.com image search engine. The age label is distributed from 1 to 80 based on age-related queries. This database is the largest one ever reported for age estimation via faces.

5.12 3D Morphable Database [25], [13]

This database contains the 3D scans of 200 adult (half males and half females) faces and 238 teenaged faces (8-16 years), with 125 males and 113 females and a resolution of 76,800 vertices per scan (320×240).

5.13 Summary

Note that the FG-NET Aging database, MORPH database, and Gallagher's Web-collected database are all publicly available. The other databases are potentially available by contacting the owners. The MORPH, YGA, LHI, Ni's, and Gallagher's Web-collected databases are large-size aging face databases which are promising for data-driven statistical algorithms of age estimation, such as AAM and age manifold with the regression model. The FG-NET is a baseline database for comparisons with many existing age synthesis and estimation techniques, such as AGES. The AI&R, Iranian, and LHI databases contain relatively high-resolution 2D images which can be effectively used for age synthesis evaluation. The 3D morphable database is specifically collected for 3D age synthesis. The other databases mentioned here were not widely used but might be useful for some specific applications.

6 EVALUATION PROTOCOLS FOR AGE SYNTHESIS AND ESTIMATION

Targeting different application purposes, age synthesis and estimation techniques are evaluated in different ways.

6.1 Age Synthesis Evaluation

The evaluations for age synthesis can be either subjective tests, evaluated aesthetically by human observers [37] and set up with age groups [6], or objective tests, evaluated quantitatively by similarity measures. In the evaluation of [122], the observers were given three sets of face images: original faces, synthesized aging faces, and target faces (ground truth). Then, they were asked to answer some questions according to the comparisons between those face images, such as "Which is older? (original or synthesized)" and "Which is more similar to the target face? (original or synthesized)." Another evaluation in [42] divided the original evaluation data into five age groups over adulthood. Then, 20 faces were picked from the youngest group and synthesized to the other four different ages. In total, there are 80 synthetic face images in a testing set. Then, 80 original face images from the corresponding age groups were selected to compose another testing set. The observers were asked to estimate the age group labels for each image in both testing sets. The comparisons of the human perception accuracy between the original set and the synthetic set indicate the age synthesis performance. In [37], subjective evaluations by 29 human observers were conducted on both pure face (gray-scale face region) and multiple cues (original color image), which showed better results in the second case.

Lanitis provided a comparative evaluation of existing age synthesis methods in [168]. Two quantitative evaluation measures, Age Similarity (age_s) and Individual Appearance Similarity (id_s), were proposed. The age_s evaluates the performance of an age estimation algorithm to synthesize aging or rejuvenating face images that can characterize traits of the target age group. The id_s evaluates the performance of the algorithm to characterize individual traits. The similarity measures are defined as the distances between synthetic images and distribution (model) of the target age group or the subject of the source image.

6.2 Age Estimation Evaluation

Since the subjective age estimation is biased [169] and costs expensive manpower, the evaluations for age estimation are usually objective tests, evaluated quantitatively, and set up with either individual images or age groups. This evaluation task often adopts two measures, MAE and Cumulative Score (CS) [17], [122], [131]. The MAE is defined as the average of the absolute errors between the estimated age labels and the ground truth age labels, that is, $\text{MAE} = \sum_{k=1}^{N} |\hat{l}_k - l_k|/N$, where l_k is the ground truth age for the kth test image, $\hat{l}_k$ is the estimated age, and N is the total number of test images. The CS [16], [36] is defined as $\text{CS}(j) = N_{e \leq j}/N \times 100\%$, where $N_{e \leq j}$ is the number of test images on which the age estimation makes an absolute error no higher than j years. As a classification problem, the age estimation performance can also be revealed by the classification accuracy.

$\text{CS}(j)$ can be viewed as the classification accuracy, which, like in face recognition, might be more important a criterion than the average performance MAE for practical applications. Here, in the age estimation field, we use both CS and MAE as metrics since different methods, databases, and systems may be too biased or unbalanced for evaluation. For example, if the ordinal label is available in a large training data set with dense distribution of different ages, CS could be the best way to reflect the performance. But, if the training data are not dense, i.e., there are a lot of missing data or missing ages, MAE could be very useful to measure the performance.

Table 1 summarizes and compares some existing age estimation techniques evaluated on several benchmark databases. We notice that comprehensive studies and active efforts have been shown in both age image representation and age estimation algorithms in the last decade. Comparisons of different existing methods shown by CS curves can be retrieved from [161], [16], [147].

7 CONCLUSIONS AND FUTURE DIRECTIONS

We have presented a complete survey of the state-of-the-art techniques for age synthesis and estimation via face images, which became fairly particular in recent decades because of their promising real-world applications in several emerging fields. The explosively comprehensive efforts from both academia and industry have been devoted recently to models and algorithms designing, face aging databases collecting, and system performances evaluation with valid protocols. Variant solutions to technical difficulties have also been provided by researchers. Table 2 summarizes the facts and characteristics versus countermeasures of age synthesis and estimation tasks. The N/A in the table can indicate possible future directions to mitigate difficulties or degrading factors in age synthesis and estimation.

In general, different age synthesis and estimation techniques and algorithms can be effectively applied to particular scenarios or applications. For face modeling, the appearance-based face model can be considered as a marriage of geometry-based model and image-based model, which shows more photorealistic effects aesthetically for age

TABLE 1
Summary and Comparison of Age Estimation Techniques on Different Databases

Technique	Database	Represent-ation	Data Description			Performance				
			Subject #	Image #	Label	Algorithm	Protocol	MAE(yrs)	CS(≤10)	Accuracy
Kwon & Lobo 1999 [128]	[128]	AM	N/A	47	Three groups	C	15 images for test	N/A	N/A	100%
Kanno et al. 2001 [156]	[156]	APM (mosaic)	110(M)	440(M)	12,15, 18,22	C (ANN)	N/A	N/A	N/A	80%(M)
Iga et al. 2003 [46]	[46]	AM	101	N/A	22–66 years	C (5 classes)	Tain on HOIP	N/A	N/A	58.4%
Lanitis et al. 2004 [17]	[17]	AAM	40	400	0–35 years	C or R	Half train half test	3.82– 5.58	N/A	N/A
Ueki et al. 2006 [50]	WIT-DB	APM (raw image)	3000(M) 2500(F)	14214(M) 12008(F)	3–85 years	C (11 classes)	2-fold cross val.	N/A	N/A	50%(M) 43%(F)
Takimoto et al. 2006 [47]	HOIP	APM	113(M) 139(F)	N/A	15–64 years	C (6 classes)	Leave-one -out CV	N/A	N/A	57.3%(M) 54.7%(F)
Takimoto et al. 2007 [48]	HOIP	AM	113(M) 139(F)	N/A	15–64 years	R (ANN)	Leave-one -out CV	3.0(M) 4.4(F)	N/A	N/A
Geng et al. 2007 [37]	MORPH	AGES	515	1430(M) 294(F)	15–68 years	R	Tain on FG-NET	8.83	≈70%	N/A
Ni et al. 2009 [41]	Web data, MORPH	APM (patches)	219892 + 515	77021 + 1690	15–68 years	R (RMIR[41])	Tain on web data	7.42	N/A	N/A
Ni et al. 2009 [41]	Web data, MORPH	APM (patches)	219892 + N/A	77021 + 55608	>10 <80	R (RMIR[41])	Tain on web data	8.60	N/A	N/A
Zhou et al. 2005 [158]	FG-NET	APM	82	1002	0–69 years	R	5-fold cross val.	5.81	N/A	N/A
Geng et al. 2006 [131]	FG-NET	AGES	82	1002	0–69 years	R	Leave one person out	6.77	≈81%	N/A
Geng et al. 2006 [131]	FG-NET	AGES	82	1002	0–69 years	C	14 ranges, 5yrs each	1.26	N/A	40.92% (hit rate)
Yan et al. 2007 [38]	FG-NET	AAM	82	1002	0–69 years	R (ranking)	Leave one person out	5.33	≈84%	N/A
Yan et al. 2007 [39]	FG-NET	AAM	82	1002	0–69 years	R (RUN[39])	Leave one person out	5.78	≈84%	N/A
Yan et al. 2008 [147]	FG-NET	APM (patches)	82	1002	0–69 years	R (kernel reg.)	Leave one person out	4.95	N/A	N/A
Guo et al. 2008 [16]	FG-NET	AAM	82	1002	0–69 years	R (LARR[31])	Leave one person out	5.07	≈88%	N/A
Guo et al. 2008 [161]	FG-NET	AAM	82	1002	0–69 years	H	Leave one person out	4.97	≈88%	N/A
Ni et al. 2009 [41]	Web data, FG-NET	APM (patches)	219892 + 82	77021 + 1002	0–69 years	R (RMIR[41])	Train on web data	9.49	N/A	N/A
Yan et al. 2009 [140]	FG-NET	AAM	82	1002	0–69 years	C (SSE[140])	Leave one person out	5.21	N/A	N/A
Xiao et al. 2009 [160]	FG-NET	AAM	82	1002	0–69 years	C (mkNN[160])	300 train 702 test	5.04	≈84%	N/A
Guo et al. 2009 [150]	FG-NET	AMF (BIF[150])	82	1002	0–69 years	R (SVR)	Leave one person out	4.77	≈89%	N/A
Fu et al. 2007 [40]	YGA	AMF (OLPP[137])	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (quadratic)	Half train half test	≈8.0(M) ≈7.8(F)	≈70%(M) ≈70%(F)	N/A
Fu & Huang 2008 [36]	YGA	AMF (CEA[170])	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (quadratic)	Half train half test	≈6.0(M) ≈5.5(F)	≈82%(M) ≈83%(F)	N/A
Yan et al. 2007 [38]	YGA	APM (raw image)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (ranking)	4-fold cross val.	6.95(M) 6.95(F)	≈79%(M) ≈78%(F)	N/A
Yan et al. 2007 [39]	YGA	APM (raw image)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (RUN[39])	1000 train 3000 test	10.36(M) 9.79(F)	≈61%(M) ≈63%(F)	N/A
Yan et al. 2008 [148]	YGA	APM (patches)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (GMM)	1000 train 3000 test	7.82(M) 8.53(F)	≈75%(M) ≈70%(F)	N/A
Yan et al. 2008 [147]	YGA	APM (patches)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (kernel reg.)	1000 train 3000 test	4.38(M) 4.94(F)	≈88%(M) ≈85%(F)	N/A
Guo et al. 2008 [16]	YGA	AMF (OLPP)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (LARR[31])	4-fold cross val.	5.30(M) 5.25(F)	≈83%(M) ≈81%(F)	N/A
Guo et al. 2008 [161]	YGA	AMF (OLPP)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	H	4-fold cross val.	5.12(M) 5.11(F)	≈83%(M) ≈82%(F)	N/A
Zhuang et al. 2008 [171]	YGA	APM (patches)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	R (HMM)	Half train half test	5.40(M) 6.33(F)	≈82%(M) ≈76%(F)	N/A
Guo et al. 2009 [150]	YGA	APM (BIF)	800(M) 800(F)	4000(M) 4000(F)	0–93 years	C (SVM)	4-fold cross val.	3.47(M) 3.91(F)	≈88%(M) ≈85%(F)	N/A
Guo et al. 2009 [155]	YGA	APM (BIF) + AMF	800(M) 800(F)	4000(M) 4000(F)	0–93 years	C (SVM)	4-fold cross val.	2.58(M) 2.61(F)	N/A	89.7%

Note N/A—**N**ot **A**vailable, M—**M**ale, and F—**F**emale. In the "Representation" column, AM—**A**nthropometric **M**odel, AAM—**A**ctive **A**ppearance **M**odel, AGES—**AG**ing patt**E**rn **S**ubspace, AMF—**A**ge **M**ani**F**old, and APM—**AP**pearance **M**odel. In the "Algorithm" column, C—**C**lassification, R—**R**egression, and H—**H**ybrid. In the "CS (percent)" column, the tolerable absolute error is set as ±10 years.

synthesis purposes. Shape synthesis is more effective for the age progression of young faces whose craniofacial growth and development are more dominant over skin aging. Texture synthesis is more effective for the face aging after adulthood when skin aging dominates and craniofacial growth slows down. Appearance synthesis is applicable to both cases since, usually, a statistical model will be available built on a large face aging database. Aging cues can be

TABLE 2
Facts/Characteristics versus Countermeasures of Age Synthesis and Estimation

Tasks	Facts and Characteristics of Face Aging						
	Two-stage Growth	Uncontrollable	Personalized	Attractiveness	Missing Data	Gender	Ethnicity
Age Synthesis	Different growth, development, and aging models: geometry-based, image-based, appearance-based.	Generic models with statistical learning: ASM [76], AAM [123], MM [99], M-face [30], muscle-wrinkle [126], dynamic model [42].	Statistical learning: implicit statistical synthesis [42], [13].	Averageness hypothesis: average face as the origin in the aging space [59]. M-face [30], MM [99], babyfaceness and average face [60].	Explicit synthesis: data driven synthesis, mechanical synthesis.	Implicit statistical synthesis: MM [99], dynamic model [42].	N/A
Age Estimation	Age group classification and aging pattern or manifold learning, [36], [150] [38], [158], [155].	Online training, universal age estimator [41].	Regression, manifold learning, [36], [16], [150].	N/A	Aging pattern and regression, AGES [37], LARR [16], web data[41].	Gender recognition, [44], [45], [43], [46] [155].	N/A

learned statistically for all aging stages and implemented to realistic age synthesis. But the most difficult part of appearance synthesis is the database collection and automatic face correspondence.

Image representation and estimation methods are two key issues in age estimation via face images. The anthropometric model and AAM provide parametric modeling for face image representation. They are flexible to handle both age synthesis and estimation. The anthropometric model focuses on shape changes, which is mainly for age estimation of young faces. The AAM can deal with all of the age estimation cases and reduce the feature dimension by model parameter representations. It is often combined with regression-based estimation methods. When sequential aging face images are available for individual subjects, the aging pattern subspace might be applied to capture individual aging patterns. Age manifold is pretty useful when a large database spanning a large age range is available. Feature extraction and dimensionality reduction are intertwined in age manifold learning. It can be combined with both regression-based and classification-based estimation methods. Appearance model is often a way to consider and handle cases when multiple facial attributes are merged. Patch-based image representation is effective to deal with slight head pose and illumination variations. Bioinspired image representation shows superior discriminant power to capture the general aging patterns.

The age estimation method can be either classification-based or regression-based, according to different image representations and databases. For large databases with sequential age labels, both can be applied, while for databases with only age group labels, classification-based methods might be more appropriate. However, it is also reasonable to consider a hybrid of these two.

Some studies [37], [17] reported human performance in age estimation. Using the whole FG-NET database for Leave-One-Person-Out (LOPO) evaluation, Geng et al. [37] achieved MAE of 6.23 years with human observer estimation and 6.22 years with their AGES algorithm. Lanitis et al. [17] used a smaller database for evaluation and achieved MAE of 3.64 years with human observer estimation and 3.82 years with their best computer-based algorithm. It is interesting to see that some computer-based algorithms can potentially exceed human ability in age estimation. This result may motivate more dedicated studies in this field.

In addition to the suggestions of Table 2, there are a couple of promising future directions as follows for age synthesis and estimation via face images:

- *Facial attributes decomposition.* When face images show multiple facial attributes [172], such as identity, expression, gender, age, ethnicity, and pose, the tensor representation (multilinear analysis) [173] can be adopted to handle uncontrollable and personalized characteristics. Then, the attributes decomposition can be handled via higher order singular value decomposition [174]. This multilinear model has more flexibility for age synthesis and estimation [175]. It also might be interesting to investigate the characteristics of aging process in different ethnicity, gender, or both.

- *Generalized aging model.* Both age synthesis and estimation share some similar ideas and can help each other. Merging the two modules for generalized aging modeling is beneficial for each one, e.g., two-stage growth and attractiveness. The AAM and "revised" cardioidal strain transformation model are possible examples of such kind of models.

- *Statistical models and databases.* It is more desirable to build statistical models for both age synthesis and estimation because of their faithful results, model extensibility and strong ability of expressiveness, which are effective to handle two-stage growth, uncontrollable, personalized, attractiveness, missing data, gender, and ethnicity characteristics. Collecting large size face aging databases with sufficient labels is crucial in this task. Web image collection is an efficient way to consider [41].

- *3D and nonfrontal view face aging.* Building 3D models for age synthesis and estimation may also be interesting and effective to combine with statistical learning strategies. How to synthesize nonfrontal view aging faces and how to estimate age from nonfrontal view face images are attractive yet open problems. The burden is still in the data collection. Fitting the 3D model with accurate alignment might be an additional difficulty.

- *Multimodality.* Data collection from multimodality imaging sensors and fusing them for age estimation might also be an interesting problem to mitigate degrading factors from uncontrollable and personalized characteristics.

- *Multilabel*. Incorporation of metadata such as life-style, geography, occupation, etc.
- *Age bias*. The bias between perceived/estimated age (from appearance age) and the actual age can be learned with data-driven methods. Sophisticated statistical models can be adopted to learn and handle the bias.

REFERENCES

[1] L.A. Zebrowitz, *Reading Faces: Window to the Soul?* Westview Press, 1997.

[2] A. Gallagher and T. Chen, "Estimating Age, Gender, and Identity Using First Name Priors," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2008.

[3] L. Zyga, Intelligent Computers See Your Human Traits, http://www.physorg.com/news131277135.html, PhysOrg.com, May 2008.

[4] N. Ramanathana, R. Chellappa, and S. Biswas, "Computational Methods for Modeling Facial Aging: A Survey," *J. Visual Languages and Computing,* vol. 20, no. 3, pp. 131-144, 2009.

[5] Aging of the Face, http://www.face-and-emotion.com/dataface/facets/aging.jsp, 2010.

[6] P.A. George and G.J. Hole, "Factors Influencing the Accuracy of Age Estimates of Unfamiliar Faces," *Perception,* vol. 24, no. 9, pp. 1059-1073, 1995.

[7] Age Progression, http://www.forensicartist.com/agepro.html, 2010.

[8] S. Mancusi, Image Modification, Aging Progression and Age Regression, http://www.forartist.com/forensic/modification/modificationpg.htm, 2010.

[9] J. Vestlund, L. Langeborg, P. Sorqvist, and M. Eriksson, "Experts on Age Estimation," *Scandinavian J. Psychology,* vol. 50, no. 4, pp. 301-307, 2009.

[10] A. Gallagher and T. Chen, "Understanding Images of Groups of People," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2009.

[11] T. Simonite,, "Virtual Face-Ageing May Help Find Missing Persons," NewScientist.com, http://technology.newscientist.com/channel/tech/forensic-science, Sept. 2006.

[12] C.M. Hill, C.J. Solomon, and S.J. Gibson, "Aging the Human Face—A Statistically Rigorous Approach," *Proc. IEE Int'l Symp. Imaging for Crime Detection and Prevention,* pp. 89-94, 2005.

[13] K. Scherbaum, M. Sunkel, H.-P. Seidel, and V. Blanz, "Prediction of Individual Non-Linear Aging Trajectories of Faces," *Proc. Ann. Conf. European Assoc. Computer Graphics,* vol. 26, no. 3, pp. 285-294, 2007.

[14] Electronic Customer Relationship Management (ECRM), http://en.wikipedia.org/wiki/ECRM, 2010.

[15] J.E. Kloeppel,, "Step Right Up, Let the Computer Look at Your Face and Tell You Your Age," http://news.illinois.edu/news/08/0923age.html, 2010.

[16] G. Guo, Y. Fu, C. Dyer, and T.S. Huang, "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression," *IEEE Trans. Image Processing,* vol. 17, no. 7, pp. 1178-1188, July 2008.

[17] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing Different Classifiers for Automatic Age Estimation," *IEEE Trans. Systems, Man, and Cybernetics Part B,* vol. 34, no. 1, pp. 621-628, Feb. 2004.

[18] A.K. Jain, S.C. Dass, and K. Nandakumar, "Soft Biometric Traits for Personal Recognition Systems," *Proc. Int'l Conf. Biometric Authentication,* pp. 731-738, 2004.

[19] K. Ricanek, E. Boone, and E. Patterson, "Craniofacial Aging on the Eigenface Biometric," *Proc. IASTED Int'l Conf. Visualization, Imaging, and Image Processing,* pp. 249-253, 2006.

[20] K. Ricanek and E. Boone, "The Effect of Normal Adult Aging on Standard PCA Face Recognition Accuracy Rates," *Proc. Int'l Joint Conf. Neural Networks,* pp. 2018-2023, 2005.

[21] E. Patterson, K. Ricanek, M. Albert, and E. Boone, "Automatic Representation of Adult Aging in Facial Images," *Proc. IASTED Int'l Conf. Visualization, Imaging, and Image Processing,* pp. 171-176, 2006.

[22] E. Patterson, A. Sethuram, M. Albert, K. Ricanek, and M. King, "Aspects of Age Variation in Facial Morphology Affecting Biometrics," *Proc. IEEE Conf. Biometrics: Theory, Applications, and Systems,* 2007.

[23] N. Ramanathan and R. Chellappa, "Face Verification across Age Progression," *IEEE Trans. Image Processing,* vol. 15, no. 11, pp. 3349-3361, Nov. 2006.

[24] J. Wang, Y. Shang, G. Su, and X. Lin, "Age Simulation for Face Recognition," *Proc. Int'l Conf. Pattern Recognition,* pp. 913-916, 2006.

[25] V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," *Proc. ACM SIGGRAPH,* pp. 187-194, 1999.

[26] M. Das and A.C. Loui, "Automatic Face-Based Image Grouping for Albuming," *Proc. IEEE Conf. Systems, Man, and Cybernetics,* vol. 4, pp. 3726-3731, 2003.

[27] Cosmetic Surgery, http://www.cosmeticsurgery.org/, 2010.

[28] R.M. Koch, M.H. Gross, F.R. Carls, D.F. Von Buren, F. Fankhauser, and Y.I.H. Parish, "Simulating Facial Surgery Using Finite Element Models," *Proc. ACM SIGGRAPH,* pp. 421-C428, 1996.

[29] Y. Fu, N. Zheng, J. Liu, and T. Zhang, "Facetransfer: A System Model of Facial Image Rendering," *Proc. IEEE Conf. Systems, Man, and Cybernetics,* pp. 4528-4534, 2004.

[30] Y. Fu and N. Zheng, "M-Face: An Appearance-Based Photo-realistic Model for Multiple Facial Attributes Rendering," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 16, no. 7, pp. 830-842, July 2006.

[31] G. Guo, Y. Fu, T.S. Huang, and C. Dyer, "Locally Adjusted Robust Regression for Human Age Estimation," *Proc. IEEE Workshop Applications of Computer Vision,* 2008.

[32] T.R. Alley, *Social and Applied Aspects of Perceiving Faces.* Lawrence Erlbaum Assoc., 1988.

[33] A. Stone, The Aging Process of the Face & Techniques of Rejuvenation, http://www.aaronstonemd.com/Facial_Aging_Rejuvenation.shtm, 2010.

[34] D.S. Berry and L.Z. McArthur, "Perceiving Character in Faces: The Impact of Age-Related Craniofacial Changes on Social Perception," *Psychological Bull.,* vol. 100, no. 1, pp. 3-18, July 1986.

[35] L. Boissieux, G. Kiss, N.M. Thalmann, and P. Kalra, "Simulation of Skin Aging and Wrinkles with Cosmetics Insight," *Proc. Euro-graphics Workshop Animation and Simulation,* pp. 15-27, 2000.

[36] Y. Fu and T.S. Huang, "Human Age Estimation with Regression on Discriminative Aging Manifold," *IEEE Trans. Multimedia,* vol. 10, no. 4, pp. 578-584, June 2008.

[37] X. Geng, Z.-H. Zhou, and K. Smith-Miles, "Automatic Age Estimation Based on Facial Aging Patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 29, no. 12, pp. 2234-2240, Dec. 2007.

[38] S. Yan, H. Wang, T.S. Huang, and X. Tang, "Ranking with Uncertain Labels," *Proc. IEEE Conf. Multimedia and Expo,* pp. 96-99, 2007.

[39] S. Yan, H. Wang, X. Tang, and T.S. Huang, "Learning Auto-Structured Regressor from Uncertain Nonnegative Labels," *Proc. IEEE Conf. Computer Vision,* 2007.

[40] Y. Fu, Y. Xu, and T.S. Huang, "Estimating Human Ages by Manifold Analysis of Face Pictures and Regression on Aging Features," *Proc. IEEE Conf. Multimedia and Expo,* pp. 1383-1386, 2007.

[41] B. Ni, Z. Song, and S. Yan, "Web Image Mining Towards Universal Age Estimator," *Proc. ACM Multimedia,* 2009.

[42] J. Suo, F. Min, S. Zhu, S. Shan, and X. Chen, "A Multi-Resolution Dynamic Model for Face Aging Simulation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2007.

[43] T. Fujiwara and H. Koshimizu, "Age and Gender Estimations by Modeling Statistical Relationship among Faces," *Proc. Int'l Conf. Knowledge-Based Intelligent Information and Eng. Systems,* V. Palade, R.J. Howlett, and L.C. Jain, eds., pp. 870-876, 2003.

[44] J. Hayashi, M. Yasumoto, H. Ito, and H. Koshimizu, "Method for Estimating and Modeling Age and Gender Using Facial Image Processing," *Proc. Seventh Int'l Conf. Virtual Systems and Multi-media,* pp. 439-448, 2001.

[45] J. Hayashi, M. Yasumoto, H. Ito, Y. Niwa, and H. Koshimizu, "Age and Gender Estimation from Facial Image Processing," *Proc. SICE Ann. Conf.,* vol. 1, pp. 13-18, 2002.

[46] R. Iga, K. Izumi, H. Hayashi, G. Fukano, and T. Ohtani, "A Gender and Age Estimation System from Face Images," *Proc. SICE Ann. Conf.,* pp. 756-761, 2003.

[47] H. Takimoto, Y. Mitsukura, M. Fukumi, and N. Akamatsu, "A Design of Gender and Age Estimation System Based on Facial Knowledge," *Proc. SICE ICASE Int'l Joint Conf.,* pp. 3883-3886, 2006.

[48] H. Takimoto, T. Kuwano, Y. Mitsukura, H. Fukai, and M. Fukumi, "Appearance-Age Feature Extraction from Facial Image Based on Age Perception," *Proc. SICE Ann. Conf.,* pp. 2813-2818, 2007.

[49] H. Fukai, H. Takimoto, Y. Mitsukura, and M. Fukumi, "Apparent Age Estimation System Based on Age Perception," *Proc. SICE Ann. Conf.,* pp. 2808-2812, 2007.

[50] K. Ueki, T. Hayashida, and T. Kobayashi, "Subspace-Based Age-Group Classification Using Facial Images under Various Lighting Conditions," *Proc. IEEE Conf. Automatic Face and Gesture Recognition,* pp. 43-48, 2006.

[51] A.M. Alberta, K. Ricanek, and E. Pattersonb, "A Review of the Literature on the Aging Adult Skull and Face: Implications for Forensic Science Research and Applications," *Forensic Science Int'l,* vol. 172, no. 1, pp. 1-9, 2007.

[52] L.S. Mark, J.B. Pittenger, H. Hines, C. Carello, R.E. Shaw, and J.T. Todd, "Wrinkling and Head Shape as Coordinated Sources of Age-Level Information," *Perception and Psychophysics,* vol. 27, no. 2, pp. 117-124, 1980.

[53] N. Ramanathan and R. Chellappa, "Modeling Age Progression in Young Faces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 387-394, 2006.

[54] L. Farkas, *Anthropometry of the Head and Face.* Raven Press, 1994.

[55] M. Gonzalez-Ulloa and E. Flores, "Senility of the Face: Basic Study to Understand Its Causes and Effects," *Plastic and Reconstructive Surgery,* vol. 36, pp. 239-246, 1965.

[56] T. Valentine, "A Unified Account of the Effects of Distinctiveness, Inversion, and Race in Face Recognition," *Quarterly J. Experimental Psychology,* vol. 43, no. 2, pp. 161-204, 1991.

[57] G. Rhodes and T. Tremewan, "Averageness, Exaggeration, and Facial Attractiveness," *Psychological Science,* vol. 7, no. 2, pp. 105-110, 1996.

[58] K.A. Deffenbacher, T. Vetter, J. Johanson, and A.J. O'Toole, "Facial Aging, Attractiveness, and Distinctiveness," *Perception,* vol. 27, pp. 1233-1243, 1998.

[59] J.H. Langlois and L.A. Roggman, "Attractive Faces Are Only Average," *Psychological Science,* vol. 1, no. 2, pp. 115-121, 1990.

[60] C. Braun, M. Gruendl, C. Marberger, C. Scherber, T. Vetter, H. Volz, and E.M. Salter, *Beautycheck-Causes and Consequences of Human Facial Attractiveness(Summary),* The German Students Award, http://www.beautycheck.de/, 2001.

[61] T.R. Alley, "Head Shape and the Perception of Cuteness," *Developmental Psychology,* vol. 17, no. 5, pp. 650-654, Sept. 1981.

[62] A.J. O'Toole, T. Price, T. Vetter, J.C. Bartlett, and V. Blanz, "3D Shape and 2D Surface Textures of Human Faces: The Role of 'Averages' in Attractiveness and Age," *Image and Vision Computing,* vol. 18, pp. 9-19, 1999.

[63] A.J. O'Toole, T. Vetter, H. Volz, and E.M. Salter, "Three-Dimensional Caricatures of Human Heads: Distinctiveness and the Perception of Facial Age," *Perception,* vol. 26, pp. 719-732, 1997.

[64] F.I. Parke, "Computer Generated Animation of Faces," *Proc. ACM SIGGRAPH,* pp. 451-457, 1972.

[65] J. Noh and U. Neumann, "A Survey of Facial Modeling and Animation Techniques," USC Technical Report 99-705, http://citeseer.ist.psu.edu/481159.htm, 1998.

[66] J. Haber and D. Terzopoulos, "Facial Modeling and Animation," *Proc. ACM SIGGRAPH Course Notes,* 2004.

[67] J. Kittler, A. Hilton, M. Hamouz, and J. Illingworth, "3D Assisted Face Recognition: A Survey of 3D Imaging, Modelling and Recognition Approachest," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* p. 114, 2005.

[68] F.I. Parke and K. Waters, *Computer Facial Animation.* A.K. Peters, Ltd. 1996.

[69] F. Pighin, R. Szeliski, and D. Salesin, "Modeling and Animating Realistic Faces from Images," *Int'l J. Computer Vision,* vol. 50, no. 2, pp. 143-169, 2002.

[70] Y.C. Lee, D. Terzopoulos, and K. Waters, "Realistic Modeling for Facial Animation," *Proc. ACM SIGGRAPH,* pp. 55-62, 1995.

[71] K. Waters, "A Muscle Model for Animation Three-Dimensional Facial Expression," *Proc. ACM SIGGRAPH,* pp. 17-24, 1987.

[72] R.L. Hsu and A.K. Jain, "Generating Discriminating Cartoon Faces Using Interacting Snakes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, no. 11, pp. 1388-1398, Nov. 2003.

[73] S.E. Brennan, "The Caricature Generator," *Leonardo,* vol. 18, pp. 170-178, 1985.

[74] P.J. Benson and D.I. Perrett, "Synthesising Continuous-Tone Caricatures," *Image Vision Computer,* vol. 9, pp. 123-129, 1991.

[75] A.J. Calder, A.W. Young, D. Rowland, and D.I. Perrett, "Computer-Enhanced Emotion in Facial Expressions," *Proc. Royal Soc. of London Series B Biological Sciences,* vol. 264, no. 1383, pp. 919-925, 1997.

[76] T.F. Cootes, D. Cooper, C.J. Taylor, and J. Graham, "Active Shape Models—Their Training and Application," *Computer Vision and Image Understanding,* vol. 61, no. 1, pp. 38-59, Jan. 1995.

[77] P. Ekman and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Consulting Psychologists Press, 1978.

[78] F. Lavagetto and R. Pockaj, "An Efficient Use of MPEG-4 FAP Interpolation for Facial Animation at 70 Bits/Frame," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 11, no. 11, pp. 1085-1097, Oct. 2001.

[79] F.I. Parke, "Parameterized Models for Facial Animation," *IEEE Computer Graphics and Applications,* vol. 2, no. 9, pp. 61-68, Nov. 1982.

[80] P. Hong, Z. Wen, and T.S. Huang, "Real-Time Speech-Driven Face Animation with Expressions Using Neural Networks," *IEEE Trans. Neural Networks,* vol. 13, no. 4, pp. 916-927, Jan. 2002.

[81] Z. Wen and T.S. Huang, *3D Face Processing: Modeling, Analysis and Synthesis.* Springer, 2004.

[82] H. Tang, Y. Fu, J. Tu, M. Hasegawa-Johnson, and T.S. Huang, "Humanoid Audio-Visual Avatar with Emotive Text-to-Speech Synthesis," *IEEE Trans. Multimedia,* vol. 10, no. 6, pp. 969-981, Oct. 2008.

[83] Z. Deng and J. Noh, "Computer Facial Animation: A Survey," *Data-Driven 3D Facial Animation,* pp. 1-28, Springer, 2007.

[84] L. Liang, H. Chen, Y.Q. Xu, and H.Y. Shum, "Example-Based Caricature Generation with Exaggeration," *Proc. Pacific Conf. Computer Graphics and Applications,* pp. 386-393, 2002.

[85] T. Kuratate and T. Nishita, "A Simple Method for Modeling Wrinkles on Human Skin," *Proc. Pacific Conf. Computer Graphics and Applications,* pp. 166-175, 2002.

[86] H. Pyun, Y. Kim, W. Chae, H. Kang, and S. Shin, "An Example-Based Approach for Cloning Facial Expressions," *Proc. 2003 ACM SIGGRAPH/Eurographics,* pp. 167-176, 2003.

[87] Y. Shan, Z. Liu, and Z. Zhang, "Image-Based Surface Detail Transfer," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 794-799, 2001.

[88] Y. Fu, "Merging Ratio Images Based Realistic Object Class Re-Rendering," *Proc. IEEE Conf. Image Processing,* pp. 3523-3526, 2004.

[89] Z. Liu, Y. Shan, and Z. Zhang, "Expressive Expression Mapping with Ratio Images," *Proc. ACM SIGGRAPH,* pp. 271-276, 2001.

[90] A. Shashua and T. Riklin-Raviv, "The Quotient Image: Class-Based Re-Rendering and Recognition with Varying Illuminations," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 2, pp. 129-139, Feb. 2001.

[91] Y. Fu, "Photorealistic Face Rendering Based on a Fusion Model of Linear and Lambertian Object Class," Master's thesis, AI&R, Xi'an Jiaotong Univ., 2004.

[92] P.J. Benson and D.I. Perrett, "Extracting Prototypical Facial Images from Exemplars," *Perception,* vol. 22, no. 3, pp. 257-262, 1993.

[93] T. Beier and S. Neely, "Feature-Based Image Metamorphosis," *Computer Graphics,* vol. 26, no. 2, pp. 35-42, 1992.

[94] A. Stoschek, "Image-Based Re-Rendering of Faces for Continuous Pose and Illumination Directions," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 582-587, 2000.

[95] S. Seitz and C. Dyer, "View Morphing," *Proc. ACM SIGGRAPH,* pp. 21-30, 1996.

[96] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin, "Synthesizing Realistic Facial Expressions from Photographs," *Proc. ACM SIGGRAPH,* pp. 75-84, 1998.

[97] D. Beymer and T. Poggio, "Image Representations for Visual Learning," *Science,* vol. 272, pp. 1905-1909, 1996.

[98] T. Vetter and T. Poggio, "Linear Object Classes and Image Synthesis from a Single Example Image," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 7, pp. 733-742, July 1997.

[99] M. Jones and T. Poggio, "Multidimensional Morphable Models," *Proc. IEEE Conf. Computer Vision,* pp. 683-688, 1998.

[100] T. Cootes, G. Edwards, and C. Taylor, "Active Appearance Models," *Proc. European Conf. Computer Vision,* vol. 2, pp. 484-498, 1998.

[101] H. Ling, S. Soatto, N. Ramanathan, and D. Jacobs, "A Study of Face Recognition as People Age," *Proc. IEEE Conf. Computer Vision,* 2007.

[102] N. Ramanathan and R. Chellappa, "Recognizing Faces Across Age Progression," *Face Biometrics for Personal Identification,* R. Hammoud, M. Abidi, and B. Abidi, eds., pp. 27-42, Springer-Verlag, 2007.

[103] D'A.W. Thompson, *On Growth and Form.* Cambridge Univ. Press, 1917.

[104] W. Arthur, "D'Arcy Thompson and the Theory of Transformations," *Nature Rev. Genetics,* vol. 7, pp. 401-406, 2006.

[105] R. Chellappa and N. Ramanathan, "Face Recognition across Age Progression," *Proc. Face Recognition Grand Challenge and ICE Workshop,* http://face.nist.gov/frgc/UMD_FRGC_Brief.pdf, Mar. 2006.

[106] J.B. Pittenger and R.E. Shaw, "Aging Faces as Viscal-Elastic Events: Implications for a Theory of Nonrigid Shape Perception," *J. Experimental Psychology: Human Perception and Performance,* vol. 1, no. 4, pp. 374-382, 1975.

[107] J.B. Pittenger, R.E. Shaw, and L.S. Mark, "Perceptual Information for the Age Level of Faces as a Higher Order Invariant of Growth," *J. Experimental Psychology: Human Perception and Performance,* vol. 5, no. 3, pp. 478-493, 1979.

[108] J.T. Todd, S.M. Leonard, R.E. Shaw, and J.B. Pittenger, "The Perception of Human Growth," *Scientific Am.,* vol. 242, no. 2, pp. 106-114, 1980.

[109] L.S. Mark, J.T. Todd, and R.E. Shaw, "Perception of Growth: A Geometric Analysis of How Different Styles of Change Are Distinguished," *J. Experimental Psychology: Human Perception and Performance,* vol. 7, no. 4, pp. 855-868, 1981.

[110] L.S. Mark and J.T. Todd, "The Perception of Growth in Three Dimensions," *Perception and Psychophysics,* vol. 33, no. 2, pp. 193-196, Feb. 1983.

[111] L.S. Mark, B.A. Shapiro, and R.E. Shaw, "Structural Support for the Perception of Growth," *J. Experimental Psychology: Human Perception and Performance,* vol. 12, no. 2, pp. 149-159, 1986.

[112] V. Bruce, M. Burton, T. Doyle, and N. Dench, "Further Experiments on the Perception of Growth in Three Dimensions," *Perception and Psychophysics,* vol. 46, no. 6, pp. 528-36, 1989.

[113] D.M. Burt and D.I. Perrett, "Perception of Age in Adult Caucasian Male Faces: Computer Graphic Manipulation of Shape and Colour Information," *Proc. Royal Soc. of London Series B Biological Sciences,* vol. 259, no. 1355, pp. 137-143, 1995.

[114] S. Mukaida, H. Ando, K. Kinoshita, M. Kamachi, and K. Chihara, "Facial Image Synthesis Using Age Manipulation Based on Statistical Feature Extraction," *Proc. Conf. Visualization, Imaging, and Image Processing,* 2002.

[115] M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 586-591, 1991.

[116] Y. Wu, N.M. Thalmann, and D. Thalmann, "A Plastic-Visco-Elastic Model for Wrinkles in Facial Animation and Skin Aging," *Proc. Second Pacific Conf. Fundamentals of Computer Graphics,* pp. 201-213, 1994.

[117] Y. Wu, N.M. Thalmann, and D. Thalmann, "A Dynamic Wrinkle Model in Facial Animation and Skin Aging," *J. Visualization and Computer Animation,* vol. 6, no. 4, pp. 195-205, 1995.

[118] Y. Wu, P. Kalra, L. Moccozet, and N.M. Thalmann, "Simulating Wrinkles and Skin Aging," *The Visual Computer,* vol. 15, pp. 183-198, 1999.

[119] S. Mukaida and H. Ando, "Extraction and Manipulation of Wrinkles and Spots for Facial Image Synthesis," *Proc. IEEE Conf. Automatic Face and Gesture Recognition,* pp. 749-754, 2004.

[120] M. Gandhi, "A Method for Automatic Synthesis of Aged Human Facial Images," MS thesis, ECE Dept., McGill Univ., 2004.

[121] F. Jiang and Y. Wang, "Facial Aging Simulation Based on Super-Resolution in Tensor Space," *Proc. IEEE Conf. Image Processing,* pp. 1648-1651, 2008.

[122] A. Lanitis, C. Taylor, and T. Cootes, "Toward Automatic Simulation of Aging Effects on Face Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 4, pp. 442-455, Apr. 2002.

[123] T. Cootes, G. Edwards, and C. Taylor, "Active Appearance Models," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 6, pp. 681-685, June 2001.

[124] C.M. Scandrett, C.J. Solomona, and S.J. Gibsona, "A Person-Specific, Rigorous Aging Model of the Human Face," *Pattern Recognition Letters,* vol. 27, no. 15, pp. 1776-1787, 2006.

[125] E. Patterson, A. Sethuram, M. Albert, and K. Ricanek, "Comparison of Synthetic Face Aging to Age Progression by Forensic Sketch Artist," *Proc. Seventh IASTED Int'l Conf. Visualization, Imaging, and Image Processing,* pp. 247-252, 2007.

[126] N. Ramanathan and R. Chellappa, "Modeling Shape and Textural Variations in Aging Faces," *Proc. IEEE Conf. Series on Automatic Face and Gesture Recognition,* 2008.

[127] Y. Kwon and N. Lobo, "Age Classification from Facial Images," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 762-767, 1994.

[128] Y. Kwon and N. Lobo, "Age Classification from Facial Images," *Computer Vision and Image Understanding,* vol. 74, no. 1, pp. 1-21, 1999.

[129] The FG-NET Aging Database, http://www.fgnet.rsunit.com/, http://www-prima.inrialpes.fr/FGnet/, 2010.

[130] A. Gunay and V.V. Nabiyev, "Automatic Detection of Anthropometric Features from Facial Images," *Proc. IEEE Conf. Signal Processing and Comm. Applications,* 2007.

[131] X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, and H. Dai, "Learning from Facial Aging Patterns for Automatic Age Estimation," *Proc. ACM Conf. Multimedia,* pp. 307-316, 2006.

[132] H.S. Seung and D.D. Lee, "The Manifold Ways of Perception," *Science,* vol. 290, no. 5500, pp. 2268-2269, 2000.

[133] Y. Fu and T.S. Huang, "Locally Linear Embedded Eigenspace Analysis," IFP-TR, Univ. of Illinois at Urbana-Champaign, Jan. 2005.

[134] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin, "Graph Embedding and Extension: A General Framework for Dimensionality Reduction," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 29, no. 1, pp. 40-51, Jan. 2007.

[135] S.T. Roweis and L.K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science,* vol. 290, no. 5500, pp. 2323-2326, 2000.

[136] J.B. Tenenbaum, V. de Silva, and J.C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science,* vol. 290, pp. 2319-2323, 2000.

[137] D. Cai, X. He, J. Han, and H.-J. Zhang, "Orthogonal Laplacianfaces for Face Recognition," *IEEE Trans. Image Processing,* vol. 15, no. 11, pp. 3608-3614, Nov. 2006.

[138] X. He and P. Niyogi, "Locality Preserving Projections," *Advances in Neural Information Processing Systems,* MIT Press, 2003.

[139] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face Recognition Using Laplacianfaces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 27, no. 3, pp. 328-340, Mar. 2005.

[140] S. Yan, H. Wang, Y. Fu, J. Yan, X. Tang, and T.S. Huang, "Synchronized Submanifold Embedding for Person Independent Pose Estimation and Beyond," *IEEE Trans. Image Processing,* vol. 18, no. 1, pp. 202-210, Jan. 2009.

[141] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[142] A. Günay and V. Nabiyev, "Automatic Age Classification with LBP," *Proc. Int'l Symp. Computer and Information Sciences,* 2008.

[143] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 22, no. 10, pp. 1090-1104, Oct. 2000.

[144] Z. Yang and H. Ai, "Demographic Classification with Local Binary Patterns," *Proc. Int'l Conf. Biometrics,* pp. 464-473, 2007.

[145] C. Liu and H. Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition," *IEEE Trans. Image Processing,* vol. 11, no. 4, pp. 467-476, Apr. 2002.

[146] F. Gao and H. Ai, "Face Age Classification on Consumer Images with Gabor Feature and Fuzzy LDA Method," *Proc. Int'l Conf. Advances in Biometrics,* pp. 132-141, 2009.

[147] S. Yan, X. Zhou, M. Liu, M. Hasegawa-Johnson, and T.S. Huang, "Regression from Patch-Kernel," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2008.

[148] S. Yan, M. Liu, and T.S. Huang, "Extracting Age Information from Local Spatially Flexible Patches," *Proc. IEEE Conf. Acoustics, Speech, and Signal Processing,* pp. 737-740, 2008.

[149] J. Suo, T. Wu, S. Zhu, S. Shan, X. Chen, and W. Gao, "Design Sparse Features for Age Estimation Using Hierarchical Face Model," *Proc. IEEE Conf. Automatic Face and Gesture Recognition,* 2008.

[150] G. Guo, G. Mu, Y. Fu, and T.S. Huang, "Human Age Estimation Using Bio Inspired Features," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 2009.

[151] M. Riesenhuber and T. Poggio, "Hierarchical Models of Object Recognition in Cortex," *Nature Neuroscience,* vol. 2, no. 11, pp. 1019-1025, 1999.

[152] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust Object Recognition with Cortex-Like Mechanisms," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 29, no. 3, pp. 411-426, Mar. 2007.

[153] V.N. Vapnik, *Statistical Learning Theory.* John Wiley, 1998.

[154] H.C. Lian and B.L. Lu, "Age Estimation Using a Min-Max Modular Support Vector Machine," *Proc. Int'l Conf. Neural Information Processing,* pp. 83-88, 2005.

[155] G. Guo, G. Mu, Y. Fu, C. Dyer, and T.S. Huang, "A Study on Automatic Age Estimation Using a Large Database," *Proc. IEEE Conf. Computer Vision,* 2009.

[156] T. Kanno, M. Akiba, Y. Teramachi, H. Nagahashi, and T. Agui, "Classification of Age Group Based on Facial Images of Young Males by Using Neural Networks," *IEICE Trans. Information and Systems,* vol. E84-D, no. 8, pp. 1094-1101, 2001.

[157] S. Weisberg, *Applied Linear Regression,* third ed. Wiley Interscience, 2004.

[158] S.K. Zhou, B. Georgescu, X.S. Zhou, and D. Comaniciu, "Image Based Regression Using Boosting Method," *Proc. IEEE Int'l Conf. Computer Vision,* pp. 541-548, 2005.

[159] P. Viola and M. Jones, "Robust Real-Time Object Detection," *Int'l J. Computer Vision,* vol. 57, no. 2, pp. 137-154, 2004.

[160] B. Xiao, X. Yang, and Y. Xu, "Learning Distance Metric for Regression by Semidefinite Programming with Application to Human Age Estimation," *Proc. ACM Conf. Multimedia,* 2009.

[161] G. Guo, Y. Fu, T.S. Huang, and C. Dyer, "A Probabilistic Fusion Approach to Human Age Prediction," *Proc. IEEE Computer Vision and Pattern Recognition–Semantic Learning and Applications in Multimedia Workshop,* 2008.

[162] K. Ricanek and T. Tesafaye, "MORPH: A Longitudinal Image Database of Normal Adult Age-Progression," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition,* pp. 341-345, 2006.

[163] MORPH Face Database, http://faceaginggroup.com/, 2010.

[164] LHI Image Database, http://www.lotushill.org/LHIFrameEn.html, 2010.

[165] Human and Object Interaction Processing (HOIP) Face Database, Softopia Japan Foundation, http://www.hoip.jp/, 2010.

[166] A. Bastanfard, M.A. Nik, and M.M. Dehshibi, "Iranian Face Database with Age, Pose and Expression," *Proc. Int'l Conf. Machine Vision,* pp. 50-55, 2007.

[167] Gallagher's Web Collected Database, http://amp.ece.cmu.edu/people/Andy/ImagesOfGroups. html, 2008.

[168] A. Lanitis, "Comparative Evaluation of Automatic Age-Progression Methodologies," *EURASIP J. Advances in Signal Processing,* vol. 2008, pp. 1-10, 2008.

[169] N. Miyamoto, Y. Jinnouchi, N. Nagata, and S. Inokuchi, "Subjective Age Estimation System Using Facial Images—How Old We Feel Compared to Others," *Proc. IEEE Conf. Systems, Man, and Cybernetics,* vol. 4, pp. 3449-3453, 2005.

[170] Y. Fu, M. Liu, and T.S. Huang, "Conformal Embedding Analysis with Local Graph Modeling on the Unit Hypersphere," *Proc. IEEE Conf. Computer Vision and Pattern Recognition–Workshop Component Analysis,* 2007.

[171] X. Zhuang, X. Zhou, M. Hasegawa-Johnson, and T.S. Huang, "Face Age Estimation Using Patch-Based Hidden Markov Model Supervectors," *Proc. Int'l Conf. Pattern Recognition,* 2008.

[172] L. Lu, Z. Xu, Y. Hu, and P. Shi, "Age Estimation Robust to Expression Variations," *Proc. Int'l Conf. Computational Intelligence for Modelling, Control, and Automation,* pp. 198-201, 2008.

[173] M.A.O. Vasilescu and D. Terzopoulos, "Multilinear Analysis of Image Ensembles: TensorFaces," *Proc. European Conf. Computer Vision,* pp. 447-460, 2002.

[174] H. Wang and N. Ahuja, "Facial Expression Decomposition," *Proc. IEEE Conf. Computer Vision,* 2003.

[175] X. Geng and K. Smith-Miles, "Facial Age Estimation by Multi-linear Subspace Analysis," *Proc. IEEE Conf. Acoustics, Speech and Signal Processing,* pp. 865-868, 2009.
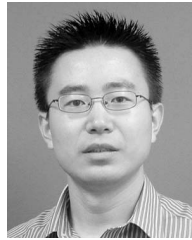
Yun Fu received the BEng degree in information engineering and the MEng degree in pattern recognition and intelligence systems, both from Xi'an Jiaotong University (XJTU), China, in 2001 and 2004 respectively; and the MS degree in statistics and the PhD degree in electrical and computer engineering, both from the University of Illinois at Urbana-Champaign (UIUC) in 2007 and 2008, respectively. From 2001 to 2004, he was a research assistant in the Institute of Artificial Intelligence and Robotics at XJTU. From 2004 to 2008, he was a graduate fellow and research assistant in the Beckman Institute for Advanced Science and Technology, ECE Department and Coordinated Science Laboratory at UIUC. He was a research intern with Mitsubishi Electric Research Laboratories, Cambridge, Massachusetts, in summer 2005; with Multimedia Research Lab of Motorola Labs, Schaumburg, Illinois, in summer 2006. He joined BBN Technologies, Cambridge, as a scientist in 2008, leading the computer vision and machine learning team. He held a part-time lecturer position in the Computer Science Department at Tufts University, Medford, Massachusetts, in spring 2009. He joined the Department of Computer Science and Engineering, University at Buffalo, The State University of New York, as an assistant professor in 2010, where he co-directs the Vision, Perceptual Machines, and Learning (VPML) Lab. His research interests include machine learning, human-computer interaction, image processing, multimedia, and computer vision. He has extensive publications in top journals, book chapters, and international conferences/workshops. He serves as an associate editor, chair, program committee member, and reviewer for many top journals and international conferences/workshops. He is the recipient of the 2002 Rockwell Automation Master of Science Award, two Edison Cups of the 2002 GE Fund "*Edison Cup*" Technology Innovation Competition, the 2003 HP Silver Medal and Science Scholarship, the 2007 Chinese Government Award for Outstanding Self-financed Students Abroad, the 2007 DoCoMo USA Labs Innovative Paper Award (IEEE ICIP '07 Best Paper Award), the 2007-2008 Beckman Graduate Fellowship, and the 2008 M.E. Van Valkenburg Graduate Research Award. He is a life member of Institute of Mathematical Statistics (IMS) and a Beckman graduate fellow. He is a member of the IEEE.

Guodong Guo received the BE degree in automation from Tsinghua University, Beijing, China, in 1991, the PhD degree in pattern recognition and intelligent control from the Chinese Academy of Sciences in 1998, and the PhD degree in computer science from the University of Wisconsin-Madison in 2006. He is currently an assistant professor in the Department of Computer Science and Electrical Engineering at West Virginia University. He visited and worked at INRIA, Sophia Antipolis, France, in 1997; Ritsumeikan University, Japan, in 1998; Nanyang Technological University, Singapore, from 1999 to 2000; and Microsoft Research, China, from 2000 to 2001. His research areas include computer vision, machine learning, and multimedia. He has authored a book, *Face, Expression, and Iris Recognition Using Learning-Based Approaches* (VDM Verlag, 2008), published more than 40 technical papers in face, iris, expression, gender recognition, age estimation, multimedia information retrieval, and image analysis, and filed three patents on iris and texture image analysis. He is a senior member of the IEEE.

Thomas S. Huang received the BS degree in electrical engineering from the National Taiwan University, Taipei, China, and the MS and ScD degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge. He was on the faculty of the Department of Electrical Engineering at MIT from 1963 to 1973, and on the faculty of the School of Electrical Engineering and the director of its Laboratory for Information and Signal Processing at Purdue University from 1973 to 1980. In 1980, he joined the University of Illinois at Urbana-Champaign, where he is now the William L. Everitt Distinguished Professor of Electrical and Computer Engineering, a research professor at the Coordinated Science Laboratory, the head of the Image Formation and Processing Group at the Beckman Institute for Advanced Science and Technology, and the cochair of the Institute's major research theme: human-computer intelligent interaction. His professional interests lie in the broad area of information technology, especially the transmission and processing of multidimensional signals. He has published 20 books and more than 500 papers in network theory, digital filtering, image processing, and computer vision. He is a member of the National Academy of Engineering, a foreign member of the Chinese Academies of Engineering and Science, and a fellow of the International Association of Pattern Recognition and the Optical Society of America, and has received a Guggenheim Fellowship, an A. von Humboldt Foundation Senior US Scientist Award, and a Fellowship from the Japan Association for the Promotion of Science. He received the IEEE Signal Processing Society's Technical Achievement Award in 1987 and the Society Award in 1991. He was awarded the IEEE Third Millennium Medal in 2000. Also in 2000, he received the Honda Lifetime Achievement Award for "contributions to motion analysis." In 2001, he received the IEEE Jack S. Kilby Medal. In 2002, he received the King-Sun Fu Prize, International Association of Pattern Recognition, and the Pan Wen-Yuan Outstanding Research Award. He is a founding editor of the *International Journal of Computer Vision, Graphics, and Image Processing* and the editor of the Springer Series in Information Sciences, published by Springer. He is a life fellow of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.