Learning Race from Face: A Survey

Siyao Fu, *Member, IEEE*, Haibo He, *Senior Member, IEEE*, and Zeng-Guang Hou, *Senior Member, IEEE*

Abstract—Faces convey a wealth of social signals, including race, expression, identity, age and gender, all of which have attracted increasing attention from multi-disciplinary research, such as psychology, neuroscience, computer science, to name a few. Gleaned from recent advances in computer vision, computer graphics, and machine learning, computational intelligence based racial face analysis has been particularly popular due to its significant potential and broader impacts in extensive real-world applications, such as security and defense, surveillance, human computer interface (HCI), biometric-based identification, among others. These studies raise an important question: How implicit, non-declarative racial category can be conceptually modeled and quantitatively inferred from the face? Nevertheless, race classification is challenging due to its ambiguity and complexity depending on context and criteria. To address this challenge, recently, significant efforts have been reported toward race detection and categorization in the community. This survey provides a comprehensive and critical review of the state-of-the-art advances in face-race perception, principles, algorithms, and applications. We first discuss race perception problem formulation and motivation, while highlighting the conceptual potentials of racial face processing. Next, taxonomy of feature representational models, algorithms, performance and racial databases are presented with systematic discussions within the unified learning scenario. Finally, in order to stimulate future research in this field, we also highlight the major opportunities and challenges, as well as potentially important cross-cutting themes and research directions for the issue of learning race from face.

Index Terms—Race classification, face recognition, image categorization, data clustering, face database, machine learning, computer vision

1 Introduction

Face explicitly provides the most direct and quickest way for evaluating implicit critical social information. For instance, face could convey a wide range of semantic information, such as race,[1] gender, age, expression, and identity, to support decision making process at different levels. Behavior research in psychology also shows that encountering a new individual, or facing a stimulus of human face normally activates three "primitive" conscious neural evaluations: race, gender, and age, which have consequential effects for the perceiver and perceived [2], [3], [4], [5], [6], [7], [8], [9] (see Fig. 1). Among which, race is arguably the most prominent and dominant personal trait, which can be demonstrated empirically by its omnirelevance with a series of social cognitive and perceptual tasks (attitude, biased view, stereotype, emotion, belief, etc.). Furthermore, it yields deep insights into how to conceptualize culture and socialization in relation to individual appearance traits, including social categorization [10], [11], [12], association [3] and communication [13]. Therefore, the estimation of racial variance by descriptive approaches for practical purposes is indeed indispensable in both social and computer science. However, while race demarcation drives the intrinsically genetic variation structure of essential facial regions to gather more explicit appearance information, the core question emerges as the computational mechanism underlying this extraordinary complexity. This raises the following fundamental multi-disciplinary conundrum: How does a computer model and categorize a racial face?

To answer this fundamental question, numerous research consortium and scholars have developed intensive investigations from different angles. For example, psychologists have studied behavior correlations of race perception such as other-race-effect (ORE) and attention model (e.g., [2], [8], [10], [14], [15]), which show existence of racially-discriminative facial features such as eye corners or nose tip (further anthropometric survey have confirmed those areas help to discriminate racial groups [16], [17], [18], [19], [20], [21]). Neurophysiologists have shown how race perception influences and regulates cognitive processes such as affection [22], [23], [24], [25] and stereotype [26], [27]. Computational neuroscientists have built models to simulate and explain

1. In general English the term "race" and "ethnicity" are often used as though they were synonymous. However, they are related to biological and sociological factors respectively. Generally, race refers to a person's physical appearance or characteristics, while ethnicity is more viewed as a culture concept, relating to nationality, rituals and cultural heritages, or even ideology. For example, detecting an Eastern Asian from Caucasian crowd is a race recognition task, while visually differentiating a German and a French belongs to ethnic category and thus requires extra ethnographically discriminative cues including dress, manner, gait, dialect, among others. Since there are over 5,000 ethnic groups all over the world [1], the idea of "ethnicity recognition" seems to be both questionable and impractical from current computer vision and pattern recognition point of view. Therefore, considering the soft biometric characteristics of distinctive human population, we prefer to use "race" as more suitable category terminology in this article. Nevertheless, for some of the existing papers in literature which have already used the term "ethnicity" but were indeed addressing "race" related issues, we have also cited and discussed those papers as well, in order to provide a comprehensive and complete survey on this topic.

- *S. Fu and H. He are with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881. E-mail: {fu, he}@ele.uri.edu.*
- *Z.-G. Hou is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. E-mail: hou@compsys.ia.ac.cn.*

Manuscript received 17 June 2013; revised 19 Mar. 2014; accepted 20 Mar. 2014. Date of publication 1 May 2014; date of current version 5 Nov. 2014. Recommended for acceptance by D. Maltoni.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2014.2321570

Fig. 1. Illustration of face-race perception. The quick glimpse of the picture will activate three "primitive" conscious evaluations of a person: a young *white* female, although this figure was a computer-generated "artificial" face with a mix of several races. (Photo source: The New Face of America, TIME Magazine, November 18, 1993.)

race perceptions (e.g., [3], [28], [29], [30], [31]). Other cognitive experiments in [11], [12], [32] have also indicated the existence of racial features as a salient visual factor. Among the general public, the validity of racial categories is often taken for granted. Converging lines of investigations provide a sterling evidence of perceptual discrimination relationship (albeit a rather quantitative one) between the implicit racial category model and the explicit face representation. Following these quantitative analysis, computer vision scholars have been motivated to tackle the inherent problem of racial demarcation by building computationally intelligent systems capable of categorizing races.

Yet, intuitively straightforward as it might seem to be, the implicit underlying algorithm implementation tends to be complicated and diversified. First, the nomenclature of race is a perplexing picture, as being defined qualitatively rather than quantitatively, painted by diverse perspectives rather than unified aspects. The ambiguity of the definition leads to the uncertainty of both problem formulation and development of methods, which differs it from other facial information analysis, such as face recognition (a specific one-to-one matching problem); facial expression (six universal emotional status [33]); and gender (only two classes). What's worse, the sensitivity to culture stereotype, racism and prejudice fuels much of the troubles, which in turn makes corresponding data collection and analysis difficult to carry on systematically. The controversy raised by definition ambiguity and the invalidity caused by data scarcity make computational classification approaches seldom work, which partially explains the scarcity of successful race learning methods and comprehensive perspectives compared to other facial data processing facets.

In spite of this, derived by burgeoning advances in applications such as security surveillance, human computer interaction, and bioinformatics, an increasing number of efforts have been reported toward race detection and categorization in the community. However, despite overwhelming theoretical principles, algorithms, applications and empirical results, there are relatively few crucial analysis that runs through these repertoire. Indeed, over the past few decades there have been several important work in face-based race perception available, such as [34], [35] in psychology, and [36], [37] in neuroscience. However, facing fast development of cutting-edge research and technology, reviewing efforts toward racial feature extraction and analysis cannot rely solely on these existing works due to following reasons: First, subjective cognitive experiments for testing human capability in race recognition cannot compare with automatic race recognition methods. Second, it is well known that automatic race recognition needs to be trained and generalized in a large-scale database, while subjects in traditional experiments can only face limited data displayed on screen. Third, emerging latest 3D scan technologies make 3D facial fiducial data far more convenient for computational recognition approaches rather than human perception, suggesting that existing survey papers may not be suitable for directing new computational research trends. Overall, the increasing contradiction between overwhelming social/scientific demographic data and scarce race mining methods calls for a comprehensive survey on computational race recognition to cover the state-of-the-art techniques and to give scholars a perspective of future advances.

Aimed at developing a unified, data-driven framework of race classification, the goal of this survey is to provide a critical and comprehensive review of the major recent advances in this field. Since "race" is a rather loosely defined and ambiguous term linked with multidisciplinary research areas, such a survey work would unavoidably be involved with a combination of both sole computer vision based analysis and psychological-physiological behavior experimental observation results. Note that while the former is the back bone of this survey, the empirical analysis and experimental support, however, come from the latter. In contrast to most previously published surveys in related fields, we focus mainly on the approaches that are capable of handling both computer vision and cognitive behaviors analysis. It is also noteworthy to point out that for face related research areas, there already exist several representative reviews and surveys, such as face recognition [31], [38], [39], [40], expression recognition [2], [38], [41], [42], [43], age estimation [44], [45], [46], gender recognition [45], [47], and even survey of face databases [38], [48]. However, to our best knowledge, there is no such comprehensive survey on race recognition, which also motivates us to present a review for completing the gap and enriching the research forefront.

The rest of this survey is organized as follows: Section 2 describes race perception formulation from multi-disciplinary perspectives, which serves as the motivations and foundations for race recognition. With these sketches of theoretical and applicational opinions, analytical feature representation techniques of identifying a racial face are detailed in Section 3, which also provides an overview of racially distinctive feature representation approaches, such as chromatic feature, local and global feature extractions, enabling us to design discriminative classification systems. Section 4 presents a comprehensive review of the state-of-the-art research on race recognition, including single-model/multi-model race recognition, 3D racial face analysis, and intra-ethnic recognition. In conjunction with reviewing of algorithms, we proceed to provide a summary and analysis of major available racial face databases in Section 5. In this section, we also provide several representative anthropometry survey data and racial face generation softwares. In Section 6, we highlight several representative applications of race recognition to demonstrate its critical potentials in a variety of domains, such as video surveillance, public safety, criminal judgement and forensic art, and human computer interaction. In order to motivate future research along the topic, we further provide insights on the challenges and opportunities of race recognition in Section 7. Finally, Section 8 concludes the survey.

Fig. 2. Illustrative example showing that skin tone is not the major determinant of perceived racial identity: observers look at a series of facial photos where a central face (racially different from the surroundings) appeared lightened or darkened, which nevertheless produce no influence on their final race identification judgement of the central face. The behavior result showed that race perception is not determined by skin color information, but by morphological characteristics of the subject (Figure source: [55]).

2 Human Race Perception Formulation

Given its both theoretical and practical purposes, race perception is inherently a multidisciplinary challenging task linking various research fields together, such as psychology, cognitive neuroscience, computer vision and graphics, pattern analysis, machine learning, among others. Therefore, the progress during the investigation is undoubtedly contingent on the progress of the research in each of these fields. Our survey thus starts from the fundamental and analytical understanding of race based on interdisciplinary expertise.

2.1 The Conceptual Description of Race

We begin by briefly introducing the basic interpretation and conceptualization of race, which are crucial for providing information about the racial features of which recognition algorithms or systems are based on, thus, "*What is race*"? From the definition of Wiki encyclopedia, the term "race" is defined as follows [49]: "Race is a classification system, used to categorize humans into large and distinct populations or groups by heritable, phenotypic characteristics, geographic ancestry, physical appearance, ethnicity, and social status."

Apparently, this definition is very "fuzzy", so apt to perpetuate confusion and engender discord for developing race recognition systems.[2] For the research purposes, it is worth highlighting several important points of consensus that have emerged from psychological-physiological research traditions:

Perhaps the most prominent, yet wrongly understood aspect of race is the common trap of confusing race with skin color. Although classification by using skin tones will simplify the problem greatly (e.g., several attempts have used color to perform race classifications, such as [51], [52], [53], [54]), skin color is such a variable visual feature within any given race that it is actually one of the least important factors in distinguishing between races and consequently, should be taken with care in applications [55], [56], [57], [58] (see Fig. 2).

Second, physical characteristics such as hairshaft morphologic characteristics and craniofacial measurements are viewed as significant indicators of race belongings. However, for computer vision methods, during preprocessing normally a face mask will be applied to cut them off, leaving only face region standing out, thus making those racial cues little assistance in applications. While in visual surveillance, those non-face visual appearances could be employed as apparent visual cues and provide salient features for subject tracking and recognition [59].

Third, in an effort to define a clear pattern classification system, such aforementioned visible physical (phenotypic) variant features must be associated with large, geographically separable populations (groups). Therefore, it is commonly acceptable that, in a rough real-world application sense, classification systems designate these continental aggregates as race labels, such as the European/Caucasian race, or the Asian/Mongolian race [60], [61], which are also known as the commonly perceived categories of race. Note, however, that such groups can range from three basic ones to more than 200 detailed branches, judged by the specific criterion [60]. Nevertheless, for practical computer vision based race classification system we suggest that seven most commonly encountered and accepted racial groups (African/African American, Caucasian, East Asian, Native American/American Indian, Pacific Islander, Asian Indian, and Hispanic/Latino, all these cover more than 95 percent world population) will be enough (see Fig. 3 for an illustration).[3]

2. We would like to note that although most people think of races as simple physically distinct populations, recent advances from bioinformatics, on the other hand, demonstrated that human physical variations do not fit an exact racial model. Thus, there are "no genes that can identify distinct groups that accord with the conventional race categories" [50]. In a rough nut shell, the concept of race has no biological validity, which is a theoretically correct viewpoint that is conclusive but actually unpersuasive for computer vision based classification systems. This survey focuses mainly on physical traits which represents basic racial groups.

3. We would like to point out that the aforementioned discussion is a general statement after comprehending views from multi-disciplinary fields such as biomedical and genetics [62], [63], [64], anthropology [60], demographic census [65], and other general perspective [66]. Apart from the five commonly accepted major race groups, Asian Indian and Hispanics/Latinos are included in our system due to their population ratio and general social/public acceptance, although they are generally not viewed as separate race groups (Also note that Hispanics and Asian Indian could be ideal for the multiracial people recognition, see Section 7.3.4 for details.). We believe that such a racial categorization system is suitable for computer vision and pattern recognition both from the research and application points of view. Note that, however, such a classification scheme can be very flexible and may vary in accordance with development of sociology and scientific discoveries in the near future.

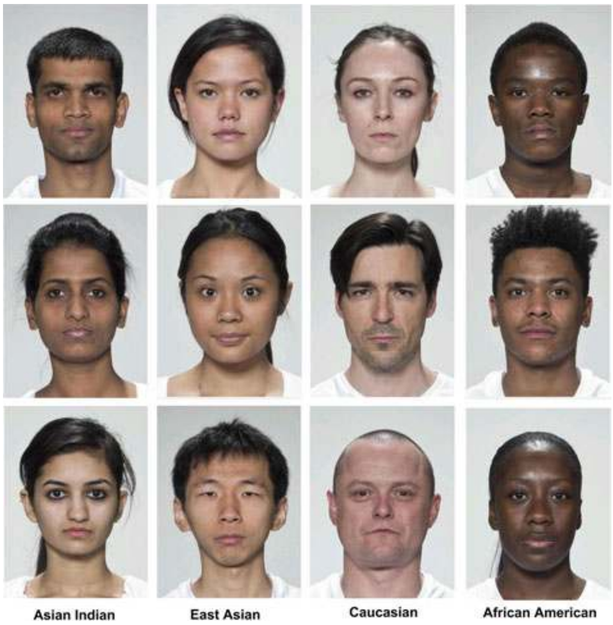

Fig. 3. Illustrative genetic variance distribution of human races, while in practice it is often accepted that 3- to 7-races classification system would be enough for regular applications (Figure source: http://www.faceresearch.org/).

Motivated by the above plethora of aspects, the definition of race classification in computer vision can be expressed as follows: Consider a data set P of all possible facial images I_i, which are assumed to belong to several pre-defined racial groups G_1 to G_n. A group G_i is called racialized if its members can be agglomerated by either explicit or implicit patterns (physical appearance observations, such as skin tones, morphology, or feature vectors, etc.) to be appropriate evidence of ancestral links to a certain geographical region.[4] Accordingly, the race classification/categorization issue can be expressed as an exploratory image analysis from a statistical pattern recognition perspective, that is, to design a classification system which is capable of determining or inferring these racial identities from given facial images with minimum ambiguity.

2.2 Race and Other-Race Effect (ORE)

Principally related with the race perceptual model is the so called other-race-effect (ORE, also called “own race bias”, “cross-race effect”). It means when given a face recognition task, human observers tend to perform significantly better with their own-race as compared to other-race faces (Fig. 4), a concept with more than 40 years’ history which has been documented repeatedly in numerous meta-analysis [34], [35]. Though varied explanations have been proposed to account for ORE by ethographers, cognitive anthropologists, sociologists, and even policy and law enforcement officer [34], [68], the fundamental reason still remains elusive [69]. Nevertheless, it is interesting to point that there lies some subtle yet complicate connections between ORE and their influence on automatic racial face processing. For example, Phillips et al. investigated the ORE effect for face recognition algorithms [70]. They found that algorithms developed by Westerners tend to recognize Caucasian faces more accurately than East Asian faces and the vice versa, indicating face recognition algorithms also favor the “majority” race data in the database, just like what human subjects have shown for the standard ORE effect. Dailey et al. also found the existence of in-group advantage for seeking evidence and a computational explanation of cultural differences in facial expression recognition [71]. The transduced conclusion is that ORE affects the training of recognition algorithm (similar to human socialization), thus a balanced face database could be essential in the race recognition in order to counter the biased results. This is also related to the imbalanced learning problem from computational analysis point of view, in which interested readers can refer to [167] for a comprehensive and critical review.

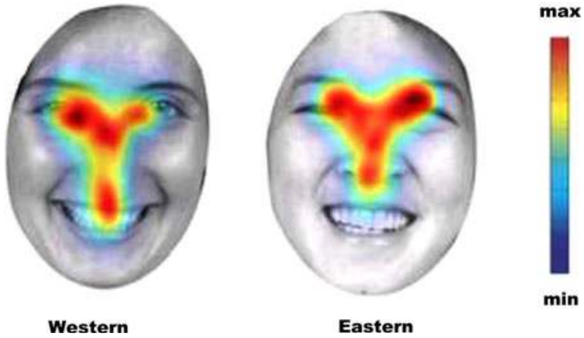

Fig. 4. Eye tracking data shows different attention regions for Western and Eastern subjects (observers) viewing same race and other race faces from JACFEE [72]. Left part: Western subject’s gazing region for Western face. Right part: Eastern subject’s gazing region for Eastern face. As can be seen from the figure, fixation distributions for both observer group varied significantly (color-area indicates the density of fixations), which can be roughly grouped into nose-centric for Western (evenly distribution across the face) and eye-centric for Eastern (biased for upper part of face). Biased attention may cause ambiguity in affective perception (Figure source: [73]).

2.3 Race and Physical Anthropometry

The definition of race emphasizes its conceptualism in an physical anthropology, which dates back even to Greek neoclassical canons (c. 450 B. C.) era. The born of anthropometry implies that race could possibly be measured and classified based on predefined standard measurements, with which a number of methods have been devised for calculating racial types [74]. Among which, L. Farkas was the known pioneer for establishing an index of 42 discriminative anthropometric facial points (ranked by mean and standard deviation) from a consolidation list of 155 cranio-facial anthropometric proportions [75], [76]. More systematic review of inter-ethnic variability in facial dimensions [77] and recent survey results in American [16], [17], [18] and Chinese demographic data

4. Note that physical morphology and geography not only play key roles as basis to justify the treatment in race classification problem formulation, they also make the definition flexible, which is in accordance with the fact that race is not rigidly defined as other facial traits such as gender or expressions. For example, given this definition, we can say a facial image I_i is of the “White” race if “White” is a racialized group G_i in data set P, and I_i is a member. As Haslanger states: “Whether a group is racialized, and so how and whether is raced, is not an absolute fact, but will depend on hierarchical context” [67]. On this view, American Indians can be racialized and labeled both anthropometrically and algorithmically insofar as there exists enough representative and annotated images in data set, otherwise they could only be demarcated as Asian due to subordinated hierarchy relationship. Nevertheless, they still belong to a different “racial” group because of their distinctive features, even the exact definitions are missing.

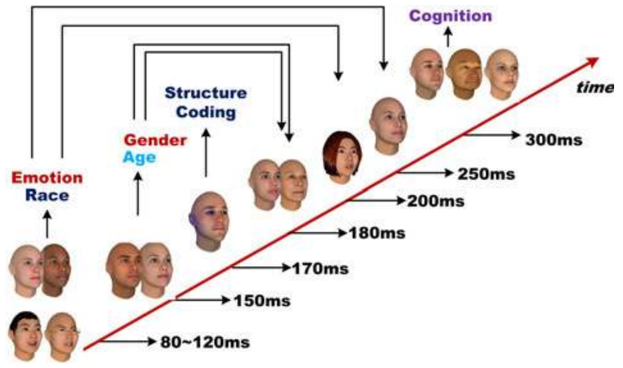

Fig. 5. Illustrative time course of face perception in human brain, recorded by ERP potential peaking. Race sensitive modulation was observed mainly in N100 (mean peak latency around 120ms) and P200. Gender modulation was observed in P200 (mean peak latency around 180ms). N170 component indicates structural encoding of face (versus non-face). Social membership cognition was observed in P300. Data source: [83].

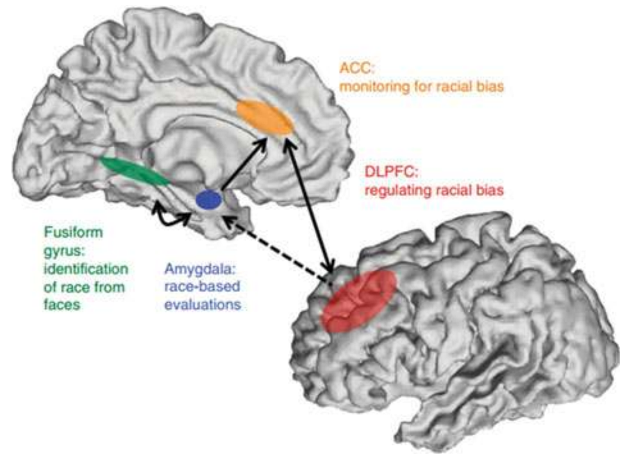

Fig. 6. Illustration of race processing regions [36].

[19], [20], [21] have also confirmed statistically significant variances in facial anthropometric dimensions between all race groups. The highlights of these variations in inter-ethnic phenotypic differences pave the way of anthropometry based automatic race recognition.

However, the drawbacks of traditional anthropometric survey sampling methods (incapable of digitalization) and data (usually univariate, unable to accurately describe 3D shape) prevent it from direct application. Fortunately, the availability of the latest 3D scan techniques (e.g., Laser scanners, CT scanner, Ultra sonic scanner, PET, etc.) is a welcome stimulus to this challenging task, offering reliable and accurate digital head and face data. Nevertheless, a fundamental question is how to incorporate those anthropometric data into computer vision based race recognition approaches because anthropometric meaningful landmarks vanish in 2D frontal images. The survey results in [78] implied that the combination of 2D and 3D images might provide a feasible solution (Detailed in Sections 4 and 5).

2.4 Race and Cognitive Neuroscience

How we perceive and categorize race and how social categories of race are processed and evaluated have long been a central topic of cognitive neuroscientists. With the development of analysis tools (e.g., fMRI, ERP, EEG) [2], [10], [14], [15], [36], [79], [80], [81], [82], [83] and theoretical cortex models embedded with computational algorithms [84], [85], [86], the research on race perception by fusing those multimodality physiological data have been greatly expanded in recent decades. For example, temporal ERP observation results in [23], [83], [87] have shown that race based modulation can be activated as early as around 100 ms, and race effects have reported repeatedly among many essential ERP components, such as N170,[5] P200, N200 and P300 (see Fig. 5 for illustration). The fast response to race label implies that race categorization might occurs at very early stages of perceptual face encoding and thus can significantly influence face processing. Studies using fMRI and ERP to investigate the neural system sensitive to race consistently report activation in an interconnected regions, including amygdala [88], [89], [90], [91], [92], anterior/posterior cingulate cortex (ACC/PCC) [87], [93], ventral/dorsolateral/Medial prefrontal cortex (VLPFC/DLPFC/MPFC), and fusiform gyrus (FFA) [94], [95], [96]. Fig. 6 presents an illustrative example of hypothetical race processing cortical areas. These fMRI results, coupled with the ERP data, confirm the key role race plays in a series of individual and social processing, such as stereotype [26], [27], prejudice [26], [97], [98], emotional and affective regulation [73], [99], [100], [101], [102], [103], personality [104], and other behavior and decision making applications [22], [23], [24], [25]. The immediate implication of these studies are: (1) Race encoding and categorization are subconscious and mandatory; (2) There exists a specific race processing unit involved with other cortical processing (similar to face processing). Apparently, these behavior findings may shed light on developing cortex-like computational intelligent systems for race recognition.

As we could see from above, recent decades have witnessed a surge in understanding of how our brain percept and recognize racial face information, and the richness of perceptual experience also highlights the pre-eminence of race in automatic human demographic information categorization. It can be generally concluded that much of current research achievements have been guided by two following distinct yet closely connected approaches: one is driven primarily by implicit neurophysiological observation data, and the other is governed by explicit physical appearance and psychophysics. This suggests that, by melding neurocomputing models and physical discriminative features, the computer vision approach could establish a unified framework for race classification. However, simple combination will not facilitate definition of algorithms, rather, it leaves crucial questions. Indeed, as much as we understand where racial face encoding and decoding process are produced cognitively, very little is known on how they are represented, processed and interpreted by our visual system. Therefore, without guiding from computer vision

5. For N170 component (modulation for face and non-face stimuli, thus, reflect structural face encoding), conflicting observations were reported concerning its sensitivity to racial faces. The reason might be due to variation of devices, subjects, and experiment conditions, the modulations may differ in response to various face stimuli, causing observation inconsistent or even fails among studies. Therefore, those results should be taken with care.

models and decision making systems, the empirical experimental data accumulated so far, the design of intelligent bioinformatics verification platform, as well as smart HCI system will remain elusive. It is also important to note that rather than view visual-based racial face analysis as simply a hyperfusion center, where different aspects of computer vision theories, models and techniques being welded together, we claim it is more plausible to treat such racial face analysis as a particular subset of the more general cognitive systems which comprehends multi-modality perception characteristics of face (see Section 7.2 for detailed discussion on this prospect). When it comes to relevant real-world application, this boils down to two key questions:

1) Which parts of the face (by either arbitrarily dividing the face into several anatomically meaningful regions such as periocular regions, nose, and mouth; or simple holistic face) are most important for a given racial face classification/recognition task?
2) How to extract meaningful features from these face regions to train a corresponding classifier or build a model?

Aiming to answer these questions, significant efforts have been made during the past decades for developing reliable appearance descriptors, compact representation methods as well as statistical discriminative classifiers, as will be discussed in following sections.

3 Racial Feature Representation

3.1 Overview

To simulate how human visual system works for race perception, the first step is to find meaningful representation of the racial face one would like to recognize. Early works involve derivation of color, texture, and shape based algorithms, also called appearance-based approaches, in which color, texture, and geometric features were extracted from an racial face and used to build a model. This model was then fitted to the testing images for comparison and recognition. Advances in pattern recognition and machine learning have made this the preferred approach in the last two decades. Although simple, computationally efficient, and the performance is acceptable from engineering practice viewpoint, these methods actually don't follow the face encoding mechanism used by human visual system, which turns out to be feature based during the rapid categorization [105]. What's more, they would perform poorly when facing image manipulation such as scale and illumination variation. Later approaches followed feature based track and recent advances seem to consider both configural and shape features, making the algorithms more robust. We start by reviewing several standard techniques for racial feature representation in this section.

3.2 Racial Feature Representation

As racially distinctive information is encoded by the visual appearance of face, racial information processing system begins with preprocessing and extraction of race-discriminative features. The goal is to find a specific representation of the face that can highlight relevant race distinctive information. According to the influential multi-level encoding model [4], [5], human faces are distinguished by their characteristic cues, ranging from the basic global/holistic level (also called first-order relations), to the detailed feature analysis (coined as second-order relations), till the final configural perceptions of both levels (for specific individual identification) [2], [106]. The diagnostic race categorization requires knowledge from second order level, which means the main work of the current researches shall consequently be focused on race sensitive feature representation and extraction techniques. In the following we present these representative feature extraction methods and sort them qualitatively as opposed to quantitatively:

- *Chromatic representation.* Two dichotomies commonly used for comparison in race classification are that of Asian versus Non-Asian, or Caucasian versus Non-Caucasian. Skin tones have long been employed as the primary features to perform such rough race classifications [51], [52], [53], [54], and the results seemed to be satisfactory. However, as illustrated previously, there exist several serious drawbacks for this rudimentary feature: First, there are many people from different racial groups who share same skin color. For instance, if judged by skin color, then all Southern Indian/Austrilian/Melanesia/Africans would be clustered together for dark-skin tones, though they apparently belong to different racial identities. What's worse, skin color is highly sensitive in uncontrolled illumination environment and direct employment may cause severe errors [55], [56], [57], [58]. To summarize, the strong claims put forward by sole skin color based race recognition can be safely refuted, not only on the basis of illumination variant characteristic of color appearance, but more compellingly because of the conclusive evidence from psychology studies [55], which demonstrates that skin color bears virtually no relationship to race perception. Nevertheless, the contribution of these aforementioned researches lies in that visual appliances of chromatic attributes still remain essential since fusion of skin tone and facial metrics can by all means boost the performance on racial prototypicality judgement (see Section 4).
- *Global feature representation.* Holistic representation is arguably the most typical technique to be used in race recognition for its capability to preserve configural (e.g., the interrelations between facial regions) information, which is essential for discriminating race identity. For example, PCA is generally preferred to reliably extract and categorize facial cues to race [5]. Phillips and O'Toole et al. have investigated PCA-based methods in [70], [107], [108], [109], [110], [111] with images from Japanese and Caucasian subjects, and their conclusion comfired that the race could be predicted with relatively good accuracy ($\geq$ 80 percent) with PCA. PCA has also been successfully used in [112] for Myanmar and Non-Myanmar classification, in [113] for Arabic, Asian, and Caucasian classification, and in [52] for Asian, African-American, and Caucasian

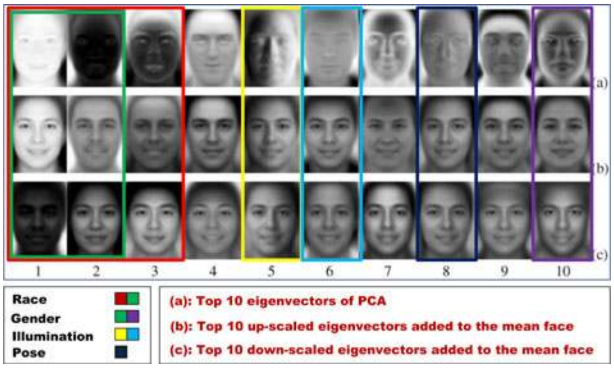

Fig. 7. Illustrative visualization representing the relationship between eigenface and facial information, note that (b) and (c) are multiplied by 3σ, the square root of eigenvalue. Clearly the facial properties such as race, pose, illumination and gender are controlled by the different eigenfaces [114].

classification, even in neural visual processing for race categorization [111]. Those studies, together with exhaustive work [114], [115], provide evidence that the eigenvectors with large eigenvalues are referred to as visually derived semantic information about racial faces (see Fig. 7). In short, PCA is still one of the most frequently used statistical feature extraction methods in racial feature representation.

- *Local feature descriptor representation.* Compared with global representation, local feature descriptor has a number of advantages in the context of unconstrained racial face recognition. For instance, being viewed as a reliable and robust appearance feature descriptor, Gabor wavelet representation is undoubtedly suitable for race categorization applications such as in [116], [117], [118], [119]. However, Gabor feature vector resides in a space of very high dimensionality, so dimension reduction techniques are needed to acquire a more sparse, decorrelated and discriminative subset. For example, Adaboost [118], optimal decision making rule [120] and QuadTree clustering [121] have been employed to select a compact subset yet preserving the discriminativeness. Other local descriptors have also been investigated by scholars. Fu et al. embedded topographic independent component analysis (TICA) to form a hierarchical multi-level cortex-like mechanism model to recognize facial expressions for subjects from different Chinese ethnic minorities [122]. Experiments show that such an ICA-based system achieves a classification rate of 82.5 percent. Weber local descriptors, wavelet, and local binary patterns (LBP) have been investigated respectively in [123], [124], [125], classification results on five race groups from FERET database showed their effectiveness and superior performance over holistic PCA. Other low-level based features such as gradient direction histograms, multiple convolution network generated features (Fig. 8), and wavelet features were also discussed in [126], [127], [128], [129].

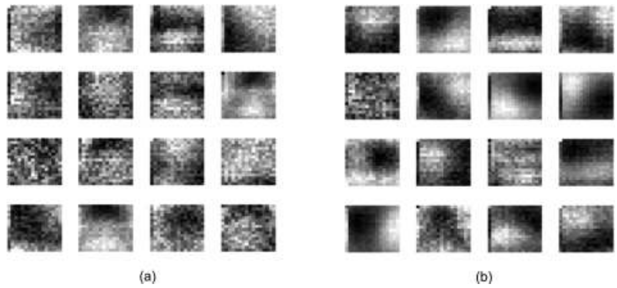

Fig. 8. Discriminative race feature representation by multiple layer convolution neural networks (CNN). (a): supervised CNN filters, (b): CNN with transfer learning filters [126].

- *Other representations.* Keeping in view of the above discussion it is hard to define the optimal representation[6] way of racial features in real-world applications, therefore several scholars have tried other alternative ways to preserve the configural racial information of the facial parts either implicitly or explicitly by compromising all those aforementioned representation methods to form a hybrid representation scheme. Such case can be exemplified by Ding et al.'s effort of boosting local texture and global shape descriptor together [130]. By doing so a holistic representation can be obtained for several local regions of face and similarly a local compact description can still be obtained by concatenating several locally processed regions of the face into one global vector. Those methods are still characterized as local since they use several local patches of the face, but they are simultaneously holistic in nature. Some of the typical combinations are fusing skin color, Gabor feature, local facial parts, and PCA together. As illustrated in the following section, this concept yields several unique approaches for race recognition tasks with satisfactory performance.

To conclude this section, it is very interesting to discuss how human use different salient facial features for race classification. Ongoing efforts within cognitive neuroscience, pattern recognition, and advanced human-machine systems have indicated the hypothesis that racial face processing is actually a tradeoff between selectivity and invariance for efficiency and accuracy, with increasing complexity to facial appearance (rotation, scale, lighting, etc.). In other words, the racial face representation is rather dynamic and adaptive, corresponding to the recognition distance and image resolution. For degraded facial images due to distance, the racial face is basically treated as a whole object, in which skin color and holistic face information would be employed, which is fast and insensitive to local distortion and partial occlusion (glasses, hair, etc.). For dealing with high resolution image or close range, geometrical method and local representation will be involved for more effective visualization, which is computationally expensive yet more robust to

6. Although several other feature extraction techniques are available. For example, anthropometric measurements can statistically reveal the most important facial feature areas (e.g., mouth, eyes and eyebrows), they are excluded because they belong to directive statistical analyzing manner rather than learning. However, the detailed discussion of these techniques will be presented in next section.

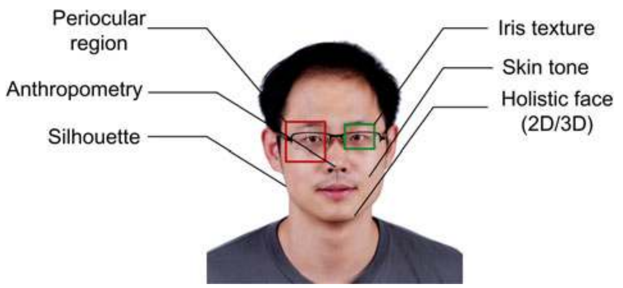

Fig. 9. Illustrative example showing racial sensitive regions.

physical variation (pose, lighting). Close range discrimination process would activate algorithms focusing on racially salient face features (eyes, mouth, chin, and nose) and extract their correspondent geometrical points, as well as more compact representations. Processing these features will require pattern recognition techniques, which will be presented in the next section.

4 Race Classification: State-of-the-Art

It is well recognized that face perception in nature is piecemeal integration process [131]. Accordingly, facial components for race classification are processed in a holistical and comprehensive way. Eye tracking results [73], [102] clearly reveal that there exists a ranking list for those racially sensitive facial components, implying a configural relationship among these features/regions. For example, the eyes and nostril part have been consistently reported to be the most robust parts of the face verification, they will naturally provide essential information that has to be taken into account of race classification. Below, we exhibit those feature's contributions from a fine-to-coarse roadmap, which corresponds to the recognition accuracy from the intrusive, close range in which single model will be sufficient, to the distant, non-intrusive range, which requires cooperation from multiple modality features (see Fig. 9).

4.1 Single-Model Race Recognition

As mentioned above, behavior observation data have shown that explicit racial categorization task is closely related to a comprehension work, in terms of both overall physical appearance featured (such as skin tones) or local discriminative regions (such as eye socket and nose). Consequently, it is reasonable that multi-modality approaches shall outperform either mode alone. However, restricted by the data set and computational burden, the majority of research has been focused on single model based race classification. Below we present several characteristic single model based approaches using various facial parts.

4.1.1 *Iris Texture*

Iris is arguably the mostly exploited biometrics in all kinds of subject verification and identification applications. Statistically significant race differences in retinal geometric characteristics have been reported in several behavioral studies [132], [133]. The study of employing iris image on race classification initiated from [134], which had used Gabor feature and Adaboost selection combination on a combined iris data set for two-races classification. Their conclusion was that iris is race-related, a result which has been further confirmed by [135], [136], [137], [138], [139], [140], all of which fit well with race recognitions (Asian/Non-Asian, Asian/Black/Cucasian). Specifically, Qiu et al. [135] investigated the relevance between race group and iris texture, with conclusion that race information is illustrated in iris texture features, with best classification rate of 88.3 percent by SVM. Similar results were obtained by Lagree and Bowyer[138] with 90.58 percent using sequential minimal optimization algorithm. Zarei and Mou [136] applied multilayer perception (MLP) neural networks to achieve the corrected recognition rate of 93.3 percent. The top result is 96.7 percent from [137] with the combination of supervised codebook optimization and locality-constrained linear codeine (LLC). However, reminding that in typical real-world applications race is always viewed as a crucial and coarse soft-biometric cue, thus normally required being identified from distance or low quality video in a non-intrusive and computationally efficient manner. Therefore, iris-based race recognition, as an intrusive and complicated acquisition procedure, is more theoretically feasible rather than practically applicable. Iris texture categories that are closely correlated with race may be of value in data retrieval system issues.

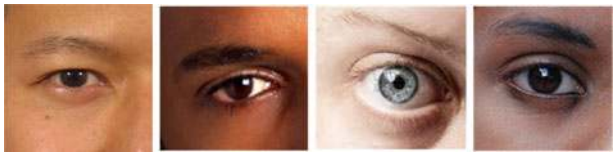

Fig. 10. Illustrative example showing different periocular region, from left to right: East Asian, African-American, Caucasian, Asian Indian. Compared with iris image, periocular region could provide more rich semantic information for race classification (Source: Google Images).

4.1.2 *Periocular Region*

Compared with iris, periocular region (including eyelid, eyelash, canthus or epicanthal, among others) is more easily acquirable, rich in texture, and more quantified as useful biometric area (e.g., canthus is essential for differing Caucasian from other races because Caucasians tend to have a distinct cusp, see Fig. 10, for example). The idea that periocular region contains discriminative racial information has received increasingly empirical confirmation recently, even in infrared images [141]. Among which, Lyle et al. [142] have used periocular region to perform race identification as Asian/non-Asian. Local binary patterns and grayscale pixel intensity have been used to evaluate their proposed system on the FRGC face data set and a baseline accuracy of 91 percent has been achieved. Li et al. [143] have investigated distinguishing features around the specific eyelash region, they located eyelash region by using active shape model (ASM) to model eyelid boundary and extracted nine local patches around it. The goals were to segment eyelash and to generate global eyelash direction distribution descriptor with which the nearest-neighbor classifier was performed. Their experimental results showed 93 percent accuracy for East-Asian/Caucasian classification. Xie et al. [52] have used kernel class-dependent feature analysis (KCFA) combined with facial color based features for large-scale ethnicity classification. Focusing on the periorbital

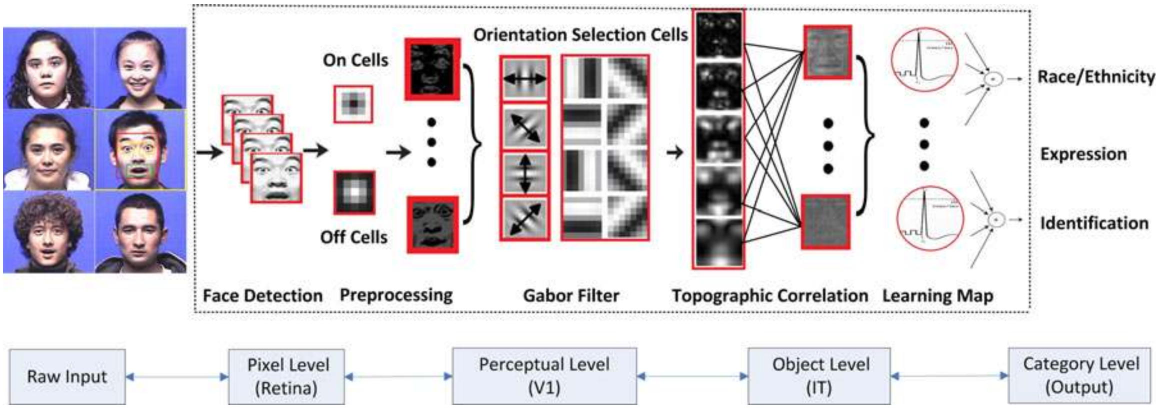

Fig. 11. Framework of a typical racial face processing system. Following the basic visual cortex scheme, the preprocessing part includes the detection of facial regions, illumination normalization, and edge detection. The second level functions like Gabor filter, sending output to the perceptual level for extracting features robust for selectivity and variance, after being grouped and classified, the category level gives the output [122].

region, the facial color based features were employed to incorporate with the filtered responses to extract the stable ethnicity features. Compared to previous approaches, their proposed method achieves the best accuracy of ethnicity classification on a large-scale face databases and the standard MBGC face database.

4.1.3 Holistic Face

Up to now, most existing approaches have been focused on holistic, frontal faces (see Fig. 11 for illustration). The early work originated from Gutta et al. [144], [145], [146] who have reported systematic investigation of using RBF neural network with decision tree for race classification on FERET database. Their best performance was 94 percent. Guo and Mu [147] have adopted biologically-inspired features in their ethnicity classification system. It showed that the proposed ethnicity classification algorithm's accuracy can be as high as 98 percent within the same gender group. They further investigated the canonical correlation analysis (CCA) in solving the joint estimation problem of three facial cues (race, gender, age) simultaneously [148]. Their results showed that CCA could derive an extremely low dimensionality in estimating these demographic information, indicating the potential for practical applications since CCA is very fast and computationally convenient. Lu and Jain's work [149] can be viewed as a classical example for performing race classification on holistic facial images, which employed LDA in an ensemble framework of multiscale analysis for Asian/non-Asian classification task. The experimental results on a manually mixed database acquired 96.3 percent accuracy overall (best 97.7 percent). A more sophisticated probabilistic graphical model based method has been proposed in patent [129], which constructed a filter pool for facial images and chose ethnicity-representative filter groups for given ethnicity class to build an ethnic class-dependent probabilistic graphical model. By computing the likelihood score from responses to each models of test image the classification result can be inferred. However, the racial group numbers and the algorithm's performance were not stated.

As mentioned above, compared with cropped face, extra frontal facial regions (such as hair) and their combinations with specific facial components (such as eyes and nose) have also been investigated to enhance the overall recognition performance. Li [150] has acclaimed a patent for ethnicity classification, in which a combination of features, namely block intensity and texture feature (BITF), and a LDA classifier were proposed to four racial groups (Asian, Caucasian, African and others). The contribution of this method is that the inventor took soft biometric features (hair color, eye color and eyebrow-to-eye metrics) into consideration, further boosting the system's performance. Lei et al. [151] have proposed another comprehensive approach using four face components (whole face, eyes, nose and mouth) and four low-level feature (HoG, color moments, Gabor, and LBP), a 16 combination (e.g., <eye, Gabor>) sets were formed. A two-level learning strategy (SVM and Adaboost) was applied to find the optimal relevance among the face region and feature (e.g., <whole face, color> is effective for African attribute). Then sparse coding has been adopted for face retrieve framework. The highlight of their work was the experiment on a considerably large-scale Flickr data set of more than 200k face images, achieving a significant improvement of hit rate@100 (from 0.036 to 0.42).

Particular attention should be paid on an interesting template-based ethnicity classification work by Manesh et al. [120], who have employed optimum decision making rule on the confidence level of automatically separated face regions using a modified Golden ratio mask. Instead of performing Gabor feature extraction and classification directly, extracted Gabor feature vector of each patch was treated as a single pattern classification problem and recognition accuracy of these classifications indicated the confidence level of each patch for ethnicity classification, which were fused together to acquire final classification results. In a combined database from FERET and CAS-PEAL for Asian/Non-Asian classification, they obtained recognition results as high as 98 percent.

The fusion of both global and local facial features for race classification has also been exploited by scholars recently. For example, Salah et al. [125] proposed a fusion scheme which used block-based uniform local binary patterns and Haar wavelet transform. K-Nearest Neighbors (KNN) classifier on EGA database for three races (European, Oriental, African) classification obtained average results of 96.9 percent.

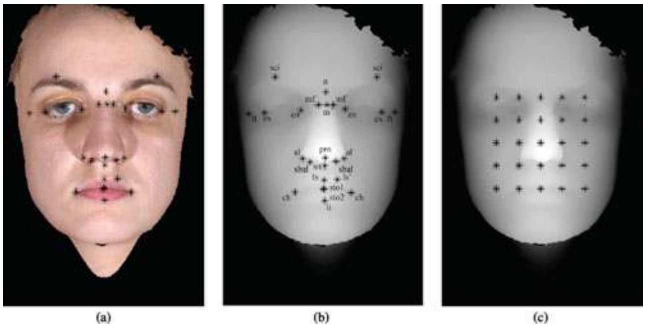

Fig. 12. The anatomically anthropometric landmarks used in [155]. It should be noted that those landmarks are highly redundant, therefore, how to chose a distinct subset according to some predefined criterion is essential.

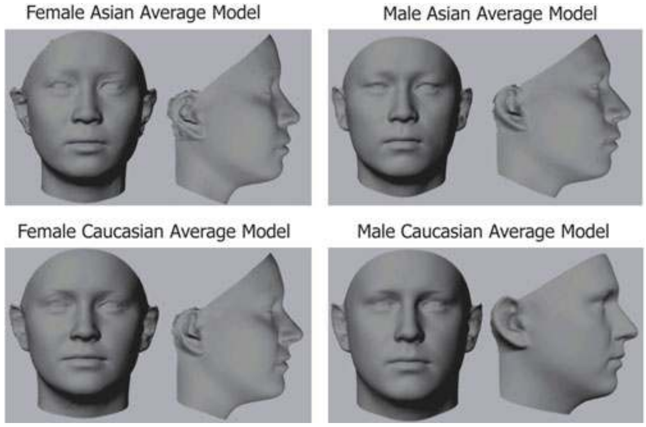

Fig. 13. 3D generic models and depth image of different races used in [119] for 3D reconstruction.

4.1.4 3D Faces

2D based racial face classifications normally encounter the difficulty when facing geometric and illumination variation, to which 3D model based approach is rather insensitive. Recently several scholars have began their research for 3D racial face classification without the associated texture or photographic information. To this end, facial geometrical structure is explored for potential discriminativeness. Lu et al. [152] has proposed an integration scheme of using both registered range and intensity images for ethnicity identifications. Toderici et al. [153] proposed a framework for both ethnicity and gender based subject retrieval. Using pure facial-structure-based metric function measured from Harr wavelet and CW-SSIM (structure similarity) coefficients, they evaluated four types of classification methods: KNN, kernelized KNN, multi-dimensional scaling (MDS) and learning based on wavelet coefficients. They reported high level of classification performance on FRGC v2.0: 99 percent mean accuracy for MDS. Their results indicated that it is possible to recognize race information with only 3D mesh of human face. Similar results have been reported by Ocegueda et al. [154], who also investigated finding the most discriminative regions of the face for 3D racial face classification. Using Gauss-Markov posterior marginals (GMPM) for computing discriminative map of subjects from BU-3DFE database (Asian/White), they performed cross validation on FRGC v2.0 for the aim of comparing with Toderci's work. The results turned out to be very competitive with a much simpler linear classifier. However, their work was based on the assumption of smooth distribution of facial discriminative information, i.e., it is unlikely that a small isolated facial region contain high discriminative information while its neighbor regions do not. From a computational point of view, there is no guarantee that this assumption can be generalized to other races. Zhong et al. incorporated Gabor features with QuadTree clustering to form a visual codebook for both western and eastern subjects [121] by using Max distance function and fuzzy membership function, they proposed a fuzzy 3D face categorization approach, which reached performance as high as 80.16 percent for eastern and 89.65 percent for westerners on FGRC 2.0 3D face database. However, a rough two-class categorization limits its further analysis and it remains unknown whether their method is suitable for 2D face problem. Ding et al. [130] proposed a combination method of boosted local texture and global shape descriptor extracted from 3D face models. Oriented gradient maps (OGMs) were used to highlight ethnicity sensitive geometry and texture feature sets. Experiments carried out on FRGC v2.0 data sets obtained performance up to 98.3 percent to distinguish Asians/non-Asians. In total, all representative achievements have been carried on FGRC for Asian and non-Asian, it is of interest if the generalization can be extent to multiple classes recognition problem.

Other than traditional approaches, recent studies on physical anthropometry have led to the investigation of using 3D anthropometric statistics for race categorization and related face analysis. From Section 2.4 we know that physical discriminative facial traits vary cross different racial groups, most of which are located in flexible areas such as mouth, nose, eyes and eyebrow. Correspondingly, building a 3D high-definition head-face model (such as the one used in [155], see Fig. 12) is a pre-stage work for reliable and accurate detection of those landmarks. An anthropometric face model has been built in [156], in which crucial facial feature region was identified by using eye-nose triangle relationship and 18 most discriminative facial feature points were detected separately by using histogram and contour following algorithm. An ethnicity specific generic elastic model from single 2D image was proposed in [119] for further synthesis and recognition (Fig. 13). The whole anthropometric discriminative structure of race and location of facial fiducial landmarks could be perfectly revealed by those models. Other from 3D race models, how to explore the feature selection during race information processing is also essential. Berretti et al. [157], [158] have investigated the individual relevance of variation of local facial regions and different ethnic groups in 3D face recognition. The aim was to identify the most relevant features for different ethnic groups. Experimental results on the FRGC v1.0 data set showed that the most relevant information for the purpose of discriminating two racial groups (Asian and Caucasian) is captured by the spatial arrangement between (3, 7) and (4, 5) pair, respectively. It represents a viable approach to improve recognition accuracy, by enabling training more accurate classifiers on specific ethnic groups (see Fig. 14). Gupta et al. [155], [159]

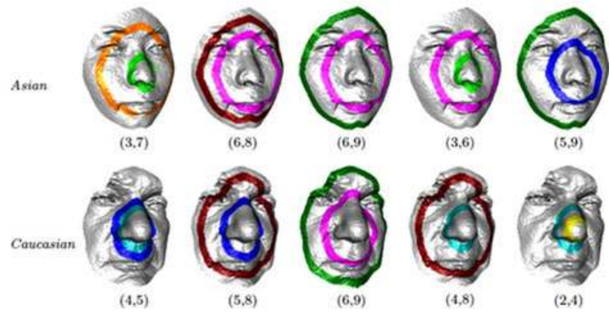

Fig. 14. The anatomical landmark regions used in [158], showing the five most relevant iso-geodesic stripe pairs for Asian and Caucasian subjects.

have extracted a highly compact anthropometric feature set including 25 facial fiducial points associated with highly variable anthropometric facial proportions. Sukumar et al. [160] have used facial anthropometric data to learn structurally discriminant features in 3D face on a racially diversified database. Their experimental results are two-fold: (1) a sparse facial fiducial distance set (e.g., the vertical profile, jaw-jaw distance, depth of nose, and depth at the eye and chin to neck distance features) which contributes most for race discrimination; and (2) the specific anthropometric features have great potential for race categorization/classification.

In summary, robust automatic location of facial fiducial landmarks (makes algorithm reliable) and effective selection of facial anthropometric proportion (makes algorithm computationally efficient) are two essential steps for guarantee of these methods' success. In fact, in vast majority of these researches, including those conducted by the authors of this survey, great care should be taken to tune the preprocessing parts, including aligning, marking, measurement, in order to minimize any difference between race categories, in terms of the low-level geometrically physical attribute of individual subjects. Therefore how to facilitate those pre-processing steps is still open question for further research exploration.

4.1.5 Dynamic Racial Face

Current methods are trained on the still images sets and thus could only apply off-line. Shakhnarovich et al. [161] have proposed a unified learning framework which could classify ethnic groups with a real-time manner. Their real-time demographic (race and gender) classifier was built on fast face detection algorithm. Harr wavelet function was used for ethnicity feature extraction which was further filtered by Adaboost. The key concept lies in evaluation of a classifier cross time by temporal integration of a decision criterion $D(t)$

$$D(t) = \frac{1}{T}\sum_{i=0}^{T} e^{-\alpha i} V(f(x_{t-i}))Q(x_{t-i}), \tag{1}$$

which allows for combining classifier output at each time frame throughout a video sequence. Experimental results on both still image sets and video clips verified their approach. Although the data set was imbalanced (non-Asians outnumber Asian) and classifier was simple binary, it offers a direction for on-line potential applications on low quality video surveillance in future.

4.2 Multi-Model Race Recognition

Up to now, lots of effort on the race classification have been focused on using single modal, which may cause uncertainty when facing degraded images. This uncertainty could arise from a number of factors, primarily with the following two: the explicit physical nature of racial face (for instance, the variation of skin color even within same race group), and the implicit inherent transformation of these physical cues into cognitively meaningful categories. Intuitively, a multi-modality fusion scheme is the straightforward way for uncertainty reduction. This suggests data association and comprehension techniques are required for measurement of multiple sources, both within and across modalities. Recently a few studies have investigated feasibility of using multiple modalities integration for the task. Zhang et al. explored the ethnic discriminability of both 2D and 3D face features by using MM-LBP (Multi-scale Multi-ratio LBP) multi-modal method [162]. Another promising approach on multi-modality based race recognition is the fusion of face and gait. Empirical results from the preliminary investigation [163] suggested significant role of gait biometrics in conjunction with face biometrics as potential candidates for more advanced race identification techniques. It should be pointed out that Zhang's team has been very active in multi-modality fusion based ethnicity classification problem. Recently they have further extended the fusion research by introducing a cascaded multi-modal biometrics system involving a fusion of frontal face and lateral gait traits in [164]. Their conclusion was that the combination of gait cues in long distance and frontal Gabor facial features in short distance could perform ethnicity recognition effectively and thus significantly improved the overall classification accuracy. For example, over 99.5 percent classification accuracy was reported on the FRGC v2.0 database in [162]. In summary, research on multi-modality race classification has witnessed significant progress in the last few years. The proliferation of literature in this direction also indicates that more various multi-modal data should be employed, such as affective cues or audio cues. Correspondingly more multi-modal data fusion methods should be investigated, such as HMM-based fusion or NN-based fusion.

4.3 Intra-Race Recognition

As far as race categorization is concerned, most of the exiting efforts studied the basic race groups (African Americans, Caucasians, Asians, in some cases with Asian Indians) due to their relative discriminativeness, their marked representation, and the availability of relevant training and testing material preparation. There are a few tentative attempts to analyze patterned difference on non-universal, sub-ethnic groups, such as Koreans, Japanese, and Chinese (including both identity and expression) [122, 178, 180, 181, 182]. Though being supported by the physical anthropometrical evidences, performing those intra-race group categorization still needs much more accurate classifiers and support from deliberately displayed, large scale, racially diversified facial datasets, including both neutral and emotional faces. Recently the researches have switched from widely separated populations, both geographically and culturally, to more closely related but ethnically distinct groups, such as

East and South East Asian [165]. Particularly, Duan et al. [117], [166] have suggested distinct differences maybe exist even in the same ethnic group geographically, this has been supported by astrometric characteristics among three Chinese ethnic groups (sub-ethnic groups in Mongoloid (Zhuang, Han, Tibetan)). With respect to these geometrical variations, three elastic templates of all ethnic groups were built. Using features extracted by Gabor wavelet and KNN classifier, performance of the approach was verified on an ethnic data set. Their results indicate that it would be of interest to sort and detect all those essential facial characteristics among intra-ethnic groups. However, since the definition of subdivisions of racial groups lack concurrence, the generalization of such methods rely on the concept from physical anthropology heavily.

To summarize this section, Table 1 provides an overview of the currently existing exemplar systems for race categorization with respect to the utilized facial features, classifier, and performance. We also mention a number of relevant aspects, including the following:

- Type of the utilized data (subject background),
- Type of the approaches applied (feature representation and classifier),
- Databases used (race category and number, other information).

Considering the benchmark results listed in the table, we could see a clear trend in increasing performance over the years as methods have become more sophisticated and training data sets have become larger. Apparently, more feature combination and more sophisticate algorithms, as well as novel algorithms which are capable of learning from imbalanced data (refer to [167] for a comprehensive survey on imbalanced learning) would allow to further improve quality. However, we already notice that existing data sets reach saturation (most results are in the range 95-98 percent of the perfect solution). We believe it is time to move towards even larger data sets, and to include subsets recorded under real-world conditions (more race categories, illumination and view variance, etc.). This is going to be discussed in next section.

5 Racial Face Databases

It is well known that for any facial information analyzing system, if the training data is not representative of the test data which an algorithm relies on, the performance of the algorithm could deteriorate. This has been repeatedly confirmed by many large-scale face recognition tournaments, such as 2002 NIST face recognition vendor test (FRVT) [28], 2006 FRVT [29], [30], and 2010 NIST multi-biometric evaluation [175]. Therefore, one can find that the majority of race face recognition researches, correspondingly, were based on those commonly accepted representative databases, such as FERET [176]. Although not intentionally designed for racial face processing, these databases's large-scale and relatively comprehensive race composition characteristics provide a more or less significant contribution for most early work on racial face recognition.

However, researchers have gradually realized that the performance could not always be guaranteed on these traditional, non race-specific face databases. Indeed, many face databases employed are actually race ill-balanced, sometimes scholars have to combine several databases together to perform multi-race classification [116], [120], [125], [177]. Therefore, a sufficiently large and fairly representative racially diverse face database, though resource-intensive, is by all means necessary and important promise for assessment and evaluation of currently developed algorithms. The merits of designing such a racially diverse face database are listed as follows:

- To be able to obtain pre-annotated, preliminary race categorization for retrieval or verification tasks while other databases are incapable of acquiring.
- To provide uniform representative racial information for discovering the deep cognitive mechanism underlying face perception.
- To pave the way for a through and systematic study of how certain factors such as race, gender, age, emotion interactively influence the automatic recognition of faces.

Due to the aforementioned merits, recently, several scholars and research institutions have began their exploration in this area by setting up racially diverse face databases. In this section we begin our discussion by briefly reviewing the current publicly available face databases[7] that are of demonstrated uses to scholars.

5.1 Major Representative Face Databases

Table 2 presents an overview of these noteworthy databases that have been ever reported in literature. For each database, we provide following information:

1) Subject background (i.e., race/ethnic groups, gender, other demographic information),
2) Capture environment (i.e., indoor/outdoor, multi-illumination, pose, accessaries),
3) Affection information (i.e., whether the facial expressions are captured and/or categorized),
4) Sample size (the number of subjects and available data samples. Such as photos or videos, whether being sampled according to some statistics or randomly).

Based on Table 2, it is apparent that most databases have considered many real world situations and have their own unique characteristics. We highlight several representative and commonly used databases with their major characteristics as follows:

1) Texas 3D face database. This database includes 3D images from the major race groups of Caucasians, Africans, Asians, East Indians, and Hispanics. This database have already been testified by some anthropometric based algorithms [155], [159].

7. Note that several databases listed in the table are actually common ones in face recognition field. However, considering that many early race recognition approaches have been evaluated on these data sets, we still list them for the aim of presenting a comprehensive review. On the other hand, for those data sets not frequently used but may have potential contribution to the field, we also list them for researcher's convenience. However, we will not introduce them in detail for the space issue.

TABLE 1
Commonly Used Methods for Racial Face Recognition

Authors	Approaches (feature+classifier)	Accuracy (%)	Database (name and numbers)	Details (specific information)
Lyle et al. [142]	Grayscale and LBP	94	FRGC (4232 faces, 404 subjects)	Asian vs. Non-Asian
Guo et al. [147]	BiFs(Biologically Inspired Features)	98.3	MORPH-II (55000 faces)	African American vs. Caucasian (only in the same gender)
Tariq et al. [165]	Shape Context matching	80.37	Silhouetted images (441 faces)	East vs. Southeastern Asian
Zhang et al. [168]	Gait Energy Image	84	Walking people from 7 cameras	East Asian vs. South American
Zhang et al. [162]	LBP based Gait Fusion	97	FRGC 2.0 (1917)	Asian vs.Caucasian
Zhang et al. [164]	LBP based Gait Fusion	95	Collected online(24)	East-Asian vs. South-American
Xie et al. [52]	KCFA and color features	98	Mugshot DB(50000 Caucasian, 50000 African, 4000 Asian)	Caucasian, Asian and African American
Xie et al. [52]	KCFA and color features	98	MBGC DB(20000 Caucasian, 10000 African, 10000 Asian)	Caucasian, Asian and African American
Qiu et al. [134]	Iris texture Gabor+Adaboost	85.9	CASIA, UPOL and UBIRIS (3982 iris images)	Asian vs. Non-Asian
Qiu et al. [134]	Iris texture Gabor+SVM	91	CASIA, BioSecure DB (2400 iris images)	Asian vs. Non-Asian
Zarei et al. [136]	Iris texture Filter + MLP neural network	93.3	ND-IRIS-0405 DB 1200 iris images (120 subjects)	Asian vs. Caucasian
Lagree et al. [138], [140]	Iris texture SMO+SVM	90.58	ND-IRIS-0405 DB 1200 iris images (120 subjects)	Asian vs. Caucasian
Zhang et al. [137]	Iris texture Codebook+SVM	96.7	CASIA+UPOL Iris DB 11320(2066 eyes)	Asian vs. Caucasian
Li et al. [143]	Eyelash region ASM+KNN	93	CMU-PIER+UBIRIS (214 iris images)	Asian vs. Caucasian
Lu et al. [152]	Range and Intensity SVM	96.8	UND and MSU (276 and 100 subjects)	Asian vs. Non-Asian
Lu et al. [149]	Gaussian+LDA	97.7	AsianPF01, NLPR AR and Yale	Asian vs. Non-Asian
Zhong et al. [121]	Gabor + K-means	89.65	FRGC 2.0	Eastern vs. Western
Akbari et al. [59]	Fuzzy Moments, Complex Geometric, Legendre	96	Ethnic DB (100 subjects)	Arabian Muslims
Klare et al. [169]	LBP+Gabor PCA+LDA	91	Operation DB (102,942 faces)	White, Hispanic African American
Gutta et al. [144], [145], [146]	RBF+Decision Tree	94	FERET DB 3006 faces (1009 subjects)	Caucasian, Asian Oriental, African
Duan et al. [117], [166]	Gabor+LDA	90.95	Chinese Ethnic DB	Tibetan, Uygur, Zhuang
Lin et al. [118]	Gabor+Adaboost+SVM	95	FERET (N/A)	African American, Caucasian, Eastern Asian
Ou et al. [115]	PCA+ICA+SVM	82.5	FERET	Asian vs. Non-Asian
Toderici et al. [153]	Harr Wavelets, kNN Kernel kNN, MDS,	99.5	FRGC 2.0 (3375 faces)	Eastern vs. Western Asian vs. White
Shakhnarovich et al. [161]	Harr Wavelets+ Adaboost	81	Mugshot from Web (3500 images +30 video clips)	Asian vs. Non-Asian
Tin et al. [112]	PCA+ KNN	96	Mugshot from Web (250 faces)	Myanmar vs. non-Myanmar Kachin, Kayah, other origins
Bellustin et al. [170]	Skin color + Adaboost	96	Indian face DB 910 Whites/646 Indians	Indian and European
Wu et al. [171]	Haar+Adaboost	96.1	FERET DB (3105 faces)	Asian, Caucasian, African Ameircan
Hosoi et al. [172]	Gabor+SVM	96	HOIP DB	Asian, Caucasian, African
Muhammad et al. [123], [124]	LBP,WLD+KNN	96	FERET DB (2368 faces)	Eastern/Middle Asian, Hispanic Caucasian, African American
Kumar et al. [127]	Color histogram SVM	94.6	PubFig DB (58797 faces)	Asian, Caucasian Indian, African American
Ding et al. [130]	Shape+ Local texture Adaboost+Decision Tree	98.3/97.8	FRGC 2.0/BU-3DFE 4007(466)/2500(100)	Asian vs. Non-Asian
Manesh et al. [120]	Facial part templates Gabor+SVM	98	FERET+PEAL (1691 faces)	Asian vs. Non-Asian
Salah et al. [125]	LBP+Harr Wavelet PCA+KNN	98.7	EGA DB (746 faces)	Asian, Caucasian, African American
Ahmed et al. [126]	CNN+ Transfer learning	93.9	FRGC 2.0 DB (14714 faces)	Asian, Caucasian, Other
Roomi et al. [53]	Skin color+ Adaboost	91.6	Yale + FERET (250 subjects)	Caucasian, Asian, African American
Demirkus et al. [173]	Skin color + Hair Color SVM	94	Online DB (600 subjects)	Caucasian, Asian, African American

Notes on the table: Not all datasets give training data and testing data, therefore they are not listed. Accuracy means the best results for all possible race groups. N/A means the information is not given in original paper.

TABLE 2
Commonly Used Racial Face Related Databases

Name	Size of subjects	Recording conditions	Race/Ethnicity background
CAS-PEAL [178]	1040 (30,900)	Expressions (6), illumination (9), occlusion (2), time (1-5), poses (21), background (4).	Chinese subjects.
IFDB [179]	616 (3,600)	Facial expressions (2), illumination (1), age (2-85), poses (4).	Iranian subjects (Middle East).
Texas 3DFRD [155], [159]	105 (1,149)	2D/3D data stereo device. Anthropometric facial fiducial points.	Caucasians, Indians, Asian.
KFDB [180], [181]	1000 (52,000)	Facial expressions (5), poses (7). illumination (16).	Korean subjects.
JAFEE [182]	216(10)	Subjects watch emotion-inducing photos and videos, displaying 7 typical expressions.	Japanese subjects.
CAFE [183]	10	posers for each of the display 7 typical expressions	Caucasians, East Asians, and Pacific Region.
FGRC 2.0 [28], [184]	4007(466)	Facial expressions (6), 3D (range and texture) , time and illumination variations (N/A), age (18-28).	Latino, Caucasian. Asian, Indian, African American.
CMU DB [185]	1.5 million	Frontal or near frontal, ground truth ethnicity labels provided.	Caucasian, African American, Asian, Hispanic.
BU-3DFE [186]	2400(100)	3D range data using 3DMD digitizer, 6 basic emotions, 4 levels of intensity	White, Black, Middle-east Asian, East-Asian, Indian, Hispanic Latino
FERET [176]	14126	Illumination (2). Facial expressions (2). Poses (9 - 20), time (2).	Caucasian, Asian, Oriental African
Ethnic DB [117]	N/A	Frontal images of Chinese ethnic minority students.	Chinese ethnic subjects.
Cohn-Kanade [187], [188]	210 (480 videos)	Grayscale values, facial Expressions (6), and AUs (action unit).	African American, Asian, Latino.
Asian PF01 [189]	103	Front face (1), illuminations (4), poses (8), facial expressions (4).	Korean subjects.
Hajj & Umrah [190] (HUDA)	N/A	Poses, facial expressions (4), illuminations, backgrounds, accessaries (N/A)	Saudi Arabia subjects. (from 25 countries)
JACFEE [72]	56	Poses, facial expressions (7), Illuminations, backgrounds, accessaries	Japanese, American subjects. (from 2 countries)
CUN [122]	112,000	Poses, facial expressions (7), Illuminations (27), backgrounds (4), accessaries (4)	Chinese ethnic subjects.
Indian DB [191]	40(710)	Pose, backgrounds, facial expressions (5),	Asian Indian subjects.
FEI [192]	2,800 (100)	Poses, facial expressions, accessaries, facial landmark pre-annotated.	Latin ethnic subjects (from Brazil)
EGA [177]	469	Poses, facial expressions, accessaries, age, illumination, gender.	5 main racial groups. (Mixed from other 6 databases)

2) The CAS-PEAL (Pose, Expression, Accessory, Lighting) database. This database has been considered as a major Chinese face database and is now mainly used for face recognition and related applications. Consequently, it is often been treated as an Asian subset to be combined with other races to form a balanced test benchmark for race recognition [120], [125], [177].
3) CUN database. This database includes 56 Chinese "nationalities" or "ethnic groups" (see Fig. 15), which covers variations in facial expressions, illumination, background, pose, accessory, among others. This database is still under construction for including 3D face subset, and will be released once finished.

We also note that there are some specific race-oriented databases such as Iranian face database (IFDB) [179], Hajj and Umrah database (HUDA) with middle-east Asian database [190], Indian face database [191], as well as FEI with Latin ethnicity database [192]. It is therefore convenient to integrate these single race data set to create a more heterogeneous and representative database for race recognition. For example, the EGA face database can be viewed as a typical example of such kind [177]. Also, BU-3DFE database [186] would be quite useful in 3D race recognition and racially diversified facial expression recognition.

We would also like to note that in addition to these aforementioned face databases, there are two additional categories of data sources which are of particular interest to the race-face learning: anthropometric demographic survey data and race face generation software. In the following two sections, we briefly review these two types of data to provide a complete survey of the racial face databases.

5.2 Anthropometric Demographic Survey Data

As mentioned in Section 2, anthropometric surveys have always been an valuable resource for carrying out race categorization. We list several representative anthropometric surveys concerned with facial anthropometric information.

- *CAESAR*, also called Civilian American and European Surface Anthropometry Resource [193]. The CAESAR project was the first anthropometric survey to provide 3D human models with spatial landmarks. Data were gathered in North America, Netherlands, and Italy (13,000 3D scans from 4,431 subjects). CAESAR has been used by Godil et al. [40] for face recognition using both 3D facial shape and color map information, satisfactory results were obtained.
- *NIOSH* (National Institute of Occupational Safety and Health) head and face database. The main body of data consisted of 947 data files in the format of a Unix-based, 3D package called INTEGRATE [16], [18]. Each file contained 3D coordinate locations of anatomical landmarks for one individual with demographic information including race, gender, age, and traditional anthropometric measures were collected.

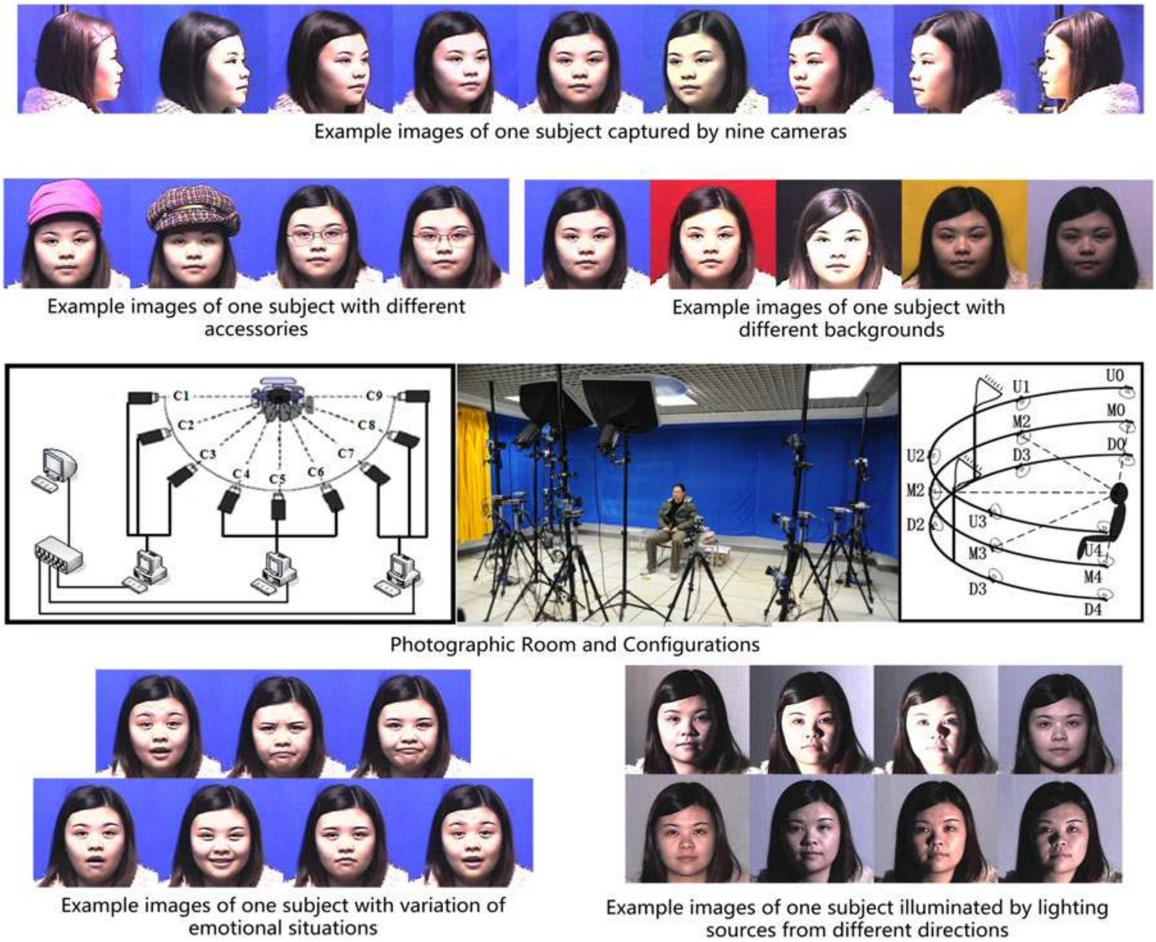

Fig. 15. Diagram showing the whole configuration of the CUN face database (with varying poses, expressions, accessories, illumination, photographic room and a multi-camera system [122], [174].

- *SizeChina*, being viewed as the first 3D head and face scan survey for the adult Chinese population [19]. It has collected high resolution 3D data from over 2,000 adult Chinese subjects. Thirty-one facial landmarks were used for statistical analysis and to create a high definition 3D anthropometric head and face model. SizeChina has been used for revealing some subtle facial dimension variations between Chinese and Caucasians in [20], [21].
- *USF*, short for the University of South Florida (USF) Human ID 3D database [194], sponsored by the US defense advanced research projects agency (DARPA). It contains 100 laser scanners aligned to a 3D reference model of 8,895 facial points. USF database has been used for several face recognition applications.

5.3 Racial Face Generation Software

Apart from these aforementioned standard databases, the emergence of artificial face generation software should also be viewed as a great supplement. They offer systematic parameters of specific features of facial structure and more general properties of the facial demographic information (race, age, gender). For example, FaceGen Modeller [195], a commercial tool originally designed for face generation in video games (Singular Inversions Inc), has been widely used for creating and handling realistic 3D facial stimuli in over 150 ways, including race, age, facial expression and gender (see Fig. 16). Another noteworthy software is FACSGen [196], [197], a synthesized lib which can import any face exported from FaceGen Modeller. With a friendly UI, users could manipulate up to 150 high level morphological topology of a face. Though being argued on the effectiveness and validity of application (generated face models in this way lacks detailed texture and scale), these softwares have already been widely used for social

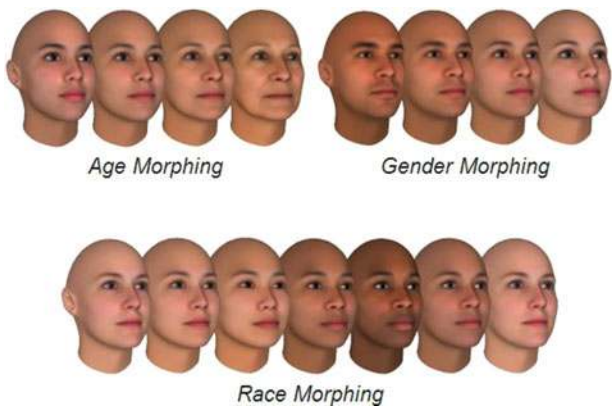

Fig. 16. 3D facial stimuli generated by FaceGen Modeller [195].

cognition, cognitive neuroscience and phycological research areas related to face perception and recognition [198], such as face sensitive perception [104], other-race-effect [103]. It is reasonable to believe that in the near future more racial face related softwares will emerge with the proliferation of racial face needed research areas.

6 Race Recognition: Real-World Applications

As an essential facial morphological trait, race perception keeps affecting our daily life. With the changing of views towards the race issue due to globalization, the ever-increasing facial data for both social and scientific research purposes, new emergence of devices and support from government agencies, increasing applications and related projects have been explored. Salutary examples of race-sensitive application include the following:

- *Biometrics based individual identification.* Racially diverse structure is inherent within facial appearance and is echoed in both human descriptions and bioinformatic representation. Therefore, race could be used to classify an individual in coarse and broad categories, helping to refine the further discriminative recognition and identification tasks. For example, in a racial/ethnic database where women with Hijab, or men with beard or mustache are crucial features in Muslim community such as Arabian countries [59]. Race could be especially useful for being incorporated into video surveillance systems [127], [173], [199], [200] for security, public safety, offender identification and forensic evidence collection. Increasingly terrorists' threats call for close integration of reliable development of such race/ethnicity sensitive information extraction and correspondingly intelligent video surveillance system, which are capable of providing meaningful analysis and extracting categorical information (such as race and gender) from poor quality video without need to recognize or identify the individual. Notable examples include Demirkus's prototype system using soft biometric features (ethnicity, skin tone, hair color, and age) to tag people in a multiple cameras network environment [173], and Bellustin's instant human attributes classification system, with embedded classifiers for race, age, and gender recognition [170]. Indeed, in certain specific applications of video surveillance where a face image is occluded or is captured in off-frontal, illumination-challenging pose, race information can offer even more valuable clues for face matching or retrieval. In short, the employment of racial trait recognition is highly expected to improve the face analysis performance when appropriately combined with a face matcher [116], [169].
- *Video security surveillance and public safety.* Race identification, along with other soft-biometric information (such as hair color, eye shape, gait, gender, age, among others), can be embedded into a high-end video surveillance analyzing system. Such an intelligent analytic solution can be applied in airports and other critical infrastructures to prevent crime by comparing images detected by the system against a preset blacklist database. Also note that fast racial group identification helps to facilitate database retrieval by discarding subjects not belonging to the specific race category. Such a video-based predictive profiling strategy has already been proven to be quite useful in aviation, maritime, mass transportation, and other open environments, such as large retail malls, parking lots, private and governmental office buildings, recreational centers and stadiums (see Fig. 17). On one side, in limited and defined circumstances, race sensitive identifier may be appropriate to protect public safety, and assist in the investigation of potential hazard activity. On the other side, the extraction of race sensitive information also helps to build more intelligent surveillance system for privacy protection, which could generate a video stream with all privacy-intrusive information encrypted or scrambled, such as PrivacyCam [201].
- *Criminal judgment and forensic art.* It is well known that race can be considered as critical ancillary information that is useful for forensic experts to give testimony in court of law where they are expected to conclusively identify suspects. For example, eyewitness identification, as one of the most important methods in the criminal scenery investigation, apprehending criminals and direct evidence of guilt, can be volatile by facial recognition deficit due to the cross-race effect or ORE (see Section 2 for ORE). It has been arguably speculated for contributing to unequal treatment or wrongful conviction of innocent ethnic minority group members [202]. Research on stereotyping in the United States also reveals that persistent racial prejudice influences the criminal justice system heavily [203], [204], [205]. The dilemma among law enforce officers and witnesses in process of identity parades addresses the necessariness of automatic race/ethnicity analysis techniques. For instance, intelligent video cameras installed in criminal scenery could offer persuasive evidence of objective race or soft-biometric information identification, which is helpful for cross validation, thus preventing false-positive identification of suspects. Also, computer-based race synthesis can significantly enhance the efficiency of the forensic artist while providing photorealistic effects needed [206]. By considering separate databases for different race groups, forensic scientists minimize the probability of unfair decisions against members of minorities [61]. Enlightened by the social and scientific significance, it is important to consider how analyzing racial identities impact policy making, law enforcement, and criminal judgement procedures. In short, computationally intelligent race recognition algorithm could certainly provide quantitative and descriptive evidence that can be used by forensic examiner in the courts.
- *Human computer interface (HCI).* Racial cues are perhaps best exemplified by their potential to the computer-consumer interaction factor. By identifying

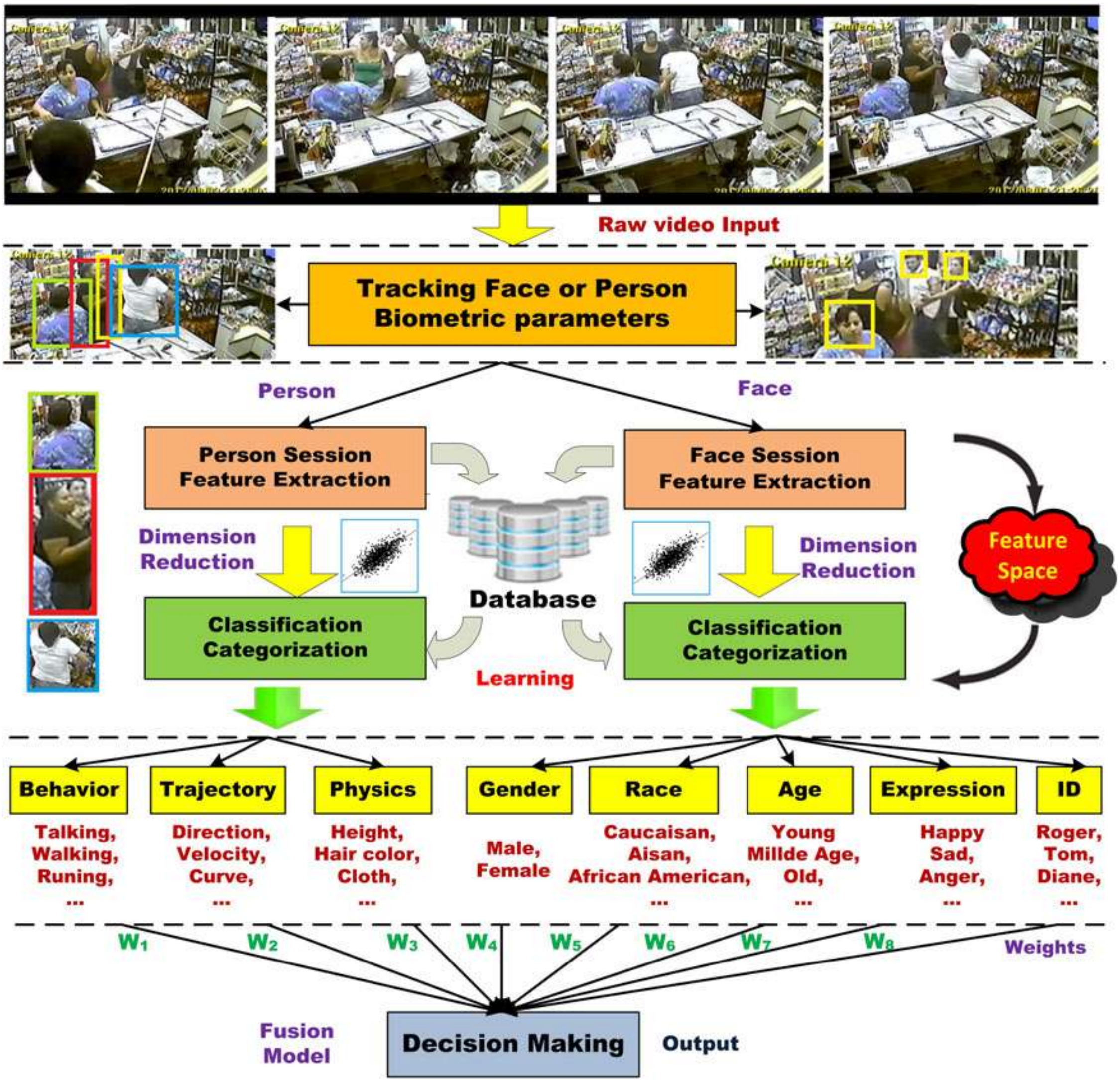

Fig. 17. Systematic illustration of a computationally intelligent biometric video surveillance system. From the top: the first level divides the input data into two branches: physical body tracking and face tracking. With interaction of databases, in the person session the system focuses on the overall biometrics such as height, behavior and trajectory following; while in the face session the system mainly concerns about detailed facial information, such as facial expressions, gender, race, identity, etc., in a coarse-to-fine order. Weighted outputs are sent to fusion center where the final decision making is performed. The system can be accomplished in a multi-camera networks. The extraction of these biometric sensitive information also helps to build intelligent CCTV for ensuring privacy.

user's race/ethnicity belongings or corresponding culture background, race/ethnicity homophyllic services (or virtual agents with sociodemographic characteristics) can offer customers racially/ethnically congruent services, thus prohibiting the potential hazards of being offended by cultural/ethnical taboos and undesirable effects [207], [208], [209]. The next generation of service robot should respond respectfully and effectively to people of all race/ethnicity belongings in a manner that recognizes, affirms, and communicates in cross-race situations. For example, in an appropriate settings to increase the quality of services (QoS), such techniques can be quite useful in public service environments (such as hospital, health care center, museum, etc.), where a smart HCI system or an avatar [210], [211], [212] can detect subtle, appearance-based or behavior-based racial (even culture specific) cues, then offer services such as the choice of speaking English versus Arabic or other alternatives. These systems will ultimately change ways in which we interact with computer vision based programs.

7 Race Recognition: Challenges and Opportunities

7.1 What Have We Learned So Far?

We have reviewed five key topics on race recognition so far: problem motivation and formulation, racial feature representation procedure, models and methods, race face databases, and applications. Based on which, lots of noteworthy discoveries have been discussed in several directions in the various branches of racial face learning. Below we briefly summarize and highlight what we have learned so far from these intensive, multidisciplinary research over the past decades.

7.1.1 Categorizing and Classifying Races

The nearly 99 percent accuracy of Toderici et al.'s work [153] indicates that common approaches such as similarity metric model could achieve satisfactory result without using anthropometric feature points, which paves the way of using traditional face recognition like method to perform race recognition, and this has been confirmed by a bunch of papers. However, it should be noted that these satisfactory performances are based on high definition 3D mesh models and on rather simple and easily separated subject groups such as Asians/Caucasians. When these two assumptions do not hold, performance of such systems cannot be guaranteed.

As a facial cue sharing both holistic attributes and localized subtle variations, the categorization of race can be accomplished both by holistic approaches and local feature extraction methods, which depends on granularity of the problem. Thus, the accuracy and importance of the racial diversity identification differs across application background. Available methods reflect these differences, but only in a rather rough way (such as Asian/Caucasian), while failing to account more diversified classes or categories. This will usually require the development of more detailed and accurate models linking discriminative feature extraction methods, multi-level racially distinct information fusion logics, and comprehensive decision making units. With the establishment of large-scale, racially diversified face databases and the development of more computationally intelligent, cortex-mechanism like approaches, more specific modeling methods and categorization applications will soon be tractable and feasible.

7.1.2 Race Categorization by Anthropometries

The next noteworthy achievement is the comprehensive framework of understanding and categorizing race in ways of physical anthropometric parameter and ad-hoc assumptions, in terms of both 2D/3D face data sets. Being viewed as a visually based stimulus category defined by the statistical variability of faces from within and across demographic categories, statistically significant differences have been observed among the race/ethnicity groups. The detection and extraction of facial features relevant to those physical differences are practically essential for building reliable anthropometric based race categorization. We list several noteworthy contributions in this field. Nevertheless, the inborn inconvenience of differences in traditional anthropometrical measurements which could be cumbersome in study design, measurement protocols, and statistical analysis, which prevents further robust analysis of the measurements, and the problems of robust and accurate location of those anthropometrically distinct facial fiducial feature landmarks. However, with fast development of 3D digital photogrammetry and scan technology with which accuracy, reliability, precision could be available [19], future improvements would be directed to the further detailed analysis of discriminative feature landmarks for race classification.

7.1.3 Racially Diversified Face Databases

Building practically useful face data set for race recognition has been actively pursued since last decade. We have listed all the representative databases in this research field, with detailed analysis in terms of parameters. More importantly, we also present the successful launching of several representative 3D human body databases, a natural wealth warehouse for carrying on anthropometric based race classification, which provides an opportunity to improve the current research standards by offering anthropometric measurements. It is reasonable to presume more joint research will be carried on both physical anthropometric data and facial image data in future. Current 3D anthropometric surveys created to date focus largely on Western populations, therefore the basic 3D head shape of a significant portion of the world's population remains unknown, which should be viewed as a future research direction [19], [20], [21].

Fig. 18. Silhouette profiles across races, from left to right: African-American (Male/Female), Asian (Male/Female), Caucasian American (Male/Female) (Figure source: [213]).

7.2 What Could We Do Next?

Based on our current understanding of this challenging topic from multiple disciplines and existing technologies, there are several important future directions to further improve general race recognition. We highlight a few of which to motivate future research and development in this critical field.

7.2.1 Real-World Learning

The real challenges in race classification, as mentioned in this survey, are how to get satisfactory classification performance in both real-world scenarios and in a massive scale. Traditional race classification is usually carried on a carefully designed face database, which provides clean and cropped frontal face images. The generalization to real world application encounters the problem of complicated, low-definition, varied illumination video cameras installed at airport, metro systems, shopping malls, etc. Extra challenge comes from the fact that traditional frontal view based methods may fail when dealing with non-frontal, multi-view face images. For example, skin tone or color attributes have long been considered as an essential part in most majority of racial face recognition. However, it must be pointed out that most crime cases empirically happen at night, in which the night-vision security camera systems could only provide inferred photo/video without color information. In such cases, more accurate identification algorithms are required. Promising directions include anthropometry and combination with other soft-biometrics, such as gait. A recent emerging direction is to use silhouette methodology (shape context based matching) (see Fig. 18). Empirical results from psychology [213] and computer vision [165], [214] have both confirmed silhouette's potential in race categorization. However, silhouettes lack texture and color information, which have been widely considered to be critical to perception and recognition of race [213]. Therefore, the best way would be a multi-modality fusion framework addressed as follows.

7.2.2 Multi-Modality Framework

Judging from Section 4, the current perspective is that human race categorization involves multiple physical facial cues integration, each modality encodes compact feature which is specific to the representation of each module. In a variation of this view, one racial face is subdivided into several racially discriminative cues, each of which is associated with particular reliability, availability, and their corresponding weights. This view to racial face classification confirms roughly to what we know of neurophysiology and psychology. Therefore, extracting racial cues through either multi-sensor system or by multi-algorithm system would be the next logical step. One advantage of such framework is that each physical trait can be assigned with different weights concerning with the contribution to the final decision level. For example, front face geometric feature usually is more informative than skin color feature while silhouette is even weaker, thus minor weights will be assigned to the latter two traits, in contrast to the more accurate information. Such framework would by all means reduce the false recognition rate. On the other hand, different with multi-modality system, the idea underlying a multi-algorithm system is that different features and/or matching algorithm will emphasize different aspects of the test subject, therefore their combination may give birth to an improved performance. Very recent empirical results from Scheirer et al. [215] have confirmed the potential of Bayesian fusion framework by combining descriptive attributes.

7.2.3 Manifold Learning

A face image lies in a high-dimensional space, but a class of images generated by latent variables (race, facial expressions, gender, pose, illumination, etc.) lies on a smooth manifold in this space [216]. Therefore, a key issue in race categorization is through projecting the face images into low-dimensional coordinates [217]. Many manifold learning models such as locally linear embedding (LLE), ISOmetric feature MAPping (ISOMAP), Laplacian Eigenmaps, Reimanian manifold learning are available. Although none of them has been applied to the issue of race classification, the methodology applies quite naturally. Generally, if we consider a set of training face sequences of C race groups, then the procedure of race manifold learning and classification can be briefed as defining the face manifold of the racial group from training data and test new data by computing the predefined projection distance. It should be pointed out that not only race category information, but other demographic soft biometric information applies as well, thus a multi-manifold learning framework could be built. It is reasonable to see more literature on this research direction in the future.

7.3 New Frontier: Emerging Opportunities and Challenges

In this section, we mainly focus on the future development of racial face recognition, referring to application potentials that come from different research directions. We would like to note, since racial face processing is a rapid developing field, we acknowledge there are intensive efforts in the community to study different aspects of racial face recognition and related tasks. For instance, 3D racial-head-face modeling in product and comics design, and the influence of racial face in social perception, are out of the core scope of this survey paper, therefore will not be discussed in this section. The goal of this section is aimed at answering the two following questions: What are (what should be) the next important research directions, and how do these fields and related methods give rise top performance in application tasks? These two questions provide the key new directions of both theory-driven study of racial face recognition, and the use of racial cues to the future more comprehensive and robust computer vision systems in various applications.

7.3.1 From 2D to 3D Facial Data

We acknowledge that current 2D racially face classification systems have already achieved “satisfactory” performance (at some extent) in constrained environments. However, just like other facial cues (express, age, gender, identify, etc.) they also face with difficulties while handling large amounts of facial variations such as head pose, lighting conditions and facial expressions, especially when all these factors combined together. For example, there is no report on race recognition with varying lighting and viewpoint situations. Real world applications, such as video security surveillance system, may require the ability of identifying the racial category of subject of interest captured in an uncooperative environment by using PTZ camera, which also poses challenges to 2D based racial face recognition methods. Three-dimensional based systems, on the other hand, are theoretically insensitive to illumination and pose changes, therefore can be viewed as potentially perfect way to further improve the classification results. Furthermore, it has been shown by several demographic surveys and statistical studies that integrating 3D face information could lead to an improved performance of race categorization. Although 3D face data (there are already several head-face demographic survey results based on both high definition 3D scanners and cameras) can make sufficient use of the comprehensive physical structure of head and face, thus offering far more reliable information and being robust to those aforementioned 2D drawbacks, they could not be simply applied directly to the algorithms. For example, even if 3D racial face models are theoretically insensitive to pose and illumination changes, they still need to be registered in an appropriate way before matching step [243]. Natural outdoor lighting has proven to be very difficult to handle, not simply because of the strong shadows cast by a light source such as the sun, but also due to distortion of face shape (e.g., irregularly squinting eyes, frowned eyebrows, and smiling lips) caused by subject facing towards direct glaring lighting source, which may fail 3D model considerably. Moreover, the issue of facial expression inuences (for example, a surprised Asian female face) is even more difficult than that in 2D modality, because 3D racial face models provide the exact shape information of facial surfaces, which might cause problems in matching [241]. Overall, feature extraction methods based on various criterion and proper registration with mapping issues further call for a comprehensive framework linking physical anthropology and geometric appearance based computer vision methods together, providing supervised guidelines for these algorithms. There is still long way to go before any satisfactory results can be obtained at meaningful scale.

7.3.2 From Isolated to Comprehensive Analysis

It must be pointed out that in an effort to study active and dynamic facial information, race, including other basic facial component analysis, must be embedded into a more comprehensive framework. In cognitive psychology, race has been reported to affect both facial expressions and emotion by observations that emotional faces are not universal and tend to be culture dependent instead. A series of investigations on the impact of recognition accuracy cross training and matching on cross races in [71], [107], [108], [109], [110], [169], [175], [218], [219] further explored and confirmed the existence of inherent bias phenomenon on face recognition. In particular, experimental results in [28], [29], [30], [31], [68], [169], [175], [220], [221], [222] have shown that both commercial and nontrainable face recognition algorithms consistently tend to perform lower on the same cohorts (females, African American, age group 18-30), which indicates that it is possible to improve overall face recognition accuracy by either separately fusing multiple weak recognition algorithms (e. g., race sensitive algorithms trained on different demographic cohorts), or by setting up a comprehensive, balanced face database. Gender perception also differs across race, indicating the existence of culture-specific stereotypes and concepts of gender [223] (But see [224]). Strong "other-race effect" was found in age estimation [225]. While in turn, all those facial cues will interact and modulate the race perception as well. For instance, Guo and Mu [147] found that race classification accuracies tend to be reduced up to 6-8 percent in average when female faces are used for training and males for testing. We shall also notice that to our best knowledge, there is no attempt in the existing literature to recognize race under emotional status. All the aforementioned discussions imply the opportunities of potential extension of racial face analysis from a single model categorization to a comprehensive multi-modality analysis, which calls for further advancements ranging from the fundamental principles, algorithms and architectures, to broader applications.

7.3.3 From Computational Intelligence to Computational Neuroscience

The inspiration of this section comes directly from the fact that recent advances in computational modeling of the primate visual system have shown deep insights of potential relevance to some of the challenges that computer vision community is facing, such as face recognition/categorization in complex environment, motion detection and human behavior analysis [230]. The problems of race perception and classification is closely related to cognitive inference and pattern recognition on these behavior data. While previous contributions are appreciated, the field of race recognition still lacks computational guidance from high level cognitive supervision. Quantitative psychophysical and physiological experimental evidences support the theory that the visual information processing in cortex can be modeled as a hierarchy of increasingly sophisticated, sparsely coded representations, along the visual pathway [174], [226]. Therefore, characterizing and modeling the functions along the hierarchy, from early or intermediate stages such as LGN, V1, are necessary steps for systematic studies in higher level, more cognitive tasks of race recognition. There are several cortex-like models proposed in recent decades [84], [85], [86], [227], [228], [229], [230], [231], combined with the illustration shown in Fig. 19. Particularly, the very recent pioneering work of Ahmed et al. [126] used hierarchial feed-forward convolutional neural network model with transfer learning to perform race classification. Being simple to be implemented and adaptable, the 3-class race classification on FRGC v2.0 data set obtained performance up to 94 percent. It is therefore very likely that near future will see the emergence of more computationally intelligent models emulating brain-like process for race perception and processing.

Another noteworthy fact is that most existing algorithms have focused on commonly used feature extraction methods. Very recent behavior experimental results [37], [232], [233], [234], [235] suggest that race category can be perceived by selective attention (see Fig. 19). A promising direction for future research, therefore, is the development of recognition models that take visual attention models [236], [237], [238] and eye tracking data [232], [234] into account. However, there is no such principled computational understanding of race by explicit attention model, which should also be clarified in the future. In an effort to explore the possibility of how computational neuroscience can help us build better computer vision and machine learning systems on race classification, the solutions go beyond the scope of computer vision field and require collaboration from multi-disciplinary communities. In summary, our work suggests a more nuance of understanding of how race classification functions both in perception and in computer vision, which may inspire new approaches to higher-level social categorization and perception.

7.3.4 From Single Race to Mixed Race

Last but not least, as worldwide racial mixing accelerates, any definitive identification of pure race based on physical appearance will undoubtedly become more problematic and less explicit. The result is the emergence of more racially-mixed people who tend to be more average looking, in aesthetical way. However, categorization of such kind of people usually leads to unexpected obfuscation, making the task even more discredited than the challenging illusive definition of race itself. That is why race recognition is much more difficult than other facial trait recognitions. The most commonly mixed-racial people we may encounter are biracial, which consist the representation or combination of two separate races. Typical examples of multiracial people include Hispanics/Latinos(admixture of Caucasian, Native American, Caribbean, African/African American, as well as Mexican-American [62], [65]), Asian Indian, and admixture of Caucasian-African American. The classification of mixed race has its unique importance to the overall race classification framework. For example, this could be of help in next generation design of human-computer-interface system. A computer trained by existing monoracial categories to categorize multiracial people may cause unanticipated trouble in such situation, which implies the importance of the

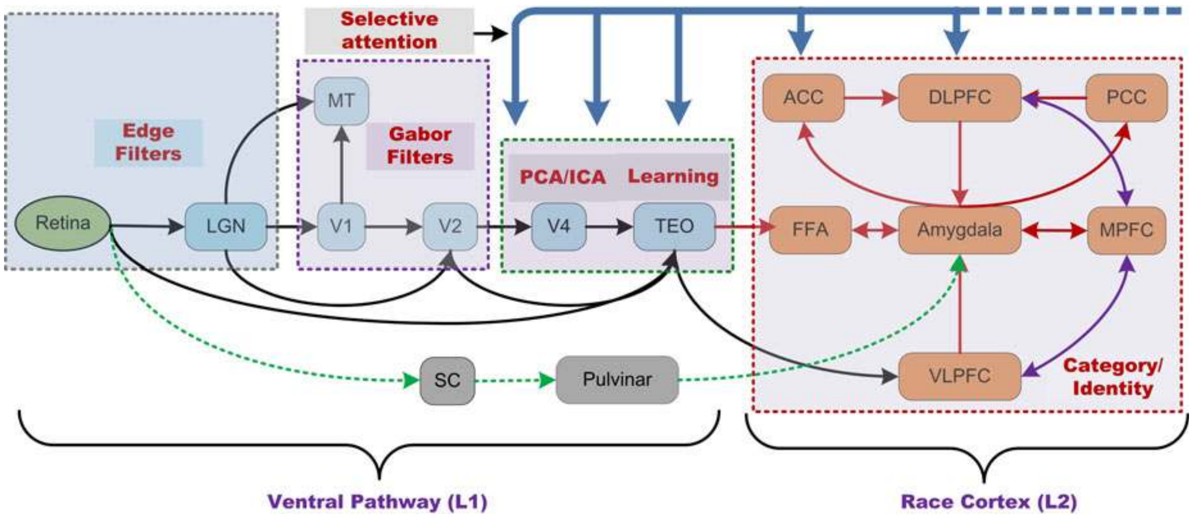

Fig. 19. Computational illustration for overall cognitive race processing regions and a hypothetical two-level working scheme hierarchy: Primary level (L1) includes ventral visual pathway (from retina to IT/TEO, for general object recognition and understanding), this level is bottom-up. Higher level (L2) includes VLPFC/DLPFC/MPFC, amygdala, ACC/PCC/FFA, it controls social interaction and application of race perception, which might exert top-down regulation to the L1 level. It is generally considered that PCC and FFA are engaged in early detection and categorization stage of race, and Amygdala is involved in the regulation of emotion and social status inference (e.g., racial attitude judgments). ACC/DLPFC/VLPFC are responsible for behavior control, such as race-biased response or stereotype, and prejudice. MPFC deals with inference with interaction among race groups. Selective attention exerts top-down administration. At present, most neuro-computational models have been focused on ventral pathway, and detailed knowledge about the precise role of cortical regions involved in racial cortex part is still missing. Note that several possible subcortical pathways also exist as shortcuts, such as the green dashed line involving the SC and the pulvinar nucleus, which provides fast and direct access to the amygdala, or LGN-MT and V2-TEO connections. Also note that the scheme is a rough hypothetical illustration and thus not strictly hierarchical. LGN, lateral geniculate nucleus; MT, medial temporal area (also known as V5); SC, superior colliculus; TEO, inferior temporal area; VLPFC, ventrolateral prefrontal cortex; DLPFC, dorsolateral prefrontal cortex; FFA, fusiform face area (fusiform gyrus).

capability of computer vision systems to make categorizations of multiracial people as multiracial [239]. Therefore, simply assigning any logically complete, consistent, and determinable categorical labels to these mixed-racial faces will fail to make any kind of objective and substantial sense [242]. As the dichotomy vanishes gradually in these multiracial subjects, the multiracial classification problem will be different from regular race recognition, since traditional racial categorization is viewed as dichotomous, as an "either/or" question (One subject could be either Asian, African American, or Caucasian, etc.). This judgment is then made difficult by the ambiguity of a multiracial person confronting them [239]. However, very recent reports [239], [240] indicate that multiracial people could be harder yet still recognizable. There are two possible ways to work along this direction to solve the problem. One is to follow traditional face recognition approaches. A multiracial data set could be established, including as many multiracial people as possible with specific labels such as Asian-Caucasian, Asian-African. Statistically discriminative features (LBP, HoG, etc.) are expected to be extracted by training enough data, then a classifier (could be a Decision Tree, kNN, Neural Networks, Bayesian method, etc.) can be learned for testing when given new images. Using fuzzy-based rules and membership functions are also promising techniques to describe the rough categorization of subjects. These categories could be included into traditional race data sets, and new trained classifier could be aware of that other than single race output (Asian, Caucasian, etc.), there could be other biracial output as well; Another possible way is to build specific 3D model of multiracial people using anthropometric data and landmarks. Although till now there has been little to no theoretical development on this direction, it is clear that either way would need strong backup from sufficiently large-scale dataset, which is also why we call for new comprehensive data set in previous section.

8 Conclusions

Race is in the face. We notice it both from embedded explicit physical appearances and implicit cognitive process. Recent developments in multidisciplinary scientific research and new technologies and tools make it possible to learn race from face by using pattern analysis and machine intelligence approaches. Over the decades, we have witnessed tremendous efforts and progresses in understanding the computational aspects of racial face learning. This survey provides a comprehensive and critical review of the research in facial image-based race recognition, perception, synthesis, and application. We begin this survey by describing the concept of race and several disciplines of relevance to provide a solid foundation for race recognition and perception. Then we present a systematical review of the state-of-the-art research and development on learning race from face, including racial feature representation, race classification, racial face databases, and real-world applications. Finally, we provide our discussions on the future research directions on this important topic, with the highlights on critical challenges, opportunities, and new frontiers of learning race from face.

Acknowledgments

This work was supported by the US National Science Foundation (NSF) under grant CAREER ECCS 1053717, Army

Research Office (ARO) under grant W911NF-12-1-0378, NSF-DFG Collaborative Research on "Autonomous Learning" (a supplement grant to CNS 1117314), and National Natural Science Foundation of China (NSFC) under grants 61225017 and 61175076. S. Fu would also like to thank G. Yang and Minzu University of China for the support in the development of this work.

References

[1] R. J. Brym and J. Lie, *Sociology: Your Compass for a New World*. Stamford, CT, USA: Cengage Learning, 2006.

[2] A. J. Calder, G. Rhodes, M. Johnson, and J. V. Haxby, *Oxford Handbook of Face Perception*. London, U.K.: Oxford Univ. Press, 2011.

[3] R. Kurzban, J. Tooby, and L. Cosmides, "Can race be erased? Coalitional computation and social categorization," *Proc. Nat. Acad. Sci.*, vol. 98, no. 26, pp. 15387–15392, 2001.

[4] V. Bruce and A. W. Young, "Understanding face recogition," *Brit. J. Psychol.*, vol. 77, pp. 305–327, 1986.

[5] A. J. Calder and A. W. Young, "Understanding the recognition of facial identify and facial expression," *Nature Rev. Neurosci.*, vol. 6, pp. 641–651, 2005.

[6] C. N. Macrae and G. V. Bodenhausen, "Social cognition: Thinking categorically about others," *Annu. Rev. Psychol.*, vol. 51, no. 1, pp. 93–120, 2000.

[7] A. J. O'Toole, D. A. Roark, and H. Abdi, "Recognizing moving faces: A psychological and neural synthesis," *Trends Cogn. Sci.*, vol. 6, no. 6, pp. 261–266, 2002.

[8] J. L. Eberhardt, N. Dasgupta, and T. L. Banaszynski, "Believing is seeing: The effects of racial labels and implicit beliefs on face perception," *Personality Soc. Psychol. Bull.*, vol. 29, pp. 360–412, 2003.

[9] O. Catz, M. Kampf, I. Nachson, and H. Babkoff, "From theory to implementation: Building a multidimensional space for face recognition," *Acta Psychol.*, vol. 131, no. 2, pp. 143–152, 2009.

[10] J. Freeman, N. Ambady, and P. Holcomb, "The face-sensitive n170 encodes social category information," *NeuroReport*, vol. 21, no. 1, pp. 24–28, 2010.

[11] D. T. Levine and B. L. Angelone, "Categorical perception of race," *Perception*, vol. 31, pp. 567–578, 2002.

[12] D. T. Levine, "Classifying faces by race: The structure of face categories," *J. Exp. Psychol.: Learn., Memory Cogn.*, vol. 22, pp. 1364–1382, 1996.

[13] M. Brewer and R. Brown, *Intergroup Relations*. New York, NY, USA: McGraw-Hill, 1998.

[14] S. Bentin and L. Y. Deouell, "Structural encoding and identification in face processing: ERP evidence for separate mechanisms," *Cogn. Neuropsychol.*, vol. 17, pp. 35–54, 2000.

[15] B. Rossion and C. Jacques, "N170: Understanding the time course of face perception in the human brain," *Int. J. Neurosci.*, vol. 112, pp. 1499–1512, 2002.

[16] B. Bradtmiller and M. Friess, "A head-and-face anthropometric survey of US respirator users," Anthrotech, Yellow Springs, OH, USA,Tech. Rep. for NIOSH/NPPTL 2004.

[17] Z. Zhuang and B. Bradtmiller, "Head-and-face anthropometric survey of U.S. respirator users," *J. Occup. Environ. Hygiene*, vol. 2, no. 11, pp. 567–576, 2005.

[18] Z. Zhuang, D. E. Slice, S. Benson, S. Lynch, and D. J. Viscusi, "Shape analysis of 3D head scan data for U.S. respirator users," *EURASIP J. Adv. Signal Process.*, vol. 3, pp. 1–10, 2010.

[19] R. M. Ball and J. F. M. Molenbroek, "Measuring Chinese heads and faces," in *Proc. 9th Int. Congr. Physiol. Anthropol. Human Diversity: Des. Life*, 2008, pp. 150–155.

[20] R. Ball, C. Shu, P. Xi, M. Rioux, Y. Luximon, and J. Molenbroek, "A comparison between Chinese and Caucasian head shapes," *Appl. Ergonom.*, vol. 41, no. 6, pp. 832–839, 2010.

[21] Y. Luximon, R. Ball, and L. Justice, "The Chinese face: A 3D anthropometric analysis," in *Proc. TMCE Symp.*, 2010, pp. 255–266.

[22] D. Schiller, J. B. Freeman, J. P. Mitchell, J. S. Uleman, and E. A. Phelps, "A neural mechanism of first impressions," *Nature Neurosci.*, vol. 12, no. 4, pp. 508–515, 2009.

[23] H. Wiese, "The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study," *Biol. Psychol.*, vol. 89, no. 1, pp. 137–147, 2012.

[24] K. Hugenberg, S. G. Young, M. J. Bernstein, and D. F. Sacco, "The categorization-individuation model: An integrative account of the other-race recognition deficit," *Psychol. Rev.*, vol. 117, no. 4, pp. 1168–87, 2010.

[25] K. B. Maddox, "Perspectives on racial phenotypicality bias," *Personality Soc. Psychol. Rev.*, vol. 8, no. 4, pp. 383–401, 2004.

[26] R. S. Bigler and L. S. Liben, "A developmental intergroup theory of social stereotypes and prejudice," *Adv. Child Develop. Behav.*, vol. 34, pp. 39–89, 2006.

[27] C. R. Jones and R. H. Fazio, "Person categorization and automatic racial stereotyping effects on weapon identification," *Personality Soc. Psychol. Bull.*, vol. 36, no. 8, pp. 1073–85, 2010.

[28] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2005, pp. 947–954.

[29] P. Phillips, W. Scruggs, A. OToole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe, "FRVT 2006 and ICE 2006 large-scale experimental results," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 831–846, May 2010.

[30] K. Bowyer, K. Chang, and P. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Comput. Vis. Image Understanding*, vol. 101, no. 1, pp. 1–15, 2006.

[31] K. Chang, K. Bowyer, and P. Flynn, "An evaluation of multimodal 2D+3D face biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 619–624, Apr. 2005.

[32] D. T. Levine, "Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit," *J. Exp. Psychol.: Gen.*, vol. 129, pp. 559–574, 2000.

[33] P. Ekman, E. R. Sorenson, and W. V. Friesen, "Pan-cultural elements in facial displays of emotion," *Science*, vol. 164, no. 3875, pp. 86–88, 1969.

[34] C. A. Meissner and J. C. Brigham, "Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review," *Psychol., Public Policy, Law*, vol. 7, pp. 3–35, 2001.

[35] S. G. Young, K. Hugenberg, M. J. Bernstein, and D. F. Sacco, "Perception and motivation in face recognition: A critical review of theories of the cross-race effect," *Personality Soc. Psychol. Rev.*, vol. 16, no. 2, pp. 116–42, 2011.

[36] J. T. Kubota, M. R. Banaji, and E. A. Phelps, "The neuroscience of race," *Nature Neurosci.*, vol. 15, no. 7, pp. 940–948, 2012.

[37] L. Cosmides, J. Tooby, and R. Kurzban, "Perceptions of race," *Trends Cogn. Sci.*, vol. 7, no. 4, pp. 173–179, 2003.

[38] S. Z. Li and A. K. Jain, *Handbook of Face Recognition*, New York, NY, USA, Springer, 2004.

[39] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognit. Lett.*, vol. 28, no. 14, pp. 1885–1906, 2007.

[40] A. Godil, S. Ressler, and P. Grother, "Face recognition using 3D surface and color map information: Comparison and combination," in *Proc. SPIE Symp. Biometrics Technol. Human Identification*, 2004, pp. 351–361.

[41] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern Recognit.*, vol. 25, no. 1, pp. 65–77, 1992.

[42] B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," *Pattern Recognit.*, vol. 36, no. 1, pp. 259–275, 2003.

[43] Z. H. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 39–58, Jan. 2009.

[44] Y. Fu, G. D. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 1955–1976, Nov. 2010.

[45] G. D. Guo, "Age estimation and sex classification,"in *Video Analytics for Business Intelligence*, New York, NY, USA: Springer, 2012, pp. 101–131.

[46] N. Ramanathan, R. Chellappa, and S. Biswas, "Age progression in human faces: A survey," *J. Vis. Lang. Comput.*, vol. 15, pp. 3349–3361, 2009.

[47] C. B. Ng, Y. H. Tay, and B. M. Goi, "Vision-based human gender recognition: A survey," in *Proc. Pacific Rim Int. Conf. Artif. Intell.: Trends Artif. Intell.*, 2012, pp. 335–346.

[48] C. Anitha, M. K. Venkatesha, and B.S. Adiga, “A survey on facial expression databases,” *Int. J. Eng. Sci. Technol.*, vol. 2, no. 10, pp. 5158–5174, 2010.

[49] Race Wiki. [Online]. Available: http://en.wikipedia.org/wiki/Race_(human_classification)

[50] Encyclopedia Britannica and “Race,” in *Encyclopedia Britannica Ultimate Reference Suite*. Chicago, IL, USA: Encyclopedia Britannica, 2012.

[51] H. Hill, V. Bruce, and S. Akamatsu, “Perceiving the sex and race of faces: The role of shape and color,” *Proc. Roy. Soc. London. Ser. B: Biol. Sci.*, vol. 261, no. 1362, pp. 367–373, 1995.

[52] Y. Xie, K. Luu, and M. Savvides, “A robust approach to facial ethnicity classification on large scale face databases,” in *Proc. IEEE Int. Conf. Biometrics: Theory, Appl. Syst.*, pp. 143–149, 2012.

[53] S. M. M. Roomi, S. Virasundarii, S. Selvamegala, S. Jeevanandham, and D. Hariharasudhan, “Race classification based on facial features,” in *Proc. Nat. Conf. Comput. Vis., Pattern Recognit., Image Process. Graph.*, 2011, pp. 54–57.

[54] C. R. Malskies, E. Eibenberger, and E. Angelopoulou, “The recognition of ethnic groups based on histological skin properties,” in *Proc. Vis., Modeling, Vis.*, 2011, pp. 353–360.

[55] K. R. Brooks and O. S. Gwinn, “No role for lightness in the perception of black and white? simultaneous contrast affects perceived skin tone, but not perceived race,” *Perception*, vol. 39, no. 8, pp. 1142–1145, 2010.

[56] L. Yin, J. Loi, and J. Jia, “Towards race-related face identification: Research on skin color transfer,” in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog.*, 2004, pp. 362–368.

[57] G. Anzures, O. Pascalis, P. C. Quinn, A. M. Slater, and K. Lee, “Minimizing skin color differences does not eliminate the own-race recognition advantage in infants,” *Infancy*, vol. 16, no. 6, pp. 640–654, 2011.

[58] M. A. Strom, L. A. Zebrowitz, S. Zhang, P. M. Bronstad, and H. K. Lee, “Skin and bones: The contribution of skin tone and facial structure to racial prototypicality ratings,” *PLoS ONE*, vol. 7, no. 7, pp. 1–8, 2012.

[59] R. Akbari and S. Mozaffari, “Fuzzy moments method for face recognition in an ethnic database,” *Int. J. Signal Process., Image Process. Pattern Recog.*, vol. 5, no. 1, pp. 85–97, 2012.

[60] C. Coon, *The Origins of Races*. New York, NY, USA: Knopf, 1962.

[61] G. Barbujani, “Human races: Classifying people vs understanding diversity,” *Current Genomics*, vol. 6, no. 4, pp. 215–226, 2005.

[62] N. Risch, E. Burchard, E. Ziv, and H. Tang, “Categorization of humans in biomedical research: Genes, race and disease,” *Genome Biol.*, vol. 3, no. 7, pp. 1–12, 2002.

[63] J. C. Stephens, J. A. Schneider, D. A. Tanguay, J. Choi, T. Acharya, S. E. Stanley, R. Jiang, C. J. Messer, A. Chew, J. H. Han, J. Duan, J. Carr, M. S. Lee, B. Koshy, K. A. Mdan, G. Zhang, W. R. Newell, A. Windemuth, C. Xu, T. S. Kalbfleisch, S. L. Shaner, K. Arnold, V. Schulz, C. M. Drysdale, K. Nandabalan, R. S. Judson, G. Ruaño, and G. F. Vovis, “Haplotype variation and linkage disequilibrium in 313 human genes,” *Science*, vol. 293, no. 5529, pp. 489–493, 2001.

[64] A. Bowcock, A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. Kidd, and L. L. Cavalli-Sforza, “High resolution of human evolutionary trees with polymorphic microsatellites,” *Nature*, vol. 368, no. 6470, pp. 455–457, 1994.

[65] E. M. Grieco and R. C. Cassidy, *Overview of Race and Hispanic Origin, 2000*. Washington, DC, USA: US Dept. Commerce, Econ. Statist. Admin., US Census Bureau, 2001, vol. 8, no. 2.

[66] “Census, race and science,” *Nature Genetics*, vol. 24, pp. 97–98, 2000.

[67] S. Haslanger, “Gender and race: (What) are they? (what) do we want them to be?” *Nous*, vol. 34, no. 1, pp. 31–55, 2000.

[68] M. H. Papesh and S. D. Goldinger, “A multidimensional scaling analysis of own- and cross-race face spaces,” *Cognition*, vol. 116, no. 2, pp. 283–288, 2010.

[69] A. J. O’Toole and V. Natu, “Computational perspectives on the other-race effect,” *Vis. Cogn.*, vol. 21, pp. 1–17, 2013.

[70] P. J. Phillips, F. Jiang, A. Narvekar, J. H. Ayyad, and A. J. O’Toole, “An other-race effect for face recognition algorithms,” *ACM Trans. Appl. Perception*, vol. 8, no. 2, pp. 1–11, 2011.

[71] M. N. Dailey, M. J. Lyons, H. Ishi, C. Joyce, M. Kamachi, J. Gyoba, and G. W. Cottrell, “Evidence and a computational explanation of cultural differences in facial expression recognition,” *Emotion*, vol. 10, no. 6, pp. 874–893, 2010.

[72] M. Biehl, D. Matsumoto, P. Ekman, V. Hearn, K. Heider, T. Kudoh, and V. Ton, “Japanese and caucasian facial expressions of emotion (jacfee) reliability data and cross-national differences,” *J. Nonverbal Behav.*, vol. 22, no. 1, pp. 3–21, 1997.

[73] R. E. Jack, C. Blais, C. Scheepers, P. G. Schyns, and R. Caldara, “Cultural confusions show that facial expressions are not universal,” *Current Biol.*, vol. 19, pp. 1543–1548, 2009.

[74] D. Enlow and M. G. Hans, *Facial Growth*. 3rd ed. Philadelphia, PA, USA: Saunders, 1990.

[75] L. G. Farkas, *Anthropometry of the Head and Face*. New York, NY, USA: Raven, 1994.

[76] L. G. Farkas, M. J. Katic, and C. R. Forrest, “International anthropometric study of facial morphology in various ethnic groups/races,” *J. Craniofac Surg.*, vol. 16, no. 4, pp. 615–46, 2005.

[77] F. Fang, P. J. Clapham, and K. C. Chung, “A systematic review of inter-ethnic variability in facial dimensions,” *Plastic Reconstr. Surg.*, vol. 127, no. 2, pp. 874–881, 2011.

[78] J. Y. Wong, A. K. Oh, E. Ohta, A. T. Hunt, G. F. Rogers, J. B. Mulliken, and C. K. Deutsch, “Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images,” *Cleft Palate-Craniofacial J.*, vol. 45, no. 3, pp. 232–239, 2008.

[79] T. A. Ito, E. Thompson, and J. T. Cacioppo, “Tracking the timecourse of social perception: The effects of racial cues on event-related brain potentials,” *Personality Soc. Psychol. Bull.*, vol. 30, no. 10, pp. 1267–1280, 2004.

[80] R. H. Ofan, N. Rubin, and D. M. Amodio, “Seeing race: N170 responses to race and their relation to automatic racial attitudes and controlled processing,” *J. Cogn. Neurosci.*, vol. 23, no. 10, pp. 3153–3161, 2011.

[81] J. L. Eberhardt, “Imaging race,” *Amer. Psychol.*, vol. 60, no. 2, p. 181, 2005.

[82] R. Love, “fMRI used to study race effect on face recognition,” *Lancet*, vol. 358, no. 9280, p. 476, 2001.

[83] T. A. Ito and B. D. Bartholow, “The neural correlates of race,” *Trends Cogn. Sci.*, vol. 13, no. 12, pp. 524–31, 2009.

[84] T. Serre, “Learning a dictionary of shape-components in visual cortex: Comparison with neurons, humans and machines,” Ph.D. dissertation, Dept. Brain Cogn. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2006.

[85] S. Ullman, M. Vidal-Naquet, and E. Sali, “Visual features of intermediate complexity and their use in classification,” *Nature Neurosci.*, vol. 5, no. 7, pp. 682–687, 2002.

[86] M. Riesenhuber and T. Poggio, “Hierarchical models of object recognition in cortex,” *Nature Neurosci.*, vol. 2, pp. 1019–1025, 1999.

[87] J. L. Brebnera, “Determinants of the own-race bias and neural correlates of own- and other-race face processing,” Ph.D. dissertation, Univ. Aberdeen, Aberdeen, U.K., 2011.

[88] E. A. Phelps, K. J. O’Connor, W. A. Cunningham, E. S. Funayama, J. C. Gatenby, J. C. Gore, and M. R. Banaji, “Performance on indirect measures of race evaluation predicts amygdala activation,” *J. Cogn. Neurosci.*, vol. 12, no. 5, pp. 729–738, 2000.

[89] E. A. Phelps, C. J. Cannistraci, and W. A. Cunningham, “Intact performance on an indirect measure of race bias following amygdala damage,” *Neuropsychology*, vol. 41, no. 2, pp. 203–208, 2003.

[90] M. D. Lieberman, A. Hariri, J. M. Jarcho, N. I. Eisenberger, and S. Y. Bookheimer, “An fMRI investigation of race-related amygdala activity in African-American and Caucasian-American individuals,” *Nature Neurosci.*, vol. 8, no. 6, pp. 720–722, 2005.

[91] J. Ronquillo, T. F. Denson, B. Lickel, Z. L. Lu, A. Nandy, and K. B. Maddox, “The effects of skin tone on race-related amygdala activity: An FMRI investigation,” *Soc. Cogn. Affect. Neurosci.*, vol. 2, no. 1, pp. 39–44, 2007.

[92] A. J. Hart, P. J. Whalen, L. M. Shin, S. C. McInerney, H. Fischer, and S. L. Rauch, “Differential response in the human amygdala to racial outgroup vs ingroup face stimuli,” *Neuroreport*, vol. 3, no. 11, pp. 2351–2315, 2000.

[93] W. A. Cunningham, M. K. Johnson, C. L. Raye, J. C. G. ad J. C. Gore, and M. R. Banaji, “Separable neural components in the processing of black and white faces,” *Psychol. Sci.*, vol. 15, no. 12, pp. 806–813, 2004.

[94] I. Gauthier, M. J. Tarr, J. Moylan, P. Skudlarski, J. C. Gore, and A. W. Anderson, "The fusiform face area is part of a network that processs faces at the individual level," *J. Cogn. Neurosci.*, vol. 12, no. 3, pp. 495–504, 2000.

[95] A. J. Golbym, J. D. Gabrieli, J. Y. Chiao, and J. L. Eberhardt, "Differential responses in the fusiform region to same-race and other-race faces," *Nature Neurosci.*, vol. 4, no. 8, pp. 845–850, 2001.

[96] C. Kaul, K. G. Ratner, and J. J. Van Bavel, "Dynamic representations of race: Processing goals shape race decoding in the fusiform gyri," *Soc. Cogn. Affect. Neurosci.*, vol. 9, pp. 326–332, 2012.

[97] R. S. Bigler and L. S. Liben, "Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice," *Current Direction Psychol. Sci.*, vol. 16, pp. 162–166, 2007.

[98] D. A. Stanley, P. Sokol-Hessner, M. R. Banaji, and E. A. Phelps, "Implicit race attitudes predict trustworthiness judgments and economic trust decisions," *Proc. Nat. Acad. Sci.*, vol. 108, no. 19, pp. 7710–7715, 2011.

[99] K. Hugenberg and D. F. Sacco, "Social categorization and stereotyping: How social categorization biases person perception and face memory," *Soc. Personality Psychol. Compass*, vol. 2, no. 2, pp. 1052–1072, 2008.

[100] K. J. Johnson and B. L. Fredrickson, "We all look the same to me, positive emotions eliminate the own-race bias in face recognition," *Psychol. Sci.*, vol. 16, no. 11, pp. 875–881, 2005.

[101] R. J. Dolan, "Emotion, cognition, and behavior," *Science*, vol. 298, no. 5596, pp. 1191–1194, 2002.

[102] R. E. Jack, O. G. B. Garrod, Y. Hui, R. Caldara, and P. G. Schyns, "Facial expressions of emotion are not culturally universal," *Proc. Nat. Acad. Sci.*, vol. 109, no. 19, pp. 7241–7244, 2012.

[103] O. Corneille, K. Hugenberg, and P. Timothy, "Applying the attractor field model to social cognition: Perceptual discrimination is facilitated, but memory is impaired for faces displaying evaluatively congruent expressions," *J. Personality Soc. Psychol.*, vol. 93, no. 3, pp. 335–352, 2007.

[104] A. T. Hirsh, A. F. Alqudah, L. A. Stutts, and M. E. Robinson, "Virtual human technology: Capturing sex, race, and age influences in individual pain decision policies," *Pain*, vol. 140, pp. 231–238, 2009.

[105] D. A. Wilbraham, J. C. Christensen, A. M. Martinez, and J. T. Todd, "Can low level image differences account for the ability of human observers to discriminate facial identity?" *J. Vis.*, vol. 8, no. 15, p. 5, 2008.

[106] G. Rhodes, "Looking at faces: First-order and second-order features as determinants of facial appearance," *Perception*, vol. 17, pp. 43–63, 1988.

[107] A. J. Otoole, H. Abdi, K. A. Deffenbacher, and J. C. Bartlett, "Classifying faces by race and sex using an autoassociative memory trained for recognition," in *Proc. 13th Annu. Conf. Cogn. Sci. Soc.*, 1991, pp. 847–851.

[108] A. J. Otoole, H. Abdi, K. A. Deffenbacher, and D. Valentin, "Low-dimensional representation of faces in higher dimensions of the face space," *J. Opt. Soc. Amer. A*, vol. 10, pp. 405–410, 1993.

[109] A. J. Otoole, K. A. Deffenbacher, H. Abdi, and J. C. Bartlett, "Simulating the 'other-race' effect as a problem in perceptual learning," *J. Neural Comput., Artif. Intell., Cogn. Res.*, vol. 3, no. 2, pp. 163–178, 1991.

[110] A. J. Otoole, K. A. Deffenbacher, H. Abdi, and J. C. Bartlett, "Structural aspects of face recognition and the other-race effect," *Memory Cogn.*, vol. 22, no. 2, pp. 208–224, 1994.

[111] V. Natu, D. Raboy, and A. J. OToole, "Neural correlates of own- and other-race face perception: Spatial and temporal response differences," *Neuroimage*, vol. 54, no. 3, pp. 2547–2555, 2011.

[112] H. H. K. Tin and M. M. Sein, "Race identification for face images," *ACEEE Int. J. Inform. Tech.*, vol. 01, no. 02, 2011.

[113] A. Ahmad, A. Ashraf, and A. S. Walid, "Arabic race classification of face images," *Int. J. Comput. Technol.*, vol. 4, no. 2, pp. 234–239, 2013.

[114] C. Liu, H. Y. Shum, and W. T. Freeman, "Face hallucination: Theory and practice," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 115–134, 2007.

[115] Y. S. Ou, X. Y. Wu, H. H. Qian, and Y. S. Xu, "A real time race classification system," in *Proc. IEEE Int. Conf. Inform. Acquisition*, pp. 1–6, 2005.

[116] B. Klare and A. K. Jain, "On a taxonomy of facial features," in *Proc. 4th IEEE Int. Conf. Biometrics: Theory Appl. Syst.*, 2010, pp. 1–8.

[117] X. D. Duan, C. R. Wang, X. D. Liu, Z. J. Li, J. Wu, and H. L. Zhang, "Ethnic features extraction and recognition of human faces," in *Proc. IEEE Int. Conf. Adv. Comput. Control*, 2010, vol. 2, pp. 125–130.

[118] H. Lin, H. C. Lu, and L. H. Zhang, "A new automatic recognition system of gender, age and ethnicity," in *Proc. World Congr. Intell. Control Autom.*, 2006, pp. 9988–9991.

[119] J. Heo and S. Marios, "Gender and ethnicity specific generic elastic models from a single 2D image for novel 2D pose face synthesis and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 12, pp. 2341–2350, Dec. 2012.

[120] F. S. Manesh, M. Ghahramani, and Y. P. Tan, "Facial part displacement effect on template-based gender and ethnicity classification," in *Proc. IEEE Conf. Control Autom. Robot. Vis.*, 2010, pp. 1644–1649.

[121] C. Zhong, Z. N. Sun, and T. N. Tan, "Fuzzy 3D face ethnicity categorization," in *Proc. IEEE Int. Conf. Adv. Biometrics*, 2009, pp. 386–393.

[122] S. Y. Fu, G. S. Yang, and Z. G. Hou, "Spiking neural networks based cortex like mechanism: A case study for facial expression recognition," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2011, pp. 1637–1642.

[123] G. Muhammad, M. Hussain, and F. Alenezy, "Race classification from face images using local descriptors," *Int. J. Artif. Intell. Tools*, vol. 21, no. 5, pp. 1–24, 2012.

[124] G. Muhammad, M. Hussain, and F. Alenezy, "Race recognition from face images using Weber local descriptor," in *Proc. IEEE Int. Conf. Syst., Signals Image Process.*, 2012, pp. 421–424.

[125] S. H. Salah, H. Du, and N. Al-Jawad, "Fusing local binary patterns with wavelet features for ethnicity identification," in *Proc. IEEE Int. Conf. Signal Image Process.*, 2013, vol. 21, no. 5, pp. 416–422.

[126] A. Ahmed, K. Yu, W. Xu, Y. Gong, and E. Xing, "Training hierarchical feed-forward visual recognition models using transfer learning from pseudo-tasks," in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 69–82.

[127] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar, "Describable visual attributes for face verification and image search," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 1962–1977, Oct. 2011.

[128] T. Berg and P. N. Belhumeur, "POOF: Part-based one-vs-one features for fine-grained categorization, face verification, and attribution estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 955–962.

[129] H. K. Moon, R. Sharma, and N. S. Jung, "Method and system for robust human ethnicity recognition using image feature-based probabilistic graphical models," US Patent 8 379 937, p. B1, 2013..

[130] H. X. Ding, D. Huang, Y. H. Wang, and L. M. Chen, "Facial ethnicity classification based on boosted local texture and shape descriptions," in *Proc. IEEE Int. Conf. Workshops Autom. Face Gesture Recog.*, 2013, pp. 1–6.

[131] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, "Face recognition by humans: Nineteen results all computer vision researchers should know about," *Proc. IEEE*, vol. 94, no. 11, pp. 1948–1962, Nov. 2006.

[132] M. Wagner-Schuman, A. M. Dubis, R. N. Nordgren, Y. Lei, D. Odell, H. Chiao, E. Weh, W. Fischer, Y. Sulai, A. Dubra, and J. Carroll, "Race-and sex-related differences in retinal thickness and foveal pit morphology," *Investigative Ophthalmol. Vis. Sci.*, vol. 52, no. 1, pp. 625–634, 2011.

[133] X. Li, W. L. Wong, C. Y.-l. Cheung, C.-Y. Cheng, M. K. Ikram, J. Li, K. S. Chia, and T. Y. Wong, "Racial differences in retinal vessel geometric characteristics: A multiethnic study in healthy asians," *Investigative Ophthalmol. Vis. Sci.*, vol. 54, no. 5, pp. 3650–3656, 2013.

[134] X. C. Qiu, Z. N. Sun, and T. N. Tan, "Global texture analysis of iris images for ethnic classification," in *Proc. Int. Conf. Adv. Biometrics*, 2005, pp. 411–418.

[135] X. C. Qiu, Z. N. Sun, and T. N. Tan, "Learning appearance primitives of iris images for ethnic classification," in *Proc. IEEE Int. Conf. Image Process.*, 2007, pp. 405–408.

[136] A. Zarei and D. Mou, "Artificial neural network for prediction of ethnicity based on iris texture," in *Proc. 11th Int. Conf. Mach. Learn. Appl.*, 2012, vol. 1, pp. 514–519.

[137] H. Zhang, Z. Sun, T. Tan, and J. Wang, "Ethnic classification based on iris images," in *Proc. 6th Chinese Conf. Biometric Recog.*, 2011, pp. 82–90.

[138] S. Lagree and K. W. Bowyer, "Ethnicity prediction based on iris texture features," in *Proc. 22th Midwest Artif. Intell. Cogn. Sci. Conf.*, 2011, pp. 225–230.
[139] L. Stark, K. W. Bowyer, and S. Siena, "Human perceptual categorization of iris texture patterns," in *Proc. 4th IEEE Int. Conf. Biometrics: Theory Appl. Syst.*, 2010, pp. 1–7.
[140] S. Lagree and K. W. Bowyer, "Predicting ethnicity and gender from iris texture," in *Proc. IEEE Int. Conf. Technol. Homeland Security*, 2011, pp. 440–445.
[141] K. Hollingsworth, K. W. Bowyer, and P. J. Flynn, "Useful features for human verification in near-infrared periocular images," *Image Vis. Comput.*, vol. 29, pp. 707–715, 2012.
[142] J. R. Lyle, P. E. Miller, S. J. Pundlik, and D. L. Woodard, "Soft biometric classification using periocular region features," in *Proc. IEEE 4th Int. Conf. Biometrics: Theory Appl. Syst.*, 2010, pp. 1–7.
[143] Y.-h. Li, M. Savvides, and T. Chen, "Investigating useful and distinguishing features around the eyelash region," in *Proc. 37th IEEE Appl. Imagery Pattern Recog. Workshop*, 2008, pp. 1–6.
[144] S. Gutta, H. Wechsler, and P. Phillips, "Gender and ethnic classification of face images," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog.*, 1998, pp. 194–199.
[145] S. Gutta and H. Wechsler, "Gender and ethnic classification of human faces using hybrid classifiers," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, vol. 6, pp. 4084–4089, 1999.
[146] S. Gutta and H. Wechsler, "Mixture of experts for classification of gender, ethnic origin, and pose of human faces," *IEEE Trans. Neural Netw.*, vol. 11, no. 4, pp. 948–960, Jul. 2000.
[147] G. D. Guo and G. W. Mu, "A study of large-scale ethnicity estimation with gender and age variations," in *Proc. IEEE Comput. Soc. Sci. Comput. Vis. Pattern Recog. Workshops*, pp. 79–86, 2010.
[148] G. Guo and G. Mu, "Joint estimation of age, gender and ethnicity: CCA vs. PLS," in *Proc. 10th IEEE Int. Conf. Workshops Autom. Face Gesture Recog.*, pp. 1–6, 2013.
[149] X. Lu and A. K. Jain, "Ethnicity identification from face images," in *Proc. SPIE Int. Symp. Defense Security : Biometric Technol. Human Identification*, pp. 114–123, 2004.
[150] J. Li, "Ethicity classification using multiple features," U.S. Patent 8 331 698 B2, Dec. 2012.
[151] Y. H. Lei, Y. Y. Chen, B. C. Chen, L. Iida, and W. H. Hsu, "Where is who: Large-scale photo retrieval by facial attributes and canvas layout," in *Proc. 35th Int. ACM SIGIR Conf. Res. Develop. Inform. Retrieval*, 2012, pp. 701–710.
[152] X. Lu, H. Chen, and A. K. Jain, "Multimodal facial gender and ethnicity identification," in *Proc. IEEE Int. Conf. Biometrics*, pp. 554–561, 2006.
[153] G. Toderici, S. M. O'Malley, G. Passalis, T. Theoharis, and I. A. Kakadiaris, "Ethnicity- and gender-based subject retrieval using 3-D face-recognition techniques," *Int. J. Comput. Vis.*, vol. 89, no. 2, pp. 382–391, 2010.
[154] O. Ocegueda, T. Fang, S. K. Shah, and I. A. Kakadiaris, "3D face discriminant analysis using Gauss-Markov posterior marginals," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 728–739, Mar. 2013.
[155] S. Gupta, K. R. Castleman, M. K. Markey, and A. C. Bovik, "Texas 3D face recognition database," in *Proc. IEEE Southwest Symp. Image Anal. Interpretation*, pp. 97–100, 2010.
[156] Z. B. Azouz, M. Rioux, C. Shu, and R. Lepage, "Characterizing human shape variation using 3d anthropometric data," *Vis. Comput.*, vol. 22, no. 5, pp. 302–314, 2006.
[157] S. Berretti, A. D. Bimbo, and P. Pala, "3D face recognition using isogeodesic stripes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2162–2177, Dec. 2010.
[158] S. Berretti, A. D. Bimbo, and P. Pala, "Distinguishing facial features for ethnicity-based 3D face recognition," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 3, p. 45, May 2012.
[159] S. Gupta, M. K. Markey, and A. C. Bovik, "Anthropometric 3D face recognition," *Int. J. Comput. Vis.*, vol. 90, no. 3, pp. 331–349, 2010.
[160] S. R. Sukumar, H. Bozdogan, D. L. Page, A. Koschan, and M. A. Abidi, "Learning structurally discriminant features in 3D faces," in *Proc. IEEE Int. Conf. Image Process.*, 2008, pp. 1912–1915.
[161] G. Shakhnarovich, P. A. Viola, and B. Moghaddam, "A unified learning framework for real time face detection and classification," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog.*, 2002, p. 16.
[162] D. Zhang, Y. H. Wang, and B. Bhanu, "Ethnicity classification based on gait using multi-view fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn. Workshops*, 2010, pp. 108–115.
[163] S. Samangooei and M. S. Nixon, "Performing content-based retrieval of humans using gait biometrics," *Multimedia Tools Appl.*, vol. 49, no. 1, pp. 195–212, 2010.
[164] D. Zhang, Y. H. Wang, and Z. X. Zhang, "Ethnicity classification based on a hierarchical fusion," *Biometric Recog.*, vol. 7701, pp. 300–307, 2012.
[165] U. Tariq, Y. Hu, and T. S. Huang, "Gender and ethnicity identification from silhouetted face profiles," in *Proc. IEEE Int. Conf. Image Process.*, 2009, pp. 2441–2444.
[166] L. Zedong, D. Xiaodong, and Z. Qingling, "A novel survey based on multiethnic facial semantic web," *TELKOMNIKA Indonesian J. Electrical Eng.*, vol. 11, no. 9, pp. 5076–5083, 2013.
[167] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sept. 2009.
[168] D. Zhang and Y. H. Wang, "Gender recognition based on fusion on face and gait information," in *Proc. IEEE Int. Conf. Mach. Learn. Cybern.*, 2008, vol. 1, pp. 62–67.
[169] B. F. Klare, M. Burge, J. Klontz, R. W. V. Bruegge, and A. K. Jain, "Face recognition performance: Role of demographic information," *IEEE Trans. Inform. Forensics Security*, vol. 7, no. 6, pp. 1789–1801, Dec. 2012.
[170] N. Bellustin, A. T. Kovalchuck, O. Shemagina, V. Yakho, Y. Kalafati, A. Vaish, P. Sharma, and S. Verma, "Instant human face attributes recognition system," *Int. J. Adv. Comput. Sci. Appl.*, Special Issue Artif. Intel., no. 3, pp. 112–120, 2011.
[171] B. Wu, H.Z. Ai, and C. Huang, "Facial image retrieval based on demographic classification," in *Proc. IEEE Int. Conf. Pattern Recog.*, 2004, vol. 3, pp. 914–917.
[172] S. Hosoi, T. Erina, and K. Masato, "Ethnicity estimation with facial images," in *Proc. IEEE 6th Int. Conf. Autom. Face Gesture Recog.*, 2004, pp. 195–200.
[173] M. Demirkus, K. Garg, and S. Guler, "Automated person categorization for video surveillance using soft biometrics," in *Proc. SPIE, Biometric Technol. Human Identification VII*, pp. 76670P, 2010.
[174] S. Y. Fu, G. S. Yang, and X. K. Kuai, "A spiking neural network based cortex-like mechanism and application to facial expression recognition," *Comput. Intell. Neurosci.*, vol. 7, pp. 19–33, 2012.
[175] P. J. Grother, G. W. Quinn, and P. J. Phillips, "Report on the evaluation of 2D still-image face recognition algorithms," NIST, Gaithersburg, MD, USA, NIST Interagency Rep. 7709, 2010.
[176] P. J. Phillips, P. J. Rauss, and S. Z. Der, "Feret (face recognition technology) recognition algorithm development and test results," Army Res. Lab., Adelphi, MD, USA, Army Res. Lab Tech. Rep. 995, 1996.
[177] D. Riccio, G. Tortora, M. Marsico, and H. Wechsler, "EGA, ethnicity, gender and age, a pre-annotated face database," in *Proc. IEEE Workshop Biometric Meas. Syst. Security Med. Appl.*, 2012, pp. 1–8.
[178] W. Gao, B. Cao, S. G. Shan, X. L. Chen, D. L. Zhou, X. H. Zhang, and D. B. Zhao, "The CAS-PEAL large-scale Chinese face database and baseline evaluations," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 38, no. 1, pp. 149–161, Jan. 2008.
[179] A. Bastanfard, M. A. Nik, and M. M. Dehshibi, "Iranian face database with age, pose and expression," in *Proc. IEEE Int. Conf. Mach. Vis.*, 2007, pp. 50–58.
[180] R. C. Roh and S. W. Lee, "Performance evaluation of face recognition algorithms on Korean face database," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 21, no. 6, pp. 1017–1033, 2007.
[181] B. W. Hwang, H. Byun, M. Roh, and S. W. Lee, "Performance evaluation of face recognition algorithms on the Asian face database. KFDB," in *Proc. 4th Int. Conf. Audio-Video-Based Biometric Person Authentication*, 2003, pp. 557–565.
[182] JAFEE dataset [Online]. Available: http://www.kasrl.org/jaffe.html
[183] CAFE dataset [Online]. Available: http://cseweb.ucsd.edu/users/gary/CAFE/
[184] FRGC face dataset [Online]. Available: http://face.nist.gov/frgc/
[185] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.

[186] L. Yin, X. Chen, Y. Sun, T. Worm, M. Reale. (2008, Sept. 17-19), "A high-resolution 3D dynamic facial expression database," *Proc. Int. Conf. Automat. Face Gesture Recog.*, pp. 1- 6. [Online]. Available: http://www.cs.binghamton.edu/~lijun/Research/3DFE/3DFE_Analysis.html

[187] Cohn-kanade database [Online]. Available: http://www.pitt.edu/~emotion/ck-spread.htm

[188] Cohn-Kanada dataset+ [Online]. Available: http://www.pitt.edu/~emotion/ck-spread.htm

[189] H. Dong and N. Gu, ''Asian face image database PF01,'' Tech. Rep., Dep. Comput. Sci. Eng., Pohang Univ. Sci. Technol., Pohang-Si, Korea, 2001.

[190] H. Zawbaa and S. A. Aly, "Hajj and umrah event recognition datasets," arXiv preprint arXiv:1205.2345, 2012.

[191] Indian face database [Online]. Available: http://vis-www.cs.umass.edu/∽vidit/IndianFaceDatabase/

[192] FEI dataset [Online]. Available: http://fei.edu.br/~cet/facedatabase.html

[193] K. Robinette, H. Daanen, and E. Paquet, "The caesar project: A 3-D surface anthropometry survey," in *Proc. IEEE Int. Conf. 3D Digital Imaging Model.*, pp. 380–386, 1999.

[194] S. Sudeep, P. J. Phillips, Z. Y. Liu, I. R. Vega, P. Grother and K. W. Bowyer, "The humanid gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162–177, Feb. 2005.

[195] FaceGen Modeller [Online]. Available: www.facegen.com/modeller.htm

[196] E. B. Roesch, T. Lucas, R. Lionel, G. Didier, S. David and S. R. Klaus, "FACSGen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units," *J. Nonverbal Behavior*, vol. 35, no. 1, pp. 1–16, 2011.

[197] E. B. Roesch, L. Tamarit, L. R.amd D. Grandjean, D. Sander, and K. R. Scherer, "Facsgen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units," *Int. J. Nonverbal Behav.*, vol. 35, no. 1, pp. 1–16, 2011.

[198] B. Bickel, M. Botsch, R. Angst, W. Matusik, M. Otaduy, and H. Pfister, "Multi-scale capture of facial geometry and motion," *ACM Trans. Graph.*, vol. 26, no. 3, p. 33, 2007.

[199] Y. Fu, G. Guo, T. S. Huang, "Soft biometrics for video surveillance, and " in, *Intelligent Video Surveillance: Systems and Technology*. Boca Raton, FL, USA: CRC Press, 2009.

[200] A. Jain, S. Dass, and K. Nandakumar, "Can soft biometric traits assist user recognition?" in *Proc. Int. Soc. Opt. Photonics: Defense Security*, 2004, pp. 561–572.

[201] A. Senior, S. Pankanti, and A. Hampapur, "Enabling video privacy through computer vision," *IEEE Security Privacy*, vol. 3, no. 3, pp. 50–57, May/Jun. 2005.

[202] S. M. Kassin, P. C. Ellsworth, and V. L. Smith, "The general acceptance of psychological research on eyewitness testimony: A survey of the experts," *Amer. Psychol.*, vol. 46, pp. 1089–1098, 1989.

[203] R. C. L. Lindsay and G. L. Wells, "What do we really know about cross-race eyewitness identification?," in *Evaluating Witness Evidence, Chichester*, Hoboken, NJ, USA: Wiley, 1983.

[204] S. J. Platz and H. M. Hosch, "Cross-racial/ethnic eyewitness identification: A field study," *J. Appl. Soc. Psychol.*, vol. 18, pp. 972–984, 1988.

[205] D. Hamilton, S. Stroessner, and D. Driscoll, "Social cognition: Its impact on social psychology," P. Devine, D. Hamilton, T. Ostrom, Eds., San Diego, CA, USA: Academic, 1994, pp. 291–321.

[206] C. J. Solomon, S. J. Gibson, and J. J. Mist, "Interactive evolutionary generation of facial composites for locating suspects in criminal investigations," *Appl. Soft Comput.*, vol. 13, no. 7, pp. 3298–3306, 2013.

[207] V. Evers, "Human-computer interfaces: Designing for culture," Ph.D. dissertation, School Inform. Syst., Univ. New South Wales, Sydney, Australia, 1997.

[208] M. C. Mackal, "The effects of ethnicity and message content on affective outcomes in a computer-based learning environment," Ph.D. dissertation, Dept. Edu. Psychol. Learn. Syst., Florida State Univ., FL, USA, 2009.

[209] J. C. Anacleto, and A. F. P. de Carvalho, "Improving human-computer interaction by developing culture-sensitive applications based on common sense knowledge," arXiv preprint arXiv:1001.0418, 2010.

[210] A. J. Cowell and K. M. Stanney, "Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility," *Int. J. Human-Comput. Studies*, vol. 62, no. 2, pp. 281–306, 2005.

[211] J. A. Pratt, K. Hauser, Z. Ugray, and O. Patterson, "Looking at human–computer interface design: Effects of ethnicity in computer agents," *Interacting Comput.*, vol. 19, no. 4, pp. 512–523, 2007.

[212] M. Allison and L. M. Kendrick, "Towards an expressive embodied conversational agent utilizing multi-ethnicity to augment solution focused therapy," in *Proc. 26th Int. Florida Artif. Intell. Res. Soc. Conf.*, 2013, pp. 332–337.

[213] N. Davidenko, "Silhouetted face profiles: A new methodology for face perception research," *J. Vis.*, vol. 7, no. 4, pp. 1–17, 2007.

[214] U. Tariq, Y. Hu, and T. S. Huang, *Gender and Race Identification by Man and Machine*. New York, NY, USA: Springer, 2011.

[215] W. J. Scheirer, N. Kumar, K. Ricanek, P. N. Belhumeur, and T. E. Boult, "Fusing with context: A Bayesian approach to combining descriptive attributes," in *Proc. IEEE Int. Joint Conf. Biometrics*, 2011, pp. 1–8.

[216] H. S. Seung and D. D. Lee, "The manifold ways of perception," *Science*, vol. 290, no. 5500, pp. 2268–2269, 2000.

[217] A. Hadid and M. Pietkäinen, "Demographic classification from face videos using manifold learning," *Neurocomputing*, vol. 100, pp. 197–205, 2011.

[218] N. Furl, P. J. Phillips, and A. J. O'Toole, "Face recogniton algorithms and the other-race effect: Computational mechanisms for a developmental contact hypothesis," *Cogn. Sci.*, vol. 26, pp. 797–815, 2002.

[219] M. N. Dailey, G. W. Cottrell, C. Padgett, and R. Adolphs, "Empath: A neural network that categorizes facial expressions," *J. Cogn. Neurosci.*, vol. 14, no. 8, pp. 1158–1173, 2002.

[220] A. Haque and G. W. Cottrell, "Modeling the other-race advantage with PCA," in *Proc. 27th Annu. Cogn. Sci. Conf.*, 2005, pp. 899–904.

[221] R. Bruyer, S. Leclere, and P. Quinet, "Ethnic categorisation of faces is not independent of face identity," *Perception*, vol. 33, no. 2, pp. 169–180, 2004.

[222] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Demographics versus biometric automatic interoperability," in *Proc. Int. Conf. Image Anal. Process.*, 2013, pp. 472–481.

[223] J. Wallies, O. V. Lipp, and E. J. Vanman, "Face age and sex modulate the other race effect in face recognition," *Attention, Perception Psychophys.*, vol. 74, no. 8, pp. 1712–1721, 2012.

[224] G. Farinella and J. L. Dugelay, "Demographic classification: Do gender and ethnicity affect each other?" in *Proc. IEEE Int. Conf. Inf., Electron. Vis.*, 2012, pp. 383–390.

[225] H. Dehon and S. Brédart, "An 'other-race effect' in age estimation from faces," *Perception*, vol. 30, pp. 1107–1113, 2001.

[226] L. Pessoa and R. Adolphs, "Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance," *Nature Rev. Neurosci.*, vol. 11, no. 11, pp. 773–783, 2010.

[227] T. J. Palmeri and I. Gauthier, "Visual object understanding," *Nature Rev. Neurosci.*, vol. 5, no. 4, pp. 291–303, 2004.

[228] K. Fukushima and S. Miyake, *Neocognitron, A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition*, New York, NY, USA: Springer, 1982 pp. 267–285.

[229] Y. LeCun and Y. Bengio, "Convolutional networks for images speech, time-series," in *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA, USA: MIT Press, 1995.

[230] N. Kruger, P. Janssen, S. Kalkan, M. Lappe, A. Leonardis, J. Piater, A. J. Rodriguez-Sanchez, and L. Wiskott, "Deep hierarchies in the primate visual cortex: What can we learn for computer vision?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1847–1871, Aug. 2013.

[231] T. Poggio and T. Serre, "Models of visual cortex," *Scholarpedia*, vol. 8, no. 4, p. 3516, 2013.

[232] S. D. Goldinger, Y. He, and M. H. Papes, "Deficits in cross-race face learning: Insights from eye movements and pupillometry," *J. Exp. Psychol.: Learn., Memory, Cogn.*, vol. 35, pp. 1105–1122, 2009.

[233] S. Trawalter, A. R. Todd, and A. A. Baird, "Attending to threat: Race-based patterns of selective attention," *J. Exp. Soc. Psychol.*, vol. 44, no. 5, pp. 1322–1327, 2008.

[234] B. Rehder and A. B. Hoffman, "Eyetracking and selective attention in category learning," *Cogn. Psychol.*, vol. 51, no. 1, pp. 1–41, 2005.

[235] I. Bülthoff, "What gives a face its ethnicity?" *J. Vis.*, vol. 12, no. 9, p. 1282, 2012.

[236] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[237] L. Itti and C. Koch, "Computational modelling of visual attention," *Nature Rev. Neurosci.*, vol. 2, no. 3, pp. 194–203, 2001.
[238] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 185–207, Jan. 2013.
[239] J. M. Chen and D. L. Hamilton, "Natural ambiguities: Racial categorization of multiracial individuals," *J. Exp. Soc. Psychol.*, vol. 48, no. 1, pp. 152–164, 2012.
[240] A. K. Ho, J. Sidanius, D. T. Levin, and M. R. Banaji, "Evidence for hypodescent and racial hierarchy in the categorization and perception of biracial individuals," *J. Personality Soc. Psychol.*, vol. 100, no. 3, pp. 492–506, 2011.
[241] D. Huang, W. B. Soltana, M. Ardabilian, Y. Wang, and L. Chen, "Textured 3d face recognition using biological vision-based facial representation and optimized weighted sum fusion," in *IEEE Comput. Soc. Conf. Comput. Vis. and Pattern Recognit. Workshops (CVPRW)*, 2011, pp. 1–8.
[242] Earnest, Les. "Can computers cope with human races?." Communications of the ACM 32.2 (1989): pp. 173–182.
[243] A. Salah, N. Alyüz, and L. Akarun, "Registration of 3D face scans with average face models," *J. Electronic Imaging*, vol. 17, no. 1, pp. 011006-1–011006-14, 2008.

Siyao Fu (M'08) received the PhD degree from the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences in 2008. From 2008 to 2012, he was an associate professor in the Department of Electrical and Computer Engineering, Minzu University, Beijing, China. He is currently a postdoctoral research fellow at the Department of Electrical, Computer, and Biomedical Engineering at the University of Rhode Island, Kingston, Rhode Island. His research interests include computer vision, intelligent robotic vision system, computational intelligence, machine learning, and data mining.

Haibo He (SM'11) received the BS and MS degrees in electrical engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1999 and 2002, respectively, and the PhD degree in electrical engineering from Ohio University, Athens, in 2006. From 2006 to 2009, he was an assistant professor in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, New Jersey. He is currently the Robert Haas endowed professor in Electrical Engineering at the University of Rhode Island, Kingston, RI. His research interests include pattern analysis, machine learning, cyber security, smart grid, and various application fields. He has published one research book (Wiley), edited one research book (Wiley-IEEE) and six conference proceedings (Springer), and authored and coauthored more than 140 peer-reviewed journal and conference papers, including Cover Page Highlighted Paper in the *IEEE Transactions on Information Forensics and Security*, Best Readings of the IEEE Communications Society on Communications and Information Systems Security and highly cited papers in the *IEEE Transactions on Knowledge and Data Engineering*, *IEEE Transactions on Neural Networks*, and IEEE Transactions on Power Delivery. His researches have been covered by national and international media such as *IEEE Smart Grid Newsletter*, *The Wall Street Journal*, Providence Business News, among others. He is the general chair of the IEEE Symposium Series on Computational Intelligence (SSCI'14) and program co-chair of the International Joint Conference on Neural Networks (IJCNN'14). He was the co founding-editor-in-chief of the *Journal of Intelligent Learning Systems and Applications*, and currently is an associate editor of the *IEEE Transactions on Neural Networks and Learning Systems* and *IEEE Transactions on Smart Grid*. He is also the chair of the IEEE Computational Intelligence Society (CIS) Neural Networks Technical Committee (NNTC) (2013 and 2014). He received the IEEE International Conference on Communications (ICC'14) Best Paper Award (2014), IEEE CIS Outstanding Early Career Award (2014), K. C. Wong Research Award, Chinese Academy of Sciences (2012), National Science Foundation (NSF) CAREER Award (2011), Providence Business News (PBN) "Rising Star Innovator" Award (2011), and Best Master Thesis Award of Hubei Province, China (2002). He is a senior member of the IEEE.

Zeng-Guang Hou (SM'09) received the BE and ME degrees in electrical engineering from Yanshan University (formerly North-East Heavy Machinery Institute), Qinhuangdao, China, in 1991 and 1993, respectively, and the PhD degree in electrical engineering from Beijing Institute of Technology, Beijing, China, in 1997. From May 1997 to June 1999, he was a postdoctoral research fellow at the Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing. He was a research assistant at the Hong Kong Polytechnic University, Hong Kong SAR, China, from May 2000 to January 2001. From July 1999 to May 2004, he was an associate professor at the Institute of Automation, Chinese Academy of Sciences, and has been a full professor since June 2004. From September 2003 to October 2004, he was a visiting professor at the Intelligent Systems Research Laboratory, College of Engineering, University of Saskatchewan, Saskatoon, SK, Canada. He is a professor and deputy director of the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His current research interests include neural networks, robotics, and intelligent control systems. He was an associate editor of the *IEEE Computational Intelligence Magazine* and *IEEE Transactions on Neural Networks*. He was/is program chair/member of several prestigious conferences. He is currently an associate editor of *Neural Networks*, and *ACTA Automatica Sinica*, etc. He is a senior member of the IEEE.