# A Face in any Form: New Challenges and Opportunities for Face Recognition Technology

**Zahid Akhtar,** University of Udine

**Ajita Rattani,** University of Missouri

*Despite new technologies that make face detection and recognition more sophisticated, long-recognized problems in security, privacy, and accuracy persist. Refining this technology and introducing it into new domains will require solving these problems through focused interdisciplinary efforts among developers, researchers, and policymakers.*

Biometrics is the science of establishing human identity according to physical or behavioral characteristics such as faces and fingerprints or the way individuals walk or sign their names. The face, in particular, has rich features that provide strong biometric cues to identify individuals, which has made face recognition suitable for a range of law enforcement applications.[1]

Face recognition has been widely adopted because the face has a significant role in conveying an individual's identity in social interaction; it is not hidden; and recognizing it requires neither advanced hardware nor physical contact. Face recognition systems (FRSs) use pattern matching to compare two faces and generate a match score that reflects the degree of their similarity.

As FRS use becomes more prevalent worldwide, resolving these accuracy issues is becoming more pressing. Many applications depend critically on recognition to ensure individuals' security and protect their identity. For example, FRSs are part of the US Visitor and Immigrant Status Indicator Technology (US-VISIT), a US Customs and Border Protection management system and of the (Unique Identification Authority of India (UIDA), which issues unique ID numbers to all Indian residents. Pervasive software, such as Microsoft Windows 10 and Kinect, use face recognition when users attempt to access the dashboard and automatically login to a profile, such as that on Xbox Live. The Toshiba YL863 TV uses face biometrics to provide customized, automatic, and advanced use settings, while the Sony HX920 TV uses face biometrics to sound an alert when viewing distance is too short or to power off the TV when it detects the absence of a viewer. Face biometrics are also used ubiquitously as a password alternative on some mobile devices. Examples include the Android KitKat mobile OS, Lenovo VeriFace, Asus Smart-Logon, and Toshiba SmartFace.

Although decades of rigorous research have produced FRSs that are accurate in constrained environments (in

0018-9162/17/$33.00 © 2017 IEEE

which face pose and illumination are controllable), in some scenarios, recognition errors are still too high and security and privacy questions remain. Unconstrained—in the wild—face detection and recognition has many challenges, including how to capture face images of sufficient quality in less-than-ideal conditions and accurately localize the spatial extent of the face in poor-quality images. Nonetheless, the promise of face recognition, driven by the universality of faces and the supporting hardware and software's auspicious capabilities, is a compelling motivation to solve these problems and broaden FRS use.

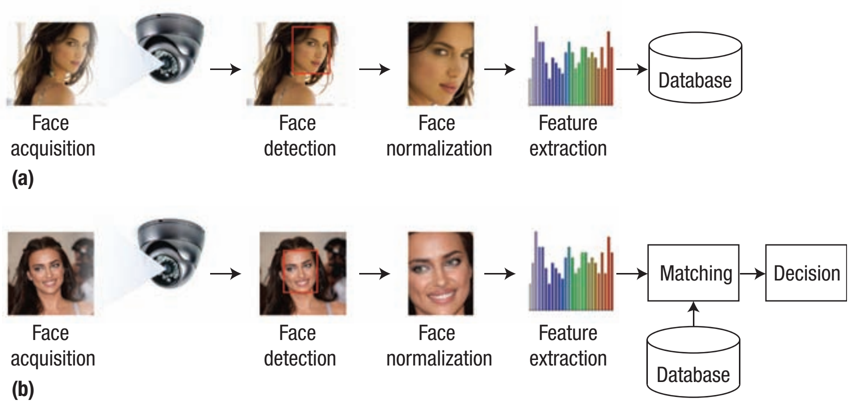


**FIGURE 1.** Tasks in a face recognition system (FRS). The FRS has seven main modules that reflect the tasks it undergoes during (a) the enrollment stage of acquiring, detecting, and normalizing an individual's face and then extracting a feature set, which it then stores as a template in the database. In (b) the recognition stage, the FRS repeats the first four tasks for a new series of face images, attempts to match the new feature set with stored templates, and makes a decision about the match on the basis of a similarity score.

## SYSTEMS AND ALGORITHMS

Existing FRSs are generally image-based, video-based, or 2D- or 3D-based, although these are broad classifications. Image-based systems use stationary face images, whereas video-based systems use videos for temporal or multiple-instance information. 2D-based systems use typical 2D imaging or image-analysis techniques, and 3D-based systems use 3D imaging or information about face shape, such as depth and curvature. In all these categories, the systems operate under either constrained sensing with cooperative subjects, such as scanning a driver's license or passport photo, or unconstrained sensing with uncooperative subjects, as in video surveillance.[1]

### Face recognition tasks

As Figure 1 shows, most FRSs perform seven main tasks. Figure 1a shows the enrollment stage, which starts with *face acquisition*, during which the FRS acquires an image of an individual's face. *Face detection* and *face normalization* involve localizing the acquired face and normalizing its appearance. Finally, in *feature extraction*, the FRS obtains a feature set to be used as a face template, which it stores in the *database* along with an identifier. In Figure 1b, which shows the recognition stage, the FRS repeats the feature acquisition, detection, normalization, and extraction steps, but this time rather than storing the feature set, it performs *matching*, in which it compares it against the stored templates and then attempts to make a *decision* about whether or not the new feature set is a match to one of the templates.

### Representative algorithms

Aside from their classification category, FRSs differ according to the face recognition methods they use, which fall roughly into four types.[1,2] Table 1 lists some examples along with the year they first appeared in the literature.[2] (Details are available at www.cs.umass.edu/lfw/results.html and www.face-rec.org.)

**Local, holistic, and hybrid.** Local methods, such as local binary patterns, use local facial features for recognition. Holistic methods like 2D principal-component analysis use the whole-face image as input. Hybrid methods, such as local probabilistic subspace techniques, employ both local and holistic features.

**Appearance- and model-based.** Appearance-based methods consider an image as a point in a high-dimensional vector space. The methods use statistical techniques such as partial least squares (PLS) to compare the sample image with the stored prototypes in the feature space. Model-based schemes, such

TABLE 1. Types of face recognition methods and sample algorithms.

| Method category | Sample algorithms: year first appeared in the literature |
|---|---|
| Local, holistic, and hybrid | Principal component analysis (Eigenfaces): 1991<br>Modular Eigenfaces: 1994<br>Linear discriminant analysis (Fisherfaces): 1997<br>Independent component analysis (ICA): 2002<br>Local binary pattern (LBP): 2006<br>Scale-invariant feature transform (SIFT): 2006<br>Speeded-up robust features (SURF): 2009<br>Learning-based descriptor (LBD): 2010 |
| Appearance- and model-based | 3D morphable model: 1999<br>Active appearance model (AAM): 2000<br>Eigen light field: 2004<br>Associate–predict model (APM): 2011 |
| Geometry- and template-based | Dynamic link architecture (DLA): 1993<br>Elastic bunch-graph matching (EBGM): 1997<br>Trace transform (TT): 2003<br>Kernel methods: 2002<br>Simulated annealing for 3D face recognition: 2009 |
| Template-matching, statistical, and neural networks | Probabilistic decision-based neural network (PDBNN): 1997<br>Genetic algorithm–evolutionary pursuit (EP): 1998<br>Wavelet packet analysis (WPA): 2000<br>Sparse representation (SR): 2009<br>Partial least squares (PLS): 2013<br>Hybrid deep learning (HDL): 2013<br>Discriminant face descriptor: 2014<br>DeepFace deep neural network: 2014<br>Deep hidden identity features (DeepID): 2014<br>FaceNet embedding: 2015 |

as active-appearance models, aim to model a face. Appearance-based methods can be further subdivided into linear and nonlinear, and model-based techniques, into 2D and 3D.

**Geometry- and template-based.** Geometry-based methods, such as elastic bunch-graph matching, analyze local facial features and their geometric relationships. Template-based methods define a face as a function to compare the input image with a template set. Template sets can be built by using statistical tools, such as kernel methods.

**Template-matching, statistical, and neural network.** Template matching depicts patterns by using models, pixels, curves, or textures. The recognition function is usually a distance measure or correlation. In the statistical approach, patterns are represented as features, and recognition is a discriminant function. The pattern representation in neural network approaches varies, but there is always a network function at some point.

## FACE DETECTION AND RECOGNITION IN THE WILD

The ability to detect a face is at the heart of any face analysis method, from identifying facial expressions to searching for individuals in images and videos. The objective is to ascertain whether any faces are in a given image or video and, if so, to return the detected face's location and size. Figure 2a shows some examples.

Face detection research is prevalent in the literature because of the complexities attributable to variations, such as pose (out-of-plane rotation), orientation (in-plane rotation), scale, location, resolution, and occlusion (part of the face is blocked). Faces can also be hard to determine if the individual is wearing glasses or has extensive facial hair. Although recent work has produced many face-detection algorithms,[2] detection in the wild remains problematic because it magnifies detection complexities. The environment might have poor lighting, for example, or background clutter; gender, ethnicity, accessories, clothing, makeup, viewing angle, and occlusion also strain the algorithm's ability to discern a face from other parts of the environment with any accuracy. Figure 2b shows some examples of faces in the wild that illustrate these challenges. Researchers have yet to design an algorithm that is robust under these arbitrary variations. Even state-of-the-art methods can attain at most 90 percent detection success in constrained environments; that percentage drops to 65 to 80 percent in an unconstrained scenario.[3,4]

Face recognition in the wild (also referred to as unconstrained face recognition) is even more challenging because once the FRS overcomes detection-in-the-wild problems, it must still match the face to images in a database. This matching requires occlusion categories that are not predefined, which in turn requires collecting vast amounts of ground-truth data. Novel frameworks are needed that adopt unsupervised or semi-supervised learning schemes to reduce the amount of data collection time and effort—a level that is currently impractical. The growing importance of face recognition in the wild must be met with ways to devise robust features and learning schemes that can deal with detection and recognition complexities. At present, the performance of most recognition algorithms degrades severely on datasets of faces in the wild, with the even top-ranked techniques achieving an accuracy of only 60–70 percent.[4]

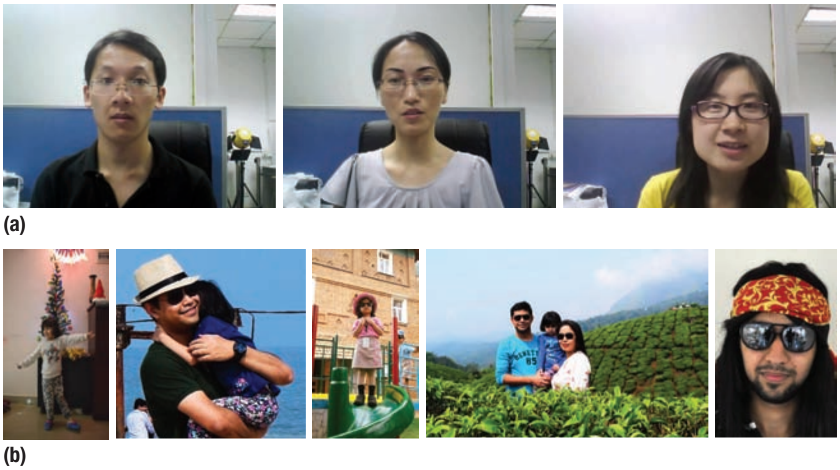


**FIGURE 2.** Examples of faces that an FRS must detect (discern from their environment) and recognize (match to a template in the database). (a) In a constrained environment, individuals are posed, but FRSs must still identify that a face appears in an image and distinguish it from the background. (b) Faces in the wild (in an unconstrained environment) are problematic for detection algorithms because too many aspects are unpredictable, such as viewing angle, resolution, background clutter, and occlusion from accessories such as glasses.

## RECOGNITION ACCURACY

For decades, FRSs have struggled to overcome the obstacles in achieving higher recognition accuracy. Despite continued work on problems such as analyzing arbitrary or invariant poses and expressions (which change facial geometry), hidden faces, and unpredictable lighting, research has yet to provide solutions without major flaws. The most promising proposals involve generating virtual views and relying on face symmetry.[1]

### Pose invariants

Arbitrary poses—in which the image pose is not controllable—remain problematic, particularly in applications such as security, online image search and tagging, and the analysis of personal photos.[5] 2D technologies can typically attain no more than 50 percent accuracy in the presence of arbitrary poses. Multiview recognition is one way to address pose variations, but the algorithms require gallery images at every pose,[6] which is impractical. Other proposed solutions include face recognition across pose, which aims to build algorithms that will recognize a face from views that have not previously been seen,[3] and the generation of virtual views.[1]

Despite the large body of work addressing pose variations, no algorithm yet exists that is highly accurate, is database independent, and can handle continuous pose variations without manual intervention. Topics that need further research include understanding the subspaces of pose-variant images, the complex face-surface-reflection mechanism, and the problem of intractability in 3D face modeling.

### Expression invariants

Facial expression also complicates recognition because it changes face geometry.[5] Proposed solutions to expression variability include learning morphable or active appearance models that capture expression variations along with identity,[7] but these methods are computationally expensive and not robust enough in the presence of illumination

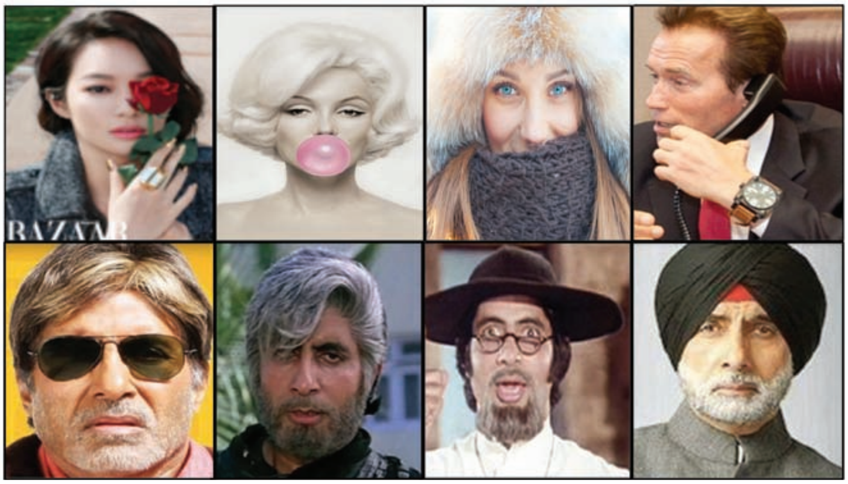


**FIGURE 3.** Examples of facial occlusion. Occlusion can be intentional, as in a disguise, or unintentional, as in sunglasses or a hat or objects in front of the face. Both occlusion types are challenging for recognition.

and pose variations. Some newly developed FRSs (including 3D methods) use video sequences and can handle a wider range of facial expressions.[5] Combining face texture and geometry could help minimize the negative impacts of expression invariants.

### Facial occlusion

Faces can be occluded by accessories worn on or near the face, such as sunglasses and hats, by hands or other objects covering an individual's face, or by passing objects, such as a tree or sign. FRSs can attain only about 10 percent accuracy under these conditions,[4] so robustness to occlusion is an important goal for practical FRSs. As Figure 3 shows, face occlusion can be either intentional (as in a disguise) or unintentional (as in a veil); in either case, identity is obfuscated, thus facial occlusion falls under the broader category of biometric obfuscation.[2]

Occlusion types—particularly disguise—has not been the focus of much work relative to other face recognition issues. Rather, the thrust seems to be on finding corruption-tolerant features or classifiers to reduce the effect of partial occlusions.[8] Recent research has shown that prior knowledge about occlusion and locally emphasized algorithms improve recognition rate.[3] Hence, explicit occlusion analysis is a crucial step toward face recognition that is robust to occlusion. Recognition could also be improved by understanding how the occlusion of individual facial parts affects facial representations and approximating a set of occlusion and disguise artifacts.

The use of face symmetry might also improve recognition accuracy,[3] although it requires further statistical analyses.[9] Work is need to prove the statistically significant relation between face symmetry and face recognition. Results could pave the way for the use of half-faces in recognition, which could help accuracy when a face is occluded and would require less storage and computation time relative to the use of full faces.

### Illumination invariants

Illumination variation is a serious concern in recognition—the pixel-value differences in different lighting conditions can actually be greater than the variations in two images of different faces in the same lighting.[9] Work so far has resulted in both passive and active approaches.[5] Passive approaches are concerned with visible-spectrum images in which lighting variations have changed the face's appearance. Active approaches, in contrast, aim to capture images either in consistent illumination conditions or in illumination-invariant modalities, such as infrared or thermal imagery. Both these approaches have drawbacks, and practical systems still require innovative and computationally efficient ways to handle illumination.

### Multibiometrics

Multibiometrics can overcome many of the restrictions in algorithms that consider only facial features, such as pose and occlusion. By consolidating evidence from multiple sources of information, such as a face and fingerprint, multibiometrics can significantly improve the performance of unimodal FRSs and recognition in the wild.[1]

Multibiometrics has much potential because of its flexible application. It could include multiple algorithms, sensors, samples, instances, modes, and hybrids, and information fusion could be at the sensor, feature, score, rank, or decision level. Fusion architectures continue to be explored. In addition, FRSs that use multibiometrics are intrinsically more robust to spoofing; however, recent work has shown that they are still vulnerable to spoofing using a single biometric trait.[10] Clearly, there is a need for novel fusion strategies that are even more robust to spoofing.[11]

## INTEROPERABILITY AND GENERALIZATION

In this category, concerns are how to make face recognition heterogeneous and applicable across modalities and databases. Scalability is also an issue.

### Heterogeneous face recognition

Heterogeneous face recognition (HFR) involves matching two face images from different imaging modalities.[12] HFR has great value in law enforcement and forensics, often requiring matching a forensic sketch to a mugshot, which is a visible-spectrum photograph. In this application, mugshots are the predominant dataset, and the forensic sketch is a probe image that must match an item in that dataset. HFR can handle any combination of imaging modalities: visible-spectrum photographs; viewed and forensic sketches; near-infrared images; short-, mid-, and long-wave infrared images; and high- and low-resolution images, such as those in Figure 4. Viewed sketches are drawn by an artist looking at a subject or a photograph, such as the top right image in the figure, and are thus more accurate than forensic sketches, which are drawn from a description.

However, even the most complex procedures proposed for HFR achieve only 50–60 percent matching accuracy, such as when visible-spectrum face images are matched with infrared images and sketches.[12] Moreover, most commercial off-the-shelf FRSs are not designed for HFR scenarios. Although much recent work has addressed HFR issues, methods are still incapable of dealing with all possible heterogeneous scenarios because of their inherently complex interrelationships. However, coupled-space learning (finding a common discriminant subspace) and nonlinear learning have proven somewhat effective.[12]

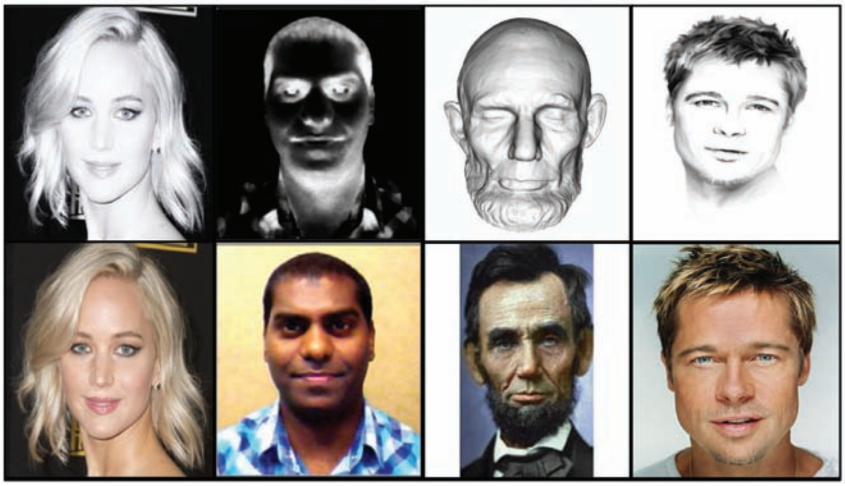


**FIGURE 4.** Images from various modalities. Heterogeneous face recognition allows matching across image modalities, such as sketches and photographs, but its accuracy is low and most FRSs are not designed for this kind of recognition. The top row contains images from (left to right) near-infrared, thermal-infrared, 3D depth, and viewed-sketch modalities. The bottom row contains the corresponding photograph of the same subject (visible-band face image).

### Cross-modality recognition

Cross-modality matching—matching images that correspond to different biometric modalities—has yet to be investigated.[12] For example, when a face image is captured but no match is found in the database, matching the face against iris images by using the periocular information in both the face and legacy database images might be a solution. Research must address the variations in modalities, sensors, resolutions, and imaging spectra, perhaps through novel deformable methods for modeling photometric and geometric variations between different modalities.

### Cross-database setting

Most recognition methods based on learning and feature design are either prone to overfitting to data samples or have low generalization ability.[12] Cross-database setting—taking training and testing sets from different sources—is one way to increase an FRS's interoperability and generalization capability, which makes sense in the real world, but has been largely ignored.[13] Developing cross-database methods will require deep-learning, concept-drift, and unsupervised learning techniques.

### Achieving web scale

With the growth of streaming and social media and the popularity of webcams for surveillance, possibly billions of videos and images are being exchanged daily. Much attention has been directed toward achieving open source, large-scale face recognition,[9] but many recognition approaches neither address nor adapt to this scale.[2] Possible directions are to combine the meta-information associated with face images and videos, or to formulate data-independent feature extraction that would rely on deep-learning and classifier-learning algorithms.[9]

## OBTAINING ANCILLARY INFORMATION

An FRS stores face images along with feature-set templates, which enables the extraction of information other than identity, such as demographics, soft attributes (for example, tattoos), and attractiveness.

### Demographics

Extracting demographic attributes such as age, gender, and race from a facial image and using them in recognition is a topic of growing interest with potential applications in multimedia communication and the beauty industry.[7] However, demographic attributes are affected not just by gender and age (internal factors), but also by place of residence and degree of multiracial heritage (external factors). Most studies show that a single demographic attribute results in less-than-satisfactory recognition performance—even in constrained environments.[2]

Considerably more research is needed in analyzing faces acquired in unconstrained conditions. The lack of public databases containing metadata with multiple labels (for example, lifestyle, geography, and occupation) has further stymied efforts to address this problem. Crowdsourcing might be useful to collect ground truth for such large datasets. Information fusion from multimodality imaging sensors might also help, as would the design of a single special-features extractor that can be used both to estimate demographic attributes and to aid recognition. However, few researchers have studied the interrelationships of age, race, and gender, which is required for any solution that fuses demographic and visual attributes (such as a pointed nose) for face recognition and search engines.[7]

### Soft attributes

Demographic information and soft attributes, such as eye and hair color and facial characteristics such as wrinkles and moles, are collectively referred to as soft biometrics.[1] Soft attributes are so named because they do not explicitly identify a person, but rather complement the identity information that primary biometrics provide. Although research interest in soft biometrics is increasing, the area is still nascent, with open issues such as the need for more accurate mechanisms to extract soft biometrics and understanding how such mechanisms will combine new soft-biometrics modalities. For example, a beard and mustache, formerly considered contributors to facial occlusion, could become soft biometric traits instead.

### Attractiveness

Relatively little machine learning research has been devoted to analyzing facial attractiveness,[14] possibly because an absolute definition of the aesthetics contributing to facial beauty remains problematic. Understanding how aesthetics are perceived among different cultures is only one aspect of defining what constitutes facial attractiveness and the relation between low-level image features and high-level aesthetics.[14]

## SECURITY AND PRIVACY

A range of security and privacy problems must be solved for FRSs to become more widespread, including vulnerability to spoof attacks, de-identification mechanisms, and template security.

### Spoof attacks

FRSs are vulnerable to spoof attacks—when an impostor tries to masquerade as a genuine user by replicating the user's face biometrics through a photo, video, or 3D face model and thus gains illegitimate access and advantages.[11,13] The quintessential anti-spoofing mechanism is face-liveness detection, which aims to disambiguate live human faces from spoof artifacts.[1] Although numerous solutions have been proposed, none have a sufficiently low error rate, and most have little interoperability.[10]

Proactive defense strategies are needed, such as security by design and security by obscurity.[2] Security by design focuses on designing a secure system from the ground up, making sure that features include both the highest generalization capability and the least vulnerability to spoofing. Security by obscurity aims to improve system security by hiding information from attackers.

### Visual privacy and de-identification

A tension against major FRS improvement is increasing concern about individual privacy and data security, specifically how biometric information can be misused to profile and track individuals against their will. Most FRSs store an original face image along with a template for the future extraction of new feature sets and templates.[1] An automated scheme could obtain demographic and privacy information from face images, but that might lead to function creep and undesirable intrusions of privacy, such as an insurance company using biometrics to deny insurance applications from individuals with risky genetic patterns or an undisclosed disease. Thus, it is vital to devise visual privacy preservation methods (also known as changeability or de-identification methods) for faces.[15]

Privacy preservation methods aim to perturb a face image so that it cannot be used to ascertain attributes, such as age, gender, and race, but still be useful in automatic face recognition. The tradeoff between privacy protection and image utility is critical. Various methods exist to ensure privacy protection, such as masking and morphing,[15] but they drastically decrease FRS performance.[2] They rely on the prior detection of pertinent regions, such as the eyes, but even a single wrong detection can seriously violate an individual's privacy.

THE TRADEOFF BETWEEN PRIVACY PROTECTION AND IMAGE UTILITY IS CRITICAL IN ADDRESSING THE POTENTIAL ABUSE OF BIOMETRIC DATA.

A better approach is to avoid dependency on the detector performance and apply privacy protection to the entire image. Efforts in this area are already evident in social networks; it is not trivial to autonomously select the proper privacy setting for any specific face, and tools are available for privacy protection. Additional work might focus on evaluating robustness to potential re-identification attacks, exploring visual privacy preservation in video recognition, and improving de-identification of facial tattoos.

### Template security

An attack on a stored face template could be extremely damaging[11] because the attacker can replace that template and then fabricate a physical spoof by using inverse biometrics—regenerating the original biometric sample from the template—or feed the stolen template to a matcher to obtain unauthorized access.[2]

An ideal scheme for protecting against such an attack should have four properties:

- *diversity*—it should prohibit cross-database matching, thereby ensuring user privacy;
- *revocability*—it should be able to revoke a compromised template and reissue a new one (referred to as cancelable biometrics);
- *security*—it should be computationally hard to obtain the original template from a secured template; and
- *performance*—it should not degrade the FRS's performance.

A scheme with all these properties remains elusive,[11] implying that a single protection approach might not be sufficient; hybrid schemes that combine various protection advantages are worth exploring.

## FORENSIC CHALLENGES

This category includes problems with using face recognition as supporting evidence. Problems center on applications such as genealogy and individuality, as well as issues such as lookalike faces and avatars.

### Genealogical application

Genealogical face recognition, also referred to as kinship verification, is a relatively new field. Kinship verification determines whether a pair of faces have a kin relation—two individuals are biologically related and have overlapping genes.[1] Kinship verification has many potential applications, including as mechanism to organize family albums, search for missing family members, and analyze social media. Once again, the lack of sufficiently large public databases has stymied research on this topic.[16] Novel genetic-invariant features, extreme learning machines, and metric- and transfer-subspace learning methods could greatly advance genealogical recognition in unconstrained scenarios.

### Face individuality

As yet, there is no formal scientific basis for evaluating face individuality (a face's uniqueness)—quantitative information on the likelihood that another person could exhibit the same facial feature set.[1,2] Consequently, court cases are challenging the identification validity of FRS-based evidence. A scientific basis for establishing face individuality will make FRS identification admissible in a court of law and help establish an upper bound on FRS performance.

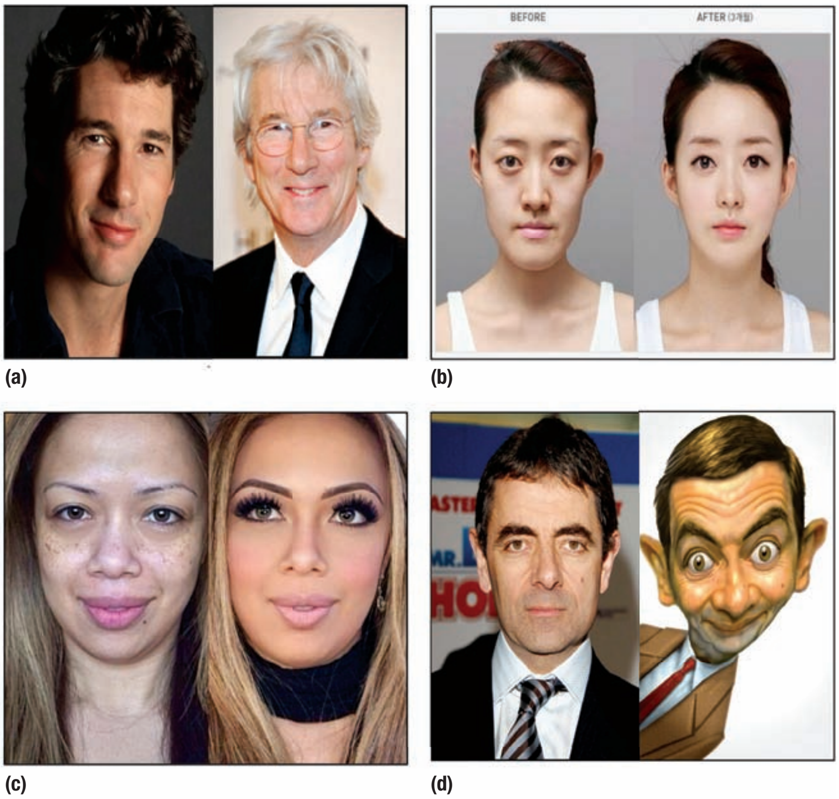
A taxonomy of available facial features could be a first step in providing this basis,[1] but taxonomy standards are essential. A standards-based taxonomy will not only facilitate an individuality measure for face images that are admissible in legal testimony but will also aid in the development of a universally applicable framework for commercial FRSs. The taxonomy should define the same features and feature levels for both recognition engines and human examiners. Thus far, only basic efforts have been made to define global-, local-, and micro-level features.[2]

### Lookalikes

A major face recognition challenge is how to design a feature extractor and matcher that increases interclass variations between two individuals with low interclass variations, such as between lookalikes or identical twins.[9] One study showed that neither human examiners nor FRSs can efficiently recognize lookalikes or twins,[2] which strongly suggests the need for methods to improve FRS performance in this context.

### Avatars

The advancement of VR technology requires new solutions for identity management across worlds populated with both humans and artificial entities. Recognizing people as avatars in virtual communities is becoming a major problem for security and anti-terrorism experts because criminal activities such as identity theft, fraud, and terrorist training can be carried out



(a)

(b)

(c)

(d)

**FIGURE 5.** Examples of face alterations. (a) Aging, (b) before and after plastic surgery, (c) effects of cosmetics on facial appearance, and (d) a photograph and a caricature of the same person. All these alterations can cause matching errors, particularly a false positive for no match.

in virtual worlds as well as the real world.[1] The ability to identify and track a person as an avatar in a virtual world is becoming critical, and some scenarios move between the physical and virtual worlds, requiring tracking in both environments. Most avatars resemble their owners because individuals tend to upload pictures of themselves as a prototype avatar. Even so, FRSs were not designed to recognize nonhuman agents, and very few efforts have focused on resolving this problem.[2]

## FACE ALTERATION

Faces can change either naturally or artificially through aging, surgery, and cosmetics. How to recognize and account for these changes remains an open problem.

### Aging

As Figure 5a shows, faces age and facial shape and texture can change. Studies have shown that such changes make it hard to match the stored templates for that individual after a certain time.[1] Despite receiving increasing attention, longitudinal studies of recognition remain scarce. The most common solution is virtual template synthesis for aging and de-aging transformations.[17] However, this method is prone to estimation errors. Genetic makeup and lifestyle contribute to aging, and aging manifests differently with and across societies. Another promising solution that is just being explored is to continuously adapt enrolled templates to the aging variations of input samples.[2]

### Plastic surgery

Plastic surgery to correct feature defects or to improve attractiveness also drastically changes skin texture, face components, and an individual's overall appearance, as Figure 5b shows. Consequently, recognition algorithms fail to recognize faces before and after plastic surgery.[18] According to the International Society of Aesthetic Plastic Surgery, more than 23 million cosmetic and nonsurgical procedures were carried out worldwide in 2013. Despite plastic surgery's rising popularity, there are few studies on recognition in this context, and proposed methods are far from satisfactory.[18] Promising research directions include finding ways to detect plastic surgery and model the resulting facial changes.

### Cosmetics

Studies show that facial makeup makes automated face recognition harder and degrades the performance of algorithms to estimate gender and age.[13] Simple cosmetic alterations, such as those shown in Figure 5c, have significantly decreased FRS accuracy. Overcoming the obstacles imposed by cosmetics, which are in widespread daily use, will require deeply understanding how they affect recognition as well as novel schemes to address the identified effects.

## EMERGING TRENDS AND APPLICATIONS

Several new developments in face recognition could improve accuracy and robustness to facial alteration. These include video-based recognition as well as the use of 3D facial geometry, mobile devices, and caricatures. Novel applications in medicine and the prevention of newborn switching and abduction are also on the horizon.

### Video recognition

Face recognition and tracking through video recordings have become pivotal in law enforcement as well as in commercial video surveillance, social networking, and movie indexing. Videos can result in more accurate recognitions than still images because FRSs use information accumulated over multiple frames.[1] Video-based recognition can be video to image, in which videos are input to the FRS which compares them to a database of still images; image to video, the reverse of video to image, and video-to-video, in which both inputs and images in the database are videos. Representative techniques with standard databases achieve recognition accuracy of 80–95 percent for video-to-image and image-to-video, but only 70–80 percent for video-to-video recognition.[19]

The latter percentage is likely due to the high number of possible appearance variations in two sets of facial sequences: for example, poor resolution greatly increases recognition error. Video databases also require tremendous amounts of storage, so few such databases exist, making it harder to conduct large-scale systematic evaluations of video-based recognition.

### 3D recognition

Studies show that 3D recognition methods, which use 3D facial geometry, can achieve better accuracy relative to 2D techniques because they are robust to illumination, pose, and cosmetic alteration.[6] Current work is addressing how to design better 3D cross-modality (matching 2D to 3D faces and vice versa) and how to incorporate 3D face modeling into 2D FRSs.[9] 3D recognition methods usually require a range camera, but several vendors are already offering low-cost 3D sensors that can be used for face recognition on laptops and mobile devices. However, 3D FRSs are not robust to expression, facial hair, and large occlusions.[5]

### Mobile recognition

The ubiquity of mobile devices with cameras has opened nearly limitless applications for face recognition technology. Nonetheless, mobile processing power is limited, and even commercial mobile FRSs are either vulnerable to spoofing or produce a high level of false positives on a large dataset.[20] The use of multiple information sources from auxiliary sensors could address some of these problems. For example, in a tilted face image, a gyroscope could be used to infer a phone pose and compensate for the tilt of a face image. Also, more efforts should focus on developing face-based continuous authentication methods (methods that constantly monitor users and periodically reauthenticate them), including making them user-friendly and unobtrusive.

### Caricature recognition

Recognition does not always involve photorealism. In some scenarios, such as hunting for a celebrity image, face images might have attributes and features that are exaggerated beyond realism, but that are recognizable enough to convey identity, as shown in Figure 5d. Existing FRSs cannot deal with such exaggerations. More work is needed to advance automatic caricature recognition, which in turn would advance facial representations, indexing, and search engines.

### Novel applications

Although many new face recognition applications have recently emerged, two notable examples are the use of face recognition technology in diagnostics and in newborn identification. In the former area, Oxford University researchers have developed recognition software for diagnosing rare genetic conditions and obtaining hints about ultra-rare genetic disorders that involve changes in face and skull shape.[21] The software can be used to determine these conditions without the individual's knowledge, which raises an interesting question: in people diagnosed with various syndromes, which facial features should an FRS pay attention to and which ones should it ignore?[21] More research is needed to answer this question.

The use of face recognition to identify newborns aims to address the global problem of newborn switching and abduction, particularly in developing countries, where hospitals lack trained personnel and technical management to oversee births and birth registrations. A few researchers have used recognition in newborn identification,[1,7] but, to our knowledge, no FRS has yet been developed that can be used in this way. The main obstacles are pose and expression covariates and the lack of large-scale databases.

Despite implementation challenges and larger societal issues, face recognition remains a promising biometric technology. Our review reveals significant progress over the past two decades as well as unresolved issues related to accuracy, security, and user privacy. We are confident that, with time and focused interdisciplinary research and development, face recognition

## ABOUT THE AUTHORS

**ZAHID AKHTAR** is a postdoctoral researcher in the National Institute for Scientific Research's Energy, Materials and Telecommunications Center (INRS-EMT) at the University of Quebec. While conducting the research reported in this article, he was a research associate in the Department of Mathematics and Computer Science at the University of Udine. His research interests include computer vision, pattern recognition, and image processing with applications in biometrics, affective computing, and security systems. Akhtar received a PhD in electronic and computer engineering from the University of Cagliari. He is a member of the Italian Association for Pattern Recognition. Contact him at zahid.eltc@gmail.com.

**AJITA RATTANI** is a postdoctoral fellow in the Department of Computer Science and Electrical Engineering at the University of Missouri. Her research interests include biometrics, image processing, and computer vision. Rattani received a PhD in electronic and computer engineering from the University of Cagliari. She is a member of the Italian Association for Pattern Recognition. Contact her at rattania@umkc.edu.

will reach its full potential in a wide range of application domains.

## REFERENCES

1. S.Z. Li and A.K. Jain, *Handbook of Face Recognition*, Springer, 2011.
2. S.Z. Li and A.K. Jain, *Encyclopedia of Biometrics*, Springer, 2014.
3. S. Zafeiriou, C. Zhang, and Z. Zhang, "A Survey on Face Detection in the Wild: Past, Present and Future," *Computer Vision and Image Understanding*, vol. 138, 2015; doi:10.1016/j.cviu.2015.03.015.
4. G.B. Huang et al., *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*, tech. report 07-49, Univ. of Mass., 2007.
5. H. Drira et al., "3D Face Recognition under Expressions, Occlusions, and Pose Variations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, no. 9, 2013, pp. 2270–2283.
6. H. Ho and R. Chellappa, "Pose-Invariant Face Recognition Using Markov Random Fields," *IEEE Trans. Image Processing*, vol. 22, no. 4, 2013, pp.1573–1584.
7. N. Kumar et al., "Describable Visual Attributes for Face Verification and Image Search," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, 2011, pp. 1962–1977.
8. J. Wright et al., "Robust Face Recognition via Sparse Representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, 2009, pp. 210–227.
9. Y. Taigman et al., "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (CVPR 14), 2014, pp. 1701–1708.
10. Z. Akhtar, *Security of Multimodal Biometric Systems against Spoof Attacks*, PhD thesis, Dept. of Electrical and Electronic Engineering, Univ. of Cagliari, 2012.
11. C. Rathgeb and A. Uhl, "A Survey on Biometric Cryptosystems and Cancelable Biometrics," *EURASIP J. Information Security*, vol. 3, 2011; doi:10.1186/1687-417X-2011-3.
12. Y. Jin, J. Lu, and Q. Ruan, "Coupled Discriminative Feature Learning for Heterogeneous Face Recognition," *IEEE Trans. Information Forensics and Security*, vol. 10, no. 3, 2015, pp. 640–652.
13. C. Chen, A. Dantcheva, and A. Ross, "An Ensemble of Patch-Based Subspaces for Makeup-Robust Face Recognition," *Information Fusion*, vol. 32, issue PB, 2015, pp. 80–92; doi:10.1016/j.inffus.2015.09.005.
14. S. Wang, M. Shao, and Y. Fu, "Attractive or Not? Beauty Prediction with Attractiveness-Aware Encoders and Robust Late Fusion," *Proc. 22nd ACM Int'l Conf. Multimedia* (MM 14), 2014, pp. 805–808.
15. E. Newton, L. Sweeney, and B. Malin, "Preserving Privacy by De-identifying Facial Images," *IEEE Trans. on Knowledge and Data Eng.*, vol. 17, no. 2, 2005, pp. 232–243.
16. S. Xia, M. Shao, and Y. Fu, "Kinship Verification through Transfer Learning," *Proc. Int'l Joint Conf. Artificial Intelligence* (IJCAI 11), 2011, pp. 2539–2544.
17. Y. Fu, G. Guo, and T.S. Huang, "Age Synthesis and Estimation via Faces: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, 2010, pp. 1955–1976.
18. R. Singh et al., "Plastic Surgery: A New Dimension to Face Recognition," *IEEE Trans. Information Forensics and Security*, vol. 5, no. 3, 2010, pp. 441–448.
19. Z. Huang et al., "A Benchmark and Comparative Study of Video-Based Face Recognition on COX Face Database," *IEEE Trans. Image Processing*, vol. 24, no. 12, 2015, pp. 5967–5981.
20. W. Chu, F. Torre, and J. Cohn, "Selective Transfer Machine for Personalized Facial Action Unit Detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (CVPR 13), 2013, pp. 3515–3522.
21. Q. Ferry et al., "Diagnostically Relevant Facial Gestalt Information from Ordinary Photos," *eLife*, 24 June 2014; doi:10.7554/eLife.02020.