A Comprehensive Database for Benchmarking Imaging Systems

Karen Panetta, *Fellow, IEEE*, Qianwen Wan, *Student Member, IEEE*, Sos Agaian, *Fellow, IEEE*, Srijith Rajeev, *Student Member, IEEE*, Shreyas Kamath, *Student Member, IEEE*, Rahul Rajendran, *Student Member, IEEE*, Shishir Paramathma Rao, *Student Member, IEEE*, Aleksandra Kaszowska, Holly A. Taylor, Arash Samani, *Senior Member, IEEE*, and Xin Yuan

Abstract—Cross-modality face recognition is an emerging topic due to the wide-spread usage of different sensors in day-to-day life applications. The development of face recognition systems relies greatly on existing databases for evaluation and obtaining training examples for data-hungry machine learning algorithms. However, currently, there is no publicly available face database that includes more than two modalities for the same subject. In this work, we introduce the Tufts Face Database that includes images acquired in various modalities: photograph images, thermal images, near infrared images, a recorded video, a computerized facial sketch, and 3D images of each volunteer's face. An Institutional Research Board protocol was obtained and images were collected from students, staff, faculty, and their family members at Tufts University. The database includes over 10,000 images from 113 individuals from more than 15 different countries, various gender identities, ages, and ethnic backgrounds. The contributions of this work are: 1) Detailed description of the content and acquisition procedure for images in the Tufts Face Database; 2) The Tufts Face Database is publicly available to researchers worldwide, which will allow assessment and creation of more robust, consistent, and adaptable recognition algorithms; 3) A comprehensive, up-to-date review on face recognition systems and face datasets.

Index Terms—The tufts face database, computerized face sketches, thermal, 3D, near infrared, face recognition, cross-modality

1 Introduction

Algorithms performing autonomous facial recognition have many profound implications in the fields of homeland security, human trafficking, law enforcement, biometric identification, and multimedia [1], [2], [3], [4]. Extensive research on the topic has led to critical developments in the area [5], [6] yet the field continues facing difficulties with heterogeneous face recognition (HFR) algorithms: matching face images across different image types, such as thermal, near-infrared (NIR), 2D, 3D, and other modalities (See Fig. 1). Often times, the motivation behind HFR is there is only one modality of face images can be acquired as probe images [7]. For instance, face sketches of a criminal are sometimes the only evidence in law enforcement applications [8]. Accurate cross-modality face matching is necessary due to the increasing variety of imaging sensors in different real-life scenarios, such as depth, optical, capacitive, and thermal: for example, night vision cameras can provide more information in low/no light military applications. Furthermore, researchers developing facial recognition systems are adopting multi-modal information to overcome the obstacles to achieve a higher accuracy. It has been shown that the combination of multiple imaging sensor information, such as $2D + 3D$, can significantly improve facial recognition rate [9].

Such cross-modality matching algorithms need to be extensively tested and validated in order to ensure their robustness in recognizing faces across images from advanced sensors and traditional surveillance cameras [10]. Therefore, it is crucial to provide researchers with the ground truth images to benchmark their recognition algorithms and aid in the development of potential solutions to the existing HFR challenges.

Currently, there is no publicly available multi-modality face database that includes visible, near-infrared, thermal, computerized sketch, video and 3D images. Researchers are limited to test their algorithms on very narrowly constrained test samples for one or two modalities. The Tufts Face Database, the most comprehensive multi-modality face database published to date, addresses this issue as it includes six different types of image modalities: photographic images, a computerized facial sketch created from a volunteer's portrait using the FACEs software, thermal images, near infrared (NIR) images, a video, and 3D images of the participant. This database will provide researchers

- *K. Panetta, Q. Wan, S. Rajeev, S. Kamath, R. Rajendran, S. P. Rao, A. Samani, and X. Yuan are with Department of Electrical and Computer Engineering, Tufts University, Medford, MA. E-mail: karen@ece.tufts.edu, {qianwen.wan, srijith.rajeev, skamat04, rahul.rajendran, shishir.rao, Xin.Yuan}@tufts.edu, samani@ieee.org.*
- *S. Agaian is with the City University of New York, New York, NY. E-mail: Sos.Agaian@csi.cuny.edu.*
- *A. Kaszowska and H. A. Taylor are with Department of Psychology, Tufts University, Medford, MA. E-mail: {aleksandra.kaszowska, holly.taylor}@tufts.edu.*

Manuscript received 28 June 2018; accepted 25 Nov. 2018. Date of publication 30 Nov. 2018; date of current version 4 Feb. 2020.
(Corresponding author: Qianwen Wan.)
Recommended for acceptance by G. Shakhnarovich.
Digital Object Identifier no. 10.1109/TPAMI.2018.2884458

Fig. 1. Heterogeneous face recognition in real-life scenarios: images captured through different sources, such as (a) social media; (b) driver's licenses, passports, and other identification documents; (c) night-vision surveillance cameras; (d) Thermal cameras; and (e) 3D cameras.

developing new recognition algorithms with a dataset to use as a benchmark to validate performance and accuracy.

An Institutional Research Board protocol was developed, and more than 10,000 raw images were collected from 113 volunteers of different nationalities, gender identities, ages, and ethnicity at Tufts University.

The paper presents the following:

- An up-to-date review of existing facial recognition algorithms and systems.
- A comparative summary of publicly available face datasets.
- Information on the Tufts Face Database, a comprehensive face database, which contains face images captured in multiple modalities.

The modalities include different variations of 2D face photographs, computerized facial sketches, thermal face photos, near infrared face images, a recorded video, and 3D face images. All 6 modalities are available for every one of 113 individuals depicted in the database. This database will be available to researchers worldwide in order to benchmark facial recognition algorithms for sketch, thermal, NIR, 3D face recognition and heterogamous face recognition.

This paper is organized as follows. Section 2 presents a brief review of the literature and related work. In Section 3, the acquisition procedure of the Tufts Face Database is described. Then, the instructions for accessing the database are outlined in Section 4. Finally, Section 5 summarizes the contributions of this work and discusses the future directions.

2 Related Work

This section outlines different types of recognition systems currently in use, their existing challenges and forthcoming advances, as well as a comparison of the Tufts Face Database to other existing databases.

2.1 Different Types of Face Recognition Systems, Their Development, and Applications

Conventional visible face recognition systems have been used in a broad range of real-life applications in fields such as law enforcement, homeland security, biometric identification, cognitive psychology, and entertainment [11]. Generally, an automatic real-time facial recognition system comprises of two steps [12]: face detection and face recognition. Despite extensive work done in the field [13], automated facial recognition systems still lack in accuracy when compared to human performance [14], [15]. The existing challenges for an automated facial recognition system include: uncontrolled conditions caused by low-quality surveillance cameras or sensors, unpredictable environments, difficulty in detecting faces in video content, emulating the effects of aging, low tolerance to the variation in facial expressions, illumination changes [16], and the difference in body positions [14].

Sketch-to-face recognition is a commonly used method to assist police to identify persons of interest, such as suspects or missing persons, especially when there is limited access to other evidence [17], [18]. Usually there are two types of sketches: hand-drawing sketches and computerized sketches, both based on an eyewitness' testimony. Most law enforcement agencies are adopting computerized sketches to identify criminals due to its cost efficiency over hand-drawn sketches [19]. The existing challenges of a sketch-to-face recognition system include: significant differences in shape and textures across modalities, and limited amount of testing and training data. Methods such as component-based MLBP, holistic SIFT feature descriptors, and Log-Gabor filtering were utilized and tested to perform promising computerized sketch-to-face recognition [20], [21], [22], [23]. Moreover, intra-modality and inter-modality algorithms for face-sketch recognition can be fused together to yield better performance [24]. Sketch photo synthesis methods [25] were introduced to improve the recognition accuracy of computerized sketch-to-face recognition by reducing the discrepancies between real facial photos and computerized sketches. Most recently, deep learning algorithms, such as DCNN [26], are being developed for face photo-sketch recognition, where the neural network is fine-tuned by using transfer learning and traditional data augmentation methods.

Thermal face recognition is widely used in biomedical applications. This type of image contains reliable and distinguishable physiological biometric features [27], [28]. The amount of radiation emitted by an object increases with temperature; therefore, thermography depicts variations in temperature. When viewed through a thermal imaging camera, warm objects stand out well against cooler backgrounds; humans become easily visible against the environment, regardless of environmental illumination. As a result, facial thermography is particularly useful to first responders during calamities and in military applications. However, thermal imaging technology has several drawbacks in the context of autonomous facial recognition system. The infrared face images displays the body heat pattern, which can be easily affected by ambient temperature, air flow conditions, exercise, illness, and drugs [29]. Also, thermal face images provide less facial features when compared to visible face images. Currently, feature extraction algorithms such as gray-scale value histogram features [30], thermal local binary pattern (LBP), SIFT features [31], and speeded up robust features (SURF) [32], [33], are widely applied in thermal face recognition. Image transformations, such as wavelet transform [34] and Gabor filtering [35], were used to enhance the robustness of extracting face appearance features. Meanwhile, in order to match the thermal face image to the visible face image, algorithms such as partial least squares-discriminant analysis (PLS-DA) [36], [37] that reduce the modality gap were proposed. Furthermore, leading machine learning techniques and classifiers were applied to construct thermal face recognition systems [7], [38]. Machine learning methods such as the Optimized super-pixel and AdaBoost classifier were combined for human thermal face recognition [39]; Generative Adversarial Network (GAN) based method [40] was developed to address challenges like

TABLE 1
Advantages and Limitations of Popular Imaging Sensors

Characteristics	Thermal/Near infrared	3D	Visible
Visibility in low/no light	yes	yes	no
Record image through translucent obstacles	yes	no	no
Provide color content information	no	yes	yes
Display surface temperatures of solid objects	yes	no	no
Distinguish objects at varying distances	no	yes	yes
Image quality	low	low	high
Presence of noise	low	low	high
Commercial cost	high	high	low
Consumer applications	limited	limited	yes
Existing tools for facial recognition	limited	limited	more

occlusions, different skin tones and limited thermal training data available for face recognition.

Near-infrared (NIR) face recognition systems were proposed to overcome the challenges of no illumination and varying pose [41]. Imaging beyond the visible spectrum has recently received a lot attention from the research community [42] as opens a new way for face detection, localization and recognition [43], [44]. Proper feature descriptors and classifiers are found in literature for visible (VIS)-to-NIR face recognition. Wasserstein convolutional neural network (WCNN) approach was proposed to learn invariant features between NIR and VIS face images [45]. Liao et al. [46] combined the difference of Gaussian (DoG) filter and multi-block local binary pattern (MBLBP) to represent face features. In [47], HoG and LBP descriptors were utilized to explore the accuracy of VIS-to-NIR face recognition system. Furthermore, a multi-view smooth discriminant analysis was proposed to learn a common discriminative feature space for matching visible and NIR face images [48]. Canonical correlation analysis estimated the similarity between NIR and VIS images by training on NIR/VIS image pairs of the same individual [49]. Chen et al. [50] converted NIR images to synthetic VIS images for matching by using local linear embedding with a dictionary of corresponding NIR/VIS face pairs. Handling cross-spectral face images is still an open research topic.

3D face recognition systems have seen a large spike in interest in the past few years. A 3D scanner can capture the surface texture and the depth information. 3D scanners can generate higher image quality with distortion-free images, can capture partial scans more effectively as they have potential for better description of facial features enable creation of synthetic face images, and are invariant to the point of view [51]. Most importantly, the depth information from a 3D image can be combined with 2D features to deliver higher recognition rates [51]. As 3D technology becomes more prevalent and affordable, the demand for real world applications such as 3D image reconstruction algorithms and 3D face recognition continue gaining more attention [52], [53], [54]. Moreover, 3D scanners have recently been recommended to address the challenges of 2D recognition systems [55]. In recent years, several 3D acquisition algorithms have been developed, among them structured light, stereo, and structure from motion. Structure from Motion (SfM) is a more practical approach for facial 3D reconstruction in a real-world environment as this algorithm constructs a 3D image from a sequence of images taken from different perspectives. As such, a 3D image of an individual can be reconstructed from multiple surveillance camera images. Surface registration-based and feature matching-based approaches are the two main categories in 3D face recognition. The surface registration-based approaches commonly use natural or variant Iterative Closest Point (ICP) algorithms [56]. Feature matching-based approaches match two 3D facial surfaces defined in different coordinate systems based on object-centric shape features. In the literature there are several methods for performing 3D face recognition. Medioni and Waupotitsch [57] performed 3D face recognition using iterative closest point algorithm for matching surface of the 3D face. Lao et al. [58] used 2D edges and iso-luminance contours to locate irises. Features are determined using the location of the iris as a reference point, enabling subsequent face standardization is performed. Chang et al. [59] used principal component analysis based recognition using a combination of 3D and 2D images. The experiment resulted in 99 percent accuracy for $3D + 2D$ modality and 94 and 89 percent for 3D and 2D alone, respectively.

Heterogonous face recognition (HFR) matches face images across different image types. As diverse imaging sensors are widely used in practical applications, heterogonous face recognition is now attracting growing attentions in both research and industry community. For instance, face recognition using thermal camera, near infrared camera and 3D depth camera emerged in literature as new modalities addressing issues of complicated light conditions, biometric identification, and military applications [60]. Furthermore, the law enforcement's reliance on computerized face sketches in the absence of photographic information results in a growing need for such modalities to be included in HFR approaches. The advances in image collection and widespread reliance on different sensors is the root of HFR's growing importance within both research and industry. Table 1 summarizes the advantages and limitations of popular imaging sensors, such as thermal/near infrared cameras, 3D camera, and visible light cameras.

Research continues facing difficulties that arise from the modality gaps between heterogeneous face images. Conventional approaches can be categorized as synthesis based methods [61], [62], [63], [64], [65], common space projection based methods [66], [67], [68], [49], [69], and feature descriptor based methods [7], [70]. Furthermore, a graphical representation based HFR method (G-HFR) was described in [71]. More works can be found in [7], [10], [72], [73], [74].

2.2 Existing Databases Summary and Comparison

Significant efforts have been put into building face databases to satisfy the basic needs of fast development of facial recognition research and applications. Table 2 provides a brief overview of the most popular databases built for researchers to evaluate facial recognition algorithms, as well as providing sufficient amount of training data for machine learning algorithms.

2D face databases are used in most face recognition research topics. The FERET [75], [76] database was a pioneering large-scale color database built in the late 90s. FRVT (Face Recognition Vendor Test) 2000 [77] and 2002 [78]

TABLE 2
A Comparison of Popular Face Datasets

Database Name & References	2D	Thermal	NIR	CS	3D	#people	Highlights
10k US Adult Faces Database [105]	✓					≈10,168	Natural face photographs and several measures for 2,222 of the faces.
The AR Face Database [81]	✓					126	Color face images with different expressions, illumination conditions, and add-ons accessories like sunglasses and scarves.
Yale Database [83]	✓					15	165 grayscale images in GIF format.
AT&T Database [81]	✓					40	10 different images (time, lighting, expression, glasses) of each person
Caltech Faces [88]	✓					27	450 frontal face images.
CAS-PEAL Face Database [89]	✓					1040	A large-scale Chinese face database
The CMU Multi-PIE Face Database [80]	✓					337	Improve version of PIE face database; images were captured under 15 view points and 19 illumination conditions.
Cohn-Kanade AU Coded Facial Expression Database [106]	✓					100	Given emotion labels for different facial expressions.
The Color FERET Database [75]	✓					1199	Large-scale, one of the face database pioneer work.
Indian Movie Face database [107]	✓					100	34512 images of Indian actors
Japanese Female Facial Expression (JAFFE) Database [90]	✓					10	7 facial expressions (6 basic facial expressions + 1 neutral) depicting 10 Japanese female subjects.
Karolinska Directed Emotional Faces (KDEF) [108]	✓					70	Suitable for perception, attention, emotion, memory and backward masking studied
Labeled Faces in the Wild [109]	✓					5749	13,000 images of faces collected from the web
NIST Mugshot Identification Database [110]	✓					1573	Large scale mugshots face images with front and side view.
PUT Face Database [111]	✓					100	High resolution images with pose variations and expression
Natural Visible and Infrared Facial Expression (USTC-NVIE) [102]	✓	✓				100	Visible and infrared facial expression database
IRIS Thermal/Visible Face Database [104]	✓	✓				30	Visible and thermal face images with expression, pose, and illumination
The Terravic Facial Infrared Database [112]		✓				20	Thermal face images for 20 people with accessories
The Hong Kong Polytechnic University NIR Face Database [113]			✓			350	Large scale NIR face database from
CASIA NIR-VIS 2.0 database [114]	✓		✓			725	Pairs of visible (VIS) and NIR images
Long Distance Heterogeneous Face Database (LDHF-DB) [115]	✓		✓			100	Visible (VIS) and near-infrared (NIR) face images at distances of 60m, 100m, and 150m outdoors and a 1m distance indoors.
The PRIP Viewed Software-Generated Composite (PRIP-VSGC) database [20], [21]				✓		123	Computerized sketch build based on 123 participants from the AR face database
The UoM-SGFS database [101]				✓		600	A composite sketch dataset based on FERET dataset
3D Mask Attack Database (3DMAD) [116]	✓				✓	17	76500 frames covering 17 people under controlled conditions, frontal view and neutral expression
3D_RMA database [117]					✓	120	Created with a 3D acquisition system using structured light.
The Basel Face Model (BFM) [118]					✓	200	Uses registered 3D scans of 100 male and 100 female faces
Binghamton University Facial Expression Database (BU-3DFE/BU-4DFE) [119]	✓				✓	100	The database has large age range, nationality diversity; high-resolution 3D dynamic facial expression database
The Bosphorus Database [120]	✓				✓	105	For 2D/3D face recognition and 3D face reconstruction
The Extended M2VTS Database [86]	✓				✓	295	32 KHz 16-bit sound files, video sequences or as a 3d Model
Texas 3D Face Recognition Database (Texas 3DFRD) [121]					✓	105	Facial color and range images of 105 adults
The EURECOM Kinect Face Dataset [122]					✓	52	A Kinect database of images from 52 people of different facial expressions/different lighting/occlusion conditions
The University of Milano Bicocca 3D face database [123]	✓				✓	143	2D/3D images from 98 males, 45 females; a pair of male twins and a baby included)
The Tufts Face Database	✓	✓	✓	✓	✓	113	The most comprehensive, large-scale (over 10,000 images, 74 females + 39 males, from more than 15 countries with an age range between 4 to 70 years old) face dataset that contains 6 image modalities: visible, near-infrared, thermal, computerized sketch, a recorded video, and 3D images.

challenges furthermore improved the face recognition system evaluation protocols and databases construction to meet real-world applications, which overall provide images for 37,437 individuals. CMU PIE [79] collected by Carnegie Mellon University consists of over 40,000 facial images of 68 people and is one of the most widely used face database for evaluation of pose and illumination variations. Eight years later, Multi-PIE [80] database improved upon the CMU PIE database in a number of categories, increasing the number of participants and facial expression variations. During 90's the development face databases were populated, such as the AR face database [81] (1998, Purdue University), the ORL [81] also known as the ATT database (1994, Cambridge University), MIT face database [82] (1991, Massachusetts Institute of Technology), Yale face database [83] and Yale-B database [84], [85] (1997-2005, Yale University), M2VTS and XM2VTSDB face database [86] (1997, University of Surrey), Georgia Tech face database [87] (1999, Georgia Institute of Technology) and Caltech face database [88] (1999, California Institute of Technology).

Other publicly accessible face databases provide specific test cases for facial recognition for a variety of nationalities, races, and sexes. For example, CAS-PEAL [89] is a large-scale Chinese face database which contains 1040 Chinese participants; The Japanese Female Facial Expression (JAFFE) database [90] contains seven facial expressions posed by 10 Japanese female models; the Indian face database collected frontal face images of 40 south-east asian participants taken in February, 2002 on the IIT Kanpur campus [91]; The Iranian Face Database (IFDB) [92] includes faces of middle-eastern Iranian participants along with information such as volunteers' occupation, skin color, fingerprint, and cosmetic markers (surgical points, fracture or laceration on face). Meanwhile, there are other databases were built for specific practical applications. For instance, SCface (surveillance cameras face database) is an image dataset of uncontrolled indoor environment using surveillance cameras of various qualities; Labeled Faces in the Wild (LFW) [93] is a database of face photographs designed for studying the problem of unconstrained environmental parameters such as position, pose, lighting, expression, background, camera quality, occlusion, age, and gender; YouTube faces database [94] contains 3,425 videos of 1,595 participants designed for studying the problem of unconstrained face recognition in videos; the plastic surgery face database [95] consists of 1800 pre- and post-surgery images depicting 900 participants, which aims to open an new dimension of facial recognition by accounting for surgical interventions and alterations; the Large Age-Gap (LAG) dataset [96] is designed to aid solving the aging problem in facial recognition. The proposed Tufts Database includes a set of pictures with occlusions (such as glasses) and exaggerated facial expressions to allow for robust validation of existing algorithms and would serve as a testbed for development of approaches allowing to minimize the impact of abovementioned noise on recognition accuracy.

Computerized Sketch (CS) face databases have been constructed for evaluation of computerized sketch (CS) to face recognition approaches. Autonomous matching of computerized sketches to face photos is becoming widely studied largely due to the demand from law enforcement agencies. A composite sketch database called The PRIP Viewed Software-Generated Composite (PRIP-VSGC) database was built in [20], [21], which contains computerized sketches for 123 participants based on the images from the AR face database. Extended PRIP (e-PRIP) [97], [98] sketch database was created based on PRIP dataset, further adding a set of composite sketches prepared by an Indian user. The UoM-SGFS database, which is a composite sketch dataset based on the images from the FERET dataset, doubled the number of participants compared to previous work from 2016, now containing 1200 images of 600 participants in 2018 [99], [100], [101]. For composite sketch datasets, photographs of participants were used as references in creating the composite sketches. While there has been a huge spike in interest for computerized sketch to face recognition systems, the availability of robust databases for testing are limited as the sketches are not based on live portrait sittings of the individual. Having a large robust database available where the computerized facial sketches are derived from live portrait sittings of an individual is critical for computerized sketch face recognition system development, validation and testing. The Tufts Face Database is the first large-scale computerized sketch dataset that was created based on live participant sittings.

Infrared thermal images, which record the temperature distribution over faces and the vein branches, are not sensitive to imaging conditions. Thus, thermal face recognition is a powerful technology for biometric identification applications. One of the most commonly used databases is the Natural Visible and Infrared facial Expression Database (USTC-NVIE) [102], which contains both spontaneous and posed expres sions of more than 100 participants, recorded simultaneously by a visible light and an infrared thermal camera. NIST Equinox [103] consists of thermal face images including: co-registered broadband-visible/LWIR(8-12 microns), MWIR(3-5 microns), and SWIR(0.9-1.7 microns); however, it is no longer available online. IRIS [104] provides visible and thermal face image pairs of 30 participants.

Other thermal face images are Carl database [124], [125] Terravic Facial IR Database [112] and Kotani Thermal Facial Emotion (KTFE) [126]. However, availability of thermal image face databases is limited: there is a lower demand for thermal images as compared to 2d databases, and thermal images are significantly more difficult to obtain. The Tufts face database contains a large-scale thermal facial database, with pose variance and facial expression, which provides a benchmark for thermal face recognition research.

A limited number of *NIR face databases* are present in the literature. The CASIA NIR-VIS 2.0 database [114] is the largest face database across NIR and VIS spectrum to date. The PolyU-NIRFD contains images from 350 participants, each contributing about 100 samples with variations of pose, expression, focus, scale, and time [113]. The University of California/Irvine collected a dataset of multispectral images under halogen ambient illumination [44]. The LDHF-DB contains face images collected in an outdoor environment at distances of 60 meters, 100 meters, and 150 meters, with both visible light (VIS) face images captured in daytime and near infrared (NIR) face images captured in nighttime [115].

The cost and time required to capture *3D images* are steadily decreasing, hence researchers have attempted to develop face recognition problems directly on a 3D face

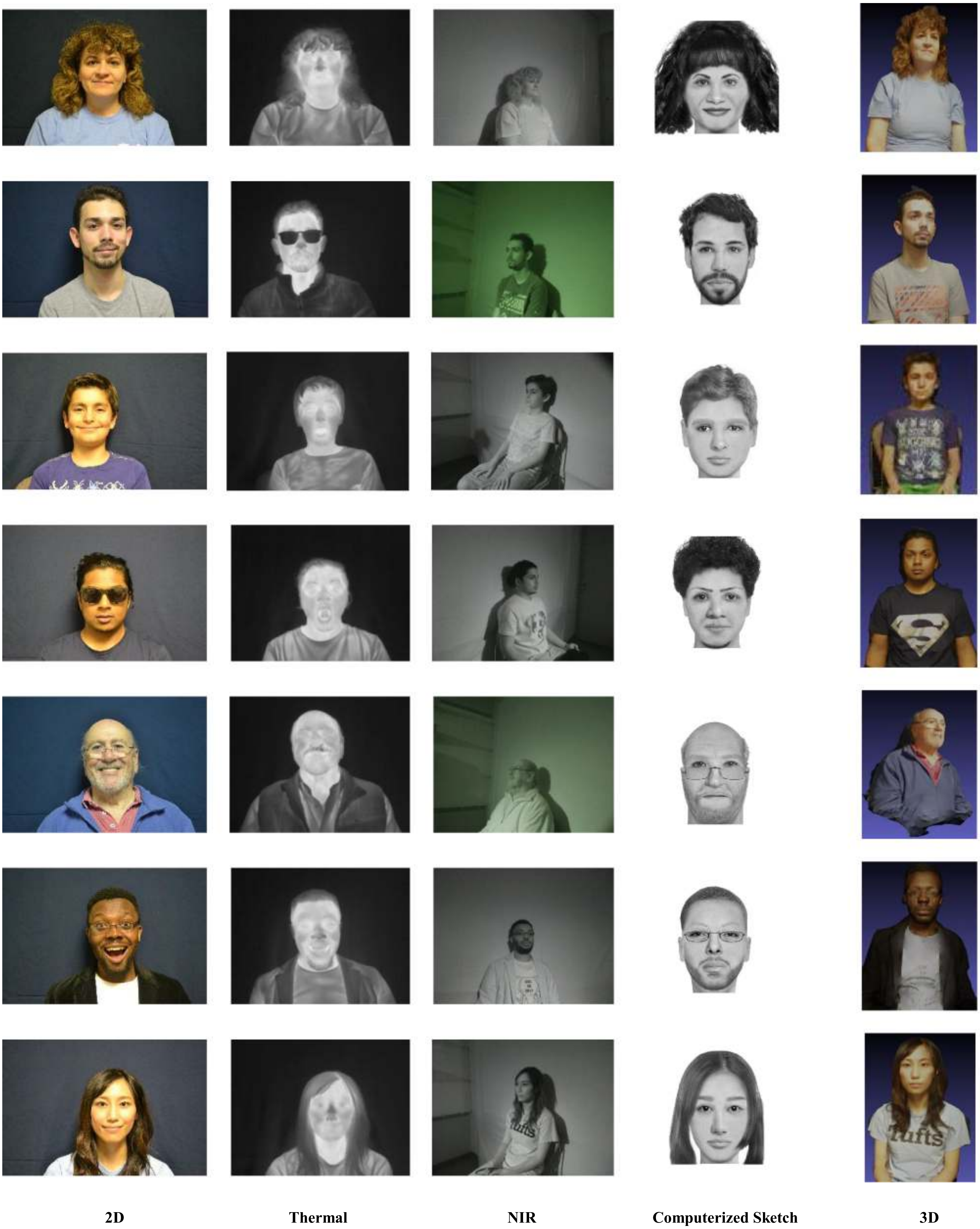
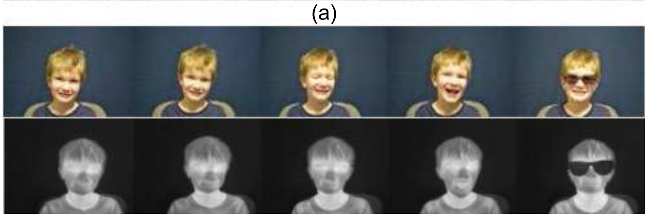


Fig. 2. Example images selected from the Tufts face database. The Tufts database has a wide range of nationalities, ages, and ethnic backgrounds.

model [127]. Part of the face recognition grand challenge (FRGC) [109] offers measures to evaluate the performance of 3D face recognition, which is one of the pioneering works on 3D face datasets. Another early 3D face dataset is 3D RMA [117] face dataset, which was collected by a 3D acquisition system based on structured light with a projector and a camera of low quality and precision, which appears sufficient for prototype testing. MIT-CBCL Face Recognition Database [108] contains synthetic 3D morphable head models of the 10 participants using high resolution 2D images. GavabDB [128] is a 3D face dataset of 61 Caucasian participants built for image applications such as pose correction

(a)



(b)

Fig. 3. (a) and (b) are two sets of frontal images of an individual with the different constraints for both visible and thermal. The images in the first row of each set illustrate visual images of a participant with various facial expressions; and the images in the second row of each set provide corresponding expressions in thermal imaging.

and 3D face model registration. Multimodal database (2D, 2.5D and 3D) FRAV3D Database [129] was captured by specific hardware devices, a Minolta VIVID 700 scanner, developed by researchers focused on RGB-D face recognition. Other well-known and publicly available 3D face databases are BU-3DFE [119], BU-4DFE [130], BP4D [131], Bosphorus [132], Texas 3D Face Recognition Database (Texas 3DFRD) [121], and the EURECOM Kinect Face Dataset (EURECOM KFD) [122]. The Tufts Face Database will provide 3D face images captured by a LYTRO camera, and a 3D face model (a high-quality 3D point cloud reconstructed facial images) for researchers to test and evaluate the robustness of 3D face recognition systems specifically against pose variations. Furthermore, the Tufts Face Database provides the 2D images and a recorded video that were used to generate the 3D face model for researchers who focus of 3D face image reconstruction.

3 The Tufts Face Database Acquisition System

Image acquisition for the Tufts Face Database took place in a 9 $\times$10 ft photographic room. 2D face images were collected from people with different expressions and accessories, under uniformed lighting and constant background conditions. FACEs sketch software was used to develop computerized sketches. Nine camera positions were derived to support the

TABLE 3
Detailed Camera Settings for the Tufts Database

Front face	RGB Thermal	Nikon D3100 FLIR Vue Pro
Around the participant	RGB+ NIR Thermal Mirrorless	QuadCam (self-crafted) FLIR Vue Pro LYTRO ILLUM 40 Megaray Light Field
Video	RGB	QuadCam
Lighting	Visible light NIR light	Diffused light with LED bulbs 850 nm Infrared 96 LED light system

Fig. 4. Selecting facial components using FACES 4.0.

3D face image acquisition and reconstruction. Moreover, a specific thermal and NIR camera systems, temperature settings, and lighting variations were configured for collecting thermal face images. Fig. 2 shows example images selected from the newly proposed Tufts Face Database.

3.1 Existing Databases Summary and Comparison

An institution research board (IRB) protocol was developed in order to collect and publish the Tufts Face Database. The IRB protocol discussed the purpose and procedure of the study, the confidentiality and public access of the database, as well as the protection of participants' identity. Each photographic session took approximately 15 minutes.

The raw image data was collected from 113 volunteers; 100 participant's images were selected and assessed for good image quality.

3.2 Hardware and Software Configuration

Front face camera setting. Acquisition of frontal images was performed using a Nikon D3100 DSLR camera (visual or VIS) and a FLIR Vue pro camera (thermal). The cameras were mounted on tripods and the height of each camera was adjusted manually to correspond to the image center. The distance to the participant was strictly controlled during the acquisition process. A constant lighting condition was maintained using diffused lights. Each participant was asked to pose with (1) a neutral expression, (2) a smile, (3) eyes closed, (4) exaggerated shocked expression, (5) sunglasses. In addition, individuals were instructed to remove any eyewear during frontal thermal acquisition. Fig. 3 shows examples of sets of frontal images of each individual with the different constraints for both visible and thermal.

Each participant was seated in front of a blue background in close proximity to the camera.

Angular acquisition camera settings. Angular acquisition was performed by using 3 different cameras: a LYTRO ILLUM 40 Megaray Light Field Camera, a FLIR Vue Pro camera, and a self-crafted QuadCam (an array of 4 cameras) NIR+RGB device. Each individual was asked to look at a fixed viewpoint while the cameras were moved to 9 equidistant positions forming an approximate semi-circle around the individual. The lighting condition for NIR imaging was maintained by using an 850nm Infrared 96 LED light system. The 3D models were reconstructed using open-source structure-from-motion algorithms. Table 3 below demonstrates the detailed camera settings.

Computerized facial sketches. were generated using software FACES 4.0 [70], one of the most widely used software packages by law enforcement agencies, the FBI, and the US Military. The software allows researchers to choose a set of candidate facial components from the database based on their observation or memory. Fig. 4 illustrates the procedure for creating a computerized facial sketch using FACES 4.0 by choosing facial components (e.g., hairstyle, head shape, eye

TABLE 4
Various Attributes of the Files in the Database

Subset	#Variations	#Participants	#Images
2D	Expressions Accessory Camera position	100	4,100
Thermal	Expressions Accessory Camera position	100	1,400
NIR	Camera position	100	3,600
3D	Point Cloud LYTRO-3D	100	200
Video	RGB	100	—
Computerized Sketch	Frontal	100	100

brows, eyes, nose, mouth, jaw shape, mustaches, beards, goatees, forehead lines, eye lines, smile lines, mouth lines, chin lines, facial markings such as scars, moles, piercings, tattoos and earrings, and using hats, headwear and eyeglasses). To the best of our knowledge, the Tufts Face Database is the first database to contain computerized facial sketches built on portrait sittings instead photographs.

4 How to Use the Database

The Tufts Face Database contains over 10,000 images of 113 participants in total. These images belong to six main subsets: the 2D subset, the thermal subset, the near infrared subset, the 3D subset, video subset and the computerized sketch subset. Raw data and the best quality 100 participants face data can be provided per request. Table 4 summarizes the contents of the best 100 participants from the Tufts database.

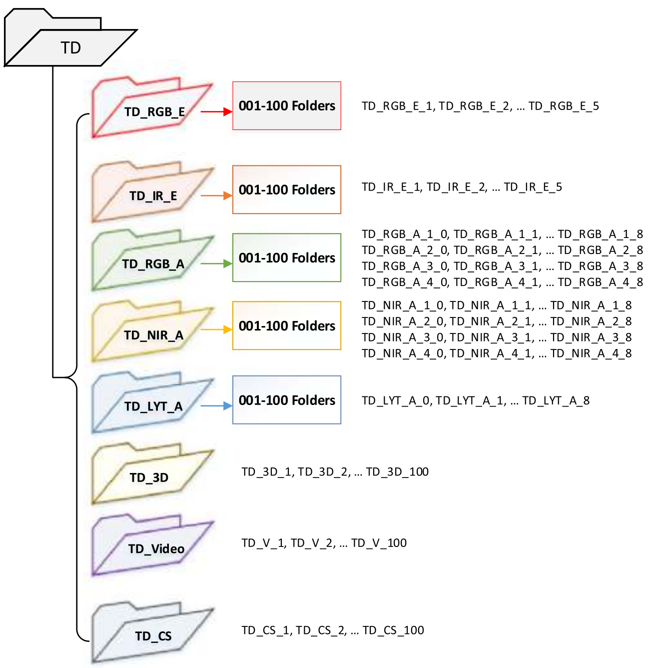


Fig. 5. Database naming convention. Detailed explanation is presented in Section 4.1. We believe our image naming convention will provide a convenient way for indexing.

TABLE 5
Contents in the Tufts Database (Best Quality)

Categories	File format	Average size per file
2D Visible	.jpg	3.45 MB
Thermal	.jpg	187 KB
NIR	.jpg	13.5 MB
3D	.ply	20.8MB
3D LYTRO	.lfr	53 MB
Sketch	.jpg	75.9 KB
Video	.h264	40 MB

4.1 Image Naming Convention

In the Tufts Face Database, the image filename encodes the image information. Its format is described in Fig. 5. It consists of 4 fields and are separated by underline marks as shown below, which provides an easy way for researcher to use.

(1) TD field: TD stands for the Tufts Face Database. It is the common filed that shared by all the files.
(2) Camera and image type fields: RGB stands for 2D visible face image; IR stands for infrared thermal face image; NIR stands for near infrared face images; LYT stands for face.
(3) images that captured by 3D LYTRO camera; 3D stands for point cloud reconstructed 3D face images; Video stands for a short video that can be used to reconstruct the 3D face image and to test face detection and recognition in video data; CS stands for the computerized sketch face images.
(4) Expression field. The initial character "E" represents 5 expression variations: natural expression, smile, open mouth, close eyes and wearing sunglasses.
(5) Camera angle field. The initial character "A" represents 9 different camera positions

4.2 Image Formatting

The raw images of the Tufts Database have different file formats that are shown in the Table 5.

4.3 How to Access

Availability of the database is critical for research transparency in the field. The Tufts Face Database will be publicly available for downloading from our research website on Tufts ECE servers. Researchers who want access the database will need to fill out a request form in order to receive the download permission.

The information on how to obtain a copy of the Tufts Face Database can be found on the project website (http://tdface.ece.tufts.edu/).

5 Conclusion

In this work, we reported our newly developed Tufts Face Database, which we believe will be a valuable resource to facilitate 2D, 3D, thermal, near-infrared, computerized sketch, and heterogeneous facial recognition research for multimedia, forensic, security, biometric, and entertainment applications. The database content and acquisition process are introduced; and the method to obtain the dataset is carefully described. In addition, a comprehensive review on face recognition systems and face datasets are presented.

To date, the Tufts Face Database is the most comprehensive large-scale face dataset. We believe that public availability of this database will enable development and evaluation of face recognition methods across multimodal images.

Acknowledgments

The authors would like to thank all the volunteers for their great support and contribution in developing the Tufts Face Database. The authors would also like to thank Dr. Lara Sloboda for helping with the IRB application process.

References

[1] I. Masi, F.-J. Chang, J. Choi, S. Harel, J. Kim, K. Kim, et al., "Learning pose-aware models for pose-invariant face recognition in the wild," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2018.2792453.

[2] L. Best-Rowden and A. K. Jain, "Longitudinal study of automatic face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 1, pp. 148–162, Jan. 2018.

[3] C. Ding and D. Tao, "Robust face recognition via multimodal deep face representation," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 2049–2058, Nov. 2015.

[4] C. Ding and D. Tao, "Trunk-branch ensemble convolutional neural networks for video-based face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 1002–1014, Apr. 2018.

[5] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surveys*, vol. 35, pp. 399–458, 2003.

[6] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[7] B. F. Klare and A. K. Jain, "Heterogeneous face recognition using kernel prototype similarities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1410–1422, Jun. 2013.

[8] J. Lu, V. E. Liong, and J. Zhou, "Simultaneous local binary feature learning and encoding for homogeneous and heterogeneous face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1979–1993, Aug. 2018.

[9] J. Serrano-Cuerda, A. Fernández-Caballero, and M. T. López, "Selection of a visible-light versus thermal infrared sensor in dynamic environments based on confidence measures," *Appl. Sci.*, vol. 4, pp. 331–350, 2014.

[10] S. Ouyang, T. Hospedales, Y.-Z. Song, X. Li, C. C. Loy, and X. Wang, "A survey on heterogeneous face recognition: Sketch, infra-red, 3d and low-resolution," *Image Vis. Comput.*, vol. 56, pp. 28–48, 2016.

[11] Y. Duan, J. Lu, J. Feng, and J. Zhou, "Context-aware local binary feature learning for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1139–1153, May 2018.

[12] S. Nercessian, "Human visual system-based multi-scale tools with biomedical and security applications," PhD dissertation, Dept. Elect. Comput. Eng., Tufts Univ., Medford, Mass., 2012.

[13] W. Deng, J. Hu, and J. Guo, "Face recognition via collaborative representation: Its discriminant nature and superposed representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 10, pp. 2513–2521, Oct. 2018.

[14] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Syst. Man Cybern. Part B*, vol. 36, no. 2, pp. 458–466, Apr. 2006.

[15] J. M. Chaves-González, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, "Detecting skin in face recognition systems: A colour spaces study," *Digit. Signal Process.*, vol. 20, pp. 806–823, 2010.

[16] W. Zhang, Z. Xi, J.-M. Morvan, and L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2018.2803179.

[17] D. McQuiston-Surrett, L. D. Topp, and R. S. Malpass, "Use of facial composite systems in US law enforcement agencies," *Psychology Crime Law*, vol. 12, pp. 505–517, 2006.

[18] K. T. Taylor, *Forensic Art and Illustration*, Boca Raton, FL, USA: CRC Press, 2000.

[19] A. Barbadekar and P. Kulkarni, "A survey of face recognition from sketches," *IJLTET*, vol. 6, no. 3, pp. 150–158, Jan. 2016.

[20] H. Han, B. F. Klare, K. Bonnen, and A. K. Jain, "Matching composite sketches to face photos: A component-based approach," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 191–204, Jan. 2013.

[21] S. J. Klum, H. Han, B. F. Klare, and A. K. Jain, "The FaceSketchID system: Matching facial composites to mugshots," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2248–2263, Dec. 2014.

[22] C. Galea and R. A. Farrugia, "Face photo-sketch recognition using local and global texture descriptors," in *Proc. 24th Eur. Signal Process. Conf.*, 2016, pp. 2240–2244.

[23] S. Klum, H. Han, A. K. Jain, and B. Klare, "Sketch based face recognition: Forensic versus composite sketches," in *Proc. Int. Conf. Biometrics*, 2013, pp. 1–8.

[24] C. Galea and R. A. Farrugia, "Fusion of intra-and inter-modality algorithms for face-sketch recognition," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, 2015, pp. 700–711.

[25] C. Peng, X. Gao, N. Wang, D. Tao, X. Li, and J. Li, "Multiple representations-based face sketch–photo synthesis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 11, pp. 2201–2215, Nov. 2016.

[26] C. Galea and R. A. Farrugia, "Matching software-generated sketches to face photos with a very deep CNN, morphed faces, and transfer learning," *IEEE Trans. Inf. Forensics Security*, Vol. 13, no. 6, pp. 1421–1431, Jul. 2017.

[27] E. Mostafa, R. Hammoud, A. Ali, and A. Farag, "Face recognition in low resolution thermal images," *Comput. Vis. Image Understanding*, vol. 117, pp. 1689–1694, 2013.

[28] M. K. Bhowmik, K. Saha, S. Majumder, G. Majumder, A. Saha, A. N. Sarma, et al., "Thermal infrared face recognition–a biometric identification technique for robust security system," in *Reviews, Refinements and New Ideas in Face Recognition*. London, United Kingdom: InTech, 2011.

[29] M. Hanif and U. Ali, "Optimized visual and thermal image fusion for efficient face recognition," in *Proc. 9th Int. Conf. Inf. Fusion*, 2006, pp. 1–6.

[30] Y. Yoshitomi, T. Miyaura, S. Tomita, and S. Kimura, "Face identification using thermal image processing," in *Proc. 6th IEEE Int. Workshop Robot Hum. Commun*, 1997, pp. 374–379.

[31] G. H. Vigneau, J. L. Verdugo, G. F. Castro, F. Pizarro, and E. Vera, "Thermal face recognition under temporal variation conditions," *IEEE Access*, vol. 5, pp. 9663–9672, 2017.

[32] T. I. Dhamecha, A. Nigam, R. Singh, and M. Vatsa, "Disguise detection and face recognition in visible and thermal spectrums," in *Proc. Int. Conf. Biometrics*, 2013, pp. 1–8.

[33] R. S. Ghiass, O. Arandjelović, A. Bendada, and X. Maldague, "Infrared face recognition: A comprehensive review of methodologies and databases," *Pattern Recognit.*, vol. 47, pp. 2807–2824, 2014.

[34] A. Srivastava, X. Liu, B. Thomasson, and C. Hesher, "Spectral probability models for infrared images and their applications to ir face recognition," in *Proc. Workshop Comput. Vis. Beyond Visual Spectr.*, 2001, pp. 1–10.

[35] P. Buddharaju, I. Pavlidis, and I. Kakadiaris, "Face recognition in the thermal infrared spectrum," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshop*, 2004, p. 133.

[36] J. Choi, S. Hu, S. S. Young, and L. S. Davis, "Thermal to visible face recognition," in *Proc. Sens. Technol. Global Health Military Med. Disaster Response Environ. Monitoring II; and Biometric Technol. Hum. Identification IX*, 2012, Art. no. 83711L.

[37] S. Hu, J. Choi, A. L. Chan, and W. R. Schwartz, "Thermal-to-visible face recognition using partial least squares," *JOSA A*, vol. 32, pp. 431–442, 2015.

[38] G. Hermosilla, J. Ruiz-del-Solar, R. Verschae, and M. Correa, "A comparative study of thermal face recognition methods in unconstrained environments," *Pattern Recognit.*, vol. 45, pp. 2445–2459, 2012.

[39] A. Ibrahim, A. Tharwat, T. Gaber, and A. E. Hassanien, "Optimized superpixel and AdaBoost classifier for human thermal face recognition," *Signal Image Video Process.*, vol. 12, pp. 711–719, 2017.

[40] T. Zhang, A. Wiliem, S. Yang, and B. C. Lovell, "TV-GAN: Generative adversarial network based thermal to visible face recognition," arXiv:1712.02514, 2017.

[41] J. Dowdall, I. Pavlidis, and G. Bebis, "Face detection in the near-IR spectrum," *Image Vis. Comput.*, vol. 21, pp. 565–578, 2003.

[42] S. Z. Li, R. Chu, S. Liao, and L. Zhang, “Illumination invariant face recognition using near-infrared images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 627–639, Apr. 2007.
[43] D. Y. Li, and W.-H. Liao, “Facial feature detection in near-infrared images,” in *Proc. 5th Int. Conf. Comput. Vis. Pattern Recognit. Image Process.*, Sept. 2003, pp. 26–30.
[44] Z. Pan, G. Healey, M. Prasad, and B. Tromberg, “Face recognition in hyperspectral images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1552–1560, Dec. 2003.
[45] R. He, X. Wu, Z. Sun, and T. Tan, “Wasserstein CNN: Learning invariant features for NIR-VIS face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, doi: 10.1109/TPAMI.2018.2842770.
[46] S. Liao, D. Yi, Z. Lei, R. Qin, and S. Z. Li, “Heterogeneous face recognition from local structures of normalized appearance,” in *Proc. Int. Conf. Biometrics*, 2009, pp. 209–218.
[47] B. Klare and A. K. Jain, “Heterogeneous face recognition: Matching nir to visible light images,” in *Proc. 20th Proc. Int. Conf. Pattern Recognit.*, 2010, pp. 1513–1516.
[48] J. Li, Y. Jin, and Q. Ruan, “Matching NIR face to VIS face using multi-feature based MSDA,” in *Proc. 12th Proc. Int. Conf. Signal Process.*, 2014, pp. 1443–1447.
[49] D. Yi, R. Liu, R. Chu, Z. Lei, and S. Z. Li, “Face matching between near infrared and visible light images,” in *Proc. Int. Conf. Biometrics*, 2007, pp. 523–530.
[50] J. Chen, D. Yi, J. Yang, G. Zhao, S. Z. Li, and M. Pietikainen, “Learning mappings for face synthesis from near infrared to visual light images,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 156–163.
[51] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J. Bone, “Face recognition vendor test (FRVT) 2002: Overview and summary,” [Online]. Available: http://www.frvt.org/FRVT2002/documents.htm, 2003.
[52] A. M. Ali, “A 3D-based pose invariant face recognition at a distance framework,” *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2158–2169, Dec. 2014.
[53] M. Emambakhsh and A. Evans, “Nasal patches and curves for expression-robust 3D face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, vol. 5, pp. 995–1007, May. 2017.
[54] K. W. 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, pp. 1–15, 2006.
[55] H. Mohammadzade and D. Hatzinakos, “Iterative closest normal point for 3D face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 2, pp. 381–397, Feb. 2013.
[56] Y. Lei, M. Bennamoun, M. Hayat, and Y. Guo, “An efficient 3D face recognition approach using local geometrical signatures,” *Pattern Recognit.*, vol. 47, pp. 509–524, 2014.
[57] G. Medioni and R. Waupotitsch, “Face modeling and recognition in 3-D,” in *Proc. IEEE Int. Workshop Anal. Model. Faces Gestures*, pp. 232–233.
[58] S. Lao, Y. Sumi, M. Kawade, and F. Tomita, “3D template matching for pose invariant face recognition using 3D facial model built with isoluminance line based stereo vision,” in *Proc. 15th Int. Conf. Pattern Recognit.*, 2000, pp. 911–916.
[59] K. Chang, K. Bowyer, and P. Flynn, “Face recognition using 2D and 3D facial data,” in *Proc. ACM Workshop Multimodal User Authentication*, 2003, pp. 25–32.
[60] P. Hegu, D. Kale, and P. Suryawanshi, “Heterogeneous face recognition,” in *Proc. Nat. Conf. Eng. Trends Med. Sci.*, 2014, pp. 15–21.
[61] J. Zhong, X. Gao, and C. Tian, “Face sketch synthesis using E-HMM and selective ensemble,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2007, pp. I-485–I-488.
[62] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, “A nonlinear approach for face sketch synthesis and recognition,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 1005–1010.
[63] X. Tang and X. Wang, “Face sketch synthesis and recognition,” in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, 2003, pp. 687–694.
[64] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, “Transductive face sketch-photo synthesis,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 9, pp. 1364–1376, Sep. 2013.
[65] X. Wang and X. Tang, “Face photo-sketch synthesis and recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 1955–1967, Nov. 2009.
[66] Z. Lei and S. Z. Li, “Coupled spectral regression for matching heterogeneous faces,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 1123–1128.
[67] D. Lin and X. Tang, “Inter-modality face recognition,” in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 13–26.
[68] A. Mignon and F. Jurie, “CMML: A new metric learning approach for cross modal matching,” in *Proc. Asian Conf. Comput. Vis.*, 2012, Art. no. 14.
[69] A. Sharma and D. W. Jacobs, “Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 593–600.
[70] H. K. Galoogahi and T. Sim, “Face sketch recognition by local radon binary pattern: Lrbp,” in *Proc. 19th IEEE Int. Conf. Image Process.*, 2012, pp. 1837–1840.
[71] C. Peng, X. Gao, N. Wang, and J. Li, “Graphical representation for heterogeneous face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 2, pp. 301–312, Feb. 2017.
[72] B. F. Klare, Heterogeneous face recognition: PhD dissertation, Dept. Comput. Sci. Eng., Michigan State Univ., East Lansing, Michigan, 2012.
[73] Y. Jin, J. Lu, and Q. Ruan, “Coupled discriminative feature learning for heterogeneous face recognition,” *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 640–652, Mar. 2015.
[74] C. A. Corneanu, M. O. Simón, J. F. Cohn, and S. E. Guerrero, “Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 8, pp. 1548–1568, Aug. 2016.
[75] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, “The FERET database and evaluation procedure for face-recognition algorithms,” *Image Vis. Comput.*, vol. 16, pp. 295–306, 1998.
[76] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, “The FERET evaluation methodology for face-recognition algorithms,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
[77] D. M. Blackburn, M. Bone, and P. J. Phillips, “Face recognition vendor test 2000: Evaluation report,” DoD Counterdrug Technology Development Program Office, Defense Advanced Research Projects Agency Arlington VA, 2001.
[78] P. J. Phillips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone, “Face recognition vendor test 2002,” in *Proc. IEEE Int. Workshop Anal. Model. Faces Gestures*, 2003, Art. no. 44.
[79] T. Sim, S. Baker, and M. Bsat, “The CMU pose, illumination, and expression (PIE) database,” in *Proc. 5th IEEE Int. Conf. Autom. Face Gesture Recognit.*, 2002, pp. 53–58.
[80] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, “Multi-pie,” *Image Vis. Comput.*, vol. 28, pp. 807–813, 2010.
[81] A. M. Martinez and R. Benavente, “The AR face database,” CVC Technical Report 24, Autonomous University of Barcelona, Barcelona, Catalonia, Spain, Jun. 1998.
[82] M. A. Turk and A. P. Pentland, “Face recognition using eigenfaces,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 1991, pp. 586–591.
[83] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces versus fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[84] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, “From few to many: Illumination cone models for face recognition under variable lighting and pose,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
[85] K.-C. Lee, J. Ho, and D. J. Kriegman, “Acquiring linear subspaces for face recognition under variable lighting,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005.
[86] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, “XM2VTSDB: The extended M2VTS database,” in *Proc. 2nd Int. Conf. Audio Video-Based Biometric Person Authentication*, 1999, pp. 965–966.
[87] A. Nefian and M. H. Hayes III, “A hidden Markov model-based approach for face detection and recognition,” School of Electrical and Computer Engineering, Georgia Institute of Technology, 1999.
[88] A. Angelova, Y. Abu-Mostafa, and P. Perona, “Pruning training sets for learning of object categories,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 494–501.
[89] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, et al., “The CAS-PEAL large-scale Chinese face database and baseline evaluations,” *IEEE Trans. Syst. Man Cybern.-Part A: Syst. Hum.*, vol. 38, no. 1, pp. 149–161, Jan. 2008.
[90] M. J. Lyons, J. Budynek, and S. Akamatsu, “Automatic classification of single facial images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 12, pp. 1357–1362, Dec. 1999.

[91] V. Jain, "The Indian face database," 2002, [Online]. Available: http://vis-www.cs.umass.edu/vidit/IndianFaceDatabase/.

[92] A. Bastanfard, M. A. Nik, and M. M. Dehshibi, "Iranian face database with age, pose and expression," in *Proc. Int. Conf. Mach. Vis.*, 2007, pp. 50–55.

[93] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Technical Report 07-49, 2007.

[94] L. Wolf, T. Hassner, and I. Maoz, "Face recognition in unconstrained videos with matched background similarity," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 529–534.

[95] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore, and S. S. Nooreyezdan, "Plastic surgery: A new dimension to face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 441–448, Sep. 2010.

[96] S. Bianco, "Large age-gap face verification by feature injection in deep networks," *Pattern Recognit. Letters*, vol. 90, pp. 36–42, 2017.

[97] P. Mittal, A. Jain, G. Goswami, R. Singh, and M. Vatsa, "Recognizing composite sketches with digital face images via SSD dictionary," in *Proc. IEEE Int. Joint Conf. Biometrics*, 2014, pp. 1–6.

[98] P. Mittal, M. Vatsa, and R. Singh, "Composite sketch recognition via deep network-a transfer learning approach," in *Proc. Int. Conf. Biometrics*, 2015, pp. 251–256.

[99] C. Galea and R. A. Farrugia, "Forensic face photo-sketch recognition using a deep learning-based architecture," *IEEE Signal Process. Lett.*, vol. 24, no. 24, pp. 1586–1590, Nov. 2017.

[100] C. Galea and R. A. Farrugia, "Matching software-generated sketches to face photographs with a very deep CNN, morphed faces, and transfer learning," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1421–1431, Jun. 2018.

[101] C. Galea and R. A. Farrugia, "A large-scale software-generated face composite sketch database," in *Proc. Int. Conf. Biometrics Special Interest Group*, 2016, pp. 1–5.

[102] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, et al., "A natural visible and infrared facial expression database for expression recognition and emotion inference," *IEEE Trans. Multimedia*, vol. 12, no. 7, pp. 682–691, Nov. 2010.

[103] Equinox Corporation, "NIST/Equinox database," (2012, Apr.). [Online]. Available: www.equinoxsensors.com/products/HID.html.

[104] University of Tennessee, "IRIS thermal/visible face database," (2012, Jan.). [Online] http://www.cse.ohio-state.edu/otcbvs-bench/

[105] W. A. Bainbridge, P. Isola, and A. Oliva, "The intrinsic memorability of face photographs," *J. Exp. Psychology: Gen.*, vol. 142, 2013, Art. no. 1323.

[106] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in *Proc. Int. Conf. Automatic Face Gesture Recognit.*, Jun. 2000, pp. 46–53.

[107] S. Setty, M. Husain, P. Beham, J. Gudavalli, M. Kandasamy, R. Vaddi, et al., "Indian movie face database: a benchmark for face recognition under wide variations," in *Proc. 4th Nat. Conf. Comput. Vis. Pattern Recognit. Image Process. Graph.*, 2013, pp. 1–5.

[108] D. Lundqvist, A. Flykt, and A. Öhman, The Karolinska directed emotional faces (KDEF), ISBN 91-630-7164-9, Karolinska Institutet, 1998.

[109] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua, "Labeled faces in the wild: A survey," in *Proc. Int. Conf Face Detection Facial Image Anal.*, 2016, pp. 189–248.

[110] R. Gross, "Face databases," in *Handbook of Face Recognition*, Berlin, Germany: Springer, 2005, pp. 301–327.

[111] A. Kasinski, A. Florek, and A. Schmidt, "The PUT face database," *Image Process. Commun.*, vol. 13, pp. 59–64, 2008.

[112] R. Miezianko, "Terravic Research Infrared Database," (2006, Jun.). [Online]. Available: http://www.cse.ohiostate.edu/otcbvs-bench/

[113] B. Zhang, L. Zhang, D. Zhang, and L. Shen, "Directional binary code with application to PolyU near-infrared face database," *Pattern Recognit. Lett.*, vol. 31, pp. 2337–2344, 2010.

[114] S. Z. Li, D. Yi, Z. Lei, and S. Liao, "The casia nir-vis 2.0 face database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2013, pp. 348–353.

[115] H. Maeng, S. Liao, D. Kang, S.-W. Lee, and A. K. Jain, "Nighttime face recognition at long distance: Cross-distance and cross-spectral matching," in *Proc. Asian Conf. Comput. Vis.*, 2012, pp. 708–721.

[116] N. Erdogmus and S. Marcel, "Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect," in *Proc. IEEE 6th Int. Conf. Biometrics: Theory Appl. Syst.*, 2013, pp. 1–6.

[117] Royal Military Academy of Belgium, "3D_RMA database," (1998, Jan.). [Online]. Available: http://www.sic.rma.ac.be/~beumier/DB/3d_rma.html

[118] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter, "A 3D face model for pose and illumination invariant face recognition," in *Proc. 6th IEEE Int. Conf. Adv. Video Signal Based Surveillance*, 2009, pp. 296–301.

[119] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3D facial expression database for facial behavior research," in *Proc. 7th Int. Conf. Autom. Face Gesture Recognit.*, 2006, pp. 211–216.

[120] A. Savran and B. Sankur, "Non-rigid registration based model-free 3D facial expression recognition," *Comput. Vis. Image Understanding*, vol. 162, pp. 146–165, 2017.

[121] 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*, 2010, pp. 97–100.

[122] R. Min, N. Kose, and J.-L. Dugelay, "Kinectfacedb: A kinect database for face recognition," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 44, no. 11, pp. 1534–1548, Nov. 2014.

[123] A. Colombo, C. Cusano, and R. Schettini, "UMB-DB: A database of partially occluded 3D faces," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2011, pp. 2113–2119.

[124] V. Espinosa-Duró, M. Faundez-Zanuy, and J. Mekyska, "A new face database simultaneously acquired in visible, near-infrared and thermal spectrums," *Cognitive Comput.*, vol. 5, pp. 119–135, 2013.

[125] V. Espinosa-Duró, M. Faundez-Zanuy, J. Mekyska, and E. Monte-Moreno, "A criterion for analysis of different sensor combinations with an application to face biometrics," *Cognitive Comput.*, vol. 2, pp. 135–141, 2010.

[126] H. Nguyen, K. Kotani, F. Chen, and B. Le, "A thermal facial emotion database and its analysis," in *Proc. Pacific-Rim Symp. Image Video Technol.*, 2013, pp. 397–408.

[127] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognit. Lett.*, vol. 28, pp. 1885–1906, 2007.

[128] A. B. Moreno and A. Sánchez, "GavabDB: A 3D face database," in *Proc. 2nd COST275 Workshop Biometrics Internet*, 2004, pp. 75–80.

[129] University of Zagreb, Face recognition homepage, [Online]. Available: http://www.face-rec.org/general-info/

[130] L. Yin, X. Chen, Y. Sun, T. Worm, and M. Reale, "A high-resolution 3D dynamic facial expression database," in *Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recognit.*, 2008, pp. 1–6.

[131] X. Zhang, L. Yin, J. F. Cohn, S. Canavan, M. Reale, A. Horowitz, et al., "Bp4d-spontaneous: A high-resolution spontaneous 3d dynamic facial expression database," *Image Vis. Comput.*, vol. 32, pp. 692–706, 2014.

[132] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, et al., "Bosphorus database for 3D face analysis," in *Proc. Eur. Workshop Biometrics Identity Manage.*, 2008, pp. 47–56.

Karen Panetta (S'84–M'85–SM'95–F'08) received the BS degree in computer engineering from Boston University, Boston, MA, and the MS and PhD degrees in electrical engineering from Northeastern University, Boston. She is currently a dean of Graduate Engineering Education and a professor in the Department of Electrical and Computer Engineering and Adjunct professor of Computer Science at Tufts University, Medford, MA, and director of the Dr Panetta's Vision and Sensing System Laboratory. She is the president-Elect 2019 of IEEE-HKN. She is the editor-in-chief of the IEEE Women in Engineering magazine. She was the IEEE-USA vice-president of Communications and Public Affairs. During 2007 to 2009, she served as the world-wide director for IEEE Women in Engineering, overseeing the world's largest professional organization supporting women in engineering and science. Her research focuses on developing efficient algorithms for simulation, modeling, signal, and image processing for biomedical and security applications. She is also the recipient of the 2012 IEEE Ethical Practices Award and the Harriet B. Rigas Award for Outstanding Educator. In 2011, she was awarded the Presidential Award for Engineering and Science Education and Mentoring by U.S. President Obama. She is a fellow of the IEEE.

Qianwen Wan received the BS degree in information engineering from the Wuhan University of Technology, Wuhan, China, the BS degree in law from Wuhan University, Wuhan, China, and the MS degree in electrical and computer engineering from Tufts University, Medford, MA, USA. She is currently working toward the PhD degree in electrical and computer engineering at Tufts University, Medford, MA, USA and she is a principle research assistant in the Dr. Panetta's Vision and Sensing lab. Her current research interests include designing automated tools for cognitive scientists, analyzing mobile eye tracking data, facial recognition, image and video processing, and machine-learning applications. She is a Student Member of the IEEE.

Sos Agaian (M'98–SM'00) is currently a distinguished professor at the City University of New York/CSI. His research interests include computational vision and machine learning, large scale data analytic analytics, multi-modal data fusion, biologically-inspired signal/image processing modeling, multimodal biometric and digital forensics, 3D imaging sensors, information processing and security, and biomedical and health informatics. He has authored more than 650 technical articles and ten books in these areas. He is also listed as a co-inventor on 44 patents/disclosures. The technologies that he invented have been adopted by multiple institutions, including the U.S. government, and commercialized by industry. He is a fellow of the SPIE, a fellow of the IEEE, fellow of the IS&T, and a fellow of the AAAS. He also serves as a foreign member of the Armenian National Academy. He received MAEStro Educator of the Year, sponsored by the Society of Mexican American Engineers. He received the Distinguished Research Award at the University of Texas at San Antonio. He is recipient of the Innovator of the Year Award (2014), of the Tech Flash Titans-Top Researcher-Award (San Antonio Business Journal, 2014), the Entrepreneurship Award (UTSA-2013 and 2016), and the Excellence in Teaching Award (2015). He is an Editorial Board member for the *Journal of Pattern Recognition and Image Analysis* and an associate editor for several journals, including the *IEEE Transactions on Image Processing*, the *IEEE Transactions on Systems, Man and Cybernetics*, *Journal of Electrical and Computer Engineering* (Hindawi Publishing Corporation), *International Journal of Digital Multimedia Broadcasting* (Hindawi Publishing Corporation) and *Journal of Electronic Imaging (SPIE, IS&T)*.

Srijith Rajeev received the BE degree in electronics and communication engineering from the Visvesvaraya Technological University, India, in 2014, and the MS degree in electrical and computer engineering from the University of Texas, San Antonio, USA, in 2016. He is currently working toward the PhD degree in electrical and computer engineering at Tufts University, USA. His current research interests include signal/image processing, video processing, machine-learning, 3D sensors and modelling, digital forensic, and biomedical applications. He is a Student Member of the IEEE.

Shreyas Kamath received the bachelor's degree (B.E.) in electronics and communication engineering from the Visvesvaraya Technological University, Belgaum, India, in 2014, the MS degree in electronic and computer engineering at the University of Texas, San Antonio, USA. He is currently working toward the PhD degree in electrical and computer engineering in Tufts University, USA. He is working as a graduate research assistant in the Visual and Sensing Lab, Tufts. His main areas of research interests include signal / image processing, 3D scanning, machine learning, automated biometric technologies particularly focusing on fingerprints and their applications. He is a Student Member of the IEEE.

Rahul Rajendran received the bachelor's degree in electrical engineering from Reva Institute of Technology and Management, Bangalore, India, in 2014, the masters' degree in electrical and computer engineering at the University of Texas, San Antonio, in 2016. He is currently working toward the PhD degree in electrical and computer science at Tufts University. His current research interests are image and video analytics, signal processing, 3D sensors, medical imaging, and Security. He is a Student Member of the IEEE.

Shishir Paramathma Rao received the BE degree in electronics and communication from the Visvesvaraya Technological University, India, in 2014, and the MS degree in electrical and computer engineering from the University of Texas, San Antonio, USA, in 2016. Currently, he is working toward the PhD degree in electrical and computer engineering at Tufts University, USA. His research interests include 3D photography, image-based modeling, multiview stereovision, image and video analytics, machine-learning and neural networks, signal/image processing, 3D sensors. He is a Student Member of the IEEE.

Aleksandra Kaszowska received the BA degree in psychology from Clark University in Worcester, MA, and the MS degree in experimental psychology from Tufts University, Medford, MA. She is currently working toward the PhD degree in Cognitive Science at Tufts University, Medford, MA. Her current research interests include image and video processing, eye tracking, problem solving, and spatial cognition.

Holly A. Taylor received the BA degree in mathematics from Dartmouth College, Hanover, NH, and the PhD degree in cognitive psychology from Stanford University, Stanford, CA. She is currently a full professor of psychology, and adjunct professor of mechanical engineering, and co-director of the Center for Applied Brain and Cognitive Science at Tufts University.

Arash Samani (S'98, M'00, SM'15) received the MS degree in electrical engineering from the University of Alabama in Huntsville (UAH). He is currently working toward the PhD degree in electrical engineering at Tufts University under the supervision of Prof. Karen Panetta. He is currently works at MITRE Corporation. He has worked in the fields of digital signal processing, image processing, digital communications, radar, and embedded systems. He is a senior member of the IEEE.

Xin Yuan received the bachelor's degree in electrical and information engineering from the Wuhan University of Science and Technology, China, in 2016, and the master's degree in electrical engineering from Tufts University, USA, in 2018. Her current research interests include image processing, facial recognition, and computer vision.