Mask Does Not Matter: Anti-Spoofing Face Authentication using mmWave without On-site Registration

Weiye Xu[1,2], Wenfan Song[1,2], Jianwei Liu[1,2], Yajie Liu[1,2], Xin Cui[4], Yuanqing Zheng[3]
Jinsong Han[1,2*], Xinhuai Wang[4], Kui Ren[1,2]
[1]Zhejiang University, Hangzhou, China
[2]ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou, China
[3]The Hong Kong Polytechnic University, HongKong, China
[4]Xidian University, Xi'an, China
{xuweiye,wenfansong,jianweiliu,yajie,hanjinsong,kuiren}@zju.edu.cn
xcui@stu.xidian.edu.cn,yqzheng@polyu.edu.hk,xinhuaiwang@xidian.edu.cn

ABSTRACT

Face authentication (FA) schemes are universally adopted. However, current FA systems are mainly camera-based and hence susceptible to face occlusion (e.g., facial masks) and vulnerable to spoofing attacks (e.g., 3D-printed masks). This paper exploits the penetrability, material sensitivity, and fine-grained sensing capability of millimeter wave (mmWave) to build an anti-spoofing FA system, named **mmFace**. It scans the human face by moving a commodity off-the-shelf (COTS) mmWave radar along a specific trajectory. The mmWave signals bounced off the human face carry the facial biometric features and structure features, which allows mmFace to achieve reliable liveness detection and FA. Due to the penetrability of mmWave, mmFace can still work well even if users wear masks. We explore a distance-resistant facial structure feature to suppress the impact of unstable face-to-device distance. To avoid inconvenient on-site registration, we also propose a novel virtual registration approach based on the core idea of cross-modal transformation from photos to mmWave signals. We implement mmFace with various antenna configurations and prototype two typical modes of mmFace. Extensive experiments show that mmFace can realize accurate FA as well as reliable liveness detection.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing**; • **Security and privacy** → **Security services**.

KEYWORDS

Millimeter Wave, Face Authentication, Biometrics

ACM Reference Format:
Weiye Xu[1,2], Wenfan Song[1,2], Jianwei Liu[1,2], Yajie Liu[1,2], Xin Cui[4], Yuanqing Zheng[3], Jinsong Han[1,2*], Xinhuai Wang[4], Kui Ren[1,2]. 2022. Mask Does Not Matter: Anti-Spoofing Face Authentication using mmWave without On-site Registration. In *The 28th Annual International Conference On Mobile Computing And Networking (ACM MobiCom '22), October 17–21, 2022, Sydney, NSW, Australia.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3495243.3560515

*J.Han is the corresponding author. W.Xu and W.Song contribute equally to this paper.

ACM MobiCom '22, October 17–21, 2022, Sydney, NSW, Australia

ACM ISBN 978-1-4503-9181-8/22/10…$15.00
https://doi.org/10.1145/3495243.3560515

1 INTRODUCTION

Face authentication (FA) has become essential to many security-critical applications, such as access control [66], mobile payment [79], and individual identification [14]. It is pervasively deployed thanks to its high accuracy and user-friendliness. Albeit these advantages, existing FA systems are commonly camera-based and hence suffer several limitations. First, camera-based FA systems are susceptible to poor light conditions, which limits their usability in practice [19]. Second, with the development of impersonation technologies such as Deepfake [45], camera-based approaches become vulnerable to emerging spoofing attacks [17, 40]. Although some latest camera-based face authentication systems (e.g., Apple Face ID [13]) take defensive measures (e.g., adopting a TrueDepth camera to defend against 2D images), they can still be fooled by 3D spoofing attacks (e.g., 3D-printed masks) as reported in [40, 49, 54]. Besides, the authentication accuracy of camera-based methods is substantially affected by face occlusion [41]. Especially, the COVID-19 pandemic makes wearing surgical masks critical to prevent the spread of contagious diseases [20]. Under this circumstance, camera-based FA systems cannot handle masks well since they can only capture partial facial features exposed to the camera [4]. For example, Face ID can only capture and recognize the features around the eyes rather than the entire face covered by masks [12].

One promising way to overcome these limitations is Radio Frequency (RF) based sensing [34, 57, 63, 65, 69]. Recent advances have witnessed the rapid development of RF-based identification methods by extracting biometric features (e.g., gait features [33, 43], cardiac features [31, 36], finger impedance features [76], respiratory features [32, 37]) based on real RF signals as well as vision-based synthesized signals [33]. RF-based sensing approaches offer appealing advantages. First, they are resilient to complex lighting conditions. Meanwhile, RF signals have brought new possibilities for mask-friendly user authentication. Unlike cameras that are susceptible to face occlusion, the penetrability of RF signals allows them to pass through masks [47] and retrieve facial features behind them. Moreover, RF signals are sensitive to the materials they

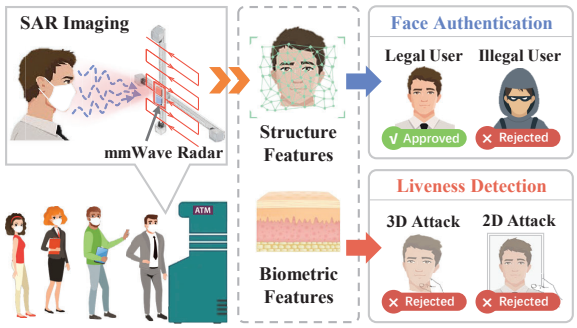


Figure 1: An anti-spoofing FA based on mmWave, which still works well under the occlusion of masks.

encounter during propagation and reflection [53]. Hence, it is possible to utilize such material sensitivity to distinguish real human faces from other materials (e.g., silicone in 3D-printed masks) and thereby defend against spoofing attacks.

Motivated by the penetrability and material-sensitivity of RF signals, we build an RF-based FA system that is friendly to mask-wearing users and enhance its security against spoofing attacks with liveness detection. However, designing a practical RF-based FA system is challenging. **(1) How to mitigate the registration overhead.** For registration, RF-based authentication systems need to collect each user's unique characteristics (e.g., facial information) to form the credential. Existing RF-based approaches usually require on-site registration [36, 56, 68] with designated RF devices at specific locations. Such on-site registration typically takes a long time [66, 76] (e.g. several minutes) to collect a large amount of data to build a representative credential. Thus, the large overhead of on-site registration prohibits the wide deployment of RF-based authentication systems. **(2) How to implement robust FA under variable face-to-device distances.** Due to the sensitivity of RF signals, the signals reflected from the same face would be quite different when the distance between the user and the sensing device varies [26, 79]. In this case, if we directly utilize the raw mmWave signal for authentication, the FA accuracy would severely degrade when the face-to-device distance during authentication is inconsistent with the one during registration.Thus, some RF-based authentication methods [10, 79] have rigorous constraints on the authentication distance, which greatly limits their applicability. **(3) How to realize reliable liveness detection for the human face with complex geometric structure.** The key to reliable liveness detection is to extract biometric features, e.g., the skin tissue, which cannot be easily imitated. Nevertheless, when RF signals impinge on a human face, they are affected not only by the material characteristic but also by the facial structure [57], i.e., its appearance. Since the structure of the human face is complex, e.g., inconsistent surface curvature, it is challenging to extract precise biometric features for liveness detection. Existing RF-based material sensing methods [15, 25, 62, 72] typically require the shape of the target object to be fixed, which is not practical in FA scenarios.

To solve these challenges and build a practical RF-based FA system, we first determine *which kind of RF signal is proper to conduct FA*. Compared with alternatives (e.g. WiFi and RFID), we choose mmWave since it is highly directional [80] and has good reflective properties [60]. However, commercial mmWave radars do not provide sufficient sensing resolution for FA [7], due to their small aperture. Although adopting specialized radars with large lenses can effectively improve the resolution [80], those devices are extremely expensive and inconvenient for wide deployment. Instead, inspired by synthetic aperture radar (SAR) [44], we explore the possibility of enhancing the sensing resolution by moving a low-cost mmWave radar to enlarge the aperture.

In this paper, we develop an anti-spoofing FA system (named **mmFace**) using a COTS mmWave radar. mmFace can work well when users wear masks, as illustrated in Fig. 1. Specifically, by moving a mmWave radar along a specific trajectory, mmFace can emulate a large aperture planar antenna array and measure the mmWave signals reflected from the human face. It extracts biometric features from the collected signals to conduct liveness detection. For the goal of FA, mmFace derives underlying facial structure features from the raw mmWave signals by reconstructing the human face with the SAR imaging algorithm [71]. Moreover, mmFace addresses the aforementioned three challenges.

Firstly, we propose a virtual registration approach based on the cross-modal transformation from photos to mmWave signals. Secondly, to extract a distance-resistant facial structure feature, we analyze the intensity values of facial images generated by the SAR imaging algorithm. We find that even though the intensity values vary with distance, the contour of the facial image's 'bright' area (the area with high-intensity values) is relatively stable. Because the contour of the bright area only depends on the user's facial surface curvature. We leverage Fourier descriptors to quantify the bright area's contour features, which are used as distance-resistant facial structure features. Finally, we suppress the impact of face structure on biometric feature extraction by selecting relatively flat regions on the face. For these selected facial regions, we extract the reflection coefficient of the face for reliable liveness detection.

We implement mmFace using mmWave radars with various configurations, including a COTS mmWave radar (with two transmitting antennas and four receiving antennas) and an advanced mmWave radar (with twelve transmitting antennas and sixteen receiving antennas). We conduct extensive experiments to evaluate mmFace's performance. The results show that mmFace can achieve over 95.9 % authentication success rate and around 4.5 % equal error rate. The experiment results demonstrate that mmFace is an accurate and robust FA system that can be applied to access control and user identification. More importantly, the results show that mmFace is resistant to spoofing attacks including 3D-printed mask attacks. As such, we envision that mmFace is promising to complement and enhance existing FA systems to better support users wearing masks.

In sum, our contributions are as follows: ❶ We propose mmFace, a practical mmWave-based FA system that can still work well under the occlusion of face masks. mmFace extracts distance-resistant facial structure features to achieve robust FA. Besides, we propose a biometric feature-based reliable liveness detection method to protect mmFace from various spoofing attacks. ❷ We design a new virtual registration approach that allows users to only provide three facial photos for registration. By building the mmWave signal propagation model, mmFace can transform the given facial photos into virtual registration signals, thus avoiding burdensome on-site

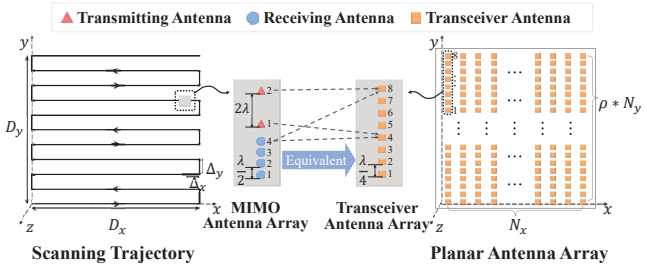


Figure 2: The planar antenna array generation.

registration. ❸ We prototype two typical modes of mmFace in terms of the radar's antenna configuration. Extensive experiments show that mmFace can realize precise and robust authentication as well as defend against spoofing attacks.

2 BACKGROUND

2.1 Planar Antenna Array

Inspired by SAR, we move a mmWave radar along a trajectory by a 2D slide rail to form a planar antenna array. In this way, the aperture size of the radar is enlarged and we can realize a millimeter-level fine-grained resolution for supporting FA. In the following, we take the COTS mmWave radar with multiple-input multiple-output (i.e., MIMO) antenna arrays as an example to demonstrate this process.

As shown in Fig. 2, under the control of the 2D slide rail, the mmWave radar first moves uniformly along the horizontal direction (X-axis) for a distance D_x and obtains N_x measurements uniformly distributed at an interval of Δx. Then, the mmWave radar moves along the vertical direction (Y-axis) for a distance Δy. These two steps are repeated N_y times in total. Besides, to make the planar antenna array's antennas uniformly distributed, the antennas on the MIMO antenna array are equivalent to ρ (i.e., eight) uniformly distributed transceiver antennas at an interval of $\frac{\lambda}{4}$ (i.e., $\frac{\Delta y}{8}$), as shown in Fig. 2. Each pair of transmitting antenna and receiving antenna are equivalent to one transceiver antenna located in the middle of them, which can both transmit and receive mmWave signals [70]. After that, we can form a planar antenna array containing $\rho * N_y$ rows with a size of $D_x * D_y$ and each row has N_x transceiver antennas, as shown in Fig. 2. In addition, we also implement a time-efficient advanced mmWave radar to achieve nearly real-time FA, which will be introduced in Sec. 7.

2.2 Attack Model

We consider two mainstream spoofing attacks (i.e., 2D attack and 3D attack) in our attack models.

■ **2D attack.** This type of attack can be further divided into static and dynamic attacks. In the former, the attacker can obtain a photo containing a legitimate user's face to spoof the FA system [51]. This attack method is effective in attacking those systems which only verify the 2D features of human faces. Recently, some FA systems are enhanced by liveness detection, where users are required to blink their eyes and move their lips [50]. However, attackers can still launch 2D attacks with videos of legitimate users [49].

■ **3D attack.** To mitigate the threats of 2D attacks, some FA systems collect the 3D structural features of human faces [58] as an additional security factor. Unfortunately, the adversary could still spoof the real users by manipulating their 3D structural features. For instance, an attacker can wear a 3D-printed mask to mimic a legitimate user's face to bypass the system's verification [54].

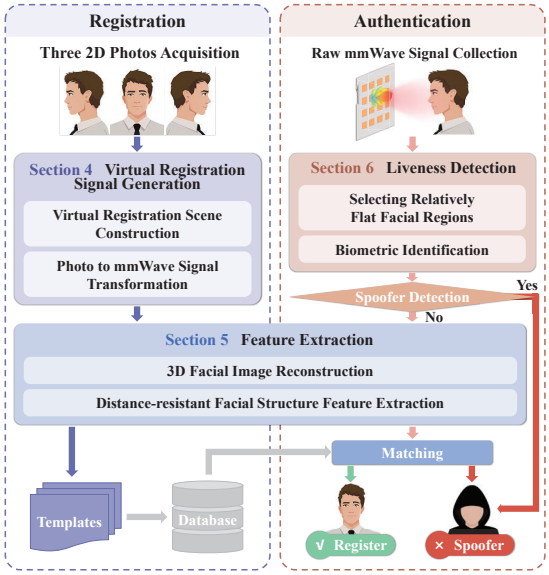


Figure 3: System architecture of mmFace.

3 SYSTEM OVERVIEW

As shown in Fig. 3, mmFace consists of two primary phases: the registration phase and the authentication phase.

In the registration phase, to avoid on-site registration, we propose a VRS generation method. Instead of directly collecting mmWave signals reflected off users' faces, mmFace only needs three 2D facial photos of each user taken from different perspectives for registration. These photos are utilized to generate VRSes. We first construct a virtual registration scene that is consistent with the authentication scene. Then, based on the virtual registration scene, we transform the photos into mmWave signals by building a theoretical model of mmWave signal propagation. Thereafter, these generated VRSes will be sent to the feature extraction module (consistent with the one used in the authentication phase and will be detailed in that phase) to obtain the facial structure features. The features will be stored in a database as templates.

In the authentication phase, users only need to place their faces in front of a mmWave radar for a few seconds. In this process, mmFace collects the mmWave signals reflected from the face by moving the radar along a specific trajectory as described in Sec. 2.1. From the collected mmWave signals, mmFace extracts biometric features to conduct liveness detection. Once the user is detected 'alive', mmFace will reconstruct the user's 3D facial image from the mmWave signals based on the SAR imaging algorithm. Then, to suppress the impact of distance variation, mmFace extracts the distance-resistant facial structure feature based on the reconstructed facial image. Next, mmFace tries to find a match in the database by calculating the distance between the extracted facial structure feature and each template stored in the database. If the smallest distance is lower than a pre-defined threshold, the user corresponding to the template with this smallest distance is regarded as the one who

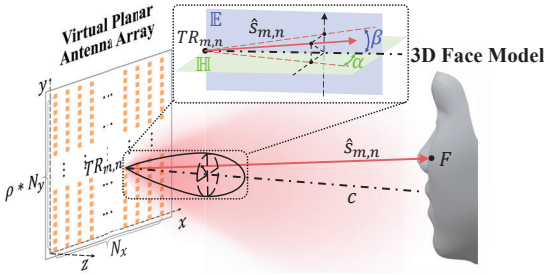


Figure 4: Virtual registration scene construction and the antenna directionality modeling.

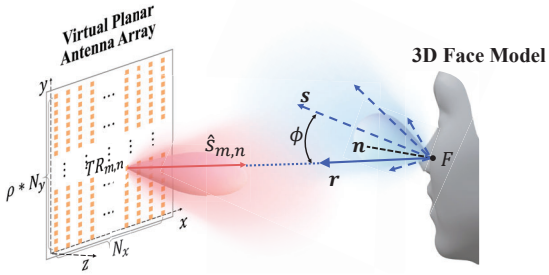


Figure 5: Modeling the reflection of Face.

initiates the authentication request. Otherwise, mmFace regards this authentication request as an illegal attempt and rejects it.

4 VIRTUAL REGISTRATION SIGNAL GENERATION

In this section, we develop a new photo-to-mmWave cross-modal registration approach to generate VRSes.

4.1 Methodology

Current RF-based authentication methods usually require users to complete time-consuming on-site registration. Inspired by [33] , we develop a cross-modal (i.e., from photos to mmWave) registration approach to reduce the registration time and effort. The prior work [33] synthesizes WiFi signals according to 3D body models reconstructed from videos. In contrast, our work aims to synthesize mmWave signals from reconstructed 3D models. Different from the prior work, our approach faces the following unique challenges: 1) The arrangement of mmWave radar antennas is more complicated. In [33], only a few WiFi transceiver antennas are placed in fixed positions to collect WiFi signals reflected by the human body. However, to perform SAR, mmFace needs to either move the mmWave radar along a trajectory or collect the signals via a planar antenna array that has hundreds of antennas. 2) The mmWave radar antennas are directional. In contrast to omnidirectional antennas in [33], the initial amplitudes of the mmWave signals emitted by the transceiver antennas are different in each direction. Therefore, in our approach, the modeling of the directionality of the mmWave antenna is more complicated. To address these two challenges and then transform facial photos into VRSes, we construct a virtual registration scene, based on which we build a mmWave signal propagation model to generate the VRSes.

■ **Virtual registration scene construction.** Identical to the authentication scene, the virtual registration scene construction consists of three parts: generating a 3D face model, constructing a virtual planar antenna array, and determining their spatially relative positions. Given a user's three facial photos taken from different perspectives, we can generate the user's precise 3D model based on the method proposed in [30]. Constructing a 3D model of the human face based on 2D photos has been extensively studied in computer vision and performs well in practice [18, 22, 30]. Then, we virtually form a planar antenna array according to the arrangement of a real one, as shown in Fig. 4. In this way, we can represent the virtual transceiver antenna located at row m and column n of the virtual planar antenna array as $TR_{m,n}$ ($m \in [1, \cdots, \rho * N_y]$, $n \in [1, \cdots, N_x]$). Finally, we rotate and align the 3D face model with the virtual planar antenna array to adjust their relative positions (the chin of the 3D face model should be aligned with the center of the bottom edge of the virtual planar antenna array). So far, as shown in Fig. 4, we virtually reconstruct the registration scene which is consistent with the real authentication scene.

■ **mmWave signal propagation model.** Based on the registration scene, we model the propagation process of mmWave signals to generate VRSes. According to the signal propagation in the real scene, we can divide the signal propagation into two stages: transmitting from the radar to the face and reflecting from the face to the radar. In the two stages, the mmWave signals are mainly affected by three factors: *the propagation distance, the directionality of antennas, and the reflection characteristic of the face.* Hence, we can represent the mmWave signal emitted by $TR_{m,n}$ (denoted as $(x_{m,n}, y_{m,n}, 0)$), reflected by one point F (denoted as (x_F, y_F, z_F)) on the 3D face model and finally received by $TR_{m,n}$, as:

$$\check{s}_{m,n}(F,t) = \frac{\epsilon \hat{s}_{m,n}(t-\tau)}{2r_{m,n}} \underbrace{p_{m,n}(F) g_{m,n}(F)}_{o_{m,n}(F)}. \quad (1)$$

ϵ represents the reflection coefficient of human face material that is regarded as a constant in our model [21]. $\frac{\hat{s}_{m,n}(t-\tau)}{2r_{m,n}}$ represents the impact of *propagation distance* on the received mmWave signals, in which $\hat{s}_{m,n}$ is the initial mmWave signal transmitted by $TR_{m,n}$, $\tau = 2r_{m,n}/c$ represents the delay for the round trip propagation and $r_{m,n}$ denotes the Euclidean distance between $TR_{m,n}$ and F. $p_{m,n}(F)$ and $g_{m,n}(F)$ indicate the amplitude attenuation caused by *the directionality of antenna* and *the reflection characteristic of face*, respectively. For the convenience of expression, $p_{m,n}(F)$ and $g_{m,n}(F)$ can be integrated as the face-mmWave response function $o_{m,n}(F)$. Based on the reflection of each point F, we can estimate the reflection of the entire face by superposing the reflection of all points on the 3D face model. Particularly, we sample sufficient points on the 3D face model to make the resolution of (x_F, y_F, z_F) reach the sub-micron level. Then, by mixing the received signal with the transmitted signal, we can get the intermediate frequency signal:

$$s_{m,n}(t) = \iiint \frac{\epsilon o_{m,n}(F)}{2r_{m,n}} e^{-j2\pi(f_0\tau + K\tau t)} dx_F dy_F dz_F, \quad (2)$$

which represents the VRS finally captured by $TR_{m,n}$. Since the virtual planar antenna array contains $N_x * N_y * \rho$ virtual transceiver antennas, we need to repeat the above estimation for each antenna.

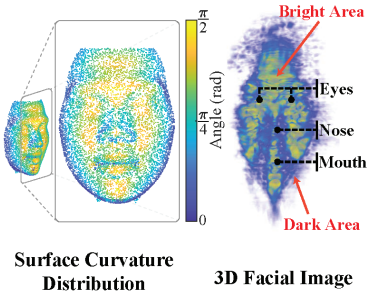


Figure 6: Surface curvature distribution and 3D facial image of the same user.

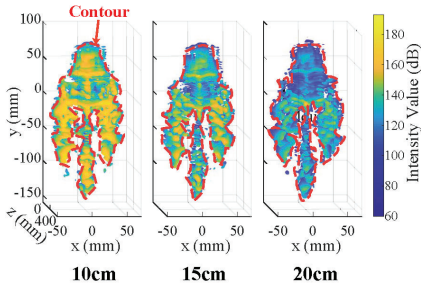


Figure 7: For the same user, the contours of bright area in different authentication distances are extremely identical.

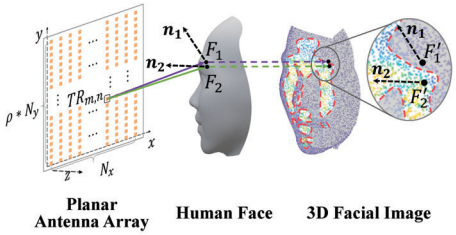


Figure 8: Distance-resistant facial structure feature exploration.

Finally, we can get a virtual mmWave signal matrix:

$$S(x,y,t) = \iiint \frac{\epsilon O(F)}{2R} e^{-j2\pi(f_0\tau + K\tau t)} dx_F dy_F dz_F, \tag{3}$$

where (x, y) represents the coordinate of the virtual transceiver antenna. O is a matrix composed of all virtual transceiver antennas' $o_{m,n}$ with the shape of $N_x * N_y * \rho$. Similar to O, R is composed of $r_{m,n}$ of all virtual transceiver antennas. Based on the above modeling, we find that the key to precise VRS estimation is to accurately characterize the impact of *the directionality of antenna* (p) and *the reflection characteristic of face* (g). Next, we will show how to model these two factors.

4.2 Modeling Directionality of Antenna

The antennas of mmWave radars are highly directional, i.e., their emitted signals are composed of a concentrated main lobe and multiple low-power side lobes [29]. The initial amplitudes of the signals emitted by the same mmWave radar antenna but in different directions would be significantly different. Additionally, since FA is near-field sensing, if we ignore the impact of antenna directionality, the VRSes cannot be accurately estimated.

To deal with the above problem, we involve the directionality of mmWave radar antenna in our signal propagation model. Since the main lobe contains much higher power than the side lobes [48], for ease of modeling, mmFace only takes the main lobe into account and models it as a fundamental Gaussian beam [23]. In this case, as shown in Fig. 4, the initial amplitude of the signal emitted from $TR_{m,n}$ to the point F on the face is determined by the deviation angles between the transmitted signal $\hat{s}_{m,n}$ and the centerline c of the main lobe on two planes (i.e., the horizontal plane (H) and the elevation plane (E)). Then, the impact of antenna directionality on the initial amplitude of the emitted signal can be expressed as:

$$p_{m,n}(F) = Ae^{[-\gamma(\frac{\alpha^2}{\alpha_h^2}+\frac{\beta^2}{\beta_h^2})]}, \tag{4}$$

where $\alpha = arctan(\frac{x_F - x_{m,n}}{z_F})$, $\beta = arctan(\frac{y_F - y_{m,n}}{z_F})$, and A represents the peak value of the main lobe amplitude. γ, α_h, and β_h are three constants determined by the physical characteristics of the antenna [48]. Among them, α_h and β_h are the angles corresponding to the *half-power point* of the main lobe on the horizontal plane and the elevation plane, respectively. They reveal the directionality of the antenna. Specifically, the *half-power point* is the point at which the antenna gain drops to half of its peak value (approximately -3 dB). It indicates that the initial amplitude of emitted signal drops significantly when α is larger than α_h or β is larger than β_h. In this way, we can precisely characterize the initial amplitude of the emitted signal by modeling the directionality of antennas.

4.3 Modeling Reflection of Face

Estimating the impact of face reflection characteristic (g in Eq. 1) requires accurately depicting the reflected signals. Previous works related to analyzing the signal reflection [2, 3] usually abstract the whole object as a single point and focus on the general position of the object, such that they are not appliable to the scenario of capturing fine-grained facial structure information. Although some works [1, 78] analyze the reflection of human body, they treat the human body as a specular reflector since the wavelength of the RF signals (e.g., WiFi) they used is much larger than the surface roughness of the human body. Considering that the wavelength of mmWave is comparable to the roughness of the human face [77], the above modeling methods can not be directly applied to mmFace, so it is challenging to accurately quantify the reflection characteristics of the human face.

We model the human face as a quasi-specular reflector that well matches the actual reflection of the face, according to [33]. Specifically, in this model, the face can reflect the incident signal in various directions with different amplitudes. We show an example in Fig. 5. For the point F, given the direction of the incident mmWave signal ($\boldsymbol{i}$) and the surface normal ($\boldsymbol{n}$) at F, the amplitudes of its reflected signals distribute as a Gaussian function centered on the specular reflection direction ($\boldsymbol{s}$). Among all the reflected signals at the point F, the amplitude of the one along the direction of the specular reflection ($\boldsymbol{s}$) is the strongest. The amplitude of the reflected signal received by the corresponding receiving antenna $TR_{m,n}$ is determined by the angle (ϕ) between the received signal vector ($\boldsymbol{r}$) and the specular reflection vector ($\boldsymbol{s}$). Therefore, we can represent the impact of face reflection characteristic on the amplitude of the reflected signal:

$$g_{m,n}(F) = exp(-\phi^2/2\sigma^2). \tag{5}$$

σ is an empirical constant related to the reflection property of human face, which is relatively stable for different users. It is obvious that the key to solving g is to accurately estimate ϕ. As shown in Fig. 5, ϕ can be calculated by $arccos(\frac{\boldsymbol{sr}}{\|\boldsymbol{r}\|})$. Thus, the problem turns into calculating the specular reflection vector $\boldsymbol{s}$. For the point F on the face, given the incident signal direction $\boldsymbol{i}$, its specular reflection vector can be calculated by $\boldsymbol{s} = \frac{\boldsymbol{i}}{\|\boldsymbol{i}\|} - \frac{2\boldsymbol{ni}}{\|\boldsymbol{i}\|}\boldsymbol{n}$. To estimate the surface

normal ($\boldsymbol{n}$) at F, we take the neighboring points of F to form a plane and regard the normal of the plane as the surface normal of the point. In this way, we can accurately estimate the impact of facial reflection characteristic g on the mmWave signal. Moreover, based on the above analysis, we can find that the face-mmWave response function O can reveal the facial surface curvature which is quantified by $\boldsymbol{n}$. We will utilize O to extract facial structure features in Sec. 5. To verify the validity of generated VSR, we will compare the facial structure features extracted from VRSes with those extracted from real signals in Sec. 5.2.

5 FEATURE EXTRACTION AND MATCHING

It is crucial for mmFace to extract representative facial structure features of each user from either the VRSes or the ones collected during authentication since FA depends on performing matching between these two.

5.1 Facial Structure Feature Extraction

Unlike photos that directly capture facial structure, it is challenging to extract facial structure features from non-semantic RF signals, e.g., mmWave. Our solution to this challenge is to reconstruct the mmWave signals into intuitive 3D facial images by using the SAR imaging algorithm and then extract distance-resistant facial structure features.

5.1.1 Imaging by SAR. With the matrix $s(x, y, t)$ formed from mmWave signals (either VRSes or really collected during authentication), we perform the Range Migration algorithm (RMA) [71] as our SAR imaging algorithm to construct a 3D facial image. The reason behind this is that RMA performs well, especially in the near-field imaging scenario. Specially, the mmWave signal matrix $s(x, y, t)$ is first converted from the time-space domain into the wavenumber domain (i.e., $s(k_x, k_y, k)$) for performing distance migration correction and target refocusing. This conversion is done by performing a 2D Fourier transformation and utilizing Weyl's Expansion expression [61], such that the matrix $s(x, y, t)$ in Eq. 3 becomes:

$$s(k_x, k_y, k) = \frac{\epsilon O(k_x, k_y, k_z)}{k_z} = \frac{\epsilon o_{1,1}(k_x, k_y, k_z) A(k_x, k_y)}{k_z}. \quad (6)$$

$o_{1,1}(k_x, k_y, k_z)$ represents the response from the whole face to the antenna in the first row and the first column on the planar antenna array. A is a constant matrix related to the arrangement of antennas on the planar antenna array. Afterwards, we transform $s(k_x, k_y, k)$ to spatial domain (i.e., xyz domain) to obtain a 3D constructed facial image, which can be expressed as:

$$w(x', y', z') = \mathrm{IFT}_{3D}^{(k_x, k_y, k_z)} \left\{ \mathrm{Stolt}^{k_z} \left(s(k_x, k_y, k) k_z \right) \right\} \quad (7)$$

$$= \epsilon o_{1,1}(x_F, y_F, z_F) * A(x, y). \quad (8)$$

According to Eq. 1 and Eq. 6, we find that $o_{1,1}(x_F, y_F, z_F)$ does not only depend on the facial surface curvature but also on the authentication distance. Thus, the above process reveals a constraint: the authentication distance from the face to the radar should be identical during the registration and each authentication attempt. Otherwise, the image $w(x', y', z')$ would be inconsistent. However, it is hard to guarantee that the user keeps an identical distance each time.

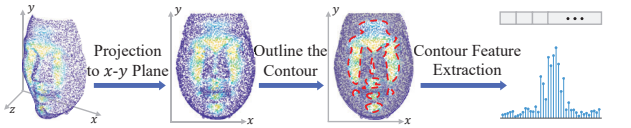
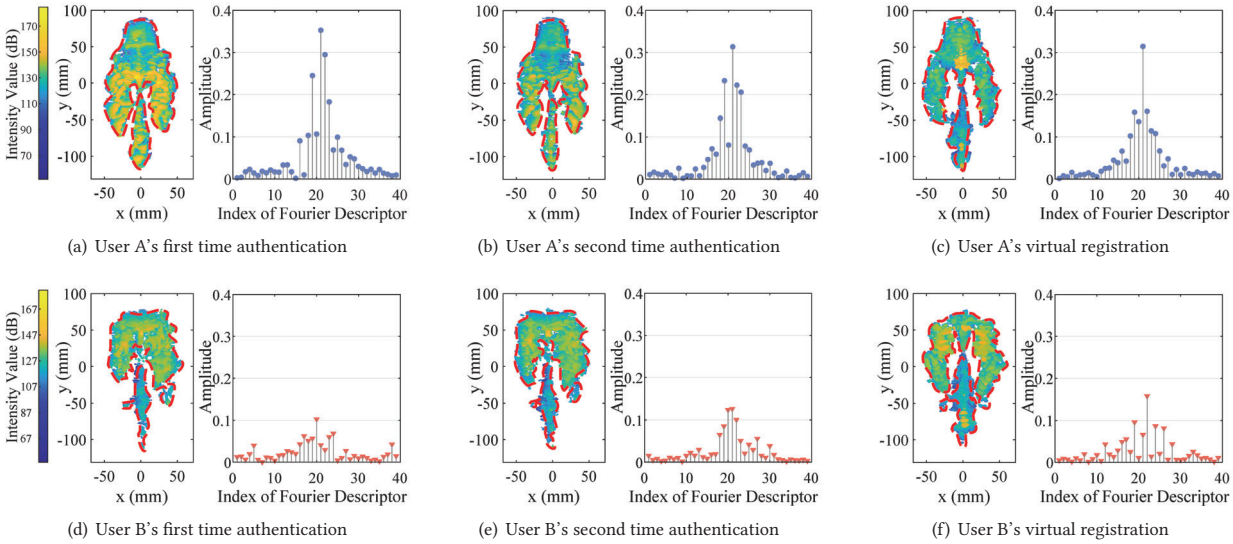


Figure 9: Workflow of distance-resistant facial structure feature extraction.

5.1.2 Distance-resistant facial structure feature. To deal with the above issue, we attempt to pursue a facial structure feature resilient to the authentication distance variation. Inspired by the phenomenon that sometimes people can be identified via their body shapes, we propose to employ the contour of the face (or some areas of the face) as the required feature. We observed an interesting fact. A 3D facial image reconstructed from mmWave signals can be distinctly divided into two areas, i.e., a 'bright' area and a 'dark' area, as shown in Fig. 6, according to its intensity values. The intensity value here is the corresponding element of $w(x', y', z')$ in Eq. 8. Our two preliminary experiments indicate two insights. 1) The bright/dark area on the facial image is corresponding to the user's facial region with high/low surface curvature, as shown in Fig. 6. Hence, the contour of the bright area can be used to represent the facial structure feature. 2) Even if varying the authentication distance, the contours of the bright area are relatively stable as demonstrated in Fig. 7. With these observations, we can safely claim that the bright area's contour is dominantly determined by the facial surface curvature, rather than the authentication distance. Therefore, it can serve as a distance-resistant facial structure feature.

The theoretical explanation of the above observations is as follows. As shown in Fig. 8, we randomly select two adjacent elements $F_1'(x_1', y_1', z_1')$ and $F_2'(x_2', y_2', z_2')$ around the contour of the 3D facial image's bright area. They belong to the dark area and the bright area, respectively. The corresponding points of F_1' and F_2' on the human face are F_1 and F_2, as shown in Fig. 8. Since F_1' and F_2' are two adjacent elements, F_1 and F_2 are very close to each other, which means that their distances to the planar antenna array are almost the same. However, the intensity values of F_1' and F_2' might be significantly different, i.e., $w(x_1', y_1', z_1')$ is much smaller than $w(x_2', y_2', z_2')$ as shown in Fig. 8. Although the authentication distances of these two points are similar, their values of facial surface curvature might be considerably different due to the face structure, like the large difference between the surface normals n_1 and n_2 in Fig. 8. In this case, such surface curvature difference plays the dominant role in determining the intensity value difference between these two points. In particular, those points along the boundary of the bright area would be more representative in reflecting the facial structure feature, since they are prone to contribute distinct differences between the bright and dark areas. Hence, the contour of the bright area on the 3D facial image is only related to the distribution of the facial surface curvature, which can be regarded as a representative distance-resistant facial structure feature.

Fig. 9 shows the steps of feature extraction. To avoid massive computational overhead, we first project the 3D facial image into a 2D facial image $w_{2D}(x, y)$ by choosing the maximum intensity value along its z-axis. Since the intensity value can reveal both the surface curvature and depth information according to Eq. 8, such 3D-to-2D



(a) User A's first time authentication (b) User A's second time authentication (c) User A's virtual registration

(d) User B's first time authentication (e) User B's second time authentication (f) User B's virtual registration

Figure 10: Contours of the bright areas on facial images and the corresponding distributions of Fourier descriptors for real collected signals and virtual registration signals.

projection can improve computational efficiency as well as retain valid information. Then, we outline the contour of the bright area in $w_{2D}(x, y)$, as shown in Fig. 9. Conventional edge detection methods always need to set an empirical threshold in advance to separate the target area (e.g., the bright area) and the background area (e.g., the dark area) [5, 9, 64]. However, allocating a specific threshold to each $w_{2D}(x, y)$ collected in each authentication attempt is time-consuming. Thus, we adopt the idea of the maximum inter-class variance method [42], such that mmFace can outline the contour of $w_{2D}(x, y)$ according to an adaptive threshold. Finally, we remove discrete points and leverage the Fourier descriptor [74] to quantify the contour of the bright area (as shown in Fig.9), which is used as the final distance-resistant facial structure features.

5.2 User matching

Before detailing the match process, we conduct a verification experiment with two volunteers (user A and user B) to confirm the feasibility of using the bright area's contour as representative facial structure features. Specifically, for these two users, we collect the real mmWave signals reflected by their faces twice a week and generate the corresponding facial images. Fig. 10 shows the contours of these facial images' bright areas and their corresponding Fourier descriptor distributions. It can be observed that the results of the same volunteer of different authentication attempts are very similar (Fig. 10(a) vs. Fig. 10(b) and Fig. 10(d) vs. Fig. 10(e)), while they are distinguishable for different volunteers (Fig. 10(a) vs. Fig. 10(d) and Fig. 10(b) vs. Fig. 10(e)). Therefore, the contour of the facial image' bright area is unique, which can serve as a representative and robust facial structure feature.

Additionally, we also conduct a preliminary experiment to validate our VRS generation approach proposed in Sec. 4. In this experiment, we aim to figure out whether user matching can be achieved by calculating the similarity between the features extracted from the VRSes and those extracted from real collected signals. Fig. 10(c) and Fig. 10(f) show the contours of the bright areas and the corresponding Fourier descriptor distributions generated from the two users' VRSes. It can be seen that the features extracted from the VRSes closely resemble the ones extracted from real signals (Fig. 10(a) vs. Fig. 10(c) and Fig. 10(d) vs. Fig. 10(f)). This result demonstrates the accuracy of our mmWave signal propagation model proposed in Sec.4 and validates that we can achieve user matching through similarity comparison.

The process of user matching is as follows. We assume that there are m users registered mmFace. Each user would possess a template stored in the database. The facial structure feature R extracted in an authentication attempt is compared with each of the m templates to obtain m Euclidean distances. mmFace chooses the smallest distance and compares it with an acceptance threshold. If it is smaller than the threshold, mmFace will regard the user as legitimate. Otherwise, mmFace will reject this authentication attempt.

6 LIVENESS DETECTION

Facial biometric features can be quantitatively characterized by the reflection coefficient of the face, which only depends on the material of the facial surface [62]. In mmFace, the amplitude of the signals reflected by the human face is correlated to the reflection coefficient. Thus, it is reasonable to extract biometric features from the amplitude.

However, it is hard to do so since the amplitude is related to multiple factors simultaneously. According to Eq. 2, the amplitude of the signal reflected by the human face and received by the antenna $TR_{m,n}$ can be calculated as:

$$A_{m,n} = \epsilon \iiint \frac{o_{m,n}(F)}{2r_{m,n}} dx_F dy_F dz_F. \tag{9}$$

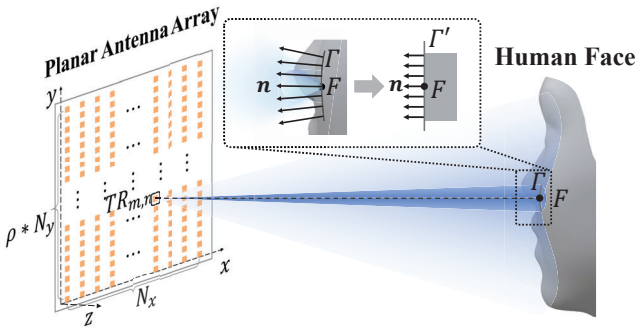


Figure 11: Equating the relatively flat region Γ on the face to a plane Γ′ parallel to the planar antenna array.

It can be observed that the amplitude $A_{m,n}$ is not only related to the reflection coefficient (ϵ) but also the facial curvature characteristic ($o_{m,n}$) as well as face-to-radar distance ($r_{m,n}$). Intuitively, we can deduce the value of ϵ if we know the values of the other two factors in advance. Nevertheless, there are a large number of points on the face. Calculating the values of these two factors for each point is significantly time-consuming. Traditional RF-based material identification works [15, 25, 62, 72] 'solve' this problem by adding constraints (e.g., fixed shape) to sensing targets. Such solutions are impractical for the FA scenario in mmFace. In this work, we propose a biometric feature extraction algorithm (as shown in Algorithm 1) that selects the regions with uniform curvature distributions (relatively flat regions parallel to the planar antenna array) on the face and further extracts biometric features. This reduces the computational complexity in calculating the curvature for each point. In the following, we will detail how we select those regions of interest and extract accurate biometric features from them.

■ **Selecting relatively flat regions on the human face.** For all antennas on the planar antenna array, we first select those that indeed detected the sensing target (i.e., face). Specifically, for each antenna on the planar antenna array, we calculate the distance between it and the target. The distances of all antennas form a distance set D. Since the target occupies most of the area covered by the planar antenna array, we select the antennas whose antenna-to-target distances equal the most common distance of the set D. Then, we form a set H with these antennas to select desirable regions. Specifically, for each antenna in the set H, we calculate the power of its received signal reflected from the target. If the signal power exceeds an empirical threshold m, we will regard the antenna's corresponding face region as a relatively flat region Γ, which can be equivalent to a plane Γ′ parallel to the planar antenna array, as shown in Fig. 11. The principle is that: 1) As illustrated in Eq. 2, although an antenna can receive the signals reflected from all points on the face, it is mainly affected by the region (Γ) closest to it, as shown in Fig. 11. Thus, each antenna can correspond to a specific facial region with a similar size. 2) According to Eq. 6 and Eq. 9, the amplitude ($A_{m,n}$) of the signal received by the antenna depends on the surface curvature ($o_{m,n}$) distribution of its corresponding face region. That is, the more the normals of the corresponding face region are perpendicular to the planar antenna array, the greater the power ($A^2_{m,n}$) of the received signal. Therefore, we can obtain the relatively flat face regions by selecting the antennas whose received signal power is higher than an empirical threshold and these antennas can compose the set G.

Algorithm 1: Biometric Feature Extraction

Input: $s_{m,n}(t)$: mmWave signals, m ∈ {1, Row}, n ∈ {1, $Column$}
Output: S: biometric feature vector
Variables: D: the set of the antennas-to-target distances,
H: the set of antennas with the most common distance,
G: the set of antennas corresponding to the relatively flat face area,
Q: the set of reflection coefficients,
Q': the set of reflection coefficients without outliers,
m: empirical threshold

```
D ← Calculate_Distance(s_{m,n}(t))
H ← Antennas_with_Most_common_Distance(D)
// select the antennas to form a set G whose corresponding face region can
   be regarded as a relatively flat region
for i ∈ H do
    if Calculate_Power(i)>m then
        G ← G ∪ i ;
    end
end
Q ← Calculate_Reflection_Coefficient(G)
Q' ← Remove_Outlier(Q)
S ← Calculate_Statistical_Features(Q')
```

■ **Biometric identification.** A region of interest can be regarded as a set of N points with similar surface curvature ($o_{m,n}$) and distance ($r_{m,n}$). Then the reflection coefficient (ϵ) can be calculated by:

$$\epsilon = A_{m,n}/(N * (\frac{o_{m,n}(F)}{2r_{m,n}})), \tag{10}$$

where N is a constant and F is a random point on the region Γ′, as shown in Fig. 11. Since Γ′ is a plane parallel to the planar antenna array, $o_{m,n}$ equals 1. $r_{m,n}$ can be calculated by the frequency of the received signal. Afterwards, for each transceiver antenna $TR_{m,n}$ in the set G, we repeat the above calculation and form a reflection coefficient set Q. Then, we remove the abnormal elements in Q based on the Pauta criterion [55] and calculate seven statistical features (*median*, *mean*, *standard deviation*, *mode*, *median absolute deviation*, *mean absolute deviation*, and *range*) to form a biometric feature vector S. Finally, we feed the vector into a pre-trained support vector machine (SVM) [11] to determine whether the target is a genuine human face or not.

7 IMPLEMENTATION

We implement mmFace on a mmWave radar module and a 2D slide rail. For probing the practicability, we prototype mmFace on two kinds of mmWave radars with different configurations, including a COTS mmWave radar, and an advanced mmWave radar. Fig. 12 shows the default setup. The data processing is completed by a PC with Intel(R) Core(TM) $i5-8250U$ CPU and 8 GB RAM.

■ **Radars.** *1. COTS mmWave radar.* As shown in Fig. 12, this radar module consists of a single-chip commercial mmWave radar board Texas Instruments (TI) IWR1642-Boost ($299) [27] and a data acquisition board TI DCA1000EVM ($499) [28]. The IWR1642 device contains two transmitting antennas and four receiving antennas. We set it to send 500 frames per second and the two transmitting antennas alternately emit one chirp signal in each frame. Each chirp signal consists of 256 sampling points and its frequency will increase from $f_0 = 77.33GHz$ to $f_T = 80.91GHz$ (with the bandwidth $B = 3.58GHz$ and frequency slope $K = 70.295MHz/\mu s$). *2. Advanced*

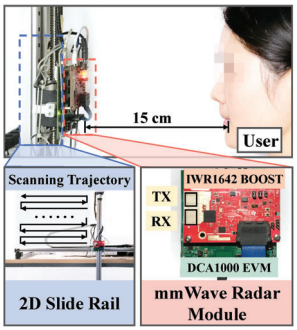


Figure 12: Setup.

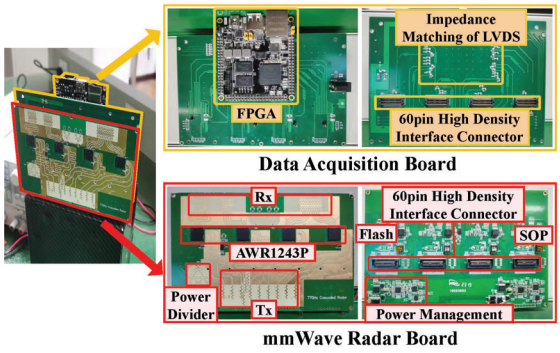


Figure 13: Advanced mmWave radar module.

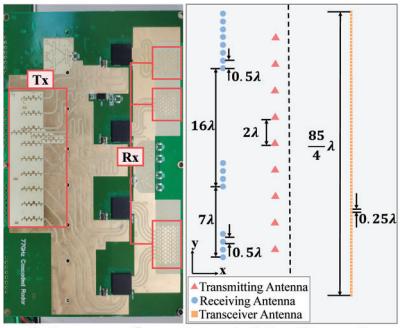


Figure 14: Antennas on the advanced radar.

mmWave radar. To shorten the scanning time, we also build a cascade mmWave radar ($1100) with more antennas, which can be a substitute for the COTS mmWave radar. As shown in Fig. 13, the advanced radar contains four TI AWR1243P radar chips. Each chip has three transmitting antennas and four receiving antennas. From all antennas, we choose nine uniformly distributed transmitting antennas and sixteen receiving antennas to be equivalent to 86 transceivers which are uniformly distributed as shown in Fig. 14. The parameters of the chirp signals (e.g., bandwidth, frequency slope) emitted by these transmitting antennas are the same as those of the COTS radar. Besides, we also build a data acquisition board to enable high-speed raw data capture ($315), as shown in Fig. 14. Compared with the COTS mmWave radar, the advanced one can construct a larger transceiver antenna array. Thus, we can greatly reduce the scanning time for signal collection.

■ **2D slide rail testbed.** In mmFace, the mmWave radar is moved with the assistance of a 2D slide rail to construct a 2D scanning plane with the size of $D_x * D_y$ (200mm*240mm). The scanning speed of the 2D slide rail is 0.5m/s. To construct such a scanning plane, the radar needs to slide N_y rows and the interval between two adjacent rows is Δy. Aligned with the timestamp of the transmitted mmWave signals, each row collects $N_x = 200$ sampling points with an interval of $\Delta_x = 1mm$. For the COTS mmWave radar, it needs to slide $N_y = 30$ rows with an interval of $\Delta y = 8mm$ and the sweeping delay is about 13 seconds. This indicates that the COTS radar is suitable for those FAs without strict requirements for real-time processing. As for the advanced cascade mmWave radar, it only needs to slide $N_y = 3$ rows with an interval of $\Delta y = 85mm$, and this sweeping process takes less than 2 seconds. The computation delays of these two mmWave radars are similar, i.e., about 1 second. In this way, mmFace can realize a nearly real-time FA.

8 EVALUATION

8.1 Data Collection and Metrics

■ **Data collection.** Our experiments are conducted in three different environments: seminar room, lab, and office. For each environment, we invite 30 volunteers including 20 males and 10 females, aged from 18 to 45 (with the height from 158 to 183 cm and the weight from 48 to 85 kg). Among the 30 volunteers, we randomly choose 5 volunteers as spoofers, and the remaining 25 volunteers registered as legitimate users. We conduct these experiments with the approval of our university's Institutional Review Board (IRB). In the registration phase, each legitimate user provides three photos to construct a 3D face model. Based on Sec.4, we generate the virtual signals for user registration. Therein, we set the distance between the 3D face model and the virtual planar antenna array as 15cm. In the authentication phase, all users are asked to place their faces around 15cm in front of the mmWave radar scanning plane. Totally, 30 volunteers launch 3600 authentication attempts over a duration of 6 months.

■ **Metrics.** We define six metrics to quantify the performance of mmFace similar to [68], including False Accept Rate (FAR), False Reject Rate (FRR), Equal Error Rate (EER), Authentication Success Rate (ASR), Receiver Operating Characteristic (ROC), and Defense Success Rate (DSR). In particular, FAR represents the probability that mmFace accepts a spoofer as a legitimate user. FRR represents the probability that mmFace rejects a legitimate user as a spoofer. EER describes the rate where FAR equals FRR. Additionally, ASR is the probability that a legitimate user is correctly authenticated by mmFace. ROC curve describes the relationship between the ASR and FAR under the various thresholds. DSR is the probability that an illegitimate authentication request initiated by an attacker is successfully detected and rejected.

8.2 Overall Performance

We first evaluate the overall performance of mmFace in three different environments. The major difference between the COTS and advanced mmWave radar is the data collection time. Since the authentication performance of these two versions is similar, in the following we focus on the performance of using the COTS mmWave radar for face authentication. To show the superiority of our facial structure feature extraction method, we compare mmFace with two baselines. Baseline 1 (BS1): we only utilize the Fourier descriptors to extract facial structure features without removing discrete points. Baseline 2 (BS2): we directly compare facial images instead of utilizing Fourier descriptors to extract features. Furthermore, to verify the validity of our virtual registration signal generation, we also compare the performance of mmFace with virtual registration and on-site registration. For the on-site registration scenario (On-site), we utilize the real mmWave signals reflected by the user's face for registration rather than the generated virtual signals. Fig. 15 shows the results of ASRs in three different environments. We can observe

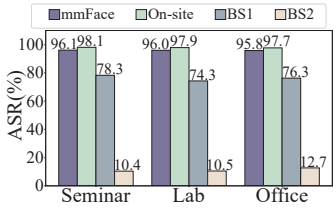


Figure 15: ASR under different environments.

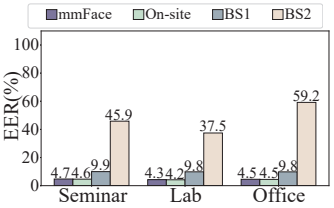


Figure 16: EER under different environments.

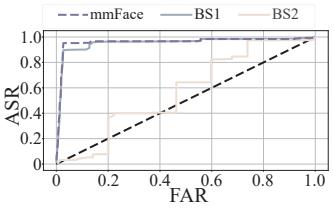


Figure 17: ROC curves under the environment of lab.

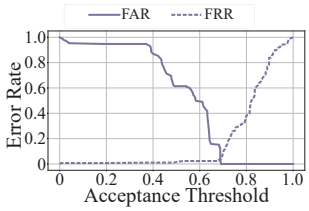


Figure 18: FAR-FRR curves of mmFace under the environment of lab.

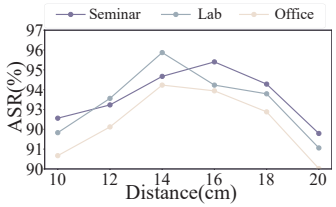


Figure 19: The impact of distance on mmFace.

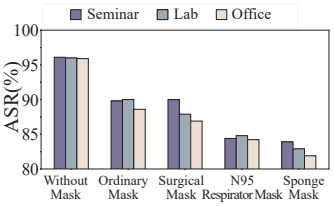


Figure 20: The impact of mask types on mmFace.

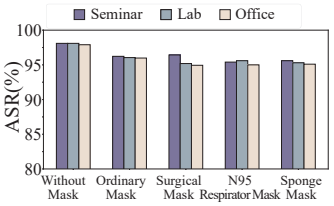


Figure 21: The impact of mask types on mmFace without conducting liveness detection.

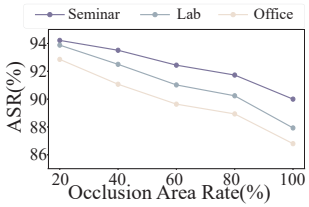


Figure 22: The impact of the mask occlusion area on mmFace.

that the average ASR of mmFace is 96%, while the average ASR of these two baselines is 76.27% and 11.19%, respectively. Additionally, as shown in Fig. 16, the average EER of mmFace is 4.52% while the EERs of these two baselines are 9.89% and 47.51%. The high ASR and low EER of mmFace indicate that it has outstanding authentication performance in different environments. These comparisons with two baselines demonstrate that utilizing Fourier descriptors can effectively extract facial structure features and removing discrete points can improve the authentication accuracy. Additionally, compared with the on-site registration, we can observe that the ASRs and EERs of these two cases are comparable, as shown in Fig. 15 and Fig. 16. This result shows that mmFace can generate virtual registration signals accurately while mitigating the overhead of user registration. Fig. 17 depicts the ROC curves and Fig. 18 shows the FAR-FRR curves under the lab environment. Specifically, the corresponding area-under-curve (AUC) of mmFace is larger than that of both Baseline 1 and Baseline 2. These results demonstrate that mmFace can authenticate users accurately and securely, and it outperforms the other two baselines.

8.3 Robustness

Robustness is a crucial issue for FA systems. Therefore, we evaluate the robustness of mmFace in two aspects: authentication distance variation and the occlusion of masks.

■ **Distance.** We evaluate the impact of distance between the face and the mmWave radar scanning plane in three different environments. We invite 12 volunteers and guide them to place their faces at various distances for authentication, i.e., from 10cm to 20cm with a step of 2cm (this range covers the typical authentication distance of conventional FA systems). For each authentication distance, every volunteer is asked to launch 45 authentication attempts. Fig. 19 shows the results of ASR with the variation of authentication distance. Although the distance variation will cause a slight decrease in ASR, the ASRs of three environments within these distances are still larger than 90%. The results indicate that distance-resistant facial structure features are effective.

■ **Mask occlusion.** To show that mmFace still performs well under the occlusion of masks, we also invite 12 volunteers to conduct two aspects of experiments: the impact of mask types and the impact of mask occlusion area. These experiments are conducted at the default distance of 15cm. We utilize four types of masks to evaluate the performance of mmFace with different materials and thicknesses, i.e., ordinary masks, surgical masks, N95 respirator masks, and sponge masks. Fig. 20 shows the ASRs of legitimate users normally wearing these four types of masks and authenticating under three environments. For each mask type and environment, every volunteer is asked to launch 45 authentication attempts. Here, we have two observations: (1) the ASRs degrade when users wear masks, and (2) different masks have varied impacts on the ASRs.

To investigate the reason for the first observation, we conduct an additional experiment to evaluate the impact of masks on mmFace without conducting liveness detection. As shown in Fig. 21, we find that while wearing masks, the ASR did not deteriorate noticeably (within 3%), compared with that without masks. Such results indicate that the masks mainly interfere with liveness detection rather than the facial structure feature extraction. Based on the above analysis, we further investigate the reason for the second observation: masks with different materials and thicknesses have different impacts on mmFace, as shown in Fig. 20. Particularly, the ordinary and the surgical masks have a smaller impact on the performance of mmFace than N95 respirator and sponge masks. The main reason is that the thicknesses and the materials of the ordinary and surgical masks make mmWave signals easier to perpetrate than N95 respirator and sponge masks. Consequently, the ordinary and surgical masks have less impact on the reflection coefficient measurement in Eq. 10, resulting in a better liveness detection performance. On the other hand, as N95 respirator and sponge masks affect the reflection coefficient measurement, mmFace can sometimes wrongly classify them as spoofing attacks since the measured coefficients differ from

Figure 23: Different occlusion areas of the surgical mask.

those expected from real users. In practice, mmFace can remind the users of its supported masks (e.g., surgical masks and ordinary masks) and advise the users to remove the unsupported masks (e.g., N95 respirator masks and sponge masks) if the authentication attempt cannot pass liveness detection.

We also explore the impact of mask occlusion area on the performance of mmFace. As shown in Fig. 23, we divide the mask occlusion area (from chin to eyes) into five categories according to the size of the occluded area: 20% (only the chin is covered), 40%, 60%, 80%, and 100% (standard wearing style). In this experiment, each volunteer is asked to use the surgical mask to occlude different areas of his face according to the aforementioned five categories and launch 45 authentication attempts for each size of occlusion area. The experiment results of the three environments are displayed in Fig. 22. We can observe that the ASR decreases slightly with the increase of the mask occlusion area rate. Even when the occlusion area rate is 100%, the average ASR is still around 90%, indicating that mmFace is robust to the occlusion of masks.

8.4 Attack and Defense

We conduct experiments under both 2D and 3D spoofing attacks. These spoofing attacks are realized at a distance of 15cm in front of the mmWave radar scanning plane.

■ **Spoofing attack realization.** As aforementioned, there are both static and dynamic 2D spoofing attacks. For realizing these two types of attacks, we use the photos and videos of six legitimate users to cheat mmFace respectively. Each 2D spoofing attack is performed 108 times. For the more challenging 3D spoofing attack, we conduct the attack by utilizing various materials to build 1:1 scale 3D-printed masks of six victims. These materials include metal, PLA, silicone, nylon, and resin. With each type of 3D-printed mask, we perform 108 times 3D spoofing attacks. The setup of these spoofing attacks is presented in Fig. 24.

■ **Defensive capability analysis.** As the 2D spoofing attacks do not involve the 3D structure features, we can easily defend them by comparing them with the registered templates in the database. Our experiment results show that all 2D spoofing attacks cannot deceive mmFace.

In the experiments of 3D spoofing attacks, we plot the confidence distributions (the outputs of one-class SVM) of both the human face and 3D-printed masks in Fig. 25. It is obvious that the values of the confidence for all 3D-printed masks are much smaller than that of human faces. Therefore, mmFace can distinguish the biometric features of human faces from other 3D-printed masks. Furthermore, we explore the impact of distance on the anti-spoofing ability of mmFace. We change the attacking distance from 10cm to 20cm with a step of 2cm. Fig. 26 shows DSRs for the 3D-printed masks of five materials under different distances. We observe that the DSRs for the 3D-printed masks made of PLA, silicone, nylon, and resin are all stable and stay around 100%. The DSR for the metal

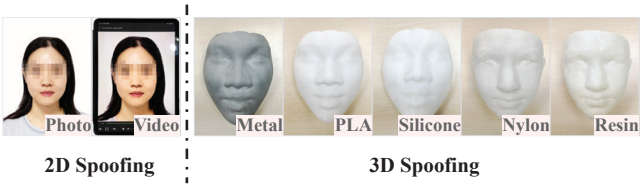


Figure 24: Experiment setup of spoofing attacks.

3D-printed mask is slightly affected when the distance decreases, but it is still higher than 90%. Therefore, mmFace can adapt to the distance variation and effectively defend against 3D spoofing attacks in real scenarios.

9 RELATED WORK

■ **Facial authentication**: Table 1 summarizes six representative FA systems that are most related to our work. Traditional FA systems mainly use RGB cameras to collect facial information [6, 24, 46, 52, 59, 67]. Those approaches have three practical limitations. First, they are susceptible to light variations. Cameras cannot capture accurate facial information in poor light conditions. Second, these FA systems are vulnerable to spoofing attacks [8, 16, 17, 46, 52, 59] since they cannot verify whether an authentication request is initiated by an alive person. Third, camera-based FA systems cannot identify users with masks well [4, 46], as they can only capture and extract unique features around the eyes rather than the entire face.

To address these limitations, recent RF-based FA systems [26, 66, 79] utilize RF signals to capture facial characteristics and achieve authentication. As summarized in Table 1, RFace [66] builds a face authentication system with passive RFID arrays. RFace is resilient to 2D and 3D spoofing attacks but it does not support masks. *Hof et al.* [26] re-purposes WiFi 802.11ad/y chipset to send mmWave signals for sensing and demonstrates that it is possible to identify a user by analyzing the reflected mmWave signals with machine learning algorithms. Similar to RFace, *Hof et al.* [26] do not address the challenges involved in face authentication with masks or spoofing attacks. Besides, these RF-based FA systems [26, 66, 79] require on-site registration which typically takes a long time. In contrast, mmFace only requires each user to provide three photos for registration. Furthermore, unlike these FA systems, mmFace can support masked FA and meanwhile defend against both 2D and 3D spoofing attacks.

■ **RF-based authentication**: RF signals in various frequency bands (RFID [38, 39, 66, 76], WiFi [56, 73, 75], and mmWave [26, 36]) have been exploited to capture user biometrics for authentication. For example, Zhao *et al.* [76] develop a finger verifier to identify users by collecting impedance features via an RFID tag array. Wang *et al.* [56] propose a WiFi-based user authentication system. Users need to stand in the sensing area for a while to collect biometrics like body fat percentage. Lin *et al.* [36] present a heart identification scheme using mmWave radar to probe cardiac characteristics. Current RF-based authentication proposals can achieve decent performance. Nevertheless, their real-world deployment is hindered by three drawbacks: burdensome on-site registration, susceptibility to distance variation, and vulnerability to spoofing attacks. Unlike these works, mmFace aims to develop a user-friendly, robust, and anti-spoofing mmWave-based FA system.

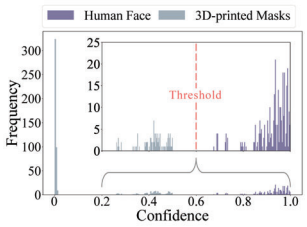


Figure 25: Confidence distribution.

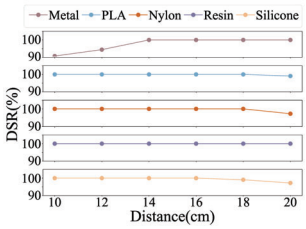


Figure 26: Impact of distance on DSR.

10 DISCUSSION

■ **Robustness:** *1) Face deflection:* mmFace is robust to slight face deflection. Particularly, we validate the face deflection within 10 degrees, and results show that the authentication performance is hardly impacted. As part of future work, we plan to extend the acceptable range of facial deflection angles, which will make mmFace more practical. *2) Authentication distance:* Currently, mmFace is applicable at an authentication distance of 10 to 20cm, which is acceptable for applications such as assess control and user identification at entrances. In our future work, we will further extend the range of authentication distance and improve the authentication robustness of mmFace. *3) Head movement:* Slight head movement during authentication would not affect mmFace. As illustrated in Section 5.1.2, mmFace utilizes the contour of the facial image's 'bright' area which only depends on the user's facial surface curvature as the facial structure features. Such robust facial structure features help mmFace perform well even if users move their heads slightly during authentication. When the user moves the head significantly during the authentication process, mmFace may not be able to capture facial features reliably. To mitigate this problem, we will design a movement detection method and remind users to remain relatively stable. *4) Environment:* Since the authentication range is short and most reflections are from the users, the surrounding environment has little impact on the performance of mmFace. To make mmFace more robust, we will add denoising technology (e.g., background subtraction) to improve mmFace's robustness against environmental noise.

■ **Resilience to signal-level attacks:** Besides the attack models mentioned in Sec. 2.2, sophisticated attackers may perform signal-level attacks, including exhibiting generated VRS to mmFace and the replay attack. However, these two attacks are hard to succeed. Since it is difficult for attackers to imitate the biometric features of the face, the former attack would be rejected by liveness detection with high probability. For the latter one, the attacker needs to capture signals during authentication, and then replay them back to mmFace. However, due to the directionality of mmWave, the receiving antennas of attacker's device must be placed near the mmWave radar, which would be detected by the user.

■ **Practicality:** *1) Accuracy and material sensing capability:* Our experiment results indicate that ordinary and surgical masks have less impact on the reflection coefficient measurement. In contrast, N95 respirator and sponge masks affect the reflection coefficient measurement and liveness detection, leading to degraded performance (e.g., 82% accuracy). That is because mmFace sometimes classifies thick sponge masks as spoofing attacks and denies legitimate users, since extracted reflection coefficients are closer to spoofing attacks and different from those of real users. In practice, mmFace can remind users to remove unsupported masks if measured coefficients are more like those of spoofing attacks. We leave the fine-grained material classification (e.g., thick sponge versus 3D printing materials) as future work. *2) Sensing resolution:* The sensing resolution of mmFace is limited by the resolution of SAR imaging, i.e., 1mm∗1mm∗4cm (length∗width∗depth). To make mmFace more practical, we plan to adopt mmWave radars with larger bandwidths, e.g., 30GHz, to improve the resolution of SAR imaging to 1mm∗1mm∗1mm (length∗width∗depth). In this case, mmFace can extract more fine-grained facial structure features and has the potential to support more users. *3) Form factor:* Due to the bulkiness of mmWave radar and antennas, the current implementation of mmFace can be deployed and used for the access control of group clients (e.g., communities, small- or moderate-scale companies). To extend to other application scenarios (e.g., user authentication for smart devices, integration into smartphones, etc.), the mmWave radar needs to be integrated into a smaller form factor such as a single chip similar to Soli [35]. We believe the sensing algorithms and authentication methods developed in this paper are not limited to the bulky mmWave radar SDK but can be applied to process the mmWave signals of compact mmWave radars in the future.

Table 1: Overview of six state-of-the-art FA systems.

System	Medium	Attack Resilience	Masked FA	On-site Registration
***EchoPrint* [79]**	RF signal	2D	Unable	Yes
***Samsung* [46]**	Camera	None	Unable	No
***Wang et al.* [59]**	Camera	None	Able	No
***Song et al.* [52]**	Camera	None	Able	No
***RFace* [66]**	RF signal	2D & 3D	Unable	Yes
***Hof et al.* [26]**	RF signal	None	Unable	Yes

11 CONCLUSION

In this paper, we develop an anti-spoofing face authentication system named mmFace using off-the-shelf mmWave radars. mmFace extracts facial biometric features as well as structure features from the mmWave signals reflected by human faces to achieve reliable liveness detection and face authentication. Leveraging the penetrability of mmWave, mmFace can still work well under the occlusion of facial masks. We also propose a virtual registration approach to avoid inconvenient on-site registration. Experiment results show that mmFace is accurate, robust, and secure in user authentication.

ACKNOWLEDGMENT

This work is supported in part by National Key R&D Program of China (2020AAA0107705), National Natural Science Foundation of China under grant U21A20462, 61872285, 62032021, Research Institute of Cyberspace Governance in Zhejiang University, Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Grant No. 2018R01005), and Ant Group Funding No.Z51202000234, Hong Kong General Research Fund (GRF) under grant PolyU 152165/19E, Key R&D program of Shaanxi Province 2020ZDXM5-01.

REFERENCES

[1] Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. Capturing the human figure through a wall. *ACM Transactions on Graphics*, 34(6):1–13, 2015.

[2] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C. Miller. 3d tracking via body radio reflections. In *Proceedings of the USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2014.

[3] Fadel Adib and Dina Katabi. See through walls with wifi! In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, 2013.

[4] Aqeel Anwar and Arijit Raychowdhury. Masked face recognition for secure authentication. *arXiv preprint arXiv:2008.11104*, 2020.

[5] John Canny. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698, 1986.

[6] Hakan Cevikalp and Golara Ghorban Dordinejad. Discriminatively learned convex models for set based face recognition. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2019.

[7] Yee Kit Chan and Voon Koo. An introduction to synthetic aperture radar (sar). *Progress In Electromagnetics Research B*, 2:27–60, 2008.

[8] Cunjian Chen, Antitza Dantcheva, Thomas Swearingen, and Arun Ross. Spoofing faces using makeup: An investigative study. In *Proceedings of the IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, 2017.

[9] Hejun Chen, Haiqiang Ding, Xiongxiong He, and Hualiang Zhuang. Color image segmentation based on seeded region growing with canny edge detection. In *Proceedings of the IEEE Conference on Signal Processing (ICSP)*, 2014.

[10] Yimin Chen, Jingchao Sun, Xiaocong Jin, Tao Li, Rui Zhang, and Yanchao Zhang. Your face your heart: Secure mobile face authentication with photoplethysmograms. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2017.

[11] Yunqiang Chen, Xiang Sean Zhou, and Thomas S Huang. One-class svm for learning in image retrieval. In *Proceedings of the IEEE International Conference on Image Processing*, 2001.

[12] Juli Clover. Face id with a mask: How it works and what you need to know. https://www.macrumors.com/guide/mask-face-id/, 2022.

[13] Apple Company. About face id advanced technology. https://support.apple.com/en-us/HT208108, 2022.

[14] Jiankang Deng, Jia Guo, Jing Yang, Alexandros Lattas, and Stefanos Zafeiriou. Variational prototype learning for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[15] Ashutosh Dhekne, Mahanth Gowda, Yixuan Zhao, Haitham Hassanieh, and Romit Roy Choudhury. Liquid: A wireless liquid identifier. In *Proceedings of the ACM Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2018.

[16] Nesli Erdogmus and Sébastien Marcel. Spoofing 2d face recognition systems with 3d masks. In *Proceedings of the International Conference of the BIOSIG Special Interest Group (BIOSIG)*, 2013.

[17] Nesli Erdogmus and Sebastien Marcel. Spoofing face recognition with 3d masks. *IEEE Transactions on Information Forensics and Security*, 9(7):1084–1097, 2014.

[18] Face++. 3d face model reconstruction. https://www.faceplusplus.com.cn/, 2021.

[19] Habiba Farrukh, Reham Mohamed Aburas, Siyuan Cao, and He Wang. Facerevelio: a face liveness detection system for smartphones with a single front camera. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2020.

[20] Centers for disaster control and prevention. Guidance for wearing masks. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-face-cover-guidance.html, 2021.

[21] Yuan Gao. *Non-invasive microwave and millimeter wave reflectometry and imaging for human skin diagnosis and materials characterization*. Missouri University of Science and Technology, 2019.

[22] Syed Zulqarnain Gilani and Ajmal Mian. Learning from millions of 3d scans for large-scale 3d face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[23] Paul F Goldsmith et al. *Quasioptical systems*. Chapman & Hall New York, 1998.

[24] Jianzhu Guo, Xiangyu Zhu, Chenxu Zhao, Dong Cao, Zhen Lei, and Stan Z. Li. Learning meta face recognition in unseen domains. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[25] Unsoo Ha, Junshan Leng, Alaa Khaddaj, and Fadel Adib. Food and liquid sensing in practical environments using rfids. In *Proceedings of the USENIX Conference on Networked Systems Design and Implementation (NSDI)*, 2020.

[26] Eran Hof, Amichai Sanderovich, Mohammad Salama, and Evyatar Hemo. Face verification using mmwave radar sensor. In *Proceedings of the IEEE Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2020.

[27] Texas Instruments Incorporated. Iwr1642: Single-chip 76-ghz to 81- ghz mmwave sensor integrating dsp and mcu. https://www.ti.com.cn/product/cn/IWR1642, 2020.

[28] Texas Instruments Incorporated. Real-time data-capture adapter for radar sensing evaluation module. https://www.ti.com/tool/DCA1000EVM, 2020.

[29] Texas Instruments. Product details about iwr1642 device. https://www.ti.com.cn/product/cn/IWR1642#tech-docs, 2022.

[30] Singular Inversions. Facegen modeller. www.facegen.com, 2021.

[31] Shekh Md Mahmudul Islam, Olga Boric-Lubecke, Yao Zheng, and Victor M. Lubecke. Radar-based non-contact continuous identity authentication. *Remote Sensing*, 12(14):2279, 2020.

[32] Shekh MM Islam, Abraham Sylvester, George Orpilla, and Victor M Lubecke. Respiratory feature extraction for radar-based continuous identity authentication. In *Proceedings of the IEEE Radio and Wireless Symposium (RWS)*, 2020.

[33] Belal Korany, Chitra R Karanam, Hong Cai, and Yasamin Mostofi. Xmodal-id: Using wifi for through-wall person identification from candidate video footage. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2019.

[34] Ke Li, Ruidong Zhang, Bo Liang, François Guimbretière, and Cheng Zhang. Eario: A low-power acoustic sensing earable for continuously tracking detailed facial movements. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(2):1–24, 2022.

[35] Jaime Lien, Nicholas Gillian, M Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics (TOG)*, 35(4):1–19, 2016.

[36] Feng Lin, Chen Song, Yan Zhuang, Wenyao Xu, Changzhi Li, and Kui Ren. Cardiac scan: A non-contact and continuous heart-based user authentication system. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2017.

[37] Jian Liu, Yudi Dong, Yingying Chen, Yan Wang, and Tianming Zhao. Poster: Leveraging breathing for continuous user authentication. In Rajeev Shorey, Rohan Murty, Yingying (Jennifer) Chen, and Kyle Jamieson, editors, *Proceedings of the ACM Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2018.

[38] Jianwei Liu, Xiang Zou, Jinsong Han, Feng Lin, and Kui Ren. Biodraw: Reliable multi-factor user authentication with one single finger swipe. In *Proceedings of the IEEE/ACM Symposium on Quality of Service (IWQoS)*, 2020.

[39] Jianwei Liu, Xiang Zou, Feng Lin, Jinsong Han, Xian Xu, and Kui Ren. Hand-key: Leveraging multiple hand biometrics for attack-resilient user authentication using COTS RFID. In *Proceedings of the IEEE Conference on Distributed Computing Systems (ICDCS)*, 2021.

[40] Daniel Oberhaus. iphone x's face id can be fooled with a 3d-printed mask. https://www.vice.com/en/article/qv3n77/iphone-x-face-id-mask-spoof, 2017.

[41] Michael Opitz, Georg Waltner, Georg Poier, Horst Possegger, and Horst Bischof. Grid loss: Detecting occluded faces. In *Proceedings of the European conference on computer vision (ECCV)*, 2016.

[42] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[43] Muhammed Zahid Ozturk, Chenshu Wu, Beibei Wang, and KJ Ray Liu. Gait-based people identification with millimeter-wave radio. In *Proceedings of the IEEE World Forum on Internet of Things (WF-IoT)*, 2021.

[44] Fabio Rocca. Synthetic aperture radar: A new application for wave equation techniques. *Stanford Exploration Project*, 56:167–189, 1987.

[45] Ian Sample. What are deepfakes and how can you spot them? https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them, 2020.

[46] SAMSUNG. Samsung fr. https://www.samsung.com/fr/smartphones/galaxy-z-fold3-5g/, 2021.

[47] David M Sheen, Douglas L McMakin, and Thomas E Hall. Three-dimensional millimeter-wave imaging for concealed weapon detection. *IEEE Transactions on Microwave Theory and Techniques*, 49(9):1581–1592, 2001.

[48] William T Silfvast. *Laser fundamentals*. Cambridge university press, 2004.

[49] Hadlee Simons. You should probably turn off the galaxy s10's face unlock if you value basic security. https://www.androidauthority.com/galaxy-s10-face-unlock-insecure-964276/, 2019.

[50] Manminder Singh and AS Arora. A novel face liveness detection algorithm with multiple liveness indicators. *Wireless Personal Communications*, 100(4):1677–1687, 2018.

[51] Thomas Smith. Replace your face with an a.i. twin to trick facial recognition. https://debugger.medium.com/replace-your-face-with-an-a-i-twin-to-trick-facial-recognition-22be6931cf1, 2020.

[52] Lingxue Song, Dihong Gong, Zhifeng Li, Changsong Liu, and Wei Liu. Occlusion robust face recognition based on mask learning with pairwise differential siamese network. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2019.

[53] Deepak Vasisht, Guo Zhang, Omid Abari, Hsiao-Ming Lu, Jacob Flanz, and Dina Katabi. In-body backscatter communication and localization. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication (SIGCOMM)*, 2018.

[54] Butao Wang. 3d printed masks are fooling facial recognition on alipay and wechat pay. https://equalocean.com/news/2019121913038/, 2019.

[55] C Wang, J Caja, and E Gómez. Comparison of methods for outlier identification in surface characterization. *Measurement*, 117:312–325, 2018.

[56] Fei Wang, Jinsong Han, Feng Lin, and Kui Ren. Wipin: Operation-free passive person identification using wi-fi signals. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, 2019.

[57] Ju Wang, Jie Xiong, Xiaojiang Chen, Hongbo Jiang, Rajesh Krishna Balan, and Dingyi Fang. Tagscan: Simultaneous target imaging and material identification with commodity rfid devices. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2017.

[58] Tao Wang, Jianwei Yang, Zhen Lei, Shengcai Liao, and Stan Z Li. Face liveness detection using 3d structure recovered from a single camera. In *Proceedings of the IEEE conference on biometrics (ICB)*, 2013.

[59] Ziyan Wang and Tae Soo Kim. Learning to recognize masked faces by data synthesis. In *Proceedings of the IEEE Conference on Artificial Intelligence in Information and Communication, (ICAIIC)*, 2021.

[60] Teng Wei and Xinyu Zhang. mtrack: High-precision passive tracking using millimeter wave radios. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2015.

[61] Hermann Weyl. Ausbreitung elektromagnetischer wellen über einem ebenen leiter. *Annalen der Physik*, 365(21):481–500, 1919.

[62] Chenshu Wu, Feng Zhang, Beibei Wang, and KJ Ray Liu. msense: Towards mobile material sensing with a single millimeter-wave radio. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, 4(3):1–20, 2020.

[63] Yi Wu, Vimal Kakaraparthi, Zhuohang Li, Tien Pham, Jian Liu, and Phuc Nguyen. Bioface-3d: Continuous 3d facial reconstruction through lightweight single-ear biosensors. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2021.

[64] Yuming Xiang, Feng Wang, and Hongjian You. Os-sift: A robust sift-like algorithm for high-resolution optical-to-sar image registration in suburban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6):3078–3090, 2018.

[65] Chenhan Xu, Zhengxiong Li, Hanbin Zhang, Aditya Singh Rathore, Huining Li, Chen Song, Kun Wang, and Wenyao Xu. Waveear: Exploring a mmwave-based noise-resistant speech sensing for voice-user interface. In *Proceedings of the ACM Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2019.

[66] Weiye Xu, Jianwei Liu, Shimin Zhao, Yuanqing Zheng, Feng Lin, Jinsong Han, Fu Xiao, and Kui Ren. RFace: anti-spoofing facial authentication using cots rfid. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2021.

[67] Xiang Xu, Nikolaos Sarafianos, and Ioannis A. Kakadiaris. On improving the generalization of face recognition in the presence of occlusions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR Workshops)*, 2020.

[68] Xiangyu Xu, Jiadi Yu, Yingying Chen, Qin Hua, Yanmin Zhu, Yi-Chao Chen, and Minglu Li. Touchpass: towards behavior-irrelevant on-touch user authentication on smartphones leveraging vibrations. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2020.

[69] Hongfei Xue, Yan Ju, Chenglin Miao, Yijiang Wang, Shiyang Wang, Aidong Zhang, and Lu Su. mmmesh: Towards 3d real-time dynamic human mesh construction using millimeter-wave. In *Proceedings of the ACM Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2021.

[70] Muhammet Emin Yanik and Murat Torlak. Near-field mimo-sar millimeter-wave imaging with sparsely sampled aperture data. *IEEE Access*, 7:31801–31819, 2019.

[71] Muhammet Emin Yanik, Dan Wang, and Murat Torlak. 3-d mimo-sar imaging using multi-chip cascaded millimeter-wave sensors. In *Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2019.

[72] Hui-Shyong Yeo, Gergely Flamich, Patrick Schrempf, David Harris-Birtill, and Aaron Quigley. Radarcat: Radar categorization for input & interaction. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2016.

[73] Yunze Zeng, Parth H. Pathak, and Prasant Mohapatra. Wiwho: Wifi-based person identification in smart spaces. In *Proceedings of the IEEE/ACM Conference on Information Processing in Sensor Networks (IPSN)*, 2016.

[74] Dengsheng Zhang and Guojun Lu. Shape-based image retrieval using generic fourier descriptor. *Signal Processing: Image Communication*, 17(10):825–848, 2002.

[75] Jin Zhang, Bo Wei, Wen Hu, and Salil S. Kanhere. Wifi-id: Human identification using wifi signal. In *Proceedings of the IEEE Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2016.

[76] Cui Zhao, Zhenjiang Li, Ting Liu, Han Ding, Jinsong Han, Wei Xi, and Ruowei Gui. Rf-mehndi: A fingertip profiled rf identifier. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2019.

[77] Mingmin Zhao, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina Katabi. Through-wall human pose estimation using radio signals. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[78] Mingmin Zhao, Yingcheng Liu, Aniruddh Raghu, Tianhong Li, Hang Zhao, Antonio Torralba, and Dina Katabi. Through-wall human mesh recovery using radio signals. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2019.

[79] Bing Zhou, Jay Lohokare, Ruipeng Gao, and Fan Ye. Echoprint: Two-factor authentication using acoustics and vision on smartphones. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2018.

[80] Yanzi Zhu, Yibo Zhu, Ben Y. Zhao, and Haitao Zheng. Reusing 60ghz radios for mobile radar imaging. In *Proceedings of the ACM Conference on Mobile Computing and Networking (MobiCom)*, 2015.