A Physical-World Adversarial Attack Against 3D Face Recognition

Yanjie Li, Yiquan Li, Bin Xiao*
Department of Computing
The Hong Kong Polytechnic University
{yanjie.li, yiquan.li}@connect.polyu.hk, csbxiao@comp.polyu.edu.hk

Abstract

3D face recognition systems have been widely employed in intelligent terminals, among which structured light imaging is a common method to measure the 3D shape. However, this method could be easily attacked, leading to inaccurate 3D face recognition. In this paper, we propose a novel, physically-achievable attack on the fringe structured light system, named structured light attack. The attack utilizes a projector to project optical adversarial fringes on faces to generate point clouds with well-designed noises. We firstly propose a 3D transform-invariant loss function to enhance the robustness of 3D adversarial examples in the physical-world attack. Then we reverse the 3D adversarial examples to the projector’s input to place noises on phase-shift images, which models the process of structured light imaging. A real-world structured light system is constructed for the attack and several state-of-the-art 3D face recognition neural networks are tested. Experiments show that our method can attack the physical system successfully and only needs minor modifications of projected images.

1 Introduction

3D face recognition systems use the human faces’ surface structural characteristics for user authentication and other tasks to compensate the disadvantages of 2D face recognition [1]. The common methods for 3D face recognition systems to get real-time 3D face data includes the structured light imaging system, the time of flight camera and the stereo vision, etc. The structured light imaging system has high measurement precision and good performance on uniform textures [2], thus has been a common method for 3D face measurement [3] and widely used in smart terminals [4] [5] [6].

However, in this paper, we demonstrate that the structured light system may have security risks and introduce a new attack surface to the 3D face recognition system. We propose a novel, physically-achievable 3D adversarial attack called structured light attack, which conceals the adversarial perturbation into the projected patterns to make it hard to perceive for human eyes. The perturbed patterns will create well-designed fake points in the face’s point cloud and spoof the face recognition system.

Fig.1 shows the main process of structured-light-based face recognition system and the structured light attack. We assume that the attacker can access the 3D data but cannot modify the 3D classifier’s inputs directly. In this scenario, some previous researches used 3D printing to deceive the imaging system [7], which is time-consuming and easy to be discovered. In this paper, we utilize the flaws of structured light imaging system itself and modify the projector’s inputs instead. We firstly put forward a 3D transform-invariant loss function to improve the physical attack’s robustness, and then inverse the adversarial noises to the input of the projector. To test the effectiveness of structured light attack, we implement experiments on the physical system and evaluate the attack on several state-of-the-art 3D deep learning models. This paper’s key contributions are:

*Corresponding author.

Preprint. Under review.
Figure 1: The main process of structured-light-based face recognition system and the structured light attack. In the normal 3D measurement process, the structured light imaging system projects specific patterns on the face through a projection device, and uses a camera and a computer to capture and recover the 3D surface. In the structure attack, the attacker adds elaborate perturbation on the projections to spoof the 3D deep learning model.

- We are the first to propose a physical-achievable adversarial attack on 3D face recognition system without using 3D printing. Our attack utilizes the flaws of structured light imaging system, which is more invisible than changing the target face itself directly.
- To satisfy the physical constraint and improve the attack’s success rate, we propose a 3D transformation-invariant loss. Contrast experiment shows the proposed loss function can greatly improve the success rate of structured light attack.
- We verify our attack’s effectiveness on a real-world structured light imaging system and test the proposed attack methods on several state-of-art 3D deep learning models.

2 Background

Structured light imaging system is an active acquisition system, which is almost insusceptible to ambient light noise and has good performance on texture-less surfaces. Hence it is widely used in 3D face scanning system [8]. structured light imaging system usually consists of one camera and one projector. The projector will project a group of specific patterns to encode every unitary position of a surface with a unique code word. Then the camera will decode the code words from captured images to get parallax or phase images used for surface reconstruction. According to the different projection patterns, structured light can be divided as fringe images [9–13], random speckles [14], and other specific patterns [15]. In this paper, we mainly discuss the fringe projection imaging.

A representative 3D reconstruction algorithm for fringe projection imaging is N-step phase-shifting algorithm, which projects N sinusoidal fringe images and gets the object’s 3D coordinates by calculating the phase map. Fig 2 shows the brief process of N-step phase-shifting algorithm. The projected fringe images (Fig 2(a)) has a phase difference of $\frac{2\pi}{N}$. The captured images (Fig 2(b)) can be written as

$$I_i(x,y) = I_a(x,y) + I_b(x,y) \cos[\varphi(x,y) - \frac{2\pi i}{N}]$$

where $i \in [0, N - 1]$, $I_a$ and $I_b$ are respectively the background image and modulation image, $(x,y)$ is the pixel coordinate in captured image. The wrapped phase map $\varphi$ can be computed by

$$\varphi(x,y) = \tan^{-1} \frac{\sum_{n=0}^{N-1} I_i(x,y) \sin(2\pi n/N)}{\sum_{n=0}^{N-1} I_i(x,y) \cos(2\pi n/N)}$$

After getting the wrapped phase map(Fig 2(c)), we can unwrap it to get the absolute phase. There are several different methods to unwrap the phase. In this paper, we use the cyclic complementary gray-code method. Finally, we can get the actual 3D coordinates through the phase-height model or the inverse-camera model [2].
Figure 2: An example of phase-shifting fringe projection: (a) The phase shift patterns to be projected. Here we take 4-step phase-shifting as an example. (b) The statue image with fringe patterns. (c) The wrapped phase map is calculated by (2). (d) The reconstructed depth image from the phase map. We can also reconstruct the point cloud from phase map.

3 Related Work

Optical Adversarial Attack  Optical adversarial attack is a subclass of physical adversarial attack. This attack uses low-cost projector to change the illumination on the target object to escape the correct classification. Compared with 2D or 3D-printed objects, this method has a lower cost on a single attack and can achieve at-scale attacks on a large amount people. This idea was firstly proposed by Nicoles et al. in 2018 [16], but their method needs iterative capturing and optimizing in each attack. Zhou et al. installed some infrared LEDs under a cap to attack the face recognition system in 2018 because infrared light is invisible to human eyes but can be captured by most cameras [17]. But their method needs additional artifacts and has a limited success rate on a large distance. Worzyk et al. projected perturbations onto the road stop signs in 2019 [18]. But they only carried out experiments on printed images rather than on real world objects. Nguyen et al. also applied optical adversarial attack to 2D face recognition system in 2020 [19]. They considered the color distortion of projector, hence only needs to project one time in each attack. In 2021, Gnanasambandam et al. improved Nguyen’s attack’s generality [20] by taking into the consideration of optics’ spatially varying and spectral nonlinear. All of these optical attacks target 2D deep learning models. However, the 2D optical attacks cannot be directly used for 3D scenarios, because their imaging principles have a large difference. In this paper, we model the structured light imaging process and introduce single direction restriction and 3D transform-invariant loss to adapt our attack for the 3D imaging scenario.

3D Adversarial Attack  3D adversarial attacks focus on generating adversarial examples for 3D deep learning models. Currently most 3D attack algorithms target point cloud models [21][29]. In 2019, Xiang et al. [21] first proposed an adversarial attack on the point cloud by manipulating the benign data in two different ways: point perturbation and point generation. Zheng et al. [22] spoofed the deep learning model by dropping some critical points based on the saliency map. To improve the transferability of the attack, Hamdi et al. [23] developed advPC to improve the success attacking rate on different models by adding an auto-encoder reconstruction module to optimize the attack in 2020. Other studies focus on improving the imperceptibility of the adversarial attack, based on the geometric properties of the adversarial point cloud [24][25]. Besides point cloud, Yang et al. [26] proposed MeshAdv attack based on 3D meshes. However, these attacks are primarily digital and few of them have discussed their applications in the real world. On the physical realizable level, in 2019, Cao et al. studied the adversarial 3D attack based on the LiDAR sensor, which is widely used in autonomous vehicles [30]. He successfully fooled the vehicle and added fake front-near obstacles through a time-lapse module and a laser emitter. To the best of our knowledge, no one has realized 3D physical adversarial attack on the face recognition system, whereas the security of face recognition system is increasingly crucial in reality.

4 Attack method

In this section, we will introduce the methodology of structured light attack. Fig.3 shows the process of structured light attack. The black block diagram is the normal process of structured light based face recognition. The red block diagram is the structured light attack, which mainly consists of the 3D transformation-invariant adversarial point generation and the inverse function to map the 3D adversarial point cloud to the 2D input of projector. In section 4.1 We will firstly model the
normal structured-light-based face recognition system process. Then, in section 4.2, we propose a 3D transform-invariant loss function to generate the adversarial point clouds with higher robust. Finally, in section 4.3, we design an inverse function to map the adversarial point clouds to the projector’s inputs.

Figure 3: Flow chart of the structured light attack method. The black block diagram is the process of structured light imaging. The red block diagram is the structured attack, which mainly consists of the adversarial point cloud generation and a inverse function to map the 3D adversarial point cloud to the projector’s inputs.

4.1 Modeling the structured light-based face recognition system

The structured light-based face recognition system consists of two parts: the fringe projection imaging system (FPP) and the face recognition algorithms. FPP system firstly projects and captures a group of phase-shift images in sequential order, then reconstructs the 3D shape using phase-shift algorithm and sends it to the face recognition system. Therefore, we model the structured-light-based face recognition system as a series system (Fig. 3), which consists of the projector, the face, the camera, and the 3D data classifier.

**Projector** Because of the hardware defect or considerations of fitting the human eye, the digital projector usually has nonlinear gray distortions. Suppose \( I_p^i \) is the projector’s \( i \)-th input phase-shift image. Then the projector’s output image \( I_p^o = f(I_p^i) \), where \( f(\cdot) \) is the nonlinear response function of projector, which can be modeled as an exponential function \([31]\)

\[
I_p^o = f(I_p^i) \approx (I_p^i)^\gamma
\]
(3)

where \( \gamma \) is the projector’s intrinsic parameter. The projector’s nonlinear grayscale distortion may introduce reconstruction error and thus cause the attack fail. Therefore, it need to be rectified.

**Object** The projector’s output will illuminate the face’s surface. Assume the ambient light and the surface reflectivity at image coordinate \((x, y)\) are \(a(x,y)\) and \(r(x,y)\). Then the fringe intensity being captured by the camera can be expressed as

\[
I_s^i(x_c, y_c) = r(x_c, y_c)I_p^o(x_p, y_p) + a(x_c, y)
\]
(4)

**Camera** The output of camera \( I_c^o = g(I_p^o) \), where \( g(\cdot) \) is the response function of CCD camera. In this paper, we assume the camera is an industrial digital camera with good linearity. This assumption is reasonable in most commercial structured light systems. Moreover, we consider the projector as an inverse camera. If the pixel at \((x_p, y_p)\) in projection image is projected on the 3D location \((x_w, y_w, z_w)\), and is imaged at \((x_c, y_c)\) in the camera output. The relationship between the \( I_c^o, I_p^o \) and 3D coordinate can be formulated as

\[
I_c^o(x_c, y_c) = \alpha I_p^o(x_p, y_p) \\
z_c[x_c \; y_c \; 1]^T = M_c T_c [x_w \; y_w \; z_w]^T \\
z_p[x_p \; y_p \; 1]^T = M_p T_p [x_w \; y_w \; z_w]^T
\]

where \( \alpha \) is a constant that approximately equals 1. The \( M_c \) and \( T_c \) are the camera’s intrinsic and extrinsic matrices. The \( M_p \) and \( T_p \) are the projector’s intrinsic and extrinsic matrices. The \( z_c \) and \( z_p \) are the camera and projector’s scale factors.
3D reconstruction algorithm For the 3D data reconstruction block, the input is a set of phase-shifted images used for phase unwrapping, and the output is the point cloud \( P \). Firstly, the wrapped phase \( \phi_w \) and the absolute phase of camera \( \phi_c \) are calculated (see Section 2). Because the camera and projector’s images’ corresponding points should have the same absolute phases, we can get the camera’s pixel coordinate \((x_c, y_c)\) by solving \( \phi_c(x_c, y_c) = \phi_p(x_p, y_p) \), where \( \phi_p \) is the absolute phase of original phase-shift images that linearly changes from 0 to \( 2\pi \). Finally, the 3D coordinates of the face is reconstructed by solving the Eq[5]. Therefore, we formulate the 3D data reconstruction block within

\[
P = h(I_c) = h_{3D}(\phi_c(\phi_w(I_c)))
\]

where \( \phi_w(\cdot) \) calculates the wrapped phase from \( I_c^t \), and \( \phi_c(\cdot) \) calculates the absolute phase from \( \phi_w \), and \( h_{3D}(\cdot) \) reconstructs the point cloud from \( \phi_c \).

3D data classifier Here, we assume that we have the knowledge of the inner parameters of the neural networks, so that we can apply the gradient method to find the better adversarial examples. We formulate 3D data classifier model with \( y = \mathcal{M}(P) \), where the input is 3D point cloud \( P \), and the output is the prediction label \( y \) for face recognition tasks.

4.2 3D transform-invariant adversarial point generation

In this subsection, we focus on generating the robust adversarial perturbations \( \Delta \) on the face point clouds. We suppose the adversarial perturbations \( \Delta \) are additive noises on the 3D coordinates of point clouds, which moves the original point clouds \( P \) to \( P' \). We restrict the perturbation on z-direction to reduce the 3D reconstruction error. The detailed reason is illustrated in Sec 3.3.2. Therefore, we formulate the adversarial point cloud with

\[
P' = P + \Delta = P(x, y, z + \Delta z)
\]

where \( P, \Delta \in \mathbb{R}^{N \times 3} \), and \( N \) is the number of points in the point cloud. The aim is to make the 3D data classifier \( \mathcal{M}(\cdot) \) have false decision. For the targeted attack, the loss function’s naive version can be written as

\[
\min_{\Delta} \left( f_{t'}(\mathcal{M}(P + \Delta)) \right) \quad \text{s. t.} \quad D(P, P') \leq \epsilon
\]

where \( t' \) is the target label, \( D(P, P') \) measures the \( l_2 \) distance between the perturbed point cloud and original point cloud. \( f(\cdot) \) is cross entropy loss function. Here we put the distance constraint as a hard constraint because when \( \Delta z \) is larger than the height corresponding to a sinusoidal period, the pixel will be decoded wrongly.

However, in the real world, because the human may have small movements and facial expressions when the attack is in progress. To solve this problem, we propose the 3D transform-invariant loss to make the adversarial point clouds more robust in physical attack. The 3D transform-invariant loss involves minor rigid transformations, normal distribution noises and a re-normalization function. The new loss function is formulated with

\[
\min_{\Delta} \sum_{i=0}^{k-1} f_{t'}(\mathcal{M}(\mathcal{N}(\mathcal{T}_i P + \Delta + \gamma \mathcal{N}_i(0, 1)))) \quad \text{s. t.} \quad D(P, P') \leq \epsilon
\]

where \( \mathcal{T}_i \) is \( i_{th} \) 3D rigid body transformation matrix, includes 3D rotation and 3D translation. We reduce the 3D rotation matrix to rotation matrix around \( x, y, z \) axis separately, according to the commonest motion of human head. \( \mathcal{N}_i(0, 1) \) is \( i_{th} \) normal distribution noise on z-axis to simulate the reconstruction error. The \( \gamma \) is a scale coefficient. In this paper, \( \gamma = 10^{-4} \). \( \mathcal{N}(\cdot) \) is a re-normalization function. Projected gradient descent (PGD) and binary search [21] are used to optimize this loss function.

4.3 Inverse function of the 3D imaging process

To reverse the 3D imaging process, firstly, we calibrate the projector and camera to get their intrinsic and extrinsic parameters. Next, we utilize the model in Sec 4.1 to reverse the point clouds to the input of projector. As we have discussed, the FPP system can be modeled as a series system. Therefore, we can inverse the FPP system blocks one by one. As shown in Fig[3], the adversarial input of projector can be formulated as

\[
I_s' = f_{-1}(r^{-1}(g^{-1}(h^{-1}(P')))))
\]
where \( f^{-1}(), r^{-1}(), g^{-1}() \) and \( h^{-1}() \) are the inverse function of projector, face, camera and 3D reconstruction system separately.

### 4.3.1 Inverse function of the 3D reconstruction algorithm

Because the original absolute phase linearly increases from 0 to \( 2\pi \), the reconstructed absolute phase can be deduced from the projector’s pixel coordinates

\[
\phi_e = h_{3D}^{-1}(P') = (y_p/w) \cdot 2\pi
\]

, where \( W \) is the projection width, and \( y_p \) is deduced by Eq.7. Similarly, we can deduce \( I'_p \) and \( I'_o \) by solving function Eq.1-2 and Eq.5-7. The whole process can be formulated as

\[
I'_o = \alpha^{-1}I'_c = \alpha^{-1}\phi^{-1}_w(h_{3D}^{-1}(P'))
\]

Here we assume \( \alpha = 1 \), according to the good linearity of industrial camera. To accelerate the calculation time, we compare the original and perturbed absolute phase, and only solve \( I'_c(i,j) \) when \( I'_c(i,j) - I_c(i,j) > \lambda \), where we think the corresponding points have been modified. In this paper, we set \( \lambda = 0.02 \).

### 4.3.2 Inverse function of the projecting and capturing process

To reconstruct fake points at \( (x, y, z + \Delta z) \) in the adversarial point cloud, we modify the corresponding pixel \( (x_p, y_p) \)'s grayscale in the projected fringe image \( I_p \) to make the 3D reconstruction algorithm get a wrong height. As shown in Fig.4(a), we assume that the camera directly faces the face, therefore the imaging plane is parallel to the reference plane. In this case, only perturbing on the \( z \)-axis has very small influence on the corresponding pixel coordinate \( (x_c, y_c) \) in camera’s view. Therefore, the \( I'_p \) can be deduced by the following relationship

\[
I'_p(x_p, y_p) = I'_c(x_c, y_c) = I_p(x_p, y'_p)
\]

It is worth mentioning that in order to get the projected image \( I'_p \), some previous 2D optical adversarial attacks estimate the ambient light \( a \) and the pixel wise reflectivity \( r(x, y) \) [20], which need to project multiple extra images beforehand. We find that in the structured light imaging case, the ambient light \( a \) and the pixel-wise reflectivity \( r(x, y) \) can be canceled out by comparing different phase-shifting images(see Eq.12), so there is no need to estimate \( a \) and \( r(x, y) \).
4.3.3 Linearizing the projector distortion

Although the ambient light and reflectivity factors can be canceled out by comparing different phase shifting images, the nonlinear of projector still influences the perturbations’ reconstruction accuracy. As mentioned before, the projector’s distortion can be modeled as an exponential function. To cancel the projector distortion out like the ambient light and reflectivity factor, we apply an inverse function on the projector’s input image to linearize the projector distortion. The method of how to estimate the distortion parameter $\gamma$ in Eq.3 is illustrated in appendix.

5 Experiment

In this section, we implement and evaluate the structured light attack on different neural networks and conduct ablation study of different components. We also test the facial expressions’ influence on the attack success rate to further evaluate the physical attack’s robustness.

Experiment Setup

We train 4 state-of-art deep learning models, including pointnet [32], pointnet++(Multi scale), pointnet++(Single scale), [33] and DGCNN [34] on Bosphorus database(about 104 faces) [35] and Eurecom face database(about 52 faces) [36]. We also collect 10 people’s faces using our own structured light system and add them into the datasets. Our structured light system includes an industry camera and a home projector. Both of their resolutions are 1600×1200 pixels. We designed two experiments of target attack and hidden attack, according to whether the attack target label is given or not.

Experiment Result

Fig.4(b) and Fig.4(c) show the generated adversarial phase-shift image and the captured face image. We can see that our attack has a good camouflage and almost invisible by human eyes. We also compare the generated point clouds in Fig.5. We can see that there are large differences between $P'$ generated by the naive version of CW attack and the reconstructed $P''$, while with single-direction restriction, fewer points are stained and $P''$ is much the same with $P'$.

The experiment results of attack success rate on 4 different deep models are shown in Table.1. We use CW attack [21] as a basic method to find adversarial examples in digital attack. In the experiments, we find that although original CW attack can achieve a high digital attack success rate, as shown in the 1st line, its attack success rate has a sharp drop after the re-normalization, as shown in the 2nd line. We think this is because of the bad robustness of original CW attack. However, normalization is a necessary step in real-world applications. The 3rd line shows that after the 3D transform-invariant loss function, the digital attack success rate has improved obviously.

Table.1’s 4th line and 5th line show the results of structured light attack with and without the 3D transform-invariant ($T$) loss. Without the $T$ loss, the average attack success rate is pretty low. After introducing the $T$ loss, the target attack success rate increases about 19% on the Pointnet, 13% on Pointnet++ MSG, 9% on Pointnet++ SSG, and 17% on DGCNN. We also find although the physical attack’s success rate improves, it is still lower than digital attack. We think this is because of the reconstruction error, diffuse reflection and other factors still influence the attack success rate.
Table 1: The result of structured light attack. We conduct attacks on 4 different models with both target and hidden attack.

| Attack Method      | Pointnet | Pointnet++(MSG) | Pointnet++(SSG) | DGCNN |
|--------------------|----------|-----------------|-----------------|-------|
|                    | target   | hidden          | target          | hidden|
| CW(without re-N)   | 0.98     | 1.00            | 0.98            | 1.00  |
| CW(with re-N)      | 0.63     | 0.95            | 0.35            | 0.72  |
| CW+$T$             | 0.95     | 1.00            | 0.67            | 0.81  |
| CW+SL attack       | 0.36     | 0.44            | 0.19            | 0.25  |
| CW+$T$+SL attack   | 0.55     | 0.58            | 0.32            | 0.41  |

Table 2: The attack success rate on different facial expression.

| Face expression   | Calm | Smile | Surpise | Frown |
|-------------------|------|-------|---------|-------|
|                    | target | hidden | target | hidden | target | hidden | target | hidden |
| CW+$T$             | 0.71  | 1.00   | 0.65    | 1.00   | 0.60   | 1.00   | 0.62   | 0.98   |
| CW+SL attack       | 0.37  | 0.77   | 0.28    | 0.62   | 0.33   | 0.72   | 0.31   | 0.68   |
| CW+$T$+SL attack   | 0.57  | 0.91   | 0.47    | 0.83   | 0.53   | 0.89   | 0.52   | 0.83   |

To test structure light attack’s robustness against the face’s small change, we evaluate the attack success rate with different facial expressions. We first reverse a face point cloud with calm expression to the adversarial phase-shifting images, and then projects these images on the same human face with different expressions. The attack result is shown in 2. The victim model is the Pointnet. From the table, we can see that our attack has a good robustness when the facial expression changes.

6 Discussion and conclusion

Countermeasures To defend against structure light attack, a possible countermeasure is to use random structured light patterns rather than fixed patterns, which is harder for attacker to reverse. Other possible defense methods include filtering the captured images or deleting outlier points in the point clouds. In our experiment, we find both of them can reduce the attack successful rate.

Conclusion In this paper, a physical-world adversarial attack against 3D face recognition system based on structured light is proposed. To increase the physical attack’s success rate, we propose 3D transform-invariant loss function and inverse techniques for structure light system. We evaluate our attack on several state-of-the-art 3D deep leaning model. The physical experiments show that our attack can successfully attack the real-world system and is robust to small face changes. Moreover, the results show that our attack can hide adversarial perturbations into structured light projections and is almost invisible to human eyes. In the future work, we plan to test our attack on more off-the-shelf devices and use infrared light to better conceal the attack.

Broader Impact

This work does not present any foreseeable societal consequence.

References

[1] X. Zhu, Z. Lei, J. Yan, D. Yi, and S. Z. Li, “High-fidelity pose and expression normalization for face recognition in the wild,” in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 787–796, 2015.

[2] S. Feng, C. Zuo, L. Zhang, T. Tao, Y. Hu, W. Yin, J. Qian, and Q. Chen, “Calibration of fringe projection profilometry: A comparative review,” Optics and Lasers in Engineering, vol. 143, p. 106622, 2021.
[3] Z. Xiong, Y. Zhang, F. Wu, and W. Zeng, “Computational depth sensing: Toward high-performance commodity depth cameras,” *IEEE Signal Processing Magazine*, vol. 34, no. 3, pp. 55–68, 2017.

[4] D. Cardinal, “How apple’s iphone x truedepth camera works.” [https://www.extremetech.com/mobile/255771-apple-iphone-x-truedepth-camera-works](https://www.extremetech.com/mobile/255771-apple-iphone-x-truedepth-camera-works) Accessed: March 22, 2022.

[5] S. Elisabeth, “Mantis vision’s 3d depth sensing system in the xiaomi mi8 explorer edition.” [https://www.systemplus.fr/reverse-costing-reports/mantis-visions-3d-depth-sensing-system-in-the-xiaomi-mi8-explorer-edition](https://www.systemplus.fr/reverse-costing-reports/mantis-visions-3d-depth-sensing-system-in-the-xiaomi-mi8-explorer-edition) Accessed: March 22, 2022.

[6] S. Elisabeth, “Orbbec’s front 3d depth sensing system in the oppo find x.” [https://www.systemplus.fr/reverse-costing-reports/orbbecs-front-3d-depth-sensing-system-in-the-oppo-find-x/](https://www.systemplus.fr/reverse-costing-reports/orbbecs-front-3d-depth-sensing-system-in-the-oppo-find-x/) Accessed: March 22, 2022.

[7] K. Yang, X.-Y. Lin, Y. Sun, T.-Y. Ho, and Y. Jin, “3d-adv: Black-box adversarial attacks against deep learning models through 3d sensors,” in *2021 58th ACM/IEEE Design Automation Conference (DAC)*, pp. 547–552, IEEE, 2021.

[8] S. Zhou and S. Xiao, “3d face recognition: a survey,” *Human-centric Computing and Information Sciences*, vol. 8, no. 1, pp. 1–27, 2018.

[9] J.-S. Hyun, G. T.-C. Chiu, and S. Zhang, “High-speed and high-accuracy 3d surface measurement using a mechanical projector,” *Optics Express*, vol. 26, no. 2, pp. 1474–1487, 2018.

[10] C. Zuo, T. Tao, S. Feng, L. Huang, A. Asundi, and Q. Chen, “Micro fourier transform profilometry (µftp): 3d shape measurement at 10,000 frames per second,” *Optics and Lasers in Engineering*, vol. 102, pp. 70–91, 2018.

[11] M. Zhang, Q. Chen, T. Tao, S. Feng, Y. Hu, H. Li, and C. Zuo, “Robust and efficient multi-frequency temporal phase unwrapping: optimal fringe frequency and pattern sequence selection,” *Optics express*, vol. 25, no. 17, pp. 20381–20400, 2017.

[12] Z. Wu, C. Zuo, W. Guo, T. Tao, and Q. Zhang, “High-speed three-dimensional shape measurement based on cyclic complementary gray-code light,” *Optics express*, vol. 27, no. 2, pp. 1283–1297, 2019.

[13] W. Yin, S. Feng, T. Tao, L. Huang, M. Trusiaq, Q. Chen, and C. Zuo, “High-speed 3d shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system,” *Optics express*, vol. 27, no. 3, pp. 2411–2431, 2019.

[14] S. Feng, Q. Chen, and C. Zuo, “Graphics processing unit—assisted real-time three-dimensional measurement using speckle-embedded fringe,” *Applied optics*, vol. 54, no. 22, pp. 6865–6873, 2015.

[15] S. Heist, C. Zhang, K. Reichwald, P. Kühnstedt, G. Notni, and A. Tünnermann, “5d hyperspectral imaging: fast and accurate measurement of surface shape and spectral characteristics using structured light,” *Optics express*, vol. 26, no. 18, pp. 23366–23379, 2018.

[16] N. Nichols and R. Jasper, “Projecting trouble: Light based adversarial attacks on deep learning classifiers,” *arXiv preprint arXiv:1810.10337*, 2018.

[17] Z. Zhou, D. Tang, X. Wang, W. Han, X. Liu, and K. Zhang, “Invisible mask: Practical attacks on face recognition with infrared,” *arXiv preprint arXiv:1803.04683*, 2018.

[18] N. Worzyk, H. Kahlen, and O. Kramer, “Physical adversarial attacks by projecting perturbations,” in *International Conference on Artificial Neural Networks*, pp. 649–659, Springer, 2019.

[19] D.-L. Nguyen, S. S. Arora, Y. Wu, and H. Yang, “Adversarial light projection attacks on face recognition systems: A feasibility study,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 814–815, 2020.

[20] A. Gnanasambandam, A. M. Sherman, and S. H. Chan, “Optical adversarial attack,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 92–101, 2021.

[21] C. Xiang, C. R. Qi, and B. Li, “Generating 3d adversarial point clouds,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9136–9144, 2019.

[22] T. Zheng, C. Chen, J. Yuan, B. Li, and K. Ren, “Pointcloud saliency maps,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1598–1606, 2019.

[23] A. Hamdi, S. Rojas, A. Thabet, and B. Ghanem, “Advpc: Transferable adversarial perturbations on 3d point clouds,” in *European Conference on Computer Vision*, pp. 241–257, Springer, 2020.
[24] Y. Wen, J. Lin, K. Chen, C. P. Chen, and K. Jia, “Geometry-aware generation of adversarial point clouds,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

[25] Q. Huang, X. Dong, D. Chen, H. Zhou, W. Zhang, and N. Yu, “Shape-invariant 3d adversarial point clouds,” arXiv preprint arXiv:2203.04041, 2022.

[26] D. Yang, C. Xiao, B. Li, J. Deng, and M. Liu, “Realistic adversarial examples in 3d meshes,” arXiv preprint arXiv:1810.05206, vol. 2, 2018.

[27] T. Tsai, K. Yang, T.-Y. Ho, and Y. Jin, “Robust adversarial objects against deep learning models,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 954–962, 2020.

[28] D. Liu, R. Yu, and H. Su, “Extending adversarial attacks and defenses to deep 3d point cloud classifiers,” in IEEE International Conference on Image Processing (ICIP), pp. 2279–2283, IEEE, 2019.

[29] D. Liu, R. Yu, and H. Su, “Adversarial shape perturbations on 3d point clouds,” in European Conference on Computer Vision, pp. 88–104, Springer, 2020.

[30] Y. Cao, C. Xiao, B. Cyr, Y. Zhou, W. Park, S. Rampazzi, Q. A. Chen, K. Fu, and Z. M. Mao, “Adversarial sensor attack on lidar-based perception in autonomous driving,” in Proceedings of the ACM SIGSAC conference on computer and communications security, pp. 2267–2281, 2019.

[31] C. Zuo, S. Feng, L. Huang, T. Tao, W. Yin, and Q. Chen, “Phase shifting algorithms for fringe projection profilometry: A review,” Optics and Lasers in Engineering, vol. 109, pp. 23–59, 2018.

[32] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 652–660, 2017.

[33] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “Pointnet++: Deep hierarchical feature learning on point sets in a metric space,” Advances in neural information processing systems, vol. 30, 2017.

[34] A. V. Phan, M. Le Nguyen, Y. L. H. Nguyen, and L. T. Bui, “Dgcnn: A convolutional neural network over large-scale labeled graphs,” Neural Networks, vol. 108, pp. 533–543, 2018.

[35] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun, “Bosphorus database for 3d face analysis,” in European workshop on biometrics and identity management, pp. 47–56, Springer, 2008.

[36] R. Min, N. Kose, and J.-L. Dugelay, “Kinectfacesdb: A kinect database for face recognition,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 44, no. 11, pp. 1534–1548, 2014.