A. Implementation Details

A.1. The Details of Commercial Services

As illustrated in Table 1, we provide the public links of APIs and the public information of practical systems used in our paper.

1) **Face recognition APIs**: We consider three popular face recognition APIs, including Amazon, Tencent, Face++. They are widely used in financial payment and automated surveillance systems.

2) **Face anti-spoofing APIs**: Currently, the mainstream face anti-spoofing solutions include cooperative living detection and non-cooperative living detection (silent living detection). Cooperative living detection requires the user to complete some specified action according to the prompt, and then output the live verification. As a comparison, silent live detection directly aims to directly judge whether the face in front of the machine is real or fake, which is widely used and studied due to convenience and practicability for users. Therefore, we mainly focus on silent living detection in this paper, *e.g.*, FaceID, SenseID, considering their practicability and *millions of* API calls every day.

3) **Practical systems**: We choose two prevailing mobile phones and two automated surveillance systems that have multiple sensors for practical testing as described in Table 1. Note that we did not disclose the manufacturer and parameters of the practical systems for preventing privacy leakage.

A.2. Some Hyperparameters

For 3D attacks, we set the number of iterations as $N = 300$, the budget of perturbation $\eta = 3$, which belongs to a balanced choice between the effective and naturalness. Besides, we utilize Adam optimizer and set the initial learning rate as $1.5 * \eta/N$.

B. More Experiments

Table 2 show the attack success rates (%) of the different face recognition models on the LFW dataset. We also conduct an ablation study as shown in Table 3 on LFW to fully investigate the coefficients of 3DMM. As illustrated in the evaluation results on CelebA-HQ, the effects on LFW present a similar overall conclusion. As a whole, AT3D-P can obviously obtain the best black-box attack success of face recognition models among all 2D and 3D attacks in most testing settings. The reason is that AD3D-P fully leverages low-dimensional optimization based on 3DMM, making generated adversarial meshes more effective and transferable for black-box models.

C. Naturalness

C.1. Evaluation of AT3D-M and AT3D-P

In this section, we quantitatively evaluate the naturalness of meshes generated by **AT3D-M** and our method **AT3D-P**. Specifically, we calculate the discrete Gaussian curvature measure [1] of the original and adversarial meshes. The Gaussian curvature measure estimates the smoothness of the surface that accounts much in visual quality according to [1].

Definition C.1 (The discrete Gaussian curvature measure)**.** *Assume P is the vertex set of a mesh M, the discrete Gaussian curvature Φ^G is the function that associates with every (Borel) set $B \subset \mathbb{R}^3$:*

$$\Phi^G(B) = \sum_{p \in B \cap P} g(p), \tag{1}$$

where $g(p)$ is the angle defect of mesh M at point p, that equals 2π minus the sum of angles between consecutive edges incident on p.

In our experiments, the set B is indeed the sphere centered at some point. Thus the discrete Gaussian curvature at point p can be re-formulated as:

$$\Phi^G\left(B(p,r)\right) = \sum_{p' \in B(p,r) \cap P} g(p'), \tag{2}$$

where r is the radius of the sphere. To evaluate the general smoothness of a mesh, we can denote an average Gaussian curvature measure of mesh M as

$$\begin{aligned} \Phi_M^G &= \frac{1}{|P|} \sum_{p \in P} |\Phi^G\left(B(p,r)\right)| \\ &= \frac{1}{|P|} \sum_{p \in P} \left| \sum_{p' \in B(p,r) \cap P} g(p') \right|. \end{aligned} \tag{3}$$

Fig.1 shows the mean value of the average Gaussian curvature measures from the original meshes, the adversarial meshes generated by **AT3D-M** and **AT3D-P** on LFW, respectively. As the perturbation η increases, the mean curvature of meshes generated by **AT3D-M** also rapidly grows, which means the method cannot reserve the smoothness of original meshes. As a comparison, our method **AT3D-P** almost keeps the initial curvature and hardly changes the smoothness, resulting in better visual quality in terms of different values of perturbation.

C.2. More Examples of AT3D-ML

AT3D-ML adopted multiple popular losses in mesh-based optimization, *e.g.*, chamfer loss, laplacian loss, and edge length loss [2], which are blended into the crafted

	Name	Public link / information
Face Recognition	Amazon	https://aws.amazon.com/rekognition/
	Face++	https://www.faceplusplus.com.cn/face-comparing/
	Tencent	https://www.tencentcloud.com/products/facerecognition
Anti-Spoofing	FaceID	https://www.faceplusplus.com.cn/faceid-solution/
	SenseID	https://www.sensetime.com/senseid/
	Tencent	https://www.tencentcloud.com/products/facerecognition
	Aliyun	https://s.alibaba.com/product/cloudauth/
Practical System	Mobile Phone 1	including RGB, depth cameras, supporting recognition and anti-spoofing
	Mobile Phone 2	including RGB, depth cameras, supporting recognition and anti-spoofing
	Access Control System 1	including RGB, depth and infrared cameras, supporting recognition and anti-spoofing
	Access Control System 2	including RGB, depth and infrared cameras, supporting recognition and anti-spoofing

Table 1. We list the public links of APIs or public information of practical systems used in this paper.

Source Model	Methods	Eye					Eye & Nose					Respirator				
		Arc.	Mob.	Cos.	Res.	API-1	Arc.	Mob.	Cos.	Res.	API-1	Arc.	Mob.	Cos.	Res.	API-1
ArcFace	2D-MIM	90.50*	15.50	13.50	8.50	2.00	99.25*	52.50	35.25	34.50	4.50	87.00*	6.75	7.00	5.00	1.50
	2D-EOT	**98.75***	37.50	30.00	22.75	10.00	99.50*	80.25	60.50	65.75	19.25	**98.25***	21.75	**22.50**	17.50	5.50
	AT3D-M	46.00*	29.25	27.75	21.75	22.50	89.75*	68.00	67.50	60.25	64.25	44.50*	18.25	19.50	13.75	29.25
	AT3D-ML	46.00*	29.00	28.50	21.00	22.25	91.00*	67.75	**68.50**	60.75	65.00	45.50*	18.75	20.00	14.00	28.50
	AT3D-P	93.75*	**59.00**	**41.00**	**52.25**	**47.50**	**99.75***	**89.75**	59.25	**86.50**	**85.25**	89.00*	**41.00**	11.50	**39.25**	**45.25**
MobileFace	2D-MIM	6.50	91.25*	30.50	28.75	3.25	35.75	**100.0***	62.75	70.25	7.50	12.75	80.75*	19.00	19.00	1.50
	2D-EOT	14.25	**100.0***	**48.25**	50.50	6.25	59.50	**100.0***	**82.25**	**90.25**	18.50	19.75	**99.50***	**38.00**	45.00	2.00
	AT3D-M	21.50	58.50*	28.25	25.75	21.75	62.50	93.75*	61.75	64.50	63.25	21.50	45.25*	19.25	14.75	26.75
	AT3D-ML	21.50	58.00*	27.50	26.25	21.50	63.75	93.25*	61.75	65.50	63.25	21.75	45.75*	19.25	15.25	26.75
	AT3D-P	**50.50**	85.75*	42.50	**54.25**	**41.00**	**82.25**	95.75*	63.00	82.75	**82.00**	**39.50**	91.00*	11.25	**46.00**	**39.25**
ResNet50	2D-MIM	5.50	33.50	27.50	89.25*	3.00	41.00	82.50	64.50	99.75*	7.00	13.75	27.50	19.75	85.25*	2.00
	2D-EOT	13.00	55.25	39.00	**99.50***	9.25	63.75	**94.75**	**79.75**	**100.0***	33.50	19.75	**48.00**	**34.00**	**99.00***	5.25
	AT3D-M	19.50	30.50	24.25	47.25*	20.75	58.50	68.75	59.50	91.75*	62.75	18.00	19.50	17.25	35.50*	25.00
	AT3D-ML	19.75	29.75	24.75	47.50*	20.50	57.25	68.00	59.75	91.25*	62.25	18.25	18.50	16.75	35.50*	24.50
	AT3D-P	**48.00**	**62.75**	**42.50**	95.00*	**47.50**	**86.25**	91.50	61.25	**100.00***	**84.75**	**33.75**	44.75	11.75	90.50*	**43.25**

Table 2. The attack success rates (%) of the face recognition models on LFW with adversarial meshes. * indicates white-box attacks.

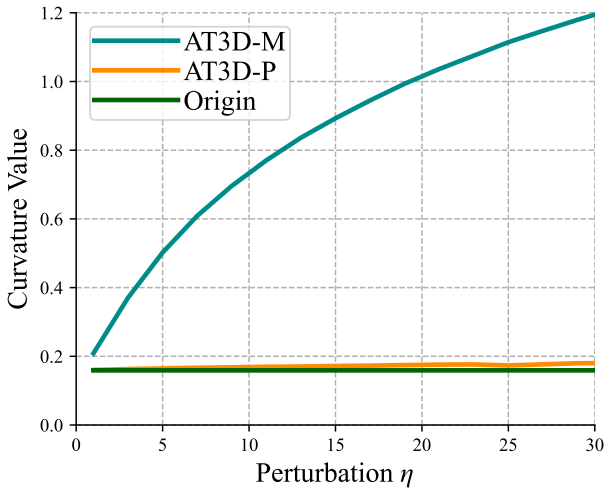

Figure 1. Comparison of naturalness metric among different methods.

AT3D to improve effectiveness and smoothness. Fig.2 shows the generated meshes of AT3D-M, AT3D-ML and AT3D-P under different perturbation η, respectively. As η increases, the results of AT3D-M and AT3D-ML have rapid changes in surface curvature with severe self-intersection. We also found that AT3D-ML can only slightly improve the visual quality of meshes. One potential reason is that the total number of vertices and faces is too large. For example, our patch covering eyes and nose contains $35,709$ vertices and $18,368$ faces, which are very typical in 3D face models due to the requirement of high-fidelity reconstruction. However, it will make AT3D-ML difficult to keep naturalness by only applying such losses. As a comparison, our proposed AT3D-P method provides better visual quality and has fewer self-intersecting faces by perturbing the 3DMM coefficients of identity and expression. Therefore, the crafted meshes can be more easily fabricated into a solid patch using 3D printers.

D. Physical Experiments

3D-printed techniques. We choose a common and popular 3D printer, *i.e.*, Stratasys J850 Prime, for printing all

Source Model	Settings	Eye					Eye & Nose					Respirator				
		Arc.	Mob.	Cos.	Res.	API-1	Arc.	Mob.	Cos.	Res.	API-1	Arc.	Mob.	Cos.	Res.	API-1
ArcFace	$\{\alpha, \beta\}$	66.50*	43.50	35.50	34.75	35.75	94.00*	77.25	57.00	72.00	75.50	62.00*	25.75	9.25	21.75	32.50
	$\{\tau\}$	72.50*	43.50	38.50	39.00	35.75	95.50*	78.00	57.00	72.50	72.75	61.25*	25.50	11.75	21.50	34.75
	$\{\alpha, \beta, \tau\}$	**93.75***	**59.00**	**41.00**	**52.25**	**47.50**	**99.75***	**89.75**	**59.25**	**86.50**	**85.25**	**89.00***	**41.00**	**11.50**	**39.25**	**45.25**
MobileFace	$\{\alpha, \beta\}$	36.75	74.75*	39.75	39.75	36.00	76.00	97.00*	58.50	74.50	78.00	24.75	61.75*	8.00	26.75	26.50
	$\{\tau\}$	39.50	83.50*	39.75	44.75	38.00	72.25	98.50*	61.75	77.50	76.50	27.25	63.50*	10.50	30.50	35.75
	$\{\alpha, \beta, \tau\}$	**50.50**	**85.75***	**42.50**	**54.25**	**41.00**	**82.25**	**95.75***	**63.00**	**82.75***	**82.00**	**39.50**	**91.00***	**11.25**	**46.00**	**39.25**
ResNet50	$\{\alpha, \beta\}$	31.25	42.75	36.75	73.50*	33.75	71.75	75.75	55.00	97.25*	74.00	22.00	28.50	10.25	62.25*	32.75
	$\{\tau\}$	34.25	47.50	39.00	82.50*	36.00	69.00	78.25	55.00	98.00*	73.75	25.00	30.25	11.25	67.50*	35.50
	$\{\alpha, \beta, \tau\}$	**48.00**	**62.75**	**42.50**	**95.00***	**47.50**	**86.25**	**91.50**	**61.25**	**100.0***	**84.75**	**33.75**	**44.75**	**11.75**	**90.50***	**43.25**

Table 3. The attack success rates (%) of different coefficients on LFW with adversarial meshes. * indicates white-box attacks.

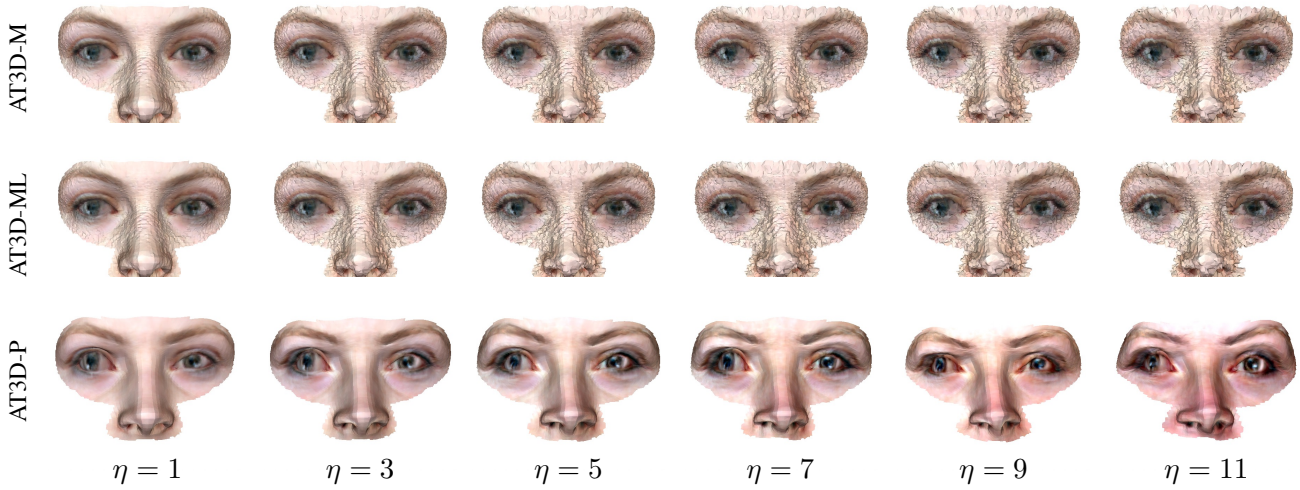

Figure 2. Experiments on how different η affects the performance on LFW.

physical adversarial meshes by using resin-based materials.

Detailed experiments. In physical experiments, we choose **50** attacker-to-victim pairs to conduct the experiments to verify the effectiveness of the proposed method in the physical world. The procedure is evaluated by: 1) taking a face photo of a volunteer with a fixed camera under natural light; 2) crafting adversarial textured meshes for each volunteer; 3) achieving 3D printing and pasting them on real faces of the volunteers; 4) testing the attack performance against practical face recognition system. For the practical device S-1, we can easily import the victims' information in batches into the system and obtain the output similarity scores when achieving attacker-to-victim adversarial testing. Therefore, we conduct the whole experiments on 50 attacker-to-victim pairs against the device S-1 and present all detailed results for every testing pair, as shown in Fig. 3. The default threshold of verification for the device S-1 is 70. If the distance of an image exceeds the threshold, the device views it as a successful verification; otherwise a failing verification. We can see that there exist 33 successful cases among all 50 examples. After considering anti-spoofing, we obtained a passing rate of $23/50$ as reported in this paper.

Finally, we also provide video demos in the supplementary material, where 3D-printed meshes can unlock one mobile phone and an automated surveillance system.

References

[1] David Cohen-Steiner and Jean-Marie Morvan. Restricted delaunay triangulations and normal cycle. In *Proceedings of the nineteenth annual symposium on Computational geometry*, pages 312–321, 2003. 1

[2] Jinlai Zhang, Lyujie Chen, Binbin Liu, Bo Ouyang, Qizhi Xie, Jihong Zhu, Weiming Li, and Yanmei Meng. 3d adversarial attacks beyond point cloud. *arXiv preprint arXiv:2104.12146*, 2021. 1

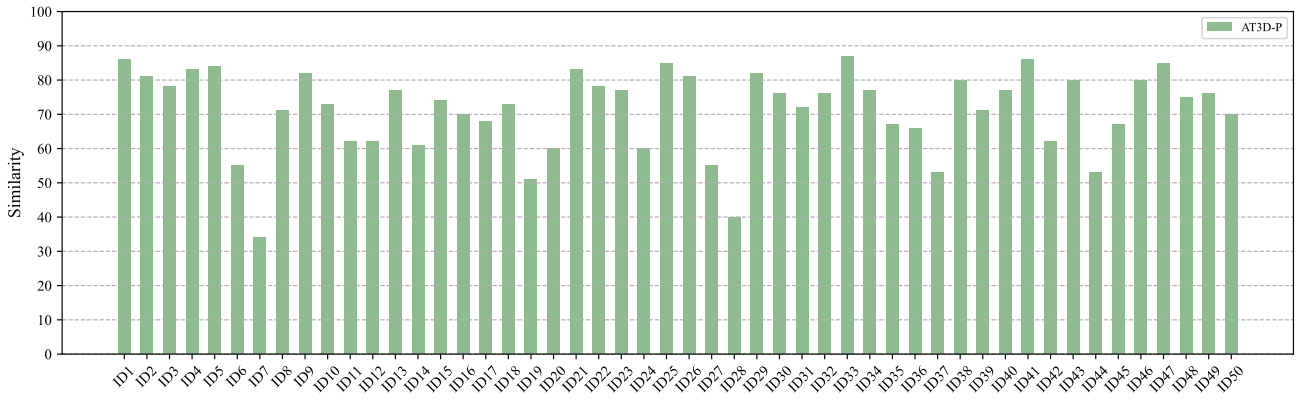

Figure 3. Detailed physical results for every testing pair against the device S-1.