·封面文章·

基于多特征融合的3D打印面具攻击检测

陆经纬[1*]，陈鹤天[2]，马肖攀[1]，陈继民[2]

[1]北京工业大学北京未来网络科技高精尖创新中心，北京 100124；

[2]北京工业大学北京市数字化医疗 3D 打印工程技术研究中心，北京 100124

摘要 针对人脸认证系统的欺骗攻击，传统欺骗攻击方式主要包括照片和视频攻击。随着三维(3D)打印技术的快速发展，使用 3D 面具进行欺骗攻击逐渐成为新威胁。在剪切波变换基础上，结合人脸 3D 几何特征和局部区域纹理变化，针对 3D 面具欺骗攻击提出一种利用多层自编码网络进行特征融合分类来识别攻击面具的方法。通过非下采样剪切波变换从目标人脸 3D 图像中提取低频子带和高频子带。在低频子带上利用尺度空间函数对特征点进行检测、定位及方向分配，生成特征描述子。将所生成的特征描述子与高频子带上提取的纹理特征输入栈式自编码器和 softmax 分类器进行瓶颈特征融合分类。在基于柔性 TPU 材质 3D 打印面具 BFFD 数据库上的实验结果表明，相比于以往单独使用纹理特征的方法，加入 3D 几何特征的多特征融合方法对反 3D 面具攻击的准确率有显著提升。

关键词 图像处理；面具攻击检测；剪切波变换；自编码器

中图分类号 TP391 **文献标识码** A **doi**：10.3788/LOP56.031002

3D Printing Mask Attacks Detection Based on Multi-Feature Fusion

Lu Jingwei[1*], Chen Hetian[2], Ma Xiaopan[1], Chen Jimin[2]

[1] *Beijing Future Network Technology Advanced Innovation Center, Beijing University of Technology, Beijing* 100124, *China*;

[2] *Beijing Digital Medical 3D Printing Engineering Technology Research Centre, Beijing University of Technology, Beijing* 100124, *China*

Abstract Aiming at the spoofing attacks for the current face authentication systems, the traditional spoofing attacks include displaying printed photos and replaying recorded videos. With the rapid development of three-dimensional (3D) printing technology, the 3D mask spoofing attack is becoming a new threat. On the basis of the shearlet transform and combining with the 3D geometric attributes and the local regional texture changes, a method by utilizing the multilayer autoencoder network to conduct the feature fusion-based classification to identify the attack mask is proposed for the 3D mask spoofing attack. The low-frequency sub-band and several high-frequency sub-bands are extracted from the 3D image of the target face by the non sub-sampled shearlet transform method. The scale space function is used to detect, locate and distribute the feature points and then to generate feature operators in the low-frequency sub-band . Then, the generated feature operators and the texture features extracted from the high-frequency sub-band are combined in series and fed into the stacked autoencoder network and the softmax classifier to conduct the bottleneck feature fusion-based classification. The experimental results in the BFFD database based on the flexible TPU material 3D print mask shows that, the multi-feature fusion method added the 3D geometric feature has an obvious improvement for the accuracy of the anti-spoofing performance against 3D mask attacks to compare with the previous method of using the texture feature alone.

Key words image processing; mask attacks detection; shearlet transform; autoencoder

OCIS codes 100.2000; 100.4994; 100.5010

收稿日期：2018-07-04；**修回日期**：2018-08-04；**录用日期**：2018-08-13

基金项目：北京市自然科学基金重大项目(Z140002)

* **E-mail**：18810815230@126.com

1 引 言

人脸识别因具有非接触性、使用方便等优势，近年来已经取得很大进展，并在个人身份认证系统中得到了广泛应用[1]。然而，各种攻击欺骗手段的多样性也使得人脸识别研究面临着巨大挑战。人脸认证系统中的欺骗攻击是指通过不正当的手段伪造真实用户面部特征，从而获得用户系统权限的一种方式[2]。传统人脸识别系统通过二维图像进行身份检测，目前二维人脸欺骗检测方法主要包括活体检测[3-4]、运动检测[5-6]和纹理分析[7-8]。活体检测通常是指通过引导用户进行眨眼、摇头、张嘴、微笑等交互式动作来判断是否为合法用户。基于运动信息分析方法认为真实人脸和打印照片或视频录屏的动作模式之间存在区别，通过建立运动模型来描述物体表面的光流场，同时利用用户头部与背景之间的光流相关性来区分真假人脸。纹理特征分析方法认为欺骗人脸经过二次采集或多次采集后，在纹理细节上存在不同程度的伪影和模糊，包括清晰度、人工痕迹、微纹理和变换域的统计信息等，例如在傅里叶频谱中伪造人脸图像的高频分量大大减少。

随着三维(3D)扫描、3D打印技术的日益成熟，3D面具的制造难度大大降低。已经存在一些网站仅需要提供一张正面照片(可选择提供侧视照片)，就可以为用户提供可穿戴3D面具的制造服务，可见通过3D打印面具的攻击方式已成为人脸识别系统面临的新挑战[9]。而传统的人脸攻击检测方法在面对3D面具的攻击时基本失效。例如，3D面具不再是二维平面运动模式，所以基于背景光流特征分析的方法便失去效果；由于3D面具极为逼真的细节且可以随时进一步改进优化，基于纹理检测的方法的效果也被极大限制。相关评估实验显示，在只使用二维(2D)人脸识别算法(ISV)[10]时，面具攻击数据集(3DMAD)数据库中有65.70%的3D面具攻击被识别为合法用户[11]。由此可见，3D面具对人脸识别系统具有很强的欺骗性。

由于此前的3D打印技术尚不成熟以及相关3D面具数据库的不完善，针对3D打印面具攻击方式的研究并不多见。Li等[12]首次提出基于近红外成像的人脸识别方法。Zhang等[13]提出一种免交互多光谱特征检测的判别方法，利用Lambertian反射模型分析人类皮肤与非皮肤的多光谱特性，选择判别波长并根据真脸和假脸的反射率数据形成数据集，训练支持向量机(SVM)分类器以学习及判别分类的多光谱分布，实验结果准确率为89.18%。但其实验数据库中所用的面具制作相对比较粗糙，仿真度很低，也没有对面具的欺骗攻击性能进行分析。Kose等[14]利用局部二值模式(LBP)分别从包含16名实验者的真实人脸及相应3D面具的Morpho数据库提取出彩色图像和深度图像的纹理特征，并利用SVM进行判别。算法分别在色彩图像和深度图像上获得88.12%和86%的准确率。在文献[14]的基础上，Wallace等[10]加入了3DMAD数据库，分别采用χ^2、线性判别分析(LDA)和SVM三种分类器进行判别。在色彩图像和深度图像上的准确率分别达到了99.05%和98.73%。Menotti等[15]提出使用深度卷积神经网络(DCNN)并对SVM进行结构优化与滤波器优化，实验在3DMAD数据库上实现0%的半错误率(HTER)。Agarwal等[16]通过Haralick纹理特征来检测攻击，本文算法从视频帧的冗余离散小波变换中按块方式提取Haralick纹理特征。使用主成分分析(PCA)减少特征向量的维度并通过SVM分类，同样在3DMAD数据库实现了0%HTER，但以上两种方法并未指出针对3D打印面具的防欺骗攻击性能。上述方法已经证明了基于二维人脸图像特征能够在一定程度上规避3D打印面具带来的威胁，不过受环境(光照、背景等)和人脸本身(表情、姿态、遮挡、年龄等)所带来的不利影响，二维图像识别精度和对3D面具识别的准确度很难进一步提高。目前，基于三维人脸扫描的特征分析方法，包括薄板样条函数插值和迭代最近点在实验测试中效果较好，但基于三维几何属性特征分析的反攻击性能尚未深入研究。此外，目前基于单一特征领域和特征维度级别上的融合研究比较广泛，但从不同范畴中选择不同活体特征进行融合的方法并不多。

本文探索了利用剪切波变换作为提取图像几何特征的滤波器和描述图像局部区域纹理变化的算子。与普遍使用的LBP不同，剪切波变换具有灵活的方向选择性，记录像素点与其周围像素点的对比信息，在表示分布不连续性方面更为有效，对光照以及噪声的干扰也有较强的稳健性。在几何特征方面，本文选择主曲率作为描述包含面部特征的几何属性。设计了基于几何特征的三维人脸识别方案，并评估了其对BFFD数据库的3D面具反攻击性能。最后，本文提出了一种基于神经网络方法的多特征融合框架。所采用的栈式自编码器不仅是一个监督分类器，还可以产生瓶颈特征，改变原始特征的空间分布并对神经网络的原始输入进行压缩和稀疏

表示，从而实现特征之间的深度融合。

2 基本原理

2.1 剪切波构造

小波变换目前已被广泛应用在图像纹理分析中。然而，二维小波变换在方向上具有局限性，缺乏捕获高维数据几何特性的能力。Guo 等[17]通过合成膨胀仿射系统构造了多尺度几何分析工具剪切波。剪切波变换可以对多维图像进行最优逼近，具有灵活的方向选择性、区域性、各向异性等优点。当维数为 2 时，根据能量有限函数生成的合成仿射系统形式为

$$\{\psi_{j,l,k}(x)\}=\{|\det \boldsymbol{A}|^{j/2}\psi(\boldsymbol{B}^l\boldsymbol{A}^j x-k):j,l\in Z,k\in Z^2\},\tag{1}$$

式中 $\psi\in L^2(R^2)$，$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 2×2 可逆矩阵，$|\det \boldsymbol{B}|=1$，j,l,k 分别是尺度、方向、平移参数，x 是输入信号。矩阵 $\boldsymbol{A}^j$ 对应尺度变换，$\boldsymbol{B}^j$ 对应剪切变换。剪切波是合成小波的一个特例，其矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的取值为

$$\boldsymbol{A}=\boldsymbol{A}_0=\begin{pmatrix}4&0\\0&2\end{pmatrix},\boldsymbol{B}=\boldsymbol{B}_0=\begin{pmatrix}1&1\\0&1\end{pmatrix},\tag{2}$$

对任意的 $\xi=(\xi_1,\xi_2)\in\hat{R}^2$，$\xi_1\neq 0$，令

$$\hat{\psi}^{(0)}(\xi)=\hat{\psi}^{(0)}(\xi_1,\xi_2)=\hat{\psi}_1(\xi_1)\hat{\psi}_2\left(\frac{\xi_1}{\xi_2}\right),\tag{3}$$

式中 $\hat{\psi}_1$、$\hat{\psi}_2$ 是 ψ_1、ψ_2 的傅里叶变换，$\hat{\psi}_1,\hat{\psi}_2\in C^\infty(\hat{R})$，$\mathrm{supp}\ \hat{\psi}_1\subset\left[-\frac{1}{2},-\frac{1}{16}\right]\cup\left[\frac{1}{16},\frac{1}{2}\right]$，$\mathrm{supp}\ \hat{\psi}_2\subset[-1,1]$。从而 $\mathrm{supp}\ \hat{\psi}^{(0)}\subset\left[-\frac{1}{2},\frac{1}{2}\right]^2$。进一步可以得到 $\psi_{j,l,k}$ 的频域支集为

$$\mathrm{supp}\ \hat{\psi}_{j,k,l}\subset\left\{(\xi_1,\xi_2):\xi_1\in[-2^{-2j-1},-2^{-2j-4}]\cup[-2^{-2j-4},-2^{-2j-1}],\left|\frac{\xi_2}{\xi_1}+l2^{-j}\right|\leqslant 2^{-j}\right\},\tag{4}$$

即 $\psi_{j,l,k}$ 是方向为 $l2^{-j}$，大小为 $2^{2j}\times 2^j$ 的梯形对。

2.2 三维人脸点云主曲率

曲率是曲面的最基本属性，不随面部动作姿态变化而改变，属于外部不变量，理论上是描述人脸三维特征最好的工具。主曲率由第一基本形式和第二基本形式定义，在曲面中的每个点邻域上都有两个主曲率，可以表示曲面在不同方向上的弯曲程度。在可微曲面 S 上任选一点 p，λ_{1_p} 和 λ_{2_p} 为对应的第二基本形式 h 的特征值（q 是与 h 有关的二次型），可以描述 S 上 p 点的处的局部弯曲信息，第二基本形式是三维欧氏空间光滑曲面切丛上的二次形式，其他点的曲率信息都可以通过欧拉定理得到。

将第二基本形式 h 及其相关的二次型 q 归纳为光滑表面 S 上对应的测量形式 $\mathbb{Z}^3$。假设 $\mathbb{Z}^3$ 中任意 Borel 子集的向量场分别为 $\boldsymbol{B}$、$\boldsymbol{X}$，$\bar{h}_B$ 和 $\bar{q}_B$ 分别为 h 和 q 的广义测度形式，即

$$\bar{h}_B(\boldsymbol{X},\boldsymbol{Y})=\int_{S\cap B}h_p(pr_{T_pS}\boldsymbol{X}_p,pr_{T_pS}\boldsymbol{Y}_p)\mathrm{d}p,\tag{5}$$

$$\bar{q}_B(x)=\int_{S\cap B}h_p(pr_{T_pS}\boldsymbol{X}_p,pr_{T_pS}\boldsymbol{X}_p)\mathrm{d}p=\int_{S\cap B}q_p(pr_{T_pS}\boldsymbol{X}_p)\mathrm{d}p,\tag{6}$$

式中：pr_{T_pS} 表示 S 上 p 点切平面 T_pS 上的正交投影；$\boldsymbol{X}$ 是 E 中的一个常数向量场；$\bar{q}_B(\boldsymbol{X})$ 是对于任何固定 Borel 集的衡量指标；$\{\lambda_{1_B},\lambda_{2_B},\lambda_{3_B}\}$ 是相关特征值的集合；$\{e_{1_B},e_{2_B},e_{3_B}\}$ 是 $\bar{h}_B$ 的特征向量集合；$\lambda_i:\boldsymbol{B}\to\lambda_{i_B}$，$i\in\{1,2,3\}$ 是 $\mathbb{Z}^3$ 的度量值，记为主曲率测量值。

然而，三维人脸样本通常描述为三角网格模型，这种离散曲面边缘点附近的曲面是分段连续不可微的，因此传统估计方法并不适用。文献[18-19]中提出并证明的一个可能的解决方法是将主曲率的定义从光滑曲面推广到离散曲面。作为一个典型的离散曲面，三角网格曲面的形状信息不能用逐点方法处理。假设 $\mathbb{Z}^3$ 中的一个三角形网格 M，根据循环周期[20]的概念，$\bar{h}$、$\bar{q}$ 的常数向量场 $\boldsymbol{X}$ 的精确计算公式为

$$\bar{h}_B(\boldsymbol{X},\boldsymbol{Y})=\sum_{e\in E}l(e\cap\boldsymbol{B})\angle(e)\langle\boldsymbol{X},e\rangle\langle\boldsymbol{Y},e\rangle,\tag{7}$$

$$\bar{q}_B(\boldsymbol{X})=\sum_{e\in E}l(e\cap\boldsymbol{B})\angle(e)\langle\boldsymbol{X},e\rangle^2,\tag{8}$$

式中：E 表示 M 中边 e 的集合；$l(e\cap\boldsymbol{B})$ 表示 e 中属于 $\boldsymbol{B}$ 的长度；$\angle e$ 表示入射面 f_1 和 f_2 至 e 的单位法线 n_1 和 n_2 之间的符号夹角。$\bar{h}_B$ 的相关矩阵 $\boldsymbol{H}_B$ 记为

$$\boldsymbol{H}_B=\sum_{e\in E}l(e\cap\boldsymbol{B})e\cdot e^t。\tag{9}$$

类似的命名 $\bar{h}_B$ 的特征值集合 $\{\lambda_{1_B},\lambda_{2_B},\lambda_{3_B}\}$ 是 M 上的 $\boldsymbol{B}$ 的主曲率测量值。相应的特征向量集合 $\{e_{1_B},e_{2_B},e_{3_B}\}$ 也可以用来估计 $\bar{h}_B$ 和 $\boldsymbol{B}$ 之间的关系。由 $\bar{h}_B$ 推导可知，三个特征向量分别是 $\boldsymbol{X}$ 在 $\boldsymbol{B}$ 上的两个主方向和一个法向方向。主曲率测量值集合 $\{\lambda_{1_B},\lambda_{2_B},\lambda_{3_B}\}$ 与三角网格中三维人脸扫描的几何特性是一致的，并适用于描述人脸表面，这是基于主曲率关键点描述算子的微分几何基础。

3 多特征融合的反面具攻击算法

3.1 整体框架

本文算法流程图如图 1 所示。首先通过非下采样剪切波变换从目标人脸 3D 图像中提取出低频子带和高频子带，在低频子带上利用尺度空间函数对特征点进行检测，然后定位特征点并分配方向，利用特征点邻域旋转不变性对特征点进行描述，生成特征描述子，即 meshSIFT 特征向量，设置不同阈值以完成特征点匹配，匹配结果作为 3D 人脸识别模型的输出。高频子带主要包含图像的纹理特征，为了实现与几何特征互补，本文采用栈式自编码器和 softmax 分类器对 meshSIFT 几何特征以及纹理特征进行融合分类。由于神经网络中使用了 sigmoid 阈值函数，瓶颈特征被自动映射到 0～1 之间，适合用于融合不同维度的特征。

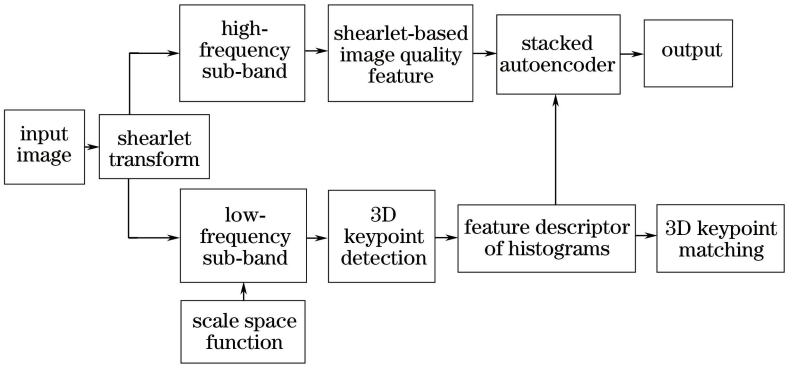

图 1 算法流程图

Fig. 1 Flowchart of the proposed algorithm

3.2 三维人脸表面剪切波变换

离散剪切波变换由多尺度分解和方向滤波两部分组成。首先初始化分解的层数为 $j=L$，通过非下采样的拉普拉斯金字塔算法，将原始表面数据 f 分解成为主要包含几何信息的低通滤波后的表面 f_a^j 和主要包含纹理信息的高通滤波后的表面 f_d^j，并在伪极向格上计算 $\hat{f}_d^j$，得到 Pf_d^j 并用二维快速傅里叶逆变换对其进行带通滤波处理，从而得到 j 层的剪切波系数。离散剪切波变换由具有平移不变性，且 j 层高频剪切波系数所包含的纹理信息会随着 j 的增加而更加细致。

3.3 三维人脸表面特征点检测

尺度不变特征变换(SIFT)特征匹配算法是一种基于尺度空间并对图像缩放、旋转甚至仿射变换保持不变性的特征匹配算法[21-22]。Smeets 等[23]把 SIFT 特征推广到三维人脸网格上，取得了较高的识别率，该特征即为 meshSIFT 特征。曲面上点的曲率表示该点的局部区域曲面的变化趋势，且曲率作为二次导数，很容易受到噪声的干扰，所以用曲率来搜索匹配点对能有效地抑制噪声，提高点云配准的效率[24]。高斯函数作为空间尺度核函数，具有描述图像数据多尺度特征的作用。首先用一系列高斯核函数滤波器 g_{σ_s} 对网格人脸扫描进行平滑来构造图像高斯尺度空间。其中 s 为尺度，σ_s 值的大小代表了图像的平滑程度，其值越小图像尺度越小，细节部分越多。指定脸部扫描中的顶点 v_i，周围的面部表面通过与其相邻顶点上卷积 g_{σ_s} 变得更平滑。中心顶点 v_i 更新为 $v_{i_{\sigma_s}}$，表达式为

$$v_{i_{\sigma_s}}=\frac{\sum_{v_j\in N(v_i,1)} g_{\sigma_s}(v_i,v_j)\cdot v_j}{\sum_{v_j\in N(v_i,1)} g_{\sigma_s}(v_i,v_j)},\tag{10}$$

式中 $N(v_i,1)$ 表示 v_i 一环邻域测地距离内的顶点集合，且高斯核函数 g_{σ_s} 为

$$g_{\sigma_s}(v_i,v_j)=\exp(-\|v_i-v_j\|^2/2\sigma_s^2),\tag{11}$$

本节估算了 1.3 小节中引入的 3D 人脸表面的每个尺度空间上的主曲率特征值 v_{i_B}，根据高斯差分计算关键点位置的曲率差，即

$$\delta_i[\lambda_i(B_{v\sigma_s})]=\lambda_i(B_{v\sigma_s})-\lambda_i(B_{v\sigma_{s-1}}),\ i=1,2,3,\tag{12}$$

式中 δ_i 表示与 B 上 i 的主曲率测量值相对应的曲率差。如果曲率差 v_i 是三个尺度 σ_{s-1}、σ_s 和 σ_{s+1} 中围绕 v_i 的一环顶点的极值，则 v_i 定义为关键点 v_k，σ_s 是其对应的检测尺度。

3.4 生成 3D 关键点特征描述子

3D 关键点描述可以分为两部分。首先是指定一个主方向来提高轻微面部姿态变化的稳健性。其次是通过创建主曲率直方图来构造特征描述符。为了获得一个方向不变的描述符，为每个关键点分配一个规范方向。利用关键点所在的表面法线构建局

部参考系，使得领域内的点不受面部表情影响。假设一个关键点定位 v_k 和它的适当尺度 σ_s 在一个网状面部扫描 F 中。与 v_k 相关的主方向 d_{v_k} 的分配由半径为 R_{σ_s} 的测地圆内的相邻顶点 $v_j \in \mathbb{N}(v_k)$ 决定，其随关键点的尺度而自适应变化，表达式为

$$\mathbb{N}(v_k)=\{v_j \mid d_g(v_k,v_j)\leqslant R_{\sigma_S}, v_j \in F\}。 \tag{13}$$

在相邻区域内，首先确定一个与 v_k 的单位法向量 $\boldsymbol{n}_{v_k}$ 正交的平面 TS_{v_k}，然后将邻域 $v_j \in \mathbb{N}(v_k)$ 的单位法向量 $\boldsymbol{n}_{v_j}$ 投影到包含关键点的网格切平面上 TS_{v_k} 上。用 2.2 节中的方法估算出 v_k 和 v_j 的所有单位法向量，这些投影的法向量集中在由 360 个直条组成的加权直方图内，每个直方图的样条数目由关键点测地距离的高斯权重组成。

通过在 TS_{v_k} 上计算的高斯加权直方图来创建 v_k 的主方向与对应于 λ_{v_k} 的最大值 e_{v_k} 之间的角度，并确定为加权方向直方图的峰值。其最高峰即被选作规范方向，高斯权重定义为

$$w(v_k,v_j)=\mathrm{mag}(v_j)\cdot g\sigma_s(v_k,v_j), \tag{14}$$

式中 mag(·)为每个特征点梯度的幅值，其表达式为

$$\mathrm{mag}(v_j)=\sqrt{n_1^x(B_{v_j})^2+n_1^y(B_{v_j})^2}。 \tag{15}$$

通过 Tola 等[25]提出的二维平面描述符可知，每个关键点的特征描述由位于 9 个重叠圆 r_1，r_2，…，r_9 中的周围顶点获得，圆半径是 $3.75\sigma_s$。r_1 位于以中心为关键点的中心部分。如图 2 所示，从主方向开始的其他 8 个圆按照时序顺序排列。从这 8 个圆的每个中心到关键点的距离是 $4.5\sigma_s$。这种花型描述符模拟人类复杂细胞在视觉皮层中的功能，并且对小幅度头部姿势变化的影响不敏感。分别建立 3 个主曲率测量直方图，每个圆弧区域标记为 $h_i^{r_j}$，$i=1,2,3$。在每个 r_j 中，第 i 个主曲率量值均等量化为 8 个直条，并用高斯核函数加权，标准偏差假设为当前点到相应圆心的欧几里德距离。按照以下规则归一化并连接与 3 个主曲率(3 个主曲率×9 个区域)相关的所有 27 个直方图为

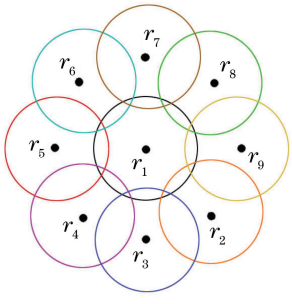

图 2 9 个区域计算局部描述符并生成 HOC

Fig. 2 Computation of local descriptor in nine regions and generating HOC

$$f_{\mathrm{HOC}}=\{h_1^{r_1},\cdots h_1^{r_9},h_2^{r_1},\cdots,h_2^{r_9},h_3^{r_1},\cdots,h_3^{r_9}\}, \tag{16}$$

式中 HOC 是三维网格人脸扫描中每个关键点的主曲率-meshSIFT 特征向量。从而得到 216 维的圆形邻域局部特征向量。

给定两个 3D 人脸扫描模型，基于每个关键点上的局部特征向量来计算它们的相似性。假设一个关键点 $v_{k_i}^1$ 属于第一个人脸扫描模型表面并且属于第二个表面的所有关键点 $\{v_{k_j}^2\}$。估计 $v_{k_i}^1$ 和 $\{v_{k_j}^2\}$ 的特征向量之间的角度集合 $\{\alpha_j^i\}$。每个角度被定义为验证集和注册集，

$$\alpha_j^i=\arccos\left(\frac{\langle H_i^1,H_j^2\rangle}{\|H_i^1\|\cdot\|H_j^2\|}\right), \tag{17}$$

式中角度 α_j^i 按升序排列，如果两个角度之间的比率小于相似性阈值 r_a，则认为匹配成功，否则拒绝。最后，将匹配关键点的数量定义为两个面部表面之间的相似性测量 μ，μ 越大表示两个 3D 人脸扫描来自同一个人的概率越大。

3.5 剪切波变换纹理特征提取

与真实人脸相比，假人脸可能具有锐度减少、纹理差异、加性噪声和伪影的特点。与 LBP 和高斯差分(DOG)算子相比，剪切波可以更好地描述曲线的奇点，包括边缘、纹理和伪影。基于剪切小波的图像纹理评估在检测虚假人脸的边缘模糊、纹理扭曲、翻录视频、打印照片等引起的各向异性伪影以及真实面部皮肤与欺骗性介质之间的纹理差异等方面具有优异的表现。同时，剪切波也可以描述各向同性噪声和 3D 面具表面伪影。对于二维图像，连续剪切波变换可通过映射定义为

$$SH_\phi f(a,s,t)=\langle f,\phi_{a,s,t}\rangle, a>0, s\in R, t\in R^2, \tag{18}$$

$$\phi_{a,s,t}(x)=|\det \boldsymbol{M}_{a,s}|^{-\frac{1}{2}}\phi(\boldsymbol{M}^{-1}x-t), \tag{19}$$

式中 $\phi \in L^2(R^2)$，t 是平移参数，x 是输入信号，$SH_\phi f(a,s,b)$ 是每个绿色块的剪切系数。矩阵 $\boldsymbol{M}_{a,s}$ 可定义为

$$\boldsymbol{M}_{a,s}=\boldsymbol{B}_s\boldsymbol{A}_a=\begin{pmatrix}a & \sqrt{a}\,s\\ 0 & \sqrt{a}\end{pmatrix}\text{and}$$

$$\boldsymbol{A}_a=\begin{pmatrix}a & 0\\ 0 & \sqrt{a}\end{pmatrix},\boldsymbol{B}_s=\begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix}, \tag{20}$$

式中 $\boldsymbol{A}_a$ 为各向异性膨胀矩阵，$\boldsymbol{B}_s$ 为剪切矩阵。图 3 总结了人脸图像进过剪切变换得到不同特征子带的剪切波系数计算过程。其中每个绿色方块中的元素被定义为

$$x(a,s,b)=\frac{\sum\left|SH_{\phi}f(a,s,b)\right|}{m},\tag{21}$$

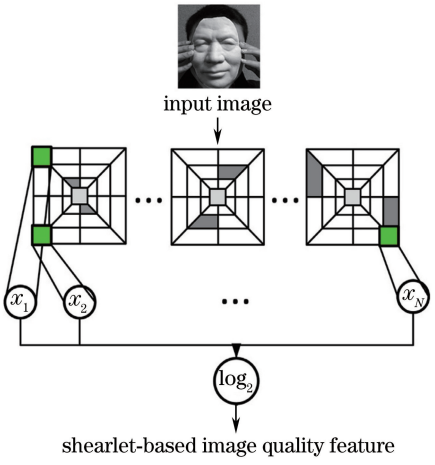

图 3 基于剪切波的图像纹理特征计算过程

Fig. 3 Calculation process of shearlet-based image texture feature

式中：$a=1,\cdots,A$ 为规模指数（不包括最小尺度）；$s=1,\cdots,S$ 为方向指数；$b=1,\cdots,(M/m)^2$ 为每个子带块索引；M 为方形图像的大小；m 为每个绿色块的大小。

3.6 特征融合

为了实现基于主曲率的 meshSIFT 特征和基于剪切波变换的图像纹理特征之间的特征融合，采用多层堆栈式自编码器以及 softmax 分类器对特征进行融合并分类。

自编码人工神经网络是一种基于无监督学习，试图学习对初始函数近似表示的神经网络，输入层和隐层组成编码器，以便输出与 x 相似的 $\hat{x}$，如图 4(a)所示，优化自编码器优化代价函数为

$$J_{\text{sparse}}(W,b)=J(W,b)+\beta\sum_{j=1}^{S_2}\mathrm{KL}(\rho\,\|\,\hat{\rho}_j),\tag{22}$$

式中：$J(W,b)$是用于学习活体信息的自编码器代价函数；$J_{\text{sparse}}(W,b)$是自动编码器的稀疏约束成本函数；ρ 是稀疏参数；$\hat{\rho}_j$ 是第 j 个隐藏层神经元的平均活跃度；KL 是用于测量 ρ 和 $\hat{\rho}_j$ 之差的相对熵散度函数；S_2 是隐藏层中神经元的数量；β 是控制稀疏性惩罚因子的权重。首先设置降维隐藏层和稀疏约束，通过反向传播算法来计算代价函数偏导数，之后通过梯度下降法，一步步迭代并更新参数 W 和 b，得到最小化的损失函数，并将输出的压缩稀疏表示结果作为瓶颈特征向量。如图 4(b)所示，将前一层稀疏自编码器的输出的结果作为后一层自编码器的输入。利用带标记数据集进行反向传播，从而对全局神经网络进行训练。自编码器训练可以看作是一个预训练过程，可以为神经网络优化提供一个很好的初始化状态。然后使用标记数据对自编码器和 softmax 分类器参数进行微调，可以进一步提高活体特征分类的瓶颈特征，减少训练时间。最后对测试组人脸特征进行提取，并传入网络中进行分类，来确定是否存在人脸攻击使用。

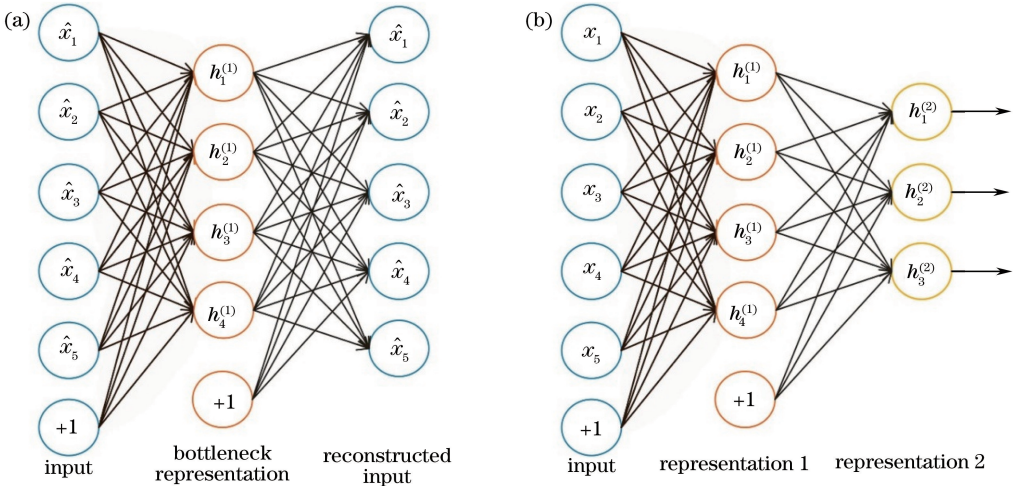

图 4 (a) 3 层网络的自编码器；(b)栈式自编码器结构

Fig. 4 (a) Autoencdoer of three layer network; (b) structure of stacked autoencoder

首先，从扫描的人脸 3D 图像中提取基于主曲率的 meshSIFT 特征向量，输入到第一个网络中得到瓶颈特征。然后，将基于剪切波变换的图像纹理特征向量输入到第二个神经网络以提取其瓶颈特

征。最后，将来自两个不同融合特征的瓶颈特征作为一个聚合的瓶颈特征向量串联起来，并进一步送入后续的活体检测神经网络，以此来尽可能去除冗余信息。如图 5 所示，其中 $x_{k,i}$ 代表第 k 个子网络的输入向量中的第 i 个元素，$h_{k,i}^{(1)}$ 是在第 k 个子网络的隐藏层中学习的第 i 个主要活体特性，$h_i^{(2)}$ 是集成神经网络中第二个隐藏层中学习的第 i 个主要活体特征，$P(y=C|x)$ 是 C 类(真/假)输入 x 的概率。

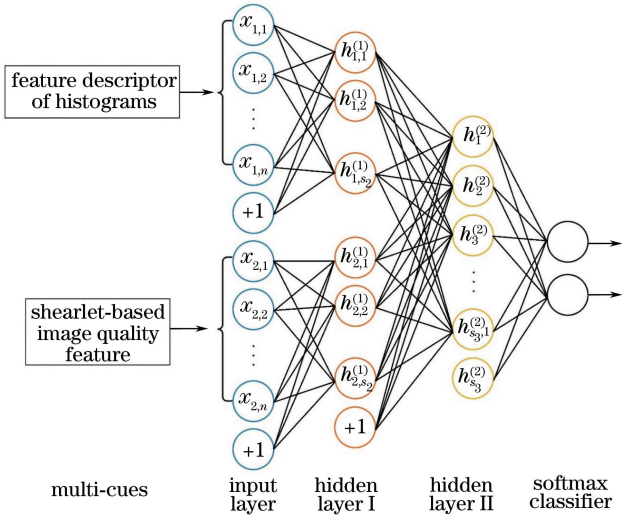

图 5 基于神经网络的 3D 打印面具攻击检测多特征融合图

Fig. 5 Flow chart of the multi-feature fusion based on 3D printing mask attack detection using neural networks

4 实验设计与分析

4.1 实验数据集

为了测试本文算法的性能，实验采用北京工业大学北京市数字化医疗 3D 打印工程技术研究中心采集的 3D 人脸数据库 Bjut-Form 3D Face Database (BFFD)和大型公开 FRGC v2.0 人脸数据库，并进行了一系列人脸识别和反攻击性能评估实验。

在 BFFD 数据库中，有一组来自 100 人，采用 TWOEYES-P 激光扫描仪捕捉的人脸模型，首次采用直径为 1.75 mm 的柔性材质热塑性聚氨酯弹性体橡胶(TPU)，通过基于熔融沉积成型(FDM)技术的弘瑞 Z500 设备进行 3D 打印并进行精细的后期制作，从而更加真实的模拟人脸的相关特征。数据库样本分为两类：a 类是 30 名测试对象的真实脸部样本数据；b 类是 30 位测试对象戴着本人或他人 3D 面具的被采集攻击脸部样本数据。由此可见，属于 b 的样品有两种情况。佩戴自己面具的人被标记为 A_A 样本，否则标记为 A_B 样本。在以下实验中，A_A 样本和 A_B 样本都计为 3D 面具攻击。图 6 示例了真实样本，类型 A_A 和类型 A_B 欺骗样本。FRGC v2.0 是目前世界上最大的三维人脸数据库，十分适合作为标准人脸识别库使用。

4.2 实验设置及评价标准

三维人脸识别系统的基本目标是正确识别身份，辅助目标是区分真假人脸。因此，在初始的系统中，仅能通过识别算法本身对于攻击面具特征的判断来区分真假人脸。首先通过 3D 关键点匹配实现人脸识别，试图证明不加入额外的攻击判别手段，人脸识别系统很容易受到 3D 面具的攻击并通过。本文加入 FRGC v2.0 数据库来大量增加测试集中真实人脸的数量来更加真实地模拟生活中实际场景，且通过数据集的交叉更能够体现算法的稳健性和泛化能力。首先定义特征相似度验证阈值 t_μ^i，用于在标准人脸识别场景中实现相应不同的识别率。其中，T_B 和 F_B 分别代表 BFFD 数据库中所有真实人脸扫描和攻击面具扫面(包括 A 和 B)，FRGC v2.0 数据库中用 T_F 表示。T_B^1 和 T_B^i 分别表示 T_B 在第一次扫描的注册集和每个测试人员后续采集的测试集，T_F 同理。实验整体可以分为两部分。第一部分首先进行两个数据库交叉下的标准人脸识别，使用真实人脸判断正确率(TAR)作为评估验证性能的指标，作为后续实验的参考标准，记录在不同错误接受率(FAR)下对应的 TAR。之后加入攻击面具，攻击面具与真实人脸的比例为 1∶20，使用攻击正确拒绝率(STRR)来评估性能。最后在完全相同的条件下测试本文方法的性能。第二部分只在 BFFD 数据库上对比了主流反 3D 打印面具攻击方法的性能，之后分别将原始特征融合，分数融合和提出的瓶颈特征融合进行比较，并分别评估反攻击性能。3D

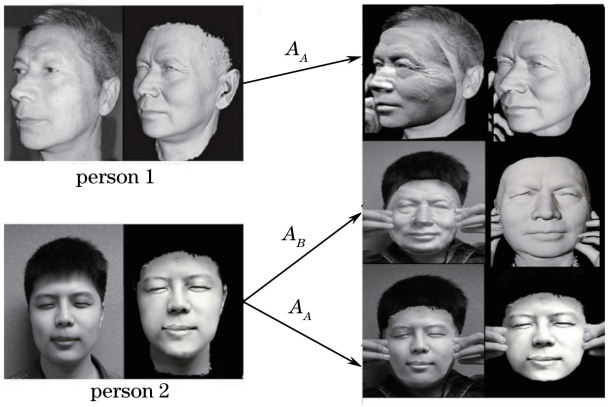

图 6 BFFD 数据库中的 2D 纹理图像和 3D 网格扫描。(a)真实人脸样本；(b) A_A，A_B 型欺骗样本

Fig. 6 2D texture image and corresponding 3D meshed scans in BFFD database. (a) Genuine faces sample; (b) A_A, A_B spoofing samples

面具欺骗攻击检测的本质上是一个二分类问题，即判断真实用户还是 3D 人脸面具。错误拒绝率(FRR)是指将真实人脸错判为 3D 面具的概率，FAR 是指将 3D 面具错判为真实人脸的概率。HTER 是 FAR 与 FRR 之和的 1/2，即

$$R_{\mathrm{HTER}}=(R_{\mathrm{FRR}}+R_{\mathrm{FAR}})/2, \tag{23}$$

HTER 的值越小，表明算法的性能越好。考虑到各个 3D 面具的攻击性能不同，实验结果取 1000 次交叉验证的平均 HTER 作为最终结果。表 1 列出了实验设置的详细信息。对基于剪切波变换的图像质量特征提取，灰度人脸图大小归一化为 256 pixel×256 pixel。分解层数、方向数和池化窗口大小分别被设置为 4、6 和 64 pixel×64 pixel，得到 384 维剪切波变换特征向量。对基于主曲率的图像几何特征提取，每个给定尺度和规范方向的关键点都对应一个由串级直方图组成的特征向量，得到一个 216 维的圆形邻域局部特征向量。

在训练样本一定的情况下，网络神经元数目越多，网络结构越复杂，也容易出现过拟合现象。因此，在达到相当精度的情况下，网络的层数越少越好。一般来说网络隐层神经元的数目与输入层和输出层数目之和相当比较合适，本文中原始输入特征为 600 维，因此输入层的数目为 600，输出层数目为 10。当特征输入栈式自编码器时，若第一隐层对数据的重构误差较大，则这种误差将会在网络中累积，网络的第一隐层神经元数目取得相对多时，能增加其对输入数据的拟合程度，有效提升网络整体的性能。其中，通过栈式自编码器构建多特征融合神经网络，分别用含有 250 个神经元的隐含层两个子网络进行训练。然后将融合的瓶颈特征输入到含有 200 个神经元的第二隐藏层进行训练，使用标记数据对整个分类网络进行微调。自编码器和 softmax 分类器的权重衰减参数是 3×10^{-5}，稀疏性参数为 0.1，稀疏性惩罚项的权重为 3。使用 SVM 进行性能对比，SVM 内核的参数使用网格搜索来设置。

表 1 实验设置参数

Table 1 Experimental setup parameters

Group	Approach	Gallery set	Probe set	
			Genuine face	Fraud mask
Base-1	3D keypoint matching	T_B^1	T_B^i	—
Base-2	3D keypoint matching	$T_F^1+T_B^1$	$T_F^i+T_B^i$	—
Base-3	3D keypoint matching	$T_F^1+T_B^1$	$T_F^i+T_B^i$	F_B
Base-4	Bottleneck feature fusion	$T_F^1+T_B^1$	$T_F^i+T_B^i$	F_B
Anti-1	Method in Ref.[15]	T_B^1	—	F_B
Anti-2	Method in Ref. [10]	T_B^1	—	F_B
Anti-3	Method in Ref. [16]	T_B^1	—	F_B
Anti-4	Raw feature fusion	T_B^1	—	F_B
Anti-5	Score fusion	T_B^1	—	F_B
Anti-6	Bottleneck feature fusion	T_B^1	—	F_B

本文所有算法均在机器软、硬件分别为 CPU Intel I7-7700K，内存 64 GB，操作系统 Ubuntu14.04 条件下，使用 Matlab R2016a 编程实现。

4.3 实验结果与分析

由表 2 可知，FAR、TAR 随不同的 t_{μ}^{i} 而变化。除了 $R_{\mathrm{FAR}}=0.01$ 时，只使用 BFFD 数据库的验证率在 92%以上。而在 Base-2 中，当添加 FRGC v2.0 扩展数据库时，即使 $R_{\mathrm{FAR}}=0.001$，验证率依然达到为 90.9%。本文比较这两种情况下的辨别能力，当 $R_{\mathrm{FAR}}=0.01$ 时，TAR 从 90.6%提升到 94.6%，说明在系统中加入更多样本，可以更准确地进行人脸检测识别。当 $R_{\mathrm{FAR}}=0.001$ 时，由于在这种情况下有 3800 个真实人脸扫描样本，$R_{\mathrm{FAR}}=0.1$ 意味着有 380 个样本被错误接受，因此第一行 Base-2 的结果保持空白。综合来看，基于 3D 人脸几何属性的特征描述算法具有较好的人脸匹配能力。

表 2 人脸验证和攻击实验结果

Table 2 Experimental results of verification and anti-spoofing performance evaluation %

R_{FAR}	Base-1 TAR	Base-2 TAR	Base-3 STRR	Base-4 STRR	Base-4 HTER
0.1	95.1	—	44.3	96.7	7.4
0.05	93.4	97.6	50.4	—	—
0.01	92.6	94.6	62.6	—	—
0.001	—	90.9	70.4	—	—

但是，在 Base-3 中，当加入了 BFFD 数据库中的攻击样本后，即使当 $R_{\mathrm{FAR}}=0.001$ 时，STRR 依然低于 67%，这意味着在正确识别率几乎无法保证的情况下，依然有 1/3 以上的攻击样本通过了系统验证，这显然是无法接受的。且人脸攻击面具样本的参与会降低系统的验证性能。而在 Base-4 中，采用本文方法，在识别前加入真假判断，实验结果显示，即使在交叉数据库下，STRR 依然达到了 96.7%，且不受识别系统阈值影响，HTER 小于 7.4%，将真实人脸误判为 3D 面具攻击的概率也比较理想。测试结果表明，本文提出的基于 meshSIFT 和剪切波纹理特征融合的的人脸识别框架可以完成反 3D 面具攻击的任务。

为验证算法的有效性，将本文算法与其他算法进行比较，在 BFFD 数据库上进行一系列对比实验。如表 3 所示，根据文献[10]、[15]、[16]中指定的实验配置，这里只使用 BFFD 数据库与其他方法进行比较。基于手动提取的多尺度 LBP 特征和 CNN 最后一个全连接层的深度信息特征在 BFFD 数据集上的 HRER 分别为 16.2%和 22.4%，效果一般。当判别制作精良、逼真程度很高的 3D 打印面具时，这些特征不具有足够的区分性以捕捉微小的纹理差异。从视频帧的冗余离散小波变换中按块方式提取 Haralick 纹理特征方法实现了 12.5%的 HTER，优于以上两种方法。更为直观的对比可从两个数据库的受试者工作特征（ROC）曲线图中看出，如图 7 所示。

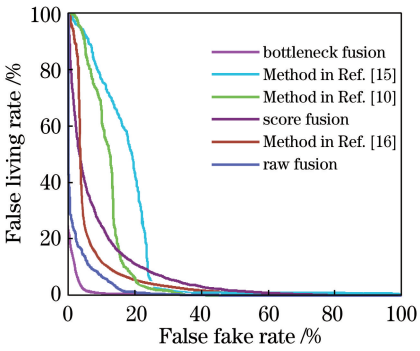

图 7 BFFD 数据库内部测试 ROC 曲线图

Fig. 7 ROC curves of intra tests for BFFD database

4.4 多特征融合算法性能对比

本文提出的多特征融合方法在神经网络中进行了瓶颈特征融合。为了评估瓶颈特征融合的有效性，对比了多种多特征融合方法对人脸反攻击的判断效果。原始特征融合是将提出的特征直接在特征层次上连接而不学习其瓶颈表示，将连接的原始特征输入 SVM 或神经网络进行分类。分数融合是把提出的两个活体特征传入两个独立的神经网络中用于人脸反欺骗攻击分类，然后使用逻辑回归将来自两个神经网络的分数融合。由于所提出的多特征集成神经网络具有两个隐藏层，因此实验统一采用两个隐藏层的自动编码器进行原始特征融合和分数融合以进行公平比较。结果如表 3 所示，与串联特征直接送入 SVM 中分类相比，基于自编码器的特征融合算法在特征融合过程中，不仅降低了特征维度，而且通过改变特征在空间上的分布，提取出了区分性更好的融合特征，所提出的瓶颈特征融合实现了 4.7%的 HTER，说明瓶颈表示的特征融合充分结合了低频子带中的几何结构信息和剪切波变换高频子带中的纹理信息，提高了融合特征的稀疏性和可

表 3 不同算法的反攻击欺骗性能对比

Table 3 Anti-attack spoofing performance comparison of different algorithms

Group	Approach	HTER
Anti-1	Method in Ref.[15]	22.4
Anti-2	Method in Ref. [10]	16.2
Anti-3	Method in Ref. [16]	12.5
Anti-4	Raw feature SVM	16.5
Anti-5	Raw feature fusion	8.8
Anti-6	Score fusion	15.3
Anti-7	Bottleneck feature fusion	4.7

区分度，有效弥补了单一特征在识别 3D 攻击面具上的劣势。与分数融合策略、原始特征融合方法相比，该方法得到了更好的结果，说明适当的特征融合策略对于人脸反攻击的多特征融合至关重要。

此外，基于瓶颈特征融合算法在时间复杂度也更具优势。直接将图像切片输入多层感知机神经网络，或者直接利用卷积神经网络提取特征，输入层所需神经元数目极其庞大，而融合算法中用到的网络输入层的数目为 600。因此，在训练过程中，本文算法需要训练的参数数目大大减少，两种算法的训练时间和测试时间如表 4 所示，本文算法在测试时间和训练时间上，速度提高了 15～20 倍。因此，在训练样本一定的情况下，基于自编码器的特征融合能有效的减少神经元的数目，简化网络结构，提高算法效率。

表 4 不同算法训练时间与测试时间对比

Table 4 Comparison between training time and testing time of different methods s

Approach	Training time	Test time
Multi-layer perceptron	2470.7	0.258
Bottleneck feature fusion	147.5	0.024

5 结 论

探索了用剪切波变换作为提取图像几何特征的滤波器和描述图像局部区域纹理变化的算子。在低频子带上利用尺度空间函数对特征点进行检测，然后对特征点进行定位及方向分配，利用特征点邻域旋转不变纹理特性对特征点进行描述，基于主曲率生成特征描述子，即 meshSIFT 特征向量，将其与高频子带上提取的纹理特征、输入栈式自编码器和 softmax 分类器进行融合分类。实验结果表明，本文方法可区分精细制作 3D 打印面具与真实人脸，且能应用于实际。针对时序数据的相关性，在下一阶段的工作中，拟将动态三维人脸图像变化特征作为网络的输入，以获得更优的反欺骗攻击效果。

参 考 文 献

[1] Zhang M G, Zhou D L, Pan Q, *et al*. Biometrics and its situation of present study[J]. Acta Biophysica Sinica, 2002, 18(2): 156-162.
张敏贵, 周德龙, 潘泉, 等. 生物特征识别及研究现状[J]. 生物物理学报, 2002, 18(2): 156-162.

[2] Nixon K A, Aimale V, Rowe R K. Spoof detection schemes[M]// Boston: Springer US, 2008: 403-423.

[3] Chetty G, Wagner M. Multi-level liveness verification for face-voice biometric authentication[C]// Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference, 2006: 1-6.

[4] Pan G, Sun L, Wu Z H, *et al*. Eyeblink-based anti-spoofing in face recognition from a generic webcamera [C] // IEEE 11th International Conference on Computer Vision, 2007: 1-8.

[5] Kollreider K, Fronthaler H, Bigun J. Evaluating liveness by face images and the structure tensor[C]// Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID′05), 2005: 75-80.

[6] de Marsico M, Nappi M, Riccio D, *et al*. Moving face spoofing detection via 3D projective invariants [C] // 2012 5th IAPR International Conference on Biometrics (ICB), 2012: 73-78.

[7] Kong Y P, Liu X, Xie X Q, *et al*. Face liveness detection method based on histogram of oriented gradient [J]. Laser & Optoelectronics Progress, 2018, 55(3): 031009.
孔月萍, 刘霞, 谢心谦, 等. 基于梯度方向直方图的人脸活体检测方法[J]. 激光与光电子学进展, 2018, 55(3): 031009.

[8] Määttä J, Hadid A, Pietikäinen M. Face spoofing detection from single images using micro-texture analysis [C] // International Joint Conference on Biometrics (IJCB), 2011: 1-7.

[9] Gu X J, Fu C Q, Gu X S. Facial vital sign based countermeasure against 3D mask attacks[J]. Journal of System Simulation, 2016, 28(2): 361-368.
谷小婧, 付传卿, 顾幸生. 基于面部生命特征的 3D 假面欺骗攻击检测方法[J]. 系统仿真学报, 2016, 28(2): 361-368.

[10] Wallace R, McLaren M, McCool C, *et al*. Inter-session variability modelling and joint factor analysis for face authentication [C] // International Joint Conference on Biometrics (IJCB), 2011: 1-8.

[11] Erdogmus N, Marcel S. Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect[C] // IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2013: 1-6.

[12] Li S Z, Chu R F, Liao S C, *et al*. Illumination invariant face recognition using near-infrared images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(4): 627-639.

[13] Zhang Z W, Yi D, Lei Z, *et al*. Face liveness detection by learning multispectral reflectance distributions[C] // Face and Gesture, 2011: 436-441.

[14] Kose N, Dugelay J L. Countermeasure for the protection of face recognition systems against mask attacks[C] // 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2013: 1-6.

[15] Menotti D, Chiachia G, Pinto A, *et al*. Deep representations for iris, face, and fingerprint spoofing detection [J]. IEEE Transactions on Information Forensics and Security, 2015, 10(4): 864-879.

[16] Agarwal A, Singh R, Vatsa M. Face anti-spoofing using Haralick features[C] // IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2016: 1-6.

[17] Guo K H, Labate D. Optimally sparse multidimensional representation using shearlets [J]. SIAM Journal on Mathematical Analysis, 2007, 39 (1): 298-318.

[18] Sun X, Morvan J M. Curvature measures, normal cycles and asymptotic cones [J]. Actes des Rencontres Du CIRM, 2013, 3(1): 3-10.

[19] Sun X, Morvan J M. Asymptotic cones of embedded singular spaces [J]. Geometry, Imaging and Computing, 2015, 2(1): 47-76.

[20] Cohen-Steiner D, Morvan J M. Restricted delaunay triangulations and normal cycle[C] // Proceedings of the nineteenth Conference on Computational Geometry - SCG ′03, the Nineteenth Conference, 2003.

[21] Lowe D G. Distinctiveimage features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.

[22] Hou Y M, Sui W X, Sun X X. SIFT feature dimension reduction method and its application in image retrieval[J]. Chinese Journal of Lasers, 2015, 42(s1): s108002.
侯一民, 隋文秀, 孙晓雪. SIFT 特征降维方法及其在图像检索中的应用[J]. 中国激光, 2015, 42(s1): s108002.

[23] Smeets D, Keustermans J, Vandermeulen D, *et al*. meshSIFT: local surface features for 3D face recognition under expression variations and partial data[J]. Computer Vision and Image Understanding, 2013, 117(2): 158-169.

[24] Xiong F G, Huo W, Han X, *et al*. Removal method of mismatching keypoints in 3D point cloud[J]. Acta Optica Sinica, 2018, 38(2): 0210003.
熊风光, 霍旺, 韩燮, 等. 三维点云中关键点误匹配剔除方法[J]. 光学学报, 2018, 38(2): 0210003.

[25] Tola E, Lepetit V, Fua P. DAISY: an efficient dense descriptor applied to wide-baseline stereo[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(5): 815-830.