DOI:10.7544/issn1000-1239.2018.20160417
55(3):643-650,2018

# 基于 LBP 和多层 DCT 的人脸活体检测算法

田 野[1] 项世军[1,2]
[1](暨南大学信息科学技术学院 广州 510632)
[2](信息安全国家重点实验室(中国科学院信息工程研究所) 北京 100093)
(shijun_xiang@qq.com)

# LBP and Multilayer DCT Based Anti-Spoofing Countermeasure in Face Liveness Detection

Tian Ye[1] and Xiang Shijun[1,2]
[1](*School of Information Science and Technology*, *Jinan University*, *Guangzhou* 510632)
[2](*State Key Laboratory of Information Security* (*Institute of Information Engineering*, *Chinese Academy of Sciences*), *Beijing* 100093)

**Abstract** As security problem has become the tightest bottleneck in the application of face recognition systems, rendering a face recognition system robust against spoof attacks is of great significance to be dealt with. In this paper, aimed at video-based facial spoof attacks, an innovative face antispoofing algorithm based on local binary patterns (LBP) and multilayer discrete cosine transform (DCT) is proposed. First, we extract face images from a target video at a fixed time interval. Second, the low-level descriptors, i.e., the LBP features are generated for each extracted face image. After that, we perform multilayer DCT on the low-level descriptors to obtain the high-level descriptors (LBP-MDCT features). To be more exact, in each layer, the DCT operation is implemented along the ordinate axis of the obtained low-level descriptors, namely the time axis of the entire target video. In the last stage, the high-level descriptors are fed into a support vector machine (SVM) classifier to determine whether the target video is a spoof attack or a valid access. In contrast to existing approaches, the outstanding experimental results attained by the proposed approach on two widely-used datasets (Replay-Attack dataset and CASIA-FASD dataset) demonstrat its performance superiority as well as its low complexity and high efficiency.

**Key words** face antispoofing; local binary patterns (LBP); multilayer discrete cosine transform (DCT); Replay-Attack database; CASIA-FASD database

**摘 要** 随着安全性成为制约人脸识别系统应用的最大瓶颈,提高人脸识别系统的抗欺骗攻击能力已成为亟待解决的问题.针对基于视频的人脸欺骗攻击,基于局部二值模式(local binary patterns, LBP)和多层离散余弦变换(discrete cosine transform, DCT)提出了一种新的人脸活体检测算法.其基本思想是首先从目标视频中每隔一定帧数提取1张人脸图像;其次对提取出的每张人脸图像进行LBP操作得到低级特征描述子(LBP算子);然后在LBP特征上进行多层DCT变换得到高级特征描述子(LBP-

**收稿日期**:2016-06-14;**修回日期**:2017-06-14
**基金项目**:国家自然科学基金面上项目(61772234,61272414);信息安全国家重点实验室开放基金项目(2016-MS-07)
This work was supported by the General Program of the National Natural Science Foundation of China (61772234, 61272414) and the Research Funding of State Key Laboratory of Information Security (2016-MS-07).

MDCT 算子);最后将得到的高级特征描述子送入支持向量机(support vector machine, SVM)中判断该视频是非法用户实施的人脸欺骗攻击还是合法用户的进入请求.通过在 Replay-Attack 和 CASIA-FASD 数据库上与现有的人脸活体检测算法做比较,验证了该算法能够取得优异的检测效果且十分简单、高效.

**关键词** 人脸活体检测;局部二值模式;多层离散余弦变换;Replay-Attack 数据库;CASIA-FASD 数据库

**中图法分类号** TP391

欺骗攻击是针对生物认证系统的一种攻击.它通过向传感器呈现合法生物特征的伪造版本,企图使生物认证系统将非法用户认证为合法用户,从而使该非法用户进入生物认证系统[1].目前,欺骗攻击可成功攻击多种生物形态已经成为不争的事实[2-6].在多种生物形态中,针对人脸的欺骗攻击尤其值得关注.一方面,无论是从经济角度还是社会角度,人脸都是最具影响力的生物特征之一[1];另一方面,由于其低成本、低技术的特征,和其他生物形态相比,针对人脸的欺骗攻击更易实施.攻击者可以轻易地在个人网站或社交网络上获得合法用户的面部特征.甚至,攻击者还可以近距离地拍摄合法用户的照片或视频.此外,随着人脸识别技术的发展,这项技术已在众多场合得到应用.大到机密场合的门禁系统,小到笔记本电脑的登录系统,甚至是移动终端的解锁系统,都能见到人脸识别技术的踪影[7].而门禁系统、登陆系统、解锁系统常常与网络系统相连,是进入网络系统的第 1 步,它们的安全与网络系统安全密切相关.如果攻击者成功攻击人脸识别系统,进入门禁系统、登陆系统或解锁系统,那么攻击者就打破了网络空间安全的第 1 道防线,极有可能进一步威胁、破坏网络空间安全.因此,应用人脸活体检测技术保障人脸识别系统的安全具有重要价值,它对保障网络空间的安全也有着十分重要的意义.

一般说来,人脸欺骗攻击可以分为 3 类:照片攻击、视频攻击和面具攻击.照片攻击是指攻击者将合法用户的照片打印在纸上或显示在电子设备的屏幕上,呈现给生物认证系统传感器的一种攻击.视频攻击也被称为重放攻击,因为该种攻击是通过重放合法用户的视频来实施的.面具攻击则是指攻击者戴上合法用户的 3D 面具,伪装成合法用户,企图进入人脸识别系统的攻击行为.

安全性已成为制约人脸识别系统应用的最大瓶颈,因此提高人脸识别系统的抗欺骗攻击能力已成为人脸认证中亟待解决的问题[8].人脸活体检测技术旨在辨别人脸的真伪,保障人脸识别系统稳定并安全地运行.具体地,它通过设置一道新的关卡,在系统进行人脸识别的同时对目标人脸进行是否为活体的判断.只有在人脸被判定为活体的情况下,识别结果才是真实有效的;否则,将其视为对人脸识别系统的一次非法攻击[7].近几年来,随着几个人脸欺骗攻击公用数据库的发布[9-12],涌现了许多人脸活体检测的方法.在不考虑面具攻击(超出本文研究范围)的前提下,现有的绝大多数人脸活体检测算法分为 2 类:基于照片的人脸活体检测算法和基于视频的人脸活体检测算法.文献[13]中提到,基于照片的人脸活体检测算法并不能直接用于检测视频攻击,尤其是高分辨率的视频攻击.一方面,视频中包含的动态信息使得生物样本更加逼真,从而增加了检测的难度.另一方面,和照片攻击相比,视频攻击包含的颜色降级、形状降级和纹理降级更少,更难被识别.另外,分辨率越高,在量化、离散过程中产生的伪迹也更少.因此,尽管迄今为止已有许多成熟的基于照片的人脸活体检测算法,针对视频攻击的人脸活体检测研究仍远未成熟.

由于其独特的能量集中特性,离散余弦变换(discrete cosine transform, DCT)在图像处理中取得了广泛的应用.然而,目前已有的应用均是利用 DCT 来提取每帧图片中的静态信息.迄今为止,DCT 从未被用来提取动态信息.基于此,本文创新性地在局部二值模式(local binary patterns, LBP)特征上实施多层 DCT 变换来表征视频中存在的时空信息.据我们所知,这是在 LBP 特征上进行 DCT 变换来提取面部动态信息,从而检测视频攻击的首次尝试.本文提出的方法不仅十分简单、省时,在公共数据库上出色的实验结果也验证了该算法的有效性.

本文首先简要介绍了已有的人脸活体检测算法;其次详细阐述了所提出的算法;然后对实验结果进行了分析和比较;最后对全文进行了总结并对将来的工作进行了展望.

## 1　相关算法回顾

根据所利用的信息类型,目前已有的人脸活体检测算法可分为两大类:利用空间信息的算法、既利用空间信息又利用时间信息的算法. 利用空间信息的人脸活体检测算法通常通过多种多样的图像处理方法对人脸样本中包含的面部纹理特征进行分析. 文献[14]是这类方法中最早的尝试之一. 在该文中,作者分析了单张人脸图像或人脸视频的傅里叶频谱. 随后,高斯差分(difference of Gaussian, DoG)被用来提取特定频域的信息[10,15]. 在文献[16]中,作者利用 Gabor 小波来加强纹理表征的效果,同时引入方向梯度直方图(histogram of oriented gradient, HOG)来描述局部形状特征. 作为一个对人脸十分有效的工具,局部二值模式及其多种变体也被许多算法[11,13,16-18]所采用.

因为同时利用了 2 种信息,人脸活体检测的第 2 类方法——利用时空信息的方法通常具有更好的性能. 但作为代价,这类方法常常需要更多的时间. 一个典型的代表是通过在 3 个正交平面上计算局部二值模式(local binary patterns from three orthogonal planes, LBP-TOP)将时空信息集中用 1 个多分辨率的纹理描述子表征[19]. 基于动态模式分解(dynamic mode decomposition, DMD)的特殊性质,文献[20]将 DMD,LBP 和支持向量机(support vector machine, SVM)结合在一起检测人脸欺骗攻击. Arashloo 等人[21]结合 2 种多尺度动态特征描述子 MBSIF-TOP(multiscale binarized statistical image features on three orthogonal planes)和 MLPQ-TOP(multi-scale local phase quantization representation on three orthogonal planes)提高了对抗欺骗攻击的检测器的鲁棒性. 在文献[22]中,作者通过提取视频中的时空信息构造了一种低级特征描述子. 除了上述这些基于纹理特征的检测算法,还可以从另一些角度融合时空信息,例如分析在二次成像过程中产生的噪声签名[23].

尽管上述利用时空信息的人脸活体检测算法取得了较好的效果,但他们也更复杂、更耗时. 针对这个缺点,本文基于 LBP 和多层 DCT 提出了一个既简单又省时的新算法. 一方面,无论是 LBP 还是 DCT 操作,实现所需的时间都很短;另一方面,本文提出的算法只需要使用视频中的少数帧,而非所有帧. 此外,根据 DCT 的能量集中特性,我们只需要提取 1 个或几个 DCT 分量来构造最终的高级特征描述子. 综上所述,本文提出的算法不仅易实现而且效率高. 在获得理想的特征描述子后,我们选择 SVM 作为后续分类器. 在严格遵循各个数据库测试协议的前提下,我们的实验结果证明所提出算法的性能超过了目前已有的所有算法. 事实上,该算法在 Replay-Attack 数据库的评估集和测试集上的半错误率均为 0,即实现了零差错的完美检测;在 CASIA-FASD 数据库测试集上的半错误率为 18.06%,低于所有其他算法且至少低 3.69%.

## 2　基于 LBP 和多层 DCT 的人脸活体检测算法

针对视频欺骗攻击,本文提出了一个新的检测算法. 该算法共包括 4 个主要步骤:人脸提取、低级特征描述子提取、高级特征描述子提取以及分类. 流程图如图 1 所示:

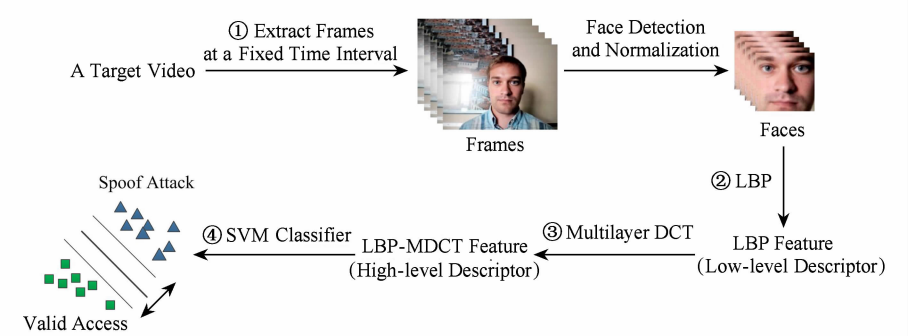


Fig. 1　Flow chart of the proposed algorithm

图 1　本文算法流程图

1) 我们从目标视频中每隔一定的帧数提取 1 张人脸图像;

2) 对提取出的每张人脸图像进行 LBP 操作得到低级特征描述子(LBP 算子);

3) 在 LBP 特征上进行多层 DCT 变换得到高级特征描述子(LBP-MDCT 算子);

4) 将得到的高级特征描述子送入 SVM 分类器中判断该视频究竟是非法用户的欺骗攻击还是合法用户的进入请求.

### 2.1 人脸提取

为了提高检测的效率,我们只使用输入视频中的少数帧进行检测. 具体来说,对每个输入视频,我们每隔一定的时间间隔 $I$ 提取 1 帧图像. 例如若 $I=1$,则提取第 1 帧图像、第 3 帧图像、第 5 帧图像,以此类推;若 $I=2$,则提取第 1 帧图像、第 4 帧图像、第 7 帧图像,以此类推. 在跳帧提取出所有需要的图像后,我们利用 Viola-Jones 算法[24] 抠出每张图像中的脸部区域,并将所有的脸部区域统一为 $64\times64$ 大小. 假设需检测的目标视频共由 $M$ 帧图像组成,则通过人脸提取步骤我们得到 $\left\lceil \frac{M}{I+1} \right\rceil$ 帧统一大小的人脸图像.

### 2.2 低级特征描述子提取

局部二值模式(local binary patterns, LBP)是一种简单但十分有效的灰度不变纹理表征. 它根据每个像素和其相邻像素灰度值的比较结果获取空间信息. 自 1994 年提出以来,许多学者对 LBP 进行了研究并提出了许多 LBP 变体. 在本文中,我们选用最多只包含 2 次 0 到 1 或 1 到 0 跳变的均匀模式 LBP 算子. 通常,我们用 $\mathrm{LBP}^{u2}_{P,R}$ 表征均匀模式 LBP 算子,其中 $P$ 和 $R$ 分别代表所使用的相邻像素个数和领域半径大小. 假设从目标视频中共提取出 $N=\left\lceil \frac{M}{I+1} \right\rceil$ 帧人脸图像,首先我们在每帧人脸图像上进行 $\mathrm{LBP}^{u2}_{8,1}$ 操作,得到 1 个 59 维的 LBP 特征向量[11];然后我们将所有帧的 LBP 特征向量并行组合在一起,得到 1 个 $N\times59$ 的 LBP 特征矩阵,即低级特征描述子.

### 2.3 高级特征描述子提取

在这个环节中,我们对得到的低级特征描述子进行 DCT 变换,从而提取出视频中的时间信息. 图 2 详细展示了高级特征描述子的构造过程. 在获得 $N\times59$ 的 LBP 特征矩阵后,我们沿着该矩阵的纵轴,即录制整个视频的时间轴,进行一维 DCT 变换,得到许多 DCT 分量. 给定输入信号 $f(n)$,其一维 DCT 变换为

$$F(k)=\alpha(k)\sum_{n=0}^{N-1}f(n)\cos\left[\frac{(2n+1)k\pi}{2N}\right], \quad (1)$$

其中,$0\leqslant k\leqslant N-1$. 根据 DCT 的能量集中特性,变换后绝大多数能量都集中在直流分量(direct component, DC)中. 因此,没有必要利用所有的 DCT 分量来构造高级特征描述子,只选用 $C$ 个 DCT 分量即可. 具体来说,若 $C=1$,只选用直流分量;若 $C=2$,则选用直流分量和第 1 个交流分量(alternating component, AC). 这样,在 DCT 变换后,我们得到 1 个 $59\times C$ 的 LBP-DCT 特征矩阵. 需要特别说明的是,通过舍弃绝大部分的 DCT 分量,我们达到了降维的效果,而降维不仅降低了计算复杂度,同时也提高了检测的效率.

为了获得更好的性能,本文采用 3 层 DCT 变换. 在第 1 层,对 LBP 特征矩阵的所有列进行 DCT 变换,得到一个 $59\times C$ 的 LBP-DCT 特征矩阵. 在第 2 层,将 LBP 特征矩阵均分为 2 部分:第 1 部分由前 $\left\lfloor \frac{N}{2} \right\rfloor$ 帧人脸图像组成;第 2 部分由随后的 $\left\lfloor \frac{N}{2} \right\rfloor$ 帧人脸图像组成. 分别沿着纵轴对每个部分进行一维 DCT 变换,得到 2 个 LBP-DCT 特征矩阵. 以此类推,在第 3 层将 LBP 特征矩阵均分为 4 部分,得到 4 个

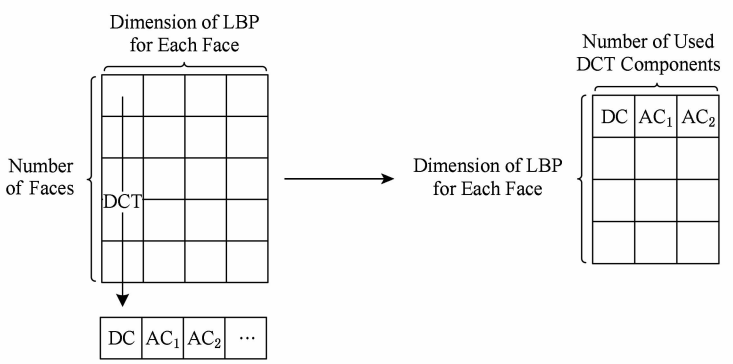


Fig. 2 Demonstration of high-level descriptor extraction

图 2 高级特征描述子构造过程

LBP-DCT 特征矩阵. 这样,在 3 层 DCT 变换后,我们一共得到 7 个 $59\times C$ 的 LBP-DCT 特征矩阵. 最后,将这 7 个矩阵连结成 1 个矩阵,得到高级特征描述子.

### 2.4　分　类

本文检测算法的最后 1 个关键步骤是通过一个高辨别力的机器学习算法找到最优分类模型,从而对目标视频进行判断:究竟该视频是非法用户的欺骗攻击,还是合法用户的进入请求. 本文选用包含径向基核函数(radial basis function, RBF)的支持向量机(support vector machine, SVM)[25] 作为分类器. 该分类器不仅具有很高的分类准确率,而且被广泛应用于人脸识别等研究课题. 将上一步得到的高级特征描述子送入 SVM,即可根据 SVM 的输出数据完成人脸活体检测. 输出数据的评价指标在 3.2 节中详细说明.

## 3　实验结果与分析

在分析实验结果之前,首先介绍本文使用的公共数据库及实验严格遵循的测试协议.

### 3.1　数据库

本文在 2 个得到广泛认可的公共数据库上进行实验.

Replay-Attack 数据库[11]:该数据库由合法用户的视频进入请求和非法用户的视频欺骗攻击组成,每个视频时长约 10 s. 在 3 种不同的情景和 2 种不同的照明条件下,该数据库对 50 个对象录制了共计 1 200 段视频,其中包括 200 段真实人脸视频及 1 000 段伪造人脸视频.

CASIA-FASD 数据库[10]:该视频库由来自 50 个对象的 600 段视频组成,其中包括 150 段合法请求及 450 段欺骗攻击. 不同于 Replay-Attack 数据库的是,CASIA-FASD 数据库的视频欺骗攻击涉及到 3 种成像质量:低质量(由 1 个分辨率为 $640\times 480$ 的旧 USB 摄像头采集)、中等质量(由 1 个分辨率为 $480\times 640$ 的新 USB 摄像头采集)、高质量(由 1 个最大分辨率为 $1920\times 1080$ 的 Sony NEX-5 摄像头采集). 此外,该数据库包含的伪造人脸视频有 3 种类型:弯曲照片攻击、裁剪照片攻击和视频回放攻击.

### 3.2　测试协议

为了公平有效地评价各个算法的性能优劣,本文选择被广泛使用的半错误率(half total error rate, *HTER*)作为性能评价指标. 如式(2)所示,半错误率指的是错误接受率(false acceptance rate, *FAR*)和错误拒绝率(false rejection rate, *FRR*)总和的一半:

$$HTER=\frac{FAR+FRR}{2}. \tag{2}$$

由于 *FAR* 和 *FRR* 都取决于阈值 $\tau$,增大其中一个会减小另一个,因此 *HTER* 值通常在接收者操作特征曲线(receiver operating characteristic curve, ROC)的一个特征点取得. 在该特征点上,*FAR* 和 *FFR* 相等,即达到等错误率(equal error rate, *ERR*)状态. 根据上述定义可知,*HTER* 值越小,错误率越小,算法的性能就越好.

测试协议Ⅰ:在该测试协议下,使用 Replay-Attack 数据库. Replay-Attack 数据库由 3 个子集组成:训练集(包含 360 段视频)、评估集(包含 360 段视频)和测试集(包含 480 段视频). 其中,训练集用来训练 SVM 分类器;评估集用来选择阈值 $\tau$;测试集则用来报告最终的 *HTER* 值.

测试协议Ⅱ:在该测试协议下,使用 CASIA-FASD 数据库. CASIA-FASD 数据库由训练集(包含 240 段视频)和测试集(包含 360 段视频)组成. 训练集用来训练 SVM 分类器,测试集用来得到最终的 *HTER* 值.

### 3.3　Replay-Attack 数据库的实验结果

本文算法在 Replay-Attack 数据库上的实验结果如表 1 所示. 需要说明的是,对每个 $C$ 值,$I$ 都有 1 个上限. 例如若 $C=1$,即只选用 DCT 变换后的直流分量,则 3 层 DCT 至少需要 4 帧人脸图像. Replay-Attack 数据库包含的所有视频最少有 221 帧,因此,当 $C=1$ 时,$I$ 的上限为 72. 表 1 展示了部分参数设置下的实验结果.

令人惊喜的是,无论 $I$ 和 $C$ 的取值为多少,所有参数组合下的 *HTER* 值均为 0,即所有参数组合都可实现完美分类. 我们知道,$I$ 越大,所提取的人脸图像就越少,检测的速度就越快,算法的复杂度也越低. 同样道理,$C$ 越小,使用的 DCT 分量越少,算法的效率就越高,算法也越简单. 出于这 2 点考虑,$I=72\&C=1$ 可使算法具有最高的效率和最低的复杂度,因此,$I=72\&C=1$ 是 Replay-Attack 数据库的最优参数.

**Table 1 *HTER* of the Proposed Algorithm on Replay-Attack Dataset and CASIA-FASD Dataset**

**表 1 本文算法在 Replay-Attack 和 CASIA-FASD 数据库上的 *HTER* 值**

| *C* | *I* | *HTER*/% | | |
|---|---|---|---|---|
| | | Replay-Attack | | CASIA-FASD |
| | | Dev | Test | Test |
| 1 | 1 | 0.00 | 0.00 | 20.00 |
| | 2 | 0.00 | 0.00 | 19.07 |
| | 3 | 0.00 | 0.00 | 20.00 |
| | 4 | 0.00 | 0.00 | 18.89 |
| 2 | 1 | 0.00 | 0.00 | 20.00 |
| | 2 | 0.00 | 0.00 | 19.07 |
| | 3 | 0.00 | 0.00 | 20.00 |
| | 4 | 0.00 | 0.00 | 19.26 |
| 3 | 1 | 0.00 | 0.00 | 18.89 |
| | 2 | 0.00 | 0.00 | 19.26 |
| | 3 | 0.00 | 0.00 | 19.07 |
| | 4 | 0.00 | 0.00 | 18.43 |
| 4 | 1 | 0.00 | 0.00 | 18.89 |
| | 2 | 0.00 | 0.00 | 20.00 |
| | 3 | 0.00 | 0.00 | 18.25 |
| | 4 | 0.00 | 0.00 | 18.06 |
| 5 | 1 | 0.00 | 0.00 | 18.89 |
| | 2 | 0.00 | 0.00 | 19.18 |
| | 3 | 0.00 | 0.00 | 19.18 |
| | 4 | 0.00 | 0.00 | 18.89 |

*C*: Number of used DCT components; *I*: Interval for extracting frames; Dev: On development set; Test: On test set.

### 3.4 CASIA-FASD 数据库的实验结果

表 1 同样展示了本文算法在 CASIA-FASD 数据库上的实验结果. 和 Replay-Attack 数据库一样, 受视频所含最少帧数影响, 给定 1 个 *C* 值, *I* 同样存在一个上限. 为节省空间, 表 1 只列出了部分参数设置下的 *HTER* 值.

从表 1 我们可以看到, CASIA-FASD 数据库上的 *HTER* 值在 19.00%左右波动; 当 $I=4\&C=4$ 时, *HTER* 值最小, 为 18.06%. 因此, $I=4\&C=4$ 是 CASIA-FASD 数据库的最优参数. 至于性能波动和参数选取的关系, 我们将在后续工作中对其进行更深入的研究. 不同于在 Replay-Attack 数据库上实现的完美分类, 本文算法在 CASIA-FASD 数据库上未能实现零差错分类. 这是因为 CASIA-FASD 数据库比 Replay-Attack 数据库难度更大, 更具挑战性. 例如, CASIA-FASD 数据库引入了裁剪照片攻击. 在实施这种攻击时, 攻击者将人脸照片的眼睛区域剪掉, 用自己的眼睛实现眨眼动作, 更加逼真, 检测的难度也更大. 此外, CASIA-FASD 数据库包含的攻击类型比 Replay-Attack 数据库更丰富. CASIA-FASD 数据库不仅包含 3 种类型的攻击(弯曲照片攻击、裁剪照片攻击和视频回放攻击), 这些攻击还是由 3 种不同的设备(低分辨率摄像头、中等分辨率摄像头和高分辨率摄像头)录制而成.

### 3.5 与现有算法的比较

我们将本文算法与文献[11,20,26]中的算法进行了比较, 比较结果如表 2 所示. 正如我们在 3.4 节中阐述的那样, CASIA-FASD 数据库比 Replay-Attack 数据库难度更大、更具挑战性. 所有算法在 Replay-Attack 数据库上的性能都明显优于在 CASIA-FASD 数据库上的性能.

**Table 2 Comparison of *HTER* on Test Sets for the Proposed Algorithm with State-of-the-art Algorithms**

**表 2 本文算法和现有算法的 *HTER* 值比较**

| Algorithm | | *HTER*/% | |
|---|---|---|---|
| | | Replay-Attack | CASIA-FASD |
| Ref[11] | LBP+LDA | 13.87 | |
| | LBP+SVM | 18.17 | |
| Ref[26] | Motion Correlation | 11.79 | 30.33 |
| | LBP | 15.45 | 23.19 |
| | LBP-TOP | 8.51 | 23.75 |
| Ref[20] | DMD+LBP+SVM$^{E}$ | 0.00 | |
| | DMD+LBP+SVM$^{F}$ | 3.75 | 21.75 |
| | DMD+SVM$^{F}$ | 7.50 | 29.50 |
| | PCA+SVM$^{F}$ | 21.50 | 33.50 |
| | PCA+LBP+SVM$^{F}$ | 17.11 | 24.50 |
| | PCA+LBP+SVM$^{E}$ | 20.50 | |
| Proposed Algorithm | LBP+DCT+SVM | 0.00 | 18.06 |

E: On entire frames; F: On face regions.

从表 2 可以看到, 本文算法的 *HTER* 值低于文献[11,20,26]中的算法, 也就是说, 本文算法性能超过了文献[11,20,26]中的算法. 对 Replay-Attack 数据库, 我们的 *HTER* 值为 0, 实现了零差错的完美检测; 对 CASIA-FASD 数据库, 我们的 *HTER* 值为 18.06%, 低于文献[11,20,26]中的算法且至少低 3.69%. Chingovska 等人[11]仅仅应用 LBP 对抗人脸欺骗攻击, 检测效果在 15.00%左右. 在 Pereira

等人[26]提出的算法中,基于 LBP-TOP 的算法性能最好,分别在 Replay-Attack 和 CASIA-FASD 数据库上取得了 8.51%和 23.75%的 *HTER* 值.尽管 Tirunagari 等人[20]提出的 DMD+LBP+SVM$^{E}$ 算法在 Replay-Attack 数据库上也实现了完美检测(*HTER*=0),但该算法需要使用一段视频中的 240 帧,而本文算法只需要 4 帧,所需图像帧数大大减小.我们知道,人脸活体检测是针对实际应用的研究,所需帧数越少,检测时间越短,算法效率越高,算法性能也就越好.因此,本文算法比 DMD+LBP+SVM$^{E}$ 算法性能更优.更重要的是,DMD+LBP+SVM$^{E}$ 算法必须使用一帧图像的全部区域,一旦仅使用人脸区域(DMD+LBP+SVM$^{F}$ 算法),其 *HTER* 值增加至 3.75%,不再是零差错检测.这是因为 Replay-Attack 数据库中不同类型欺骗攻击的背景内容存在差异,而这种差异有利于分类器分辨真假人脸视频.然而,在现实生活中,对不可随身携带的相对固定的人脸识别系统而言,如门禁系统,视频背景内容的差异性将不复存在.从这一点上考虑,DMD+LBP+SVM$^{E}$ 算法并不能在所有人脸识别系统上获得完美的检测结果,而本文算法仅使用人脸区域,适用于所有类型的人脸识别系统.综上所述,针对视频欺骗攻击,本文提出的算法在目前已有的人脸活体检测算法中具有最出色的性能.

除了出色的性能,本文算法还具备低复杂度、高效率的优点.一方面,我们只需要使用视频中的少数帧而非所有帧;另一方面,LBP 和 DCT 操作均只需很短的时间来完成.对 Replay-Attack 或 CASIA-FASD 数据库中的视频而言,在每帧图像上进行 LBP 操作只需要 0.12 s.在获得 LBP 特征矩阵后,无论参数取值为多少,计算 LBP-MDCT 特征矩阵都只需要 0.02 s.此外,根据 DCT 的能量集中特性,我们只需要 1 个或几个 DCT 分量来构造高级特征描述子.舍弃绝大部分的 DCT 分量意味着降维,而降维不仅降低了计算复杂度,同时也提高了算法的效率.基于上述 3 个原因,本文算法不仅简单、易实现,而且实时性好、效率高.兼顾优异性能、低复杂度和高效率,本文算法对人脸活体检测在实际生活中的应用有着十分重要的意义.

## 4　结论与展望

针对基于视频的人脸欺骗攻击,本文利用 LBP 和多层 DCT 提出了一种新的人脸活体检测算法.为了有效提取静态空间信息,我们对选中的人脸图像进行均匀模式 LBP 操作,得到低级特征描述子(LBP 算子).在此基础上,为了提取动态时间信息,我们沿着 LBP 算子的纵轴,即录制整个视频的时间轴,进行 3 层的 DCT 变换.这样,最终得到的高级特征描述子(LBP-MDCT 算子)既包含了静态图像的空间信息,又包含了帧与帧之间的动态时间信息.优异的实验结果验证了多层 DCT 确实能够有效捕捉面部动态信息,对正确判断真假人脸视频起到了重要的作用.

在严格遵循各个数据库测试协议的前提下,我们在 2 个广泛应用的公共数据库上进行了大量的实验,验证了本文算法的有效性,表明了本文算法相比于现有算法的性能优越性.在 Replay-Attack 数据库上,本文算法的 *HTER*=0,实现了零差错的完美检测;在 CASIA-FASD 数据库上,本文算法的 *HTER*=18.06%,低于其他所有算法且至少低 3.69%.我们将出色的性能归因于 3 个方面:1)LBP 能够有效提取出每帧图像的静态纹理信息;2)多层 DCT 变换能够有效捕捉面部动态信息;3)在 LBP 算子上进行多层 DCT 变换来同时表征时空信息的创新性的结合方式.值得注意的是,本文算法之所以能取得如此优异的性能,最主要的原因不是 LBP 或 DCT 单独的能力,而是将两者结合在一起的独特方式.在性能优异的同时,本文算法简单易实现,而且高效省时.从每帧人脸图像中提取 LBP 向量只需要 0.12 s;根据低级特征描述子得到高级特征描述子只需要 0.02 s.兼顾卓越的性能、低复杂度和高效率,本文算法具有很好的实际应用前景.

将来工作的首要方向是对本文算法在 CASIA-FASD 数据库上参数选取和性能波动的关系进行进一步的研究并进行跨数据库实验.另一个研究方向是选用其他类型的 LBP 算子,比较各自的性能.当然,用其他图像处理工具代替 LBP 或 DCT 来同时表征时空信息并比较性能优劣也是将来工作的内容之一.

## 参　考　文　献

[1] Galbally J, Marcel S, Fierrez J. Biometric antispoofing methods: A survey in face recognition [J]. IEEE Access, 2014, 2: 1530-1552

[2] Anjos A, Marcel S. Counter-measures to photo attacks in face recognition: A public database and a baseline [C] //Proc of 2011 IEEE Int Joint Conf on Biometrics. Piscataway, NJ: IEEE, 2011: 1-7

[3] Galbally J, Fierrez J, Alonso-Fernandez F, et al. Evaluation of direct attacks to fingerprint verification systems [J]. Telecommunication Systems, 2011, 47(3/4): 243-254

[4] Mjaaland B B, Bours P, Gligoroski P. Walk the walk: Attacking gait biometrics by imitation [G] //LNCS 6531: Proc of the 13th Int Conf on Information Security. Berlin: Springer, 2010: 361-380

[5] Akhtar Z, Fumera G, Marcialis G L, et al. Evaluation of serial and parallel multibiometric systems under spoofing attacks [C] //Proc of the 5th IEEE Int Conf on Biometrics: Theory, Applications and Systems. Piscataway, NJ: IEEE, 2012: 283-288

[6] Tome P, Vanoni M, Marcel S. On the vulnerability of finger vein recognition to spoofing [C] //Proc of Int Conf of the Biometrics Special Interest Group. Piscataway, NJ: IEEE, 2014: 1-10

[7] Yang Jianwei. Study on face antispoofing methods from the perspective of face recognition [D]. Beijing: Beijing University of Posts and Telecommunications, 2014 (in Chinese)
(杨健伟. 面向人脸识别的人脸活体检测方法研究[D]. 北京: 北京邮电大学, 2014)

[8] Sun Lin. Research on anti-spoofing in face recognition [D]. Hangzhou: Zhejiang University, 2010 (in Chinese)
(孙霖. 人脸识别中的活体检测技术研究[D]. 杭州: 浙江大学, 2010)

[9] Tan Xiaoyang, Li Yi, Liu Jun, et al. Face liveness detection from a single image with sparse low rank bilinear discriminative model [G] //LNCS 6316: Proc of the 11th European Conf on Computer Vision. Berlin: Springer, 2010: 504-517

[10] Zhang Zhiwei, Yan Junjie, Liu Sifei, et al. A face antispoofing database with diverse attacks [C] //Proc of the 5th IAPR Int Conf on Biometrics. Piscataway, NJ: IEEE, 2012: 26-31

[11] Chingovska I, Anjos A, Marcel S. On the effectiveness of local binary patterns in face anti-spoofing [C] //Proc of Int Conf of Biometrics Special Interest Group. Piscataway, NJ: IEEE, 2012: 1-7

[12] Erdogmus N, Marcel S. Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect [C] //Proc of the 6th IEEE Int Conf on Biometrics: Theory, Applications and Systems. Piscataway, NJ: IEEE, 2013: 1-6

[13] Pinto A, Schwartz W R, Pedrini H, et al. Using visual rhythms for detecting video-based facial spoof attacks [J]. IEEE Trans on Information Forensics and Security, 2015, 10(5): 1025-1038

[14] Li Jiangwei, Wang Yunhong, Tan Tieniu, et al. Live face detection based on the analysis of Fourier spectra [G] //SPIE 5404: Biometric Technology for Human Identification. Bellingham, WA: SPIE, 2004: 296-303

[15] Peixoto B, Michelassi C, Rocha A. Face liveness detection under bad illumination conditions [C] //Proc of the 18th IEEE Int Conf on Image Processing. Piscataway, NJ: IEEE, 2011: 3557-3560

[16] Maatta J, Hadid A, Pietikainen M. Face spoofing detection from single images using texture and local shape analysis [J]. IET Biometrics, 2012, 1(1): 3-10

[17] Kose N, Dugelay J L. Classification of captured and recaptured images to detect photograph spoofing [C] //Proc of 2012 Int Conf on Informatics, Electronics & Vision. Piscataway, NJ: IEEE, 2012: 1027-1032

[18] Maatta J, Hadid A, Pietikainen M. Face spoofing detection from single images using micro-texture analysis [C] //Proc of 2011 Int Joint Conf on Biometrics. Piscataway, NJ: IEEE, 2011: 1-7

[19] de Freitas Pereira T, Anjos A, De Martino J M, et al. LBP-TOP based countermeasure against face spoofing attacks [G] //LNCS 7728: Proc of ACCV 2012 Int Workshops. Berlin: Springer, 2013: 121-132

[20] Tirunagari S, Poh N, Windridge D, et al. Detection of face spoofing using visual dynamics [J]. IEEE Trans on Information Forensics and Security, 2015, 10(4): 762-777

[21] Arashloo S R, Kittler J, Christmas W. Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features [J]. IEEE Trans on Information Forensics and Security, 2015, 10(11): 2396-2407

[22] Pinto A, Pedrini H, Schwartz W R, et al. Face spoofing detection through visual codebooks of spectral temporal cubes [J]. IEEE Trans on Image Processing, 2015, 24(12): 4726-4740

[23] Pinto A d S, Pedrini H, Schwartz W, et al. Video-based face spoofing detection through visual rhythm analysis [C] //Proc of the 25th SIBGRAPI Conf on Graphics, Patterns and Images. Piscataway, NJ: IEEE, 2012: 221-228

[24] Viola P, Jones M J. Robust real-time face detection [J]. International Journal of Computer Vision, 2004, 57(2): 137-154

[25] Cortes C, Vapnik V. Support-vector networks [J]. Machine Learning, 1995, 20(3): 273-297

[26] de Freitas Pereira T, Komulainen J, Anjos A, et al. Face liveness detection using dynamic texture [J]. EURASIP Journal on Image and Video Processing, 2014, 2014(1): Article No. 2

**Tian Ye**, born in 1992. Master of Jinan University. Her main research interests include face recognition and reversible data hiding.

**Xiang Shijun**, born in 1974. Professor of Jinan University. His main research interests include information hiding and multimedia information security.