

基于曲波纹理分析和 SVM-KNN 分类的假指纹检测算法

张永良 刘超凡 肖刚 方珊珊

(浙江工业大学计算机科学与技术学院 杭州 310023)

摘要 假指纹攻击作为破解指纹识别的一种简单实用的方式,被某些不法分子非法使用。目前假指纹检测的主流方法是纹理分析,但是单纯的纹理分析不包含对因假指纹材质与人体皮肤有异而产生的噪声分析。提出一种利用曲波系数特征及曲波重构图像纹理特征进行 SVM-KNN 分类的假指纹检测算法。先对指纹图像进行曲波变换,提取各尺度各方向域的系数特征,重构指纹图像并提取一阶统计量、灰度共生矩阵(GLCM)和马尔科夫随机场(MRF)等纹理特征与系数特征组成特征向量,然后通过 SVM 进行训练,引入 SVM-KNN 分类对假指纹进行检测。在第二届全球假指纹检测竞赛(LivDet2011)官方数据库上的测试结果表明,该算法对假指纹检测有很好的效果。

关键词 曲波变换,灰度共生矩阵,马尔科夫随机场,SVM-KNN

中图分类号 TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.12.065

Fake Fingerprint Detection Algorithm Based on Curvelet Texture Analysis and SVM-KNN Classification

ZHANG Yong-liang LIU Chao-fan XIAO Gang FANG Shan-shan

(College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract Fake fingerprint attack, as a kind of simple and practical way of cracking fingerprint identification, is used by some outlaws. The current mainstream method of fake fingerprint detection is texture analysis, but the original texture analysis doesn't include coarseness analysis caused by the difference between the materials of fake fingerprint and real fingerprint. In this paper, a novel method was proposed based on the curvelet transform and image texture feature with support vector machine and K-nearest neighbor(SVM-KNN) classification. Firstly, curvelet coefficient features with different scales and directions are extracted. Secondly, texture features are extracted from first order statistics, gray level co-occurrence matrix(GLCM) and Markov random field(MRF) in curvelet reconstructed image, and then fingerprint images are trained to obtain the classification criterion by SVM. Lastly, SVM-KNN classification is used for fake fingerprint detection. The experimental results in the databases of the Liveness Detection Competition 2011 (LivDet2011) show that the proposed method is effective and superior.

Keywords Curvelet transform, Gray level co-occurrence matrix, Markov random field, SVM-KNN

随着经济和社会的发展,生物特征识别技术,特别是指纹技术得到了越来越广泛的应用^[1]。但一些廉价材料制作的假指纹的出现,及其制作技术的提高,为指纹特征识别带来了安全隐患。有研究表明,可以根据用户留下的指纹,利用硅胶、明胶、橡皮泥等廉价材料制成假指纹膜骗过传统的指纹采集仪^[2]。用含有授权用户身份信息的假指纹,试图创造未经授权的访问,被称为欺骗^[3]。人们在希望隐瞒真实身份或者获得已授权用户的特权时,经常采用这类针对采集仪的攻击方法。为了抵抗假指纹对指纹识别系统的攻击,假指纹检测技术应运而生。假指纹检测技术是判断指纹样本是否来自活体指纹的方法,检测方法可分为两类:第一类是通过在指纹采集仪上加入额外的硬件设备来检测得到手指温度、皮肤导电性、脉搏血氧等特性,但是其会增加采集仪的成本,称为基于硬件的假指纹检测方法^[4];第二类方法是对指纹样本图片做额外的处理,比如灰度级分析、纹理特征分析等,以检测指纹图像

的活性信息,这类方法称为基于软件的方法。基于软件的方法包括静态和动态特性的检测方法,静态特性从一张或多张图像中提取得到(手指一次或多次放到采集仪上进行采集),动态特性从多幅指纹图像帧中提取得到(手指放在采集仪上一段时间,获取一段图像序列进行分析)。基于软件的方法成本低,对用户的侵入性较小,且能用于现有的指纹采集仪^[5]。因此它对基于软件的假指纹检测方法的研究具有重大的实用价值和推广意义。

由于指纹脊线谷线交替,使指尖表面呈现一定模式的粗糙度, Moon 提出了一种分析高分辨率指尖图像表面粗糙程度的假指纹检测方法^[6]。这种方法只对高分辨率采集仪(1000dpi及以上)有效,对于市面上流行的普通的500dpi采集仪适用性不高。汗孔的存在和汗液现象,使得真指纹脊线灰度级变化明显,而人造材料如明胶和硅胶制作的假指纹,脊线灰度级变化缓慢。基于上述思想, Nikam 和 Agrwal 提出

到稿日期:2014-01-21 返修日期:2014-03-15 本文受浙江省科技厅国际科技合作专项项目(2012C24009)资助。

张永良(1977-),男,博士,副教授,主要研究方向为生物特征识别和视频侦查, E-mail: titanzhang@zjut.edu.cn(通信作者);刘超凡(1988-),男,硕士生,主要研究方向为图形图像处理、指纹识别;肖刚(1965-),教授,主要研究方向为图形图像处理、智能信息系统;方珊珊(1989-),硕士生,主要研究方向为指纹识别及防伪技术。

了一种基于纹理的方法,其用指纹像素相关的灰度级分析指纹图像的活性^[7]。这种方法在中心点(core point)准确定位时表现出了良好的性能,然而现有的中心点检测算法在处理低质量图像和太干或太湿的指纹图像时性能不佳。在文献[8]中,Abhyankar 和 Schuckers 提出了一种基于多分辨率纹理分析和纹路频率分析的方法,其用不同的纹理特性量化物理结构变化时的灰度级分布变化。但此方法在实际应用中有一定的局限性,因为局部脊线频率的计算会受天气影响,也与不同的皮肤状况有关。

目前,纹理分析是假指纹检测的主流方法,在其之前一般都会先对指纹图像进行滤波去噪,常用的纹理特征包含一阶统计量、Local Binary Pattern(LBP)等。如文献[9]中所述的方法就用到了 LBP,但 LBP 算子受小空间支撑区域的限制,并且进行单纯的纹理分析,而摒弃对真假指纹由于材料不同所造成的采集图像噪声差异的分析,因此是不可取的。综上所述,本文提出了一种基于曲波纹理分析和 SVM-KNN 分类的假指纹检测算法。该算法适用于分辨率为 500dpi 的各类指纹采集仪,分别对图像高频噪声信息以及去噪后的纹理信息进行分析,量化真假指纹之间的噪声以及纹理差异。相较 LBP 算法,本文算法能更好地表现由于假指纹制作材料引起的与真指纹的噪声分布间的微小偏差。本文首先介绍曲波变换系数的提取以及图像重构;再详细介绍所提出的假指纹检测算法,包括特征提取、特征分析及分类等过程;最后为实验和总结。

1 曲波变换

自然图像中包含有大量的纹理特征,线奇异性表现比较突出。小波变换不能达到最优的逼近,为了克服小波的这种不足,提出了脊波变换^[10]。直线奇异性特征可以由脊波变换表示,对于具有直线奇异性的多变量函数,脊波变换有良好的逼近性能,但是由于图像曲线边缘的逼近性能只接近于小波变换,不具有最优的非线性逼近误差衰减阶,与小波相同,脊波对于曲线的奇异性表示不是稀疏的。为了解决这些问题,Candes 和 Donoho 于 1999 年在脊波变换的基础上提出了曲波变换并且构造了曲波变换的紧框架;在 2003 年,Candes 等提出了一种新的、具有较快变换速度的框架体系,即第二代曲波变换^[11,12]。对于具有光滑奇异性曲线的目标函数,曲波提供了稳定的、高效的和近乎最优的表示。

设 x 代表空间位置参量, r 和 θ 代表频率域中的极坐标, ω 代表频率域参量,以下定义均在二维空间 R^2 中给出^[12]。

定义 1^[11,12] 假设存在平滑、实值、非负的半径窗 $W(r)$, $r \in (1/2, 2)$,和角窗 $V(t)$, $t \in [-1, 1]$,且满足条件:

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, r \in (3/4, 3/2) \quad (1)$$

$$\sum_{l=-\infty}^{\infty} V^2(t-l) = 1, t \in (-1/2, 1/2) \quad (2)$$

在频域中定义频窗 U_j 如下:

$$U_j(r, \theta) = 2^{-3j/4} W(2^{-j} r) V\left(\frac{2^{j/2} \theta}{2\pi}\right) \quad (3)$$

其中, $\lfloor j/2 \rfloor$ 是 $j/2$ 的整数部分。 U_j 的支撑区间是受 W 和 V 支撑区间限制获得的楔形区域,楔形区域符合各向异性尺度的特性。

定义 2^[11,12] 令母曲波为 $\varphi_j(x)$,且傅里叶变换为 $\hat{\varphi}_j(x) = U_j(x)$,则在尺度 2^{-j} 上的所有曲波都可由 φ_j 旋转和平移

获得。令 R_θ 表示以 θ 为弧度的旋转,引入:

(1)均匀的旋转角度序列

$$\theta_l = 2\pi \times 2^{-\lfloor l/2 \rfloor} \times l, l=0, 1, 2, \dots, 0 \leq \theta_l < 2\pi \quad (4)$$

(2)平移参数

$$k = (k_1, k_2) \in Z^2 \quad (5)$$

综合以上概念,定义在尺度 2^{-j} 、方向 θ_l 、平移参数 (k_1, k_2) 处的曲波为:

$$\varphi_{j,l,k}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)})) \quad (6)$$

其中, $x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j/2})$, R_{θ_l} 由 θ_l 旋转获得。曲波变换可表示为:

$$c(j, l, k) := \langle f, \varphi_{j,l,k} \rangle = \int_{R^2} f(x) \overline{\varphi_{j,l,k}(x)} dx \quad (7)$$

通过 Plancherel 理论,由上式可推得:

$$\begin{aligned} c(j, l, k) &:= \frac{1}{2\pi^2} \int \hat{f}(\omega) \overline{\varphi_{j,l,k}(\omega)} d\omega \\ &= \frac{1}{2\pi^2} \int \hat{f}(\omega) U_j(R_{\theta_l}\omega) e^{j(x_k^{(j,l)}, \omega)} d\omega \end{aligned} \quad (8)$$

1.1 离散曲波变换

基于连续曲波变换,Candes 和 Donoho 提出两种快速离散曲波变换的实现方法:第一种是 USFFT 算法;另外一种为 Wrap 算法。两种算法的分解结果相同,然而第二种算法的效率高、运算速度快,因此本文采用第二种算法^[11,13]。经过离散曲波变换后得到 $c(j, l, k)$ 结构的系数, j 表示尺度, l 表示角度, k 表示第 k 个系数。

1.2 曲波重构

曲波重构的过程是一个曲波逆变换的过程,在进行曲波逆变换之前要进行去噪处理。

$$\tilde{c}(j, l, k) = \omega * c(j, l, k) \quad (9)$$

其中, $\omega = \begin{cases} 1, & c(j, l, k) > \bar{\omega} \\ 0, & \text{其他} \end{cases}$, $\bar{\omega}$ 为去噪阈值。对 $\tilde{c}(j, l, k)$ 进行曲波逆变换得到重构图像 f 。

2 假指纹检测

假指纹检测算法的整体流程包括:特征提取、特征分析及真假判别,如图 1 所示。指纹图像经过曲波变换之后得到曲波系数 $c(j, l, k)$ 和曲波重构图像 f 。

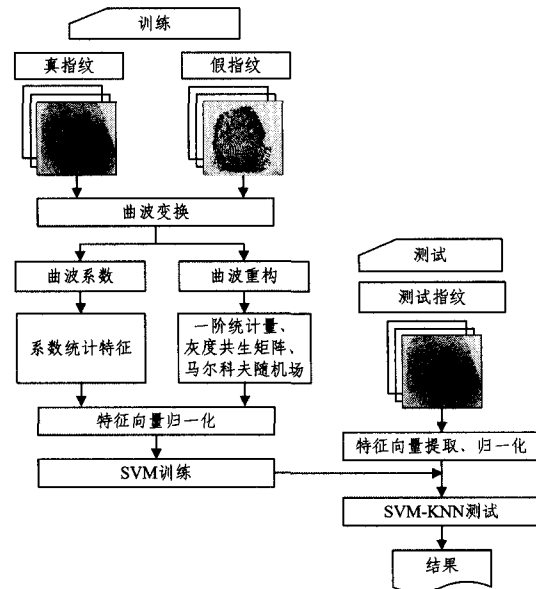


图 1 假指纹图像检测算法流程

2.1 系数特征提取

图像经过曲波变换后,被划分成多个尺度层。最内层也就是第一层称为 Coarse 尺度层,是由低频系数组成的一个矩阵;最外层也称为 Fine 尺度层,是由高频系数组成的一个矩阵;中间的几层称为 Detail 尺度层,每层系数是由中高频系数组成的矩阵。其中,Fine 层与 Detail 层中的高频以及中高频系数包含了图像的噪声信息,对 Fine 层与 Detail 层提取系数特征,由图 2 可知真假指纹之间的特征值有比较大的差异。

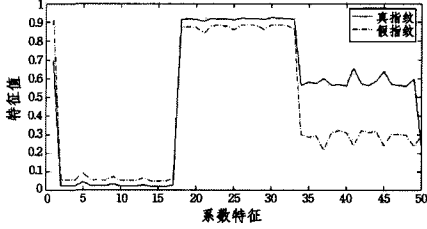


图 2 曲波系数特征

统计每层系数的能量 E 、最大值 MAX 、最小值 MIN 、均值 $Mean$ 和方差 Var ,其中计算每层系数的绝对值的平方和作为能量。

对每层系数进行一阶统计量特征提取,假设 $H(n)$ 为归一化直方图, N 表示最大灰度级, μ 为灰度均值,一阶统计量计算如下:

$$\text{能量: } e = \sum_{n=0}^{N-1} H(n)^2 \quad (10)$$

$$\text{熵: } s = - \sum_{n=0}^{N-1} H(n) \log H(n) \quad (11)$$

$$\text{方差: } \sigma^2 = \sum_{n=0}^{N-1} (n - \mu)^2 H(n) \quad (12)$$

$$\text{偏度: } \gamma_1 = \frac{1}{\sigma^3} \sum_{n=0}^{N-1} (n - \mu)^3 H(n) \quad (13)$$

$$\text{峰度: } \gamma_2 = \frac{1}{\sigma^4} \sum_{n=0}^{N-1} (n - \mu)^4 H(n) \quad (14)$$

2.2 纹理特征提取

本文对曲波重构图像提取纹理特征,包括 3 种纹理特征,分别是一阶统计量、灰度共生矩阵 (GLCM) 和马尔科夫随机场 (MRF)。

2.2.1 一阶统计量

一阶统计量可以应用于一般的一维数据特征提取,也可以应用于类似图片的二维数据矩阵,可以衡量图像中随机位置某一灰度值出现的概率,像素之间相关性可以表明指纹的真假性^[14-16]。通过直方图计算像素之间的变化程度,并提取一阶统计量,目标是当图像的物理结构发生变化时,量化灰度级分布的变化,由此判别真假指纹。

2.2.2 灰度共生矩阵 (GLCM)

图像的灰度共生矩阵能反映出图像灰度关于变化幅度、方向、相邻间隔的综合信息,它是分析图像局部模式及其排列规则的基础。在纹理确定上图像的灰度共生矩阵已被证明是一个很好的方法,广泛用于将灰度值转换为纹理信息。Haralick 于 1973 年研究了图像纹理中灰度级的空间依赖关系,提出灰度共生矩阵的实质。图像中灰度为 x 的像素,其位置为 (i, k) ,统计与其距离为 d 、方向为 θ 、灰度为 y 的像素 $(i + D_i, k + D_k)$,出现次数 $p(x, y, d, \theta)$ 的数学表达式为:

$$p(x, y, d, \theta) = \{ [(i, k), (i + D_i, k + D_k) | f(i, k) = x, f(i + D_i, k + D_k) = y] \} \quad (15)$$

式中, $x, y = 1, 2, \dots, L$ 表示图像中的灰度级; $i, k = 1, 2, \dots, K$

表示像素坐标; d 为生成灰度共生矩阵的步长; D_i, D_k 是位置偏移量;生成方向 θ 可取任意方向,从而生成不同方向的共生矩阵。对灰度共生矩阵进行归一化:

$$G(x, y) = \frac{p(x, y)}{\sum_{x=1}^L \sum_{y=1}^L p(x, y)} \quad (16)$$

将共生矩阵描述纹理的状况量化,由于指纹库中的指纹纹线宽度不一,为了提高假指纹的识别率,本文算法提取了距离 d 为 $1 \sim 4$, θ 分别取 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 4 个方向,共计 16 个灰度共生矩阵;然后分别计算对应的能量、对比度、熵、局部平稳、自相关以及不相似性,计算公式如下:

能量:

$$ASM = \sum_{x=1}^L \sum_{y=1}^L (G(x, y))^2 \quad (17)$$

对比度:

$$CON = \sum_{n=0}^{L-1} n^2 \left\{ \sum_{|x-y|=n} G(x, y) \right\} \quad (18)$$

熵:

$$ENT = - \sum_{x=1}^L \sum_{y=1}^L G(x, y) \log G(x, y) \quad (19)$$

局部平稳:

$$IDM = \sum_{x=1}^L \sum_{y=1}^L \frac{G(x, y)}{1 + (x - y)^2} \quad (20)$$

自相关:

$$COR = \sum_{x=1}^L \sum_{y=1}^L \frac{(x, y)G(x, y) - \mu_x \mu_y}{s_x s_y} \quad (21)$$

不相似性:

$$DIS = \sum_{x=1}^L \sum_{y=1}^L \frac{|x - y|}{G(x, y)} \quad (22)$$

其中,

$$\mu_x = \sum_{x=1}^L \sum_{y=1}^L x \cdot G(x, y) \quad (23)$$

$$\mu_y = \sum_{x=1}^L \sum_{y=1}^L y \cdot G(x, y) \quad (24)$$

$$s_x^2 = \sum_{x=1}^L \sum_{y=1}^L G(x, y) (x - \mu_x)^2 \quad (25)$$

$$s_y^2 = \sum_{x=1}^L \sum_{y=1}^L G(x, y) (y - \mu_y)^2 \quad (26)$$

这几个纹理特征可有效描述指纹图像的纹理特征,具有较好的鉴别能力。

2.2.3 马尔科夫随机场 (MRF)

马尔科夫随机场模型是由 Besag 在 1973 年提出,它可以表达图像数据的空间概率关系的建模并且用于纹理特征模型,已被证明对于图像特征提取是有效的^[17,18]。从结构分析角度来看,MRF 纹理特征就是要找出纹理基元以及纹理基元间的相互依赖关系,而其基元关系可以用条件概率模型表示。随机图像的纹理区域可看成是二维随机过程的有限取样,不同的统计参数表示不同的随机过程。纹理基元之间展现出的依赖关系,反映了纹理基元的不同聚集,而不同聚集之间的纹理又对应着不同的统计参数,应用数学上的 MRF 能够很好地描述纹理的随机特征。

MRF 是一个二维点阵,可以用概率模型来描述每个点,MRF 的假设前提是点阵中的每个点的像元值仅仅依赖于其邻域中像元的像元值。马尔科夫随机场可用如下局部条件概率密度 (PDF) 来描述:

$$p(z(s) | z(m), m = 1, 2, \dots, N \times M, s \neq m) = p(z(s) | z(m), m \in N(s)) \quad (27)$$

其中, $z(s)$ 是在点阵 $N \times M$ 中点 s 的像素值, $N(s)$ 是以 s 为中心的领域像素点集, $p(z(s))$ 的值受 $z(m)$ 的影响。如果 PDF 服从高斯分布, 那么马尔科夫随机场为高斯-马尔科夫随机场 (GMRF)。用邻域信息来估计像素点灰度值的对称差分方程:

$$z(s) = \sum_{\beta, m} [\beta_m z(s+m) + \beta_{-m} z(s-m)] + e_s \quad (28)$$

其中, β_m 为每个邻域像素点贡献于中心像素点灰度值的权值, e_s 为均值为 0 的高斯分布噪声, m 是离中心点 s 的偏差, 表示成矩阵符号公式为:

$$z(s) = \beta^T Q_s + e_s \quad (29)$$

其中, β 是由 β_m 组成的向量, Q_s 的定义如下:

$$Q_s = \begin{bmatrix} z(s+m_1) + z(s-m_1) \\ z(s+m_2) + z(s-m_2) \\ z(s+m_3) + z(s-m_3) \\ \dots \end{bmatrix} \quad (30)$$

利用最小二乘法计算得到纹理特征量:

$$\beta = [\sum_{s \in J} Q_s Q_s^T]^{-1} [\sum_{s \in J} Q_s z(s)] \quad (31)$$

用 GMRF 对指纹图像进行二阶参数估计, 图像中任一 3×3 窗口为图像采样模板范围, 由此确定 Q_s 。 f 代表指纹图像, 对于每一个 3×3 的窗口, β 就是特征值, MRF 特征值对灰度级变化不敏感。

2.3 特征分析

目前假指纹用到的特征量主要有一阶统计量、傅里叶变换, 但是测试效果并不好, 新的特征量的选择分析对于假指纹的检测有着决定性的意义, 本节将对文中所提到的特征做相关的性能分析^[14,19]。

本文提取了两类特征: 曲波系数统计特征和曲波重构纹理特征。曲波系数包含曲波变换后的多尺度信息, 其中某些尺度含有高频的噪声信息, 这些噪声信息是提取曲波系数特征的关键, 曲波系数特征主要表现了图像各尺度混合信息中的能量、熵等信息; 而纹理特征主要是相对于曲波去噪重构之后的图像而言的, 在去除高斯噪声等高频信息的前提下, 对指纹图像提取纹理特征进行分析。由于不同材料的分子集聚性不同, 导致制作出来的假指纹在纹理表现上有着细微的差别, 如图 3(a) 所示, 真指纹的脊线宽度变化较快, 脊线灰度变化剧烈且灰度变化是随机的, 再之灰度值随机分布且灰度级较大; 而图 3(b)~3(d) 中假指纹脊线灰度变化比较规律, 波动较小, 脊线宽度较统一, 灰度级较小, 随机变化程度不显著。



图 3 真假指纹的纹理区别, 包括脊线灰度变化、脊线宽度变化以及灰度分布

对于这些特征的具体量化分析如图 4 所示, 图中分别显示了硅胶、乳胶、木胶假指纹的灰度共生矩阵、一阶统计量以及 MRF 3 类特征的特征值, 以及与真指纹的 3 类特征值的差值: 1) 由图 4(a)、(b) 可知, 各类假指纹的特征值分布都与真指纹不同, 相对而言, 木胶的特征值分布比较接近真指纹, 而乳胶假指纹的一阶统计量特征值分布在较大程度上能区别于真指纹。由一阶统计量的定义和乳胶指纹的表现性状可知,

由于乳胶指纹脊线灰度分布较均匀且脊线宽度变化缓慢, 导致一阶统计量的方差较小而能量、熵等参量较大。2) 图 4(c)、(d) 表示 MRF 的特征量以及与真指纹的差值, 由图可知 MRF 特征对于木胶假指纹有更好的分辨效果, 这反映了纹理基元的不同聚集, 而不同聚集之间的纹理对应着不同的统计参数。木胶的分子集聚性较强, 导致木胶的纹理表现较为分散, 平滑性不高, 因此 MRF 特征值曲线的随机跳跃性比较强; 而真指纹较为平缓的纹理表现使得 MRF 特征值曲线表现较为平缓。3) 图 4(e)、(f) 表示灰度共生矩阵的特征量以及与真指纹的差值, 对硅胶指纹有较好的分辨效果, 硅胶制作出来的指纹与真指纹的区别在于脊线宽度的变化快慢程度, 硅胶指纹脊线宽度变化缓慢且灰度级较小。我们知道硅胶材料是比较接近于人体皮肤材质的一种材料, 灰度共生矩阵的多尺度多方向的特性对于分辨细微纹理差别有一定的优势。图 4(e)、(f) 中, 硅胶指纹小范围内的灰度级变化相对大范围内的灰度级变化较小, 导致硅胶指纹能量相对于真指纹较为集中, 而其对比度由于灰度级变化缓慢较真指纹小; 硅胶材质指纹由于制作工艺的原因, 一般其采集得到的指纹较真指纹粗糙, 这也是其熵大于真指纹的原因; 另外局部平稳、自相关以及不相似性等值由于真假指纹纹理分布的不一致性在量化表现上都会产生一定的差别。尤其是自相关, 硅胶指纹制作难度较高, 指纹纹线完整性不足, 其纹理一致性明显低于真指纹, 导致真指纹的自相关值较高。

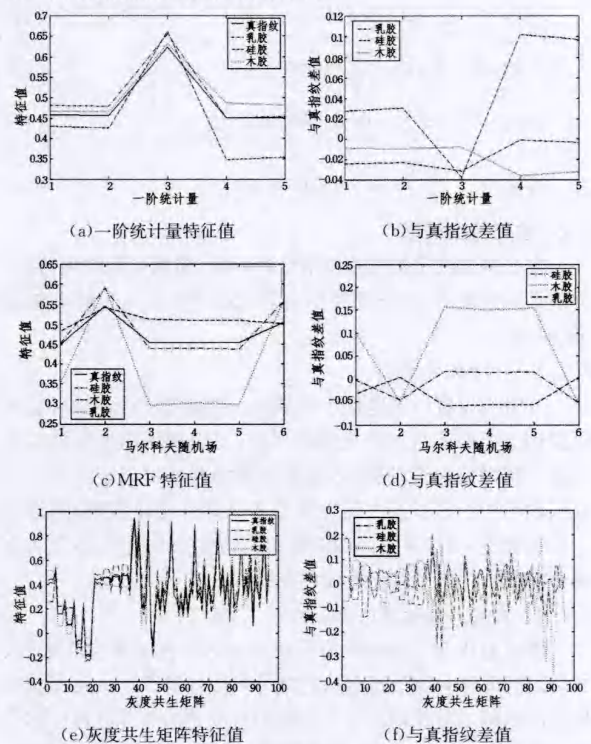


图 4 灰度共生矩阵、一阶统计量以及 MRF 的特征量值及与真指纹的差值

综上所述, 各类特征对于真假指纹的鉴别都表现了优异的性能, 并显示了对于各类假指纹的辨识能力各有所长。比如, 灰度共生矩阵对硅胶假指纹具有较好地辨识能力, 而马尔科夫随机场、一阶统计量分别对木胶以及乳胶有更好的分辨效果。

2.4 SVM-KNN 分类

2.4.1 SVM 分类机理

SVM 方法是通过一个非线性映射 ρ , 把样本空间映射到一个高维乃至无穷维的特征空间中 (Hilbert 空间), 使得原来样本空间中非线性可分的问题转化为在特征空间中的线性可分的问题, 并寻找样本在此特征空间中的最优线性分类超平面。

2.4.2 SVM-KNN 分类器形成

传统 SVM 分类算法存在一个缺点: 当样本距离分类超平面小于一个给定的阈值 ϵ 时, 其分类准确率会降低。SVM 分类过程中, 对于取得的每类支持向量的代表点的误差度决定了 SVM 的分类准确率, 可通过 KNN 来对容易出现偏差的样本进行分类, 以提高分类准确率。具体地, 对于待识别样本 α , 计算 α 与两类支持向量代表点 α^+ 和 α^- 的距离差, 当距离差小于给定的阈值, 即 α 离分界面较近时, 样本将落入区域 III (见图 5)。SVM 只计算与两类所取的一个代表点的距离来进行分类, 这样比较容易分错; 而采用 KNN 对测试样本分类, 计算待识别样本和每个支持向量的距离以得出分类结果。

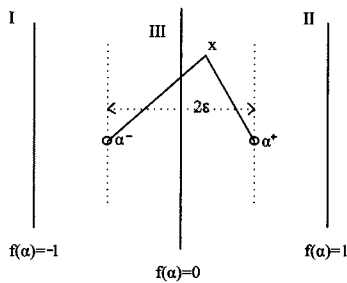


图 5 SVM-KNN 分类原理

2.4.3 分类

通过分析已知类别的训练集, 估计得到分类模型的相关参数, 据此推断测试数据的类别^[20]。本文使用 LIBSVM 进行训练, 利用特征向量得到训练模型 A, 通过训练模型利用 SVM-KNN 分别测试样本进行分类得到分类结果 ∂ 。

3 实验结果

3.1 实验数据库

本文使用的指纹库均由第二届全球假指纹检测竞赛 (the second edition of the Fingerprint Liveness Detection Competition, LivDet2011) 提供。LivDet 系列比赛是由 Cagliari 大学电子与电气工程系和 Clarkson 大学电子与计算机工程系联合举办的假指纹检测领域的权威赛事。指纹库主要由 4 类采集设备 (Biometrika, Digital Persona, Italdata and Sagem) 采集得到的指纹组成, 采集到的指纹图像有真指纹、硅胶指纹、乳胶指纹等。每个指纹库分别有 2000 枚真指纹和 2000 枚假指纹, 如表 1 所列^[21]。

表 1 LivDet2011 指纹数据库

	指纹数据库			
	Sagem	Biometrika	Italdata	Digital Persona
采集仪型号	MSO300	FX2000	ET10	400B
分辨率	500	500	500	500
图像尺寸	352 * 384	312 * 372	640 * 480	355 * 391
真指纹	2000	2000	2000	2000
假指纹	2000	2000	2000	2000

3.2 实验结果

在实验过程中, 将指纹图像数据库分成两份, 50% 用于训练, 50% 用于测试。经过测试, 因数据库不同 ϵ 值有一定差别, 本文中取 $\epsilon=0.2$, 利用 SVM-KNN 分类器进行训练分类, 在一定程度上提升了单纯 SVM 分类器的分类性能, 如表 2 所列。另外, 本文提出的假指纹检测方法的实验结果如表 3 一表 5 所列。

表 2 SVM 与 SVM-KNN 分类器分类性能的比较

算法	Livedet2011 数据库			
	Biometrika	Digital Persona	Italdata	Sagem
SVM	14.15%	5.9%	14.14%	6.45%
SVM-KNN	13.3%	4.5%	13.2%	6.15%

表 3 本算法对各假指纹数据的认假率

算法	Livedet2011 数据库				
	Biometrika	Digital Persona	Italdata	Sagem	四库均值
Winner2011	29%	6.20%	28.50%	12.50%	19.05%
SSCA ^[22]	—	—	—	11.30%	11.30%
本文算法	15.9%	6.6%	5.9%	5.2%	8.4%

表 4 本算法对真指纹数据的拒真率

算法	Livedet2011 数据库				
	Biometrika	Digital Persona	Italdata	Sagem	四库均值
Winner2011	11.00%	66.00%	15.10%	15.10%	26.80%
SSCA	—	—	—	14.40%	14.40%
本文算法	10.7%	2.4%	20.6%	7.2%	10.23%

表 5 本文算法假指纹平均错误接受率

算法	Livedet2011 数据库				
	Biometrika	Digital Persona	Italdata	Sagem	四库均值
multiresolution LBP ^[23]	10.8%	7.1%	16.6%	6.4%	10.23%
Original LBP ^[9]	13.0%	10.8%	24.1%	11.5%	14.85%
Winner2011	20.00%	36.1%	21.8%	13.80%	22.93%
SSCA	—	—	—	12.85%	12.85%
本文算法	13.3%	4.5%	13.2%	6.15%	9.28%

在上述的数据库中进行算法测试, FerrLive、FerrFake、ACE 分别是用来衡量识别效率的参量。其中 FerrLive 指的是真指纹的错误拒绝率, FerrFake 表示假指纹的错误接受率, ACE 是 FerrLive 和 FerrFake 的均值, 表示真假指纹的平均错误接受率。本文的算法同时也与另外 4 个算法进行了比较, 这 4 个算法分别是: LivDet2011 中检测精度最高的 Winner2011 算法、Pereira 提出的 SSCA 算法以及 Original LBP 和 multiresolution LBP 算法。通过与这些实测结果优秀的算法进行比较来验证本文算法的有效性。本文的算法在 Biometrika、Digital Persona、Italdata 和 Sagem 这 4 个数据库下的平均错误率比以上算法都要低。但是, 本文算法对于 Digital Persona 数据库的 FerrFake 以及 Italdata 数据库的 FerrLive 均高于 Winner2011 算法, 另外对 Biometrika 数据库的 ACE 表现也没有 multiresolution LBP 算法好。究其原因主要是, 由于不同的指纹采集仪得到的指纹数据库, 其指纹图像质量有所不同, 纹线尺度 (纹线宽度及纹线间间距) 也有一定的差异, 这些差异的产生主要源自于采集仪的元器件以及采集仪不同的采集原理。

如图 6 所示, Digital Persona 数据库中的图像边缘背景较多, 对边缘背景处理前后的 FerrFake (处理前: 7.9%; 去除边缘背景处理后: 6.6%) 进行对比发现, 边缘背景会对分类精度

产生影响;Italdata 采集仪的规格为 640×480 ,是 4 个数据库中最大的,本文算法所使用的灰度共生矩阵(4 步长,4 方向)以及 MRF(3×3 模板,即三阶 MRF)对于该规格图片可以加大灰度共生矩阵步长、方向,以及 MRF 的模板大小可以进行分类精度调整,如表 6 Italdata 数据库的 ACE 所列,这其实也是由纹线尺度差异引起的。在 LBP 多尺度问题的研究上, multiresolution LBP 算法对于不同采集仪数据库的兼容是一种很好的解决方法。一般情况下,更多的 LBP 尺度会带来更好的分类精度,但是尺度增长也可能对分类精度产生消极影响,即在到达一定阈值时,精度不会比峰值更好。本文算法对于多尺度还未做深入研究,因为多尺度在一定程度上将极大影响算法的时间复杂度,本文算法在保证较低的时间复杂度的前提下,达到较高的分类精度,因此在一定程度上还有很大的改进空间。



(a) Digital Persona 数据库 (b) Italdata 数据库

图 6 数据库样例图

表 6 Italdata 数据库的 ACE

	灰度共生矩阵及 MRF 尺度			
	步长	方向	MRF(阶)	ACE
Italdata	1~4	4	3	13.2%
	1~5	4	4	13.0%
	1~6	4	5	11.8%
	1~6	4	6	11.0%

结束语 本文以假指纹检测方法为研究内容,通过对曲波系数特征以及曲波重构图像纹理特征提取与特征分析、分类等步骤,实现了对指纹图像真假的判别。本文主要研究了特征提取、分析及训练分类过程,提出了一种基于曲波纹理分析和 SVM-KNN 的假指纹检测算法,能对单帧图像进行真假指纹判别。该算法在 LivDet2011 提供的数据库上进行测试,测试结果表明,本文算法对 4 个数据库的平均错误接受率为 9.28%,性能优于 Winner2011、Original LBP、multiresolution LBP、SSCA 4 个算法。

参 考 文 献

[1] 刘舒,于瑞华. 生物特征识别中的关键技术与发展趋势[J]. 中国人民公安大学学报:自然科学版,2006,47(1):63-65

[2] Galbally J, Fierrez J, Ortega-Garcia J. Vulnerabilities in biometric systems; attacks and recent advances in liveness detection [C]//Proc. Spanish Workshop on Biometrics(SWB). June 2007

[3] Nixon K, Aimala V, Rowe R. Spoof detection schemes[M]// Jain P F A, Ross A, et al. Handbook of Biometrics. Springer, 2007

[4] Kim H, Jin C, Elliott S. Liveness detection of fingerprint based on band-selective Fourier spectrum[J]. Information Security and Cryptology, 2007, 4817:168-179

[5] Nikam S B, Agarwal S. Ridgelet-based fake fingerprint detection

[J]. Neurocomputing, 2009, 72:2491-2506

[6] Moon Y S, Chen J S, Chan K C, et al. Wavelet based fingerprint liveness detection[J]. Electronic Letters, 2005, 41:1112-1113

[7] Mainguet J-F, Pégulu M, Harris J B. Fingerprint recognition based on silicon chips[J]. Future Generation Comp. Syst., 2000, 16(4):403-415

[8] Abhyankar, Schuckers S. Integrating a wavelet based perspiration liveness check with fingerprint recognition[J]. Pattern Recognition, 2009, 42:452-464

[9] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions[J]. Pattern Recognition, 1996, 29:51-59

[10] 徐学斌,张德运,张新曼,等. 基于离散曲波变换和支持向量机的掌纹识别方法[J]. 红外与毫米波学报, 2009, 28(6):456-460

[11] Candes E J, Demanet L, Donoho D L. Fast Discrete Curvelet Transforms[J]. Multiscale Modeling and Simulation, 2006, 5(3):861-899

[12] 李晖晖,郭雷,刘航. 基于二代 Curvelet 变换的图像融合研究[J]. 光学学报, 2006, 26(5):657-662

[13] 邓艾,吴谨,杨莘,等. 基于二代 Curvelet 变换和区域匹配度的图像融合算法[J]. 计算机科学, 2012, 39(6A):513-548

[14] Marasco E, Sansone C. An anti-spoofing technique using multiple textural features in fingerprint scanners [C]// 2010 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS). 2010:8-14

[15] Galbally J, Alonso-Fernandez F, Fierrez J, et al. A high performance fingerprint liveness detection method based on quality related features[J]. Future Generation Computer Systems, 2012, 28:311-321

[16] Marasco E, Sansone C. Combining perspiration-and morphology-based static feature for fingerprint liveness detection[J]. Pattern Recognition Letters, 2012, 33:1148-1156

[17] 李旭超,朱善安. 图像分割中的马尔可夫随机场方法综述[J]. 中国图象图形学报, 2007, 12(5):789-798

[18] 汪闽,骆剑承,周成虎,等. 结合高斯马尔可夫随机场纹理模型与支撑向量机在高分辨率遥感图像上提取道路网[J]. 遥感学报, 2005, 9(3):271-276

[19] Hyun-suk Lee, Hyun-ju Maeng, You-suk Bae. Fake Finger Detection Using the Fractional Fourier Transform[J]. Biometric ID Management and Multimodal Communication, 2009, 5707:318-324

[20] 李玲俐. 数据挖掘中分类算法综述[J]. 重庆师范大学学报:自然科学版, 2011, 28(4):44-47

[21] Yambay D, Ghiani L, Denti P, et al. LivDet 2011-Fingerprint Liveness Detection Competition 2011 [C]// IAPR/IEEE Int. Conf. on Biometrics. 2012:208-215

[22] Pereira L F A, Pinheiro H N B, Cavalcanti G D C, et al. Spatial surface coarseness analysis; technique for fingerprint spoof detection[J]. Institution of Engineering and Technology, 2013, 49(4):260-261

[23] Ojala T, Pietikäinen M, Maenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24:971-987