中图法分类号: TP309.2 **文献标志码**: A **文章编号**: 1006-8961(2011)05-0800-07
论文索引信息:周玲丽, 赖剑煌. 人脸特征的 SIFT 保护算法 [J]. 中国图象图形学报,2011,16(5): 800-806

人脸特征的 SIFT 保护算法

周玲丽[1)], 赖剑煌[2)]
[1)](中山大学数学与计算科学学院, 广州 510275) [2)](中山大学信息科学与技术学院, 广州 510275)

摘 要: 人脸识别技术在门禁、视频监控等公共安全领域中的应用日益广泛,人脸特征数据的安全性和隐私性问题成为备受关注的焦点。提出一种基于 SIFT 的人脸特征安全保护新算法,首次将随机投影应用到对人脸特征数据的保护中。该算法首先利用 SIFT 特征对旋转、尺度缩放、光照变化等具有较好稳定性的特点,提取有较好鲁棒性的人脸数据;然后根据用户密钥对 SIFT 特征进行不可逆变换,生成具有可重建性的人脸特征模板数据,从而实现对人脸特征数据的保护。实验表明,该算法在 CMU、AR 和 Feret 人脸数据库中均取得较高的识别率,不仅对人脸特征数据具有保护作用,并且对姿势、遮挡和表情的变化具有较高的鲁棒性。

关键词: 人脸识别; SIFT; 安全性; 隐私性; 鲁棒性

Protection algorithm of face feature using SIFT

Zhou Lingli[1)], Lai Jianhuang[2)]
[1)] (*School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou* 510275 *China*)
[2)] (*School of Information Science and Technology Mathematics Department, Sun Yat-Sen University, Guangzhou* 510275 *China*)

Abstract: With the growing use of face recognition in security and video control domain, there are growing concern about security and privacy of biometrics data. This paper proposes a security algorithm about face feature, which bases on the SIFT feature and random projection. Because the SIFT features are invariant to image rotation, scale and change in illumination, feature extraction is first performed on face images by SIFT algorithm. The SIFT features are transformed using invertible transformation which is generated by user specific key. All the above successive procedures produce the cancelable and non-invertible template feature data, which can achieve the protection of face data. In extensive experiments with publicly available face datasets CMU, AR and Feret, higher recognition accuracy is reached, which demonstrates that the proposed approach is not only able to protect the face data, but also robust to various complex conditions, such as changes in the pose, occlusions and expression, etc.

Keywords: face recognition; SIFT; security; privacy; robust

0 引 言

生物特征识别技术经过数十年的研究,目前已进入实用阶段,但是对生物特征数据安全性和隐私性的研究相对滞后,存在的问题日益突出[1]。由于生物特征具有唯一性,个人特征数据的丢失就意味着个人身份的丢失。特征数据一旦被破坏或窃取,无法像密码和 IC 卡那样取消或重新更新[2],个人生物特征的安全性显得特别重要。特征模板数据库是生物特征识别系统的重要组成部分,其中存储着注册用户已提取的特征数据。研究表明,运用“Hill

收稿日期:2010-01-21;**修回日期**:2010-03-16
基金项目:国家自然科学基金项目(60803083);国家自然科学基金-广东省联合项目(U0835005)。
第一作者简介:周玲丽(1973—),女,工程师。应用数学专业博士,主要研究方向模式识别、信息安全。
E-mail:mcszll@ mail. sysu. edu. cn。

Climbing Attacks"[3]技术,能够通过特征模板数据库中的数据重建原始的生物特征。因此,生物特征模板数据的安全性研究是一个重要而紧迫的研究课题,成为生物特征识别领域的研究热点。

目前,针对生物特征模板数据的安全性技术主要有两大类算法。一类是 Ratha 等人[1]提出的可重建生物特征(cancelable biometric)算法,另一类是 Uludag 等人[4-5]提出的生物特征加密系统(biometric cryptosystem)算法。Ratha 等人[1]根据生物特征具有的唯一性(invariable)、无法撤销性(irrevocable)和隐私性(privacy)特点,对生物特征进行人为的、不可逆的变换,使得特征模板数据库中保存的不再是原始的生物特征,而是生物特征的变换形式。由于变换的不可逆性,即使攻击者获取了特征模板数据库中的数据,也无法得到原始的生物特征。可以通过修改不可逆变换的参数,得到新的特征模板数据,使得模板数据库中的数据具有可重建性(cancelable),进而克服生物特征的无法撤销性所带来的不足。Ratha 等人[6]提出 3 种不可逆变换函数,用来生成可重建的指纹模板数据。生物特征加密系统算法[4-5]将生物特征识别系统与密码学结合起来。在这类算法中,特征模板数据库中保存的不再是原始的生物特征,而是其在密码学架构下的一种变换形式,主要应用纠错码(error correct code)来处理类内差异问题。由于生物特征数据中噪声的存在,引入纠错码对噪声导致的误差进行处理。Juels 等人[7]提出 Fuzzy Commitment 算法,在此基础上 Juels 等人[8]进行改进提出 Fuzzy Vault 算法。

目前,已有的关于生物特征模板数据安全性方面的研究大都是关于指纹和虹膜的,由于人脸特征容易受到光照、表情及姿势变化等外在条件的影响,同一人的图像表现差别很大,人脸特征之间具有较大的类内差异,使得人脸数据方面的保护算法较少涉及。例如,Savvides 等人[9]利用随机卷积核对人脸图像进行加密。Teoh 等人[10]提出一种基于两种因素的人脸特征识别 Biohashing 算法。人脸识别相对于虹膜和指纹识别,具有自然性和不容易被被测个体察觉等优点,越来越多地应用到身份识别中,本文主要考虑人脸特征数据的安全。

纠错码一般是对二值的串进行纠错编码,人脸特征数据是实值的,将实值的特征转化为二值的串,会丢失很多有用的信息;人脸特征的维数较大;人脸特征较易受到光照、表情或姿势等外在因素的影响,具有较大的类内差异,因此,人脸特征不适合应用纠错码技术来处理,本文算法按照可重建生物特征的研究思路提出。

已有的关于人脸特征数据的保护算法,人脸特征大都采用全局特征,而人脸的全局特征,较易受到光照、表情及姿势等外部条件的影响。局部特征则可以很好地解决这些问题。局部特征主要是对以特征点为中心的局部区域的描述,综合了特征点本身及其邻域的信息。SIFT(scale invariant feature transform)算法[11]提取的特征就属于这类局部特征,通过 SIFT 算法提取的特征对旋转、尺度缩放、亮度变化等保持不变性,对仿射变化、噪声也保持一定程度的稳定性。因此,提出的人脸特征保护算法,采用人脸的 SIFT 特征。

针对人脸识别,从保护生物特征模板数据的角度出发,提出一种基于 SIFT 对人脸特征数据进行保护的安全算法。

1　SIFT 算法

SIFT 算法是 Lowe[11]在 2004 年提出的一种对尺度空间、图像缩放、旋转甚至仿射不变的图像局部特征描述算子。SIFT 算法利用不同尺度的高斯核函数对图像连续滤波和下采样,形成高斯金字塔图像,根据相邻尺度的高斯图像得到 DOG(different of Gaussian)金字塔多尺度空间。将 DOG 尺度空间的每个点与相邻尺度和相邻位置的点逐个进行比较,获得局部极值点,再通过局部极值点确定关键点,最终得到 SIFT 特征。

1.1　尺度空间的极值检测

由于高斯卷积核是实现尺度变化的唯一线性核[12-13],因此,采用尺度可变高斯函数 $G(x,y,\sigma)$ 生成图像 $I(x,y)$ 的尺度空间 $L(x,y,\sigma)$,将不同尺度的高斯函数和人脸图像 $I(x,y)$ 进行卷积运算,即

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \tag{1}$$

式中,$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$,$\sigma$ 表示高斯函数的方差。

为了能够有效地在尺度空间中检测到稳定的关键点,利用不同尺度的高斯差分核与图像卷积生成高斯差分尺度空间(DOG scale-space)(图 1)。

$$D(x,y,\sigma) = [G(x,y,k\sigma) - G(x,y,\sigma)] * I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \tag{2}$$

式中,k 是常数因子。

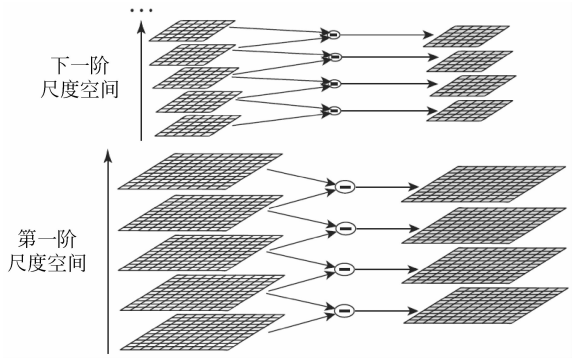

图 1 Gaussian 金字塔和 DOG 金字塔的建立

Fig. 1 The set of scales space, and the computation of DOG

为了检测 $D(x,y,\sigma)$ 的局部极值点,采样点不仅需要与它同一尺度的 8 个邻域点进行比较,还要和上下相邻尺度的 9×2 个邻域像素共 26 个相邻像素进行比较,得到局部极值点(见图 2),所有局部极值点构成了 SIFT 候选关键点。

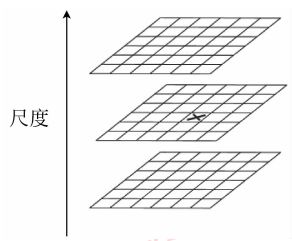

图 2 尺度空间极值点的确定

Fig. 2 Key-point detection

1.2 关键点的确定

为了进一步增强匹配的稳定性,提高抗噪声能力,极值检测得到的所有候选关键点必须排除低对比度的点和边缘点,才能确定是关键点。

通过拟合 3 维二次曲线找出低对比度的点,将 $D(x,y,\sigma)$ 进行泰勒展开为二次项

$$\boldsymbol{D}(x) = \boldsymbol{D} + \frac{\partial \boldsymbol{D}^{\mathrm{T}}}{\partial x}x + \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\frac{\partial^2 \boldsymbol{D}}{\partial x^2}\boldsymbol{x} \tag{3}$$

式中,$\boldsymbol{D}$ 是 DOG 计算的结果。由 D 和 x 可以找到一个偏移量

$$\hat{x} = -\frac{\partial^2 \boldsymbol{D}^{-1}}{\partial x^2}\frac{\partial \boldsymbol{D}}{\partial x} \tag{4}$$

将式(4)代入式(3)中,如果 $|D(\hat{x})| < 0.03$,则该点是低对比度的点[11]。

由于 DOG 函数在跨越边缘的地方存在大的主曲率,因此,通过主曲率查找边缘点。Hessian 矩阵的特征值正比于 D 的主曲率,设 α 是最大特征值,β 是最小特征值,令 $\alpha = r\beta$。

$$\boldsymbol{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{5}$$

则

$$\begin{aligned} \mathrm{tr}(\boldsymbol{H}) &= D_{xx} + D_{xy} = \alpha + \beta \\ \mathrm{Det}(\boldsymbol{H}) &= D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \\ \frac{\mathrm{tr}(\boldsymbol{H})^2}{\mathrm{Det}(\boldsymbol{H})} &= \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \end{aligned} \tag{6}$$

当 $\alpha = \beta$ 时,式(6)取最小值。通过 $\frac{\mathrm{tr}(\boldsymbol{H})^2}{\mathrm{Det}(\boldsymbol{H})} < \frac{(r+1)^2}{r}$ 检测主曲率是否小于某一个阈值 r。若不满足该条件,则该点可能是边缘点。

1.3 SIFT 特征描述子

利用关键点邻域像素的梯度方向的分布特性为每个关键点指定方向参数,使算子具备旋转不变性。点 (x,y) 处梯度的模值 $m(x,y)$ 和方向 $\theta(x,y)$ 的计算公式分别为

$$\begin{aligned} m(x,y) &= \\ &\sqrt{(L(x+1,y)-L(x-1,y))^2 + (L(x,y+1)-L(x,y-1))^2} \\ \theta(x,y) &= \arctan\left(\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)}\right) \end{aligned} \tag{7}$$

$L(x,y+1)$,$L(x,y-1)$,$L(x-1,y)$,$L(x+1,y)$ 分别表示点 (x,y) 上下左右 4 个像素点的灰度值,此时的关键点具有位置、大小和方向 3 个参数。

进一步将坐标轴旋转为关键点的方向,确保旋转不变性。然后,对任意一个关键点,在其尺度空间,取以关键点为中心的 16×16 大小的区域。在每个 4×4 的小块上计算 8 个方向的梯度向量直方图,绘制每个梯度方向的累加值,形成一个种子点(见图 3)。一个关键点使用 4×4 共 16 个点来描述,最终形成 128 维的特征向量。该向量就是 SIFT 特征描述子,即 SIFT 特征向量。SIFT 特征向量已经去除了尺度变化、旋转等几何因素的影响。将特征向量的长度归一化,则可以进一步去除光照变化的影响。

当两幅图像的 SIFT 特征向量生成后,采用关键点特征向量的欧氏距离作为关键点的相似性判定度量。取一幅图像中的某个关键点,找出另一幅图像中与该点欧氏距离最近的前两个关键点。在这两个关键点中,如果最近距离除以次近距离小于某个比

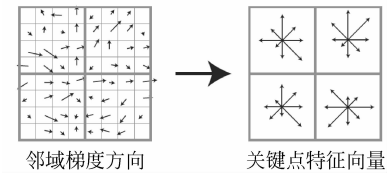

图 3　SIFT 特征描述子的生成

Fig. 3　Computation of the keypoint descriptor

例阈值,则认为这两个关键点是匹配的。

2　基于 SIFT 特征的保护算法

2.1　算法的基本结构

本文提出的基于 SIFT 特征的人脸特征保护算法的基本结构如图 4 所示。算法主要由两部分组成。

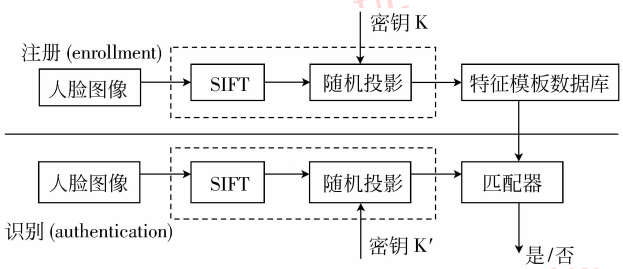

图 4　算法基本结构

Fig. 4　Algorithm structure

1) 对通过特征采集仪采集到的人脸图像,提取 SIFT 特征。

在已有的关于生物特征数据的保护算法中,大都是采用 PCA,LDA 等全局的特征提取方法。这类方法容易受到光照、姿势和表情等的影响,从而导致识别结果的偏差。SIFT 算法提取图像的局部特征,该特征对旋转、尺度缩放和亮度变化等保持不变性,对噪声也保持一定的稳定性。因此,采用 SIFT 算法作为人脸特征提取方法,后面的实验结果也进一步说明,采用这种特征,不仅对光照、遮挡及姿势的变化等的影响具有较高的鲁棒性,而且还能进一步提高识别率。

2) 对提取的 SIFT 人脸特征向量进行随机投影。

$\boldsymbol{X}_{n\times M}$为对人脸图像提取的 M 个 n 维的 SIFT 特征向量(其中 $n=128$),将其投影到 d 维子空间中($d<n$)。

$$\boldsymbol{X}_{d\times M}^{RP}=\boldsymbol{R}_{d\times n}\boldsymbol{X}_{n\times M} \tag{8}$$

式中,$\boldsymbol{R}_{d\times n}$是 $d\times n$ 维随机矩阵,$\boldsymbol{X}_{d\times M}^{RP}$是投影后得到的 d 维子空间中的特征向量,其中 RP(random projection)表示随机投影。随机矩阵 $\boldsymbol{R}_{d\times n}$是由用户拥有的密钥 K 通过随机数发生器产生的,每个用户拥有各自不同的密钥,只有合法的密钥才能唯一决定随机矩阵的生成。另外,匹配过程是在 $\boldsymbol{X}_{d\times M}^{RP}$之间进行的。

在注册和识别阶段,随机投影以相同的方式进行,只是用户各自拥有的密钥不同。经过随机投影后,特征模板数据库中保存的是 $\boldsymbol{X}_{d\times M}^{RP}$,而不再是 $\boldsymbol{X}_{n\times M}$。如果特征模板数据库中的 $\boldsymbol{X}_{d\times M}^{RP}$被非法窃取,但是真正的人脸特征 $\boldsymbol{X}_{n\times M}$仍然安全,使得特征数据的隐私性得到了保护。

随机投影与用户的密钥 K 密切相关,通过修改用户密钥 K 可以生成不同的特征模板数据。如果特征模板数据被窃取或遭到了篡改,则修改用户的密钥可以重新生成新的特征模板数据,从而保证了特征模板数据的可重建性。

2.2　算法分析

随机投影[14]是一种有效而且精确的降维技术,与其他降维方法相比,具有计算简单等优点,特别是通过随机投影进行降维并不会导致数据有较大的变形。本文应用随机投影来保护人脸的 SIFT 特征,其理论基础是 Johnson-Lindenstrauss 引理[15],这也是随机投影得到广泛应用的主要原因。

Johnson-Lindenstrauss 引理　对于任意 $0<\varepsilon<1$ 与整数 n,设 d 为正整数,且使得 $d\geqslant d_0=O(\varepsilon^{-2}\ln N)$,则对于 $\mathbf{R}^n$ 中 N 个点的集合 $\boldsymbol{W}$,存在一个映射 $f:\mathbf{R}^n\rightarrow\mathbf{R}^d$,使得对所有的 $u,v\in\boldsymbol{W}$,有

$$(1-\varepsilon)\|u-v\|^2\leqslant\|f(u)-f(v)\|^2\leqslant(1+\varepsilon)\|u-v\|^2 \tag{9}$$

Johnson-Lindenstrauss 引理确保一个向量空间中的数据点被随机投影到具有一定维数的子空间中,向量数据间的相似性近似保持不变。即同类数据投影后,仍为同一类数据,不同类的数据也是如此,从而满足生物特征识别的可辨识性要求。

由于随机投影中 $d\ll n$,因此式(8)可以看成是一个欠定方程式,因此,这是一个“多到一”的投影,是不可逆的。也就是说,即使非法用户窃取了 $\boldsymbol{X}_{d\times M}^{RP}$和 $\boldsymbol{R}_{d\times n}$,也无法获知原始的特征数据 $\boldsymbol{X}_{n\times M}$,从而保证了原始特征数据的安全性。在进行识别时,只有特征数据和用户的密钥两者都正确,才能通过识别。

3 实验与结果分析

3.1 人脸数据库

在人脸识别的应用中,人脸图像的采集容易受到姿势、遮挡和表情变化等的影响。因此,本文提出的保护算法主要考虑带有姿势、遮挡和表情变化的人脸识别情况。算法实验主要在 CMU[16]、AR[17]和 Feret[18]人脸库进行。CMU 人脸库中,人脸姿势变化较大;AR 人脸库中的人脸带有遮挡;Feret 人脸库中的人脸具有较大的表情变化。

CMU 人脸库有 68 人,每人选取 6 幅具有姿势变化的人脸图像,图像大小为 92×112。本文算法以正面姿势图像(Pose 27)为模板数据,其他图像用来测试(见图 5)。

图 5 CMU 人脸库

Fig. 5 CMU face database

AR 人脸库选取 119 人,每人有 26 幅分两个不同时间采集的,具有不同光照、不同表情和带遮挡(围巾和墨镜)的人脸图像,图像大小为 92×112。排除光照和表情的影响,本文算法只考虑带遮挡的情况,则每人有 4 个测试图像。取光照均匀、没有遮挡的图像作为模板数据(见图 6)。

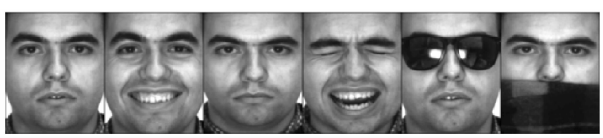

图 6 AR 人脸库

Fig. 6 AR face database

Feret 人脸库共有 255 人,每人有 4 幅图像,同一人的图像具有不同表情、光照、姿态、年龄和眼镜等的变化。本文算法选取每个人的第一幅图像作为模板图像,对其他图像进行测试,图像大小为 92×112(见图 7)。

3.2 实验分析

人脸识别主要包含两个方面的应用,即认证和识别。认证是指验证用户是否是他所声明的身份,通过把待认证用户的特征数据和特征模板数据库中

图 7 Feret 人脸库

Fig. 7 Feret face database

所声明的模板数据进行“一对一”的比较,来确定身份;识别则是一个“一对多”的过程,通过将待识别用户的特征数据与特征模板数据库中的模板逐一比较,从而找到一个最匹配的模板,来确认身份。本文的实验是针对识别来进行的,如果测试图像与某个模板数据的 SIFT 特征匹配点数最多,则判定测试图像为该模板数据所属的那一类。为了减少随机性的影响,实验反复循环 10 次,从而得到平均识别率。

图 8—10 分别是当用户持有合法密钥时,在 CMU、AR 和 Feret 人脸库上得到的识别率。维数为 128 的识别率为原始 SIFT 算法的识别率。从图中可以看到,本文的安全算法(图中“SIFT”表示本文算法)的平均识别率与原始 SIFT 算法识别率相比,均有一定的提高,特别是 AR 库和 Feret 库。实验结果表明,本文算法对姿势、遮挡和表情的变化具有较好的鲁棒性。

另外,在进行随机投影时,特征空间的维数会改变,从图 8—10 中可以看到,本文算法的识别率不会随着维数的变小有较大的降低。因此,本文算法对特征空间维数的改变有较好的鲁棒性。

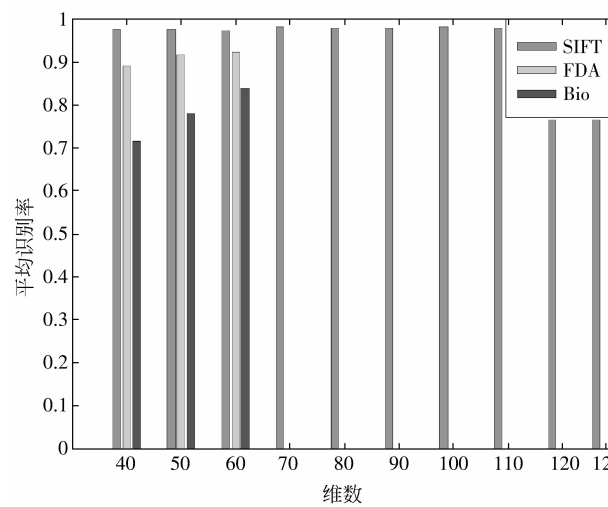

图 8 CMU 库识别率(密钥相同)

Fig. 8 CMU database (same key)

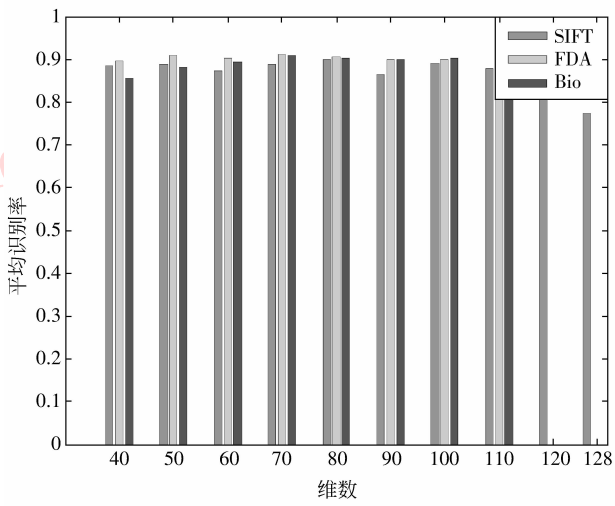

图 9　AR 库识别率(密钥相同)

Fig. 9　AR database (same key)

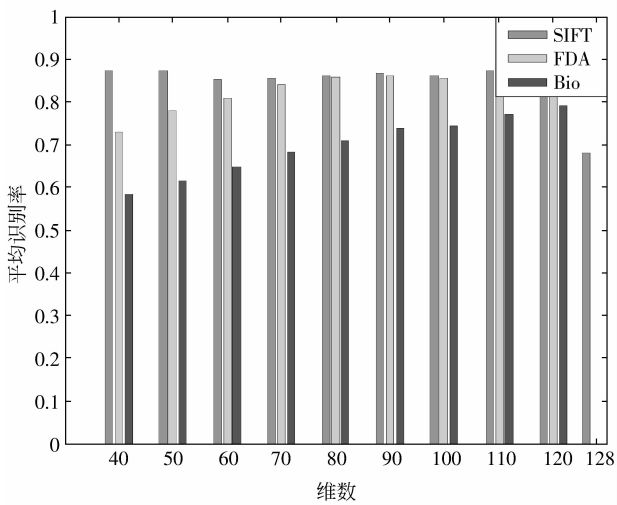

图 10　Feret 库识别率(密钥相同)

Fig. 10　Feret database (same key)

进一步考虑不同人脸特征对本文算法的影响,在本文算法中采用 FDA 特征,特征的匹配采用欧氏距离。在图 8—10 中,"FDA"表示这类情况的实验结果。CMU 库的 FDA 特征向量的维数为 67,AR 库的 FDA 特征向量维数为 118,Feret 库的 FDA 特征向量维数为 254。因此,图 8 中采用 FDA 特征只有维数为 60、50 和 40 的实验结果。图 9 中则没有维数为 120 的识别率。从图 8—10 可以看到,在 CMU 和 Feret 人脸库上,采用 SIFT 特征的识别率均高于采用 FDA 特征的识别率。在 AR 人脸库,本文算法的识别率稍差于采用 FDA 特征的结果。但是图 8—10 中,采用 FDA 特征,算法的识别率随着子空间维数的减少有明显的降低。因此,本文的保护算法采用 SIFT 特征比采用 FDA 特征的识别效果要好,并且对特征空间维数的改变具有更好的鲁棒性。

3.3　算法比较

Biohashing 算法是 Teoh 等人[10]提出的一种基于两因素的生物特征识别算法,Biohashing 算法是将提取的特征向量与 Token(或智能卡)中存储的随机序列进行迭代内积运算,从而得到基于用户的编码 BioHashCode,BioHashCode 之间通过汉明距离(Hamming distance)来度量相似性。Biohashing 算法针对人脸识别取得较好的识别效果,并且该算法也引入了与特征数据无关的随机因素 Key。因此,本文与 Biohashing 算法进行相关的算法比较和分析。

图 8—10 中,"Bio"表示 Biohashing 算法的实验结果。用户拥有合法密钥时,在 CMU 库和 Feret 库上,本文提出的算法识别率要高于 Biohashing 算法的识别率,在 AR 库上的识别率略低于 Biohashing 算法的结果。在这 3 个人脸数据库上,Biohashing 算法的识别率随着子空间维数的减少有较大的降低,特别是 CMU 和 Feret 库。实验结果表明,Biohashing 算法对子空间维数的改变较为敏感,对子空间维数的改变鲁棒性较差。

用户所拥有的密钥代表用户的身份,因此,密钥的安全性对整个算法的安全具有至关重要的作用。进一步考虑当用户密钥丢失情况下,本文算法的安全性。

表 1—3 是 CMU 库、AR 库和 Feret 库当待识别用户不具有合法密钥时的平均识别率。即非法用户冒充合法用户,但是没有合法用户的密钥,即图 4 中 $K \neq K'$ 的情形,此时的识别率是越低越好。

表 1　CMU 库识别率(密钥不同)

Tab. 1　CMU database (different key)

	维数	120	110	100	90	80	70	60	50	40
识别率	SIFT	0.008 8	0.010 0	0.014 1	0.002 9	0.005 8	0.009 4	0.008 2	0.005 8	0.008 2
	Bio	—	—	—	—	—	—	0.018 6	0.015 1	0.017 6
	FDA	—	—	—	—	—	—	0.017 6	0.016 7	0.023 5

表 2 AR 库识别率(密钥不同)

Tab.2 AR database (different key)

	维数	120	110	100	90	80	70	60	50	40
识别率	SIFT	0.006 3	0.007 6	0.007 6	0.006 7	0.008 4	0.006 7	0.008 4	0.008 0	0.005 0
	Bio	—	0.011 3	0.010 2	0.006 3	0.007 7	0.008 8	0.007 7	0.007 9	0.009 6
	FDA	—	0.010 5	0.007 6	0.012 6	0.011 3	0.011 8	0.008 4	0.011 3	0.011 3

表 3 Feret 库识别率(密钥不同)

Tab.3 Feret database (different key)

	维数	120	110	100	90	80	70	60	50	40
识别率	SIFT	0.000 5	0.001 8	0.002 4	0.001 3	0.001 0	0.002 4	0.001 6	0.001 3	0.001 3
	Bio	0.002 7	0.004 5	0.002 9	0.003 5	0.004 3	0.003 7	0.003 9	0.006 5	0.003 7
	FDA	0.004 7	0.003 5	0.004 7	0.002 0	0.005 5	0.003 5	0.004 3	0.005 1	0.003 9

从表 1—3 可以看出,在 CMU、AR 和 Feret 人脸数据库中,本文算法的识别率与采用 FDA 特征和 Biohashing 算法相比,是最低的,由此可见,本文算法的安全性也是最高的。

4 结 论

随着人脸识别的广泛应用,特征模板数据的安全性和隐私性问题尤显重要。提出一种与用户密钥相关,能够生成具有可重建性和不可逆性的人脸特征数据的保护算法。采用 SIFT 算法对人脸图像提取特征,然后利用随机投影,引入随机因素(key),利用不同的密钥生成不同的模板数据,从而实现对生物特征数据的保护。实验结果表明,本文提出的算法不仅能进一步提高识别率,而且对人脸图像的姿势、遮挡及表情等变化具有较高的鲁棒性。

参考文献(References)

[1] Ratha N K, Connell J H, Bolle R M. Enhancing security and privacy in biometrics-based authentication systems [J]. IBM Systems Journal, 2001, 40(3):614-634.

[2] Schneier B. Inside risks: the uses and abuses of biometrics [J]. Communications of the ACM, 1999, 42(8):136.

[3] Adler A. Images can be regenerated from quantized biometric match score data [C]//Proceedings of the Canadian Conference on Electrical and Computer Engineering. Washington USA: IEEE, 2004:469-472.

[4] Uludag U, Pankanti S, Prabhakar S, et al. Biometric cryptosystems: issues and challenges [C]//Proceedings of IEEE Conference on Multimedia Information Retrieval. Los Alamitos CA: IEEE Press, 2004, 92(6):948-960.

[5] Cavoukian A, Stoianov A. Biometric Encryption: A Positive-Sum Technology that Achieves Strong Authentication, Security and Privacy [R]. Toronto, Ontario: [s.n.], 2007.

[6] Ratha N K, Chikkerur S, Connell J H, et al. Generating cancelable fingerprint templates [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 9(4):561-572.

[7] Juels A, Wattenberg M. A fuzzy commitment scheme [C]// Proceedings of the 6th ACM Conference on Computer and Communications Security. New York, US: ACM Press, 1999:28-36.

[8] Juels A, Sudan M. A fuzzy vault scheme [C]//Proceedings of IEEE International Symposium on Information Theory. Los Alamitos, CA: IEEE Press, 2002: 408-409.

[9] Savvides M, Kumar B V K V, Khosla P K. Cancelable biometric filters for face recognition [C]//Proceedings of IEEE International Conference Pattern Recognition. Los Alamitos, CA: IEEE Press, 2004(3):922-925.

[10] Teoh A B J, Goh A, Ngo D C L. Random multispace quantization as an analytic mechanism for Biohashing of biometric and random identity inputs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(12):1892-1901.

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

[12] Koenderink J J. The structure of images [J]. Biological Cybernetic, 1984, 50(5): 363-370.

[13] Lindeberg T. Scale-space theory: a basic tool for analyzing structures at different scales [J]. Journal of Applied Statistics, 1994, 21(2): 224-270.

[14] Goal N, Bebis G, Nefian A. Face recognition experiments with random projection [C]//SPIE Defense and Security Symposium. Bellingham WA, US: SPIE Press, 2005:426-437.

[15] Johnson W B, Lindenstrauss J. Extensions of Lipschitz mappings into a Hilbert space [C]//Proceedings of Conference on Modern Analysis and Probability, Contemporary Mathematics. Rhode Island, US: American Math Society, 1984, (26):189-206.

[16] Sim T, Baker S, Bsat M. The CMU Pose, Illumination, and Expression (PIE) Database of Human Faces, CMU-RI-TR-01-02 [R]. Pittsburgh: Robotics Institute, Carnegie Mellon University, 2001.

[17] Martinez A M, Benavente R. The AR Face Database [R]. Barcelona: Computer Vision Center (CVC) Technical Report, 1998.

[18] Phillips P J, Moon H, Rauss P, et al. The facial recognition technology (FERET) database [J]. Lecture Notes in Computer Science, 1993, 42(3): 300-311.