

基于多模板特征点融合的指纹识别

叶学义¹⁾ 庄镇泉¹⁾ 张云超¹⁾ 水孝敏²⁾

¹⁾(中国科学技术大学电子科学与技术系,合肥 230026) ²⁾(信息产业部第38研究所,合肥 230031)

摘要 低质量指纹图像通常会生成较多的伪特征点,或者当待识指纹图像与模板图像之间存在较大平移和旋转时,识别算法的性能会受到很大的影响。利用多匹配算法的融合技术可以提高系统的识别率,但是这种数据融合的方法既增大了计算量,同时又难以充分利用原始指纹图像的信息。针对上述问题,提出了一种评估特征点置信度的方法,并在匹配算法中根据特征点置信度来修正匹配分数;然后运用特征层数据融合技术,将一个指纹的多个模板的特征点融合成一个特征点的本征模板,再与待识指纹图像的特征点进行匹配。实验结果表明,该方法在不明显增加存储空间和在线计算量的同时,可以达到很好的识别性能。

关键词 指纹识别 特征层数据融合 伪特征点 本征模板

中图分类号: TP391.4 **文献标识码**: A **文章编号**: 1006-8961(2005)02-0170-05

Fingerprint Identification Based on Multi-templet Characters Fusion

YE Xue-yi¹⁾, ZHUANG Zhen-quan¹⁾, ZHANG Yun-chao¹⁾, SHUI Xiao-ming²⁾

¹⁾(The Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230026)

²⁾(The Institute 38 of the Information Industry Bureau of China, Hefei 230031)

Abstract Presently, most fingerprint identification algorithms can not effectively work on some occasions, for instance, in which there are many false characters in a low quality fingerprint image, or existing place large motivation and rotation between a template fingerprint image and a compared fingerprint image. Although the technology of multi-algorithm fusion can improve the true rate of an identification system, it will occupy much more calculating time and chunk away some information of an original fingerprint. This paper, firstly, for false characters in a fingerprint image, proposes a method to evaluate the credibility of each character and continuously to modify their matching scores by the character's credibility in the identification process and secondly, for large motivation and rotation, fuses a set of templates into a character's eigen template before the matching process. The experimental results show that, by the method of character's layer data-fusion, the algorithm can achieve a better identification performance, while it does not apparently increase the memory occupancy and the online-calculating-time.

Keywords fingerprint identification, character's layer data-fusion, false characters, eigen template

1 引言

指纹图像在采集、增强、二值化、细化等过程中都可能由于采集设备、噪声或者其他原因而产生伪特征点。图1是标注了特征点的低质量指纹图像,其中,用浅色方框标出的都是伪特征点。

指纹图像的伪特征点目前并没有严格的定义。一般认为,由于各种原因产生的位置错误、类型不对的特征点是伪特征点。出于对伪特征点产生的最大概率考虑,伪特征点定义为在图像采集和处理过程中受到随机因素影响而产生的、不能稳定重复出现的特征点。伪特征点对识别算法的不利影响主要表现在伪特征点可能会导致错误的中心点对齐;伪特

基金项目:国家自然科学基金项目(90104030)

收稿日期:2004-02-25; **改回日期**:2004-06-15

第一作者简介:叶学义(1973~),男,1995年于电子科技大学获得自动控制专业学士学位。现为中国科学技术大学电子科学与技术系博士研究生。主要研究方向为图像处理、生物特征识别、信息安全。E-mail: xueyi_ye@ustc.edu

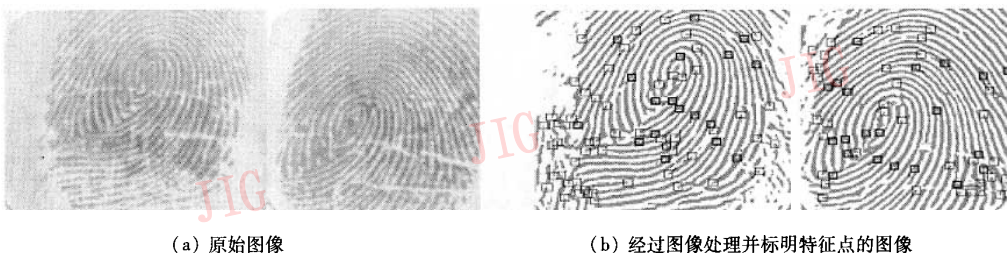


图1 标注了特征点的低质量指纹图像

Fig.1 Characters of the low-quality fingerprint

征点会影响待识别者的匹配分数,导致错误识别结果;伪特征点给匹配算法增加了不必要的计算量。

所以,一些识别算法采取各种思路来考虑对伪特征点的处理,如在提取特征后,利用统计和结构信息的方法来删除虚假特征^[1]。然而,这些方法在删除虚假特征点的同时也会删除一些真正的特征点,而且对于模糊和低质量的指纹图像鲁棒性差。

另外,在待识图像与模板图像之间存在较大平移和旋转的时候,由于待识图像与模板图像之间重合的特征点很少,识别算法会产生较高的误拒率(FRR)。有人提出用多种匹配算法融合的方法来提高这种情况下识别系统的性能,或者尝试利用多印痕指纹识别的数据融合技术来解决该问题^[2]。但是,一方面这两类方法都会不同程度地增加识别的计算量;另一方面,它们都是决策层水平的数据融合技术^[3],难以充分利用原始图像信息。为了充分利用图像信息,可以在特征层应用数据融合技术。但是在指纹识别的研究中,特征层的数据融合技术一直缺乏有效的方法。

针对上述问题,采用将多个模板的特征点融合成一个特征点模板并且存储起来,然后再与待识图像的特征点进行匹配。而对于伪特征点的问题,提出了一种利用投票表决方法来评估特征点置信度的方法,并在匹配算法中根据特征点置信度来修正待识别图像的匹配分数,从而提高算法对于低质量指纹图像的识别性能。实验结果证明,这种特征层数据融合技术,在不明显增加存储空间和在线计算量^[4]的情况下,大大提高了系统的识别性能。

2 多特征模板的融合

令 I 表示待识指纹图像, $T_i (i=1, 2, \dots, n)$ 表示第 i 幅模板指纹图像。首先选取 n (在试验中取 $n =$

2) 幅来自同一指纹的多次采集的图像作为模板;再以其中第 1 幅模板图像 T_1 (通常选择比较居中,没有很大旋转的图像) 作为对齐的基模板,其他 $n-1$ 幅图像与第 1 幅图像对齐;完成对齐以后,经过特征点的融合处理,以及计算各特征点的置信度;最后形成标识该指纹的特征点图像本征模板 T_0 。

2.1 对齐

要实现特征模板的融合,首先要把 $T_i (i=2, \dots, n)$ 与 T_1 对齐。对齐通常是以一个假定的特征点对 (P_{ref}, Q_{ref}) 来实现的,称为基准点对(reference point pair)或参考特征点对^[5]。针对寻找基准点对存在各种不同的方法,有些算法首先定位奇异点(例如中心点和三角点),并以为之基准点。这类算法的缺点是:依赖于中心点定位的准确;对于某些指纹采集器所采集的指纹来说,由于采集区域相对较小,常常会出现无法找到中心点的情况。而一个优秀、健壮的识别算法应当与指纹图像的奇异点无关。

另外,还有一些算法通过比较与特征点相关联的脊线来寻找参考特征点,可以称之为基于脊线对齐的基准点寻找方法^[6]。它可以通过比较两条脊线,并按下式计算相似度 S ,再和给定的阈值比较,确定基准点对。

$$S = \frac{\sum_{i=1}^L d_i D_i}{\sqrt{\sum_{i=1}^L d_i^2 D_i^2}} \quad (1)$$

式中, L 是脊线上所取像素的个数, d_i 和 D_i 分别表示进行对齐的模板脊线上第 i 个像素到坐标系 x 轴的距离上,如图 2(a) 所示。

也可以利用一种更简单直观的公式来比较两条脊线的差异:

$$Diff_{dist} = \frac{1}{L} \sum_{i=0}^L |R(d_i) - r(d_i)| \quad (2)$$

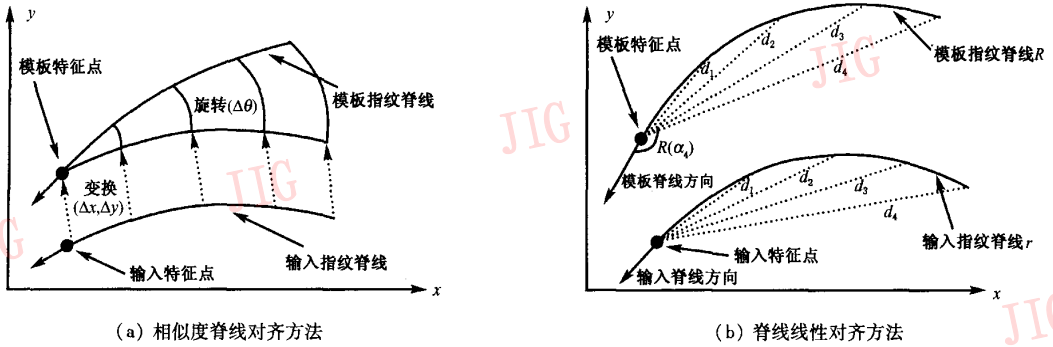


图 2 两种基于脊线标准的对齐方法
Fig. 2 Two ridge-adjusted normalization methods

$$Diff_{ang} = \frac{1}{L} \sum_{i=0}^L |R(\alpha_i) - r(\alpha_i)| \quad (3)$$

式(2), (3)中 L 是记录脊线中点的个数, $R(d_i)$ 和 $r(d_i)$ 分别表示从脊线 R 与 r 上的点 i 到各自对应的特征点距离, $R(\alpha_i)$ 和 $r(\alpha_i)$ 分别表示连接脊线 R 与 r 上的点 i 和对应特征点的直线同相应特征点方向的夹角, 如图 2(b) 所示。

基于脊线校准的对齐算法的缺点在于在对齐阶段中需要用到相关的脊线信息, 对于每个特征点都至少需要附加保存 L 个脊线上像素点的坐标, 导致存储模板所需空间较大。

利用特征点模板来代替指纹图像, 只描述特征点的位置、方向、类型等参数信息, 并由此寻找两个特征点模板的对齐方法。

当特征提取过程完成后, 得到的是一些散布在 2 维空间上的点集。如图 3 所示, 一个特征点可以用一个特征向量 (x, y, θ, t) 表示, 其中, (x, y) 是坐标, θ 是特征点方向, t 是特征点类型, 则特征点集合可以看成是一个四元组 (x, y, θ, t) 的集合。

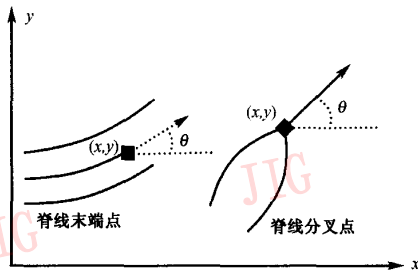


图 3 特征点信息

Fig. 3 Two types of fingerprint characteristics

首先考察特征点 k 周围的 l 个邻域特征点的集合, 并根据下面的式(4) ~ 式(11)计算它们之间的

关系:

$$d_{ki} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \quad (4)$$

$$\varphi_{ki} = \tan^{-1} \left(\frac{y_k - y_i}{x_k - x_i} \right) \quad (5)$$

$$\Phi_{ki} = D(\theta_k, \theta_i) \quad (6)$$

$$D(\theta_k, \theta_i) = \begin{cases} |\theta_k - \theta_i| & -\pi < \theta_k - \theta_i \leq \pi \\ 2\pi + (\theta_k - \theta_i) & \theta_k - \theta_i \leq -\pi \\ 2\pi - (\theta_k - \theta_i) & \pi < \theta_k - \theta_i \end{cases} \quad (7)$$

式中, d_{ki} 是特征点 k 与特征点 i 的距离, φ_{ki} 和 Φ_{ki} 描述了特征点 k 与特征点 i 的角度差别。为了描述简练, 当取 $l=2$ (i 和 j) 时, 则对每个特征点, 可以用如下的邻域关系向量描述:

$$F_k^{(l)} = (d_{ki}, d_{kj}, \varphi_{ki}, \varphi_{kj}, \Phi_{ki}, \Phi_{kj}, t_k, t_i, t_j) \quad (8)$$

并且用下式来比较两个特征点的相似程度:

$$S^{(l)}(i, j) = \begin{cases} \text{如果 } W \cdot |F_k^{(l)}(T_i) - F_k^{(l)}(T_j)| < b^{(l)} \\ \frac{b^{(l)} - W \cdot |F_k^{(l)}(T_i) - F_k^{(l)}(T_j)|}{b^{(l)}} \\ \text{其他} \\ 0 \end{cases} \quad (9)$$

$$W = (w_d, w_d, w_\varphi, w_\varphi, w_\Phi, w_\Phi, w_t, w_t, w_t) \quad (10)$$

W 为加权向量, 表示邻域关系向量中不同分量在比较中所占权重, 依据经验取值为

$$w_d = 1; w_\varphi = w_\Phi = 0.3 \times \frac{180}{\pi}; w_t = 1 \quad (11)$$

阈值 $b^{(l)}$ 设为 $6 \times m$, 其中, m 设为邻域关系向量的维数。 $S^{(l)}(i, j)$ 的值描述了两个模板假定对应特征点匹配成功的程度。

如果特征点 b_1, b_2 是使得 $S^{(l)}(b_1, b_2) = \max(S^{(l)}(i, j))$ 的点, 则以点对 (b_1, b_2) 为找到的基准点对 (P_{ref}, Q_{ref}) 。

2.2 特征点融合

找到基准点对($P_{\text{ref}}, Q_{\text{ref}}$)以后,根据下式求出平移因子和旋转因子:

$$\Delta x = x_p - x_q \quad (12)$$

$$\Delta y = y_p - y_q \quad (13)$$

$$\Delta \theta = \theta_p - \theta_q \quad (14)$$

式中 $x_p, y_p, \theta_p, x_q, y_q, \theta_q$ 分别是 P_{ref} 和 Q_{ref} 的 x 轴坐标, y 轴坐标, 脊线方向。然后根据下式把 $T_i (i=2, \dots, n)$ 中的特征点信息转换到 T_1 的坐标系中。

$$\begin{pmatrix} x'_{ij} \\ y'_{ij} \end{pmatrix} = \begin{pmatrix} \Delta x_i \\ \Delta y_i \end{pmatrix} + \begin{pmatrix} \cos(\Delta \theta_i) & \sin(\Delta \theta_i) \\ \sin(\Delta \theta_i) & -\cos(\Delta \theta_i) \end{pmatrix} \begin{pmatrix} x_{ij} \\ y_{ij} \end{pmatrix} \quad (15)$$

$$\theta'_{ij} = \theta_{ij} - \Delta \theta_i \quad (16)$$

式中, $\begin{pmatrix} \Delta x_i \\ \Delta y_i \end{pmatrix}$ 表示图像模板 $T_i (i=2, \dots, n)$ 与图像模板 T_1 之间的平移因子, $\Delta \theta_i$ 表示 T_i 与 T_1 之间的旋转因子, $\begin{pmatrix} x_{ij} \\ y_{ij} \end{pmatrix}$ 表示 $T_i (i=2, \dots, n)$ 中特征点 j 的坐标, $\begin{pmatrix} x'_{ij} \\ y'_{ij} \end{pmatrix}$ 表示 $T_i (i=2, \dots, n)$ 中特征点 j 的转换后的坐标, θ_{ij} 和 θ'_{ij} 分别是转换前后 $T_i (i=2, \dots, n)$ 中特征点 j 的脊线方向。

转换以后,如果 $\begin{pmatrix} x'_{ij} \\ y'_{ij} \end{pmatrix}$ 与 T_1 中某个特征点 k 的距离小于界限盒阈值 B_g (取 $B_g = 10$), 则认为 $T_i (i=2, \dots, n)$ 中特征点 j 与 T_1 中某个特征点 k 是同一特征点, 用它们的坐标和角度平均值作为该特征点 k 的坐标和脊线方向。多模板的特征点融合虽然需要相当于进行匹配的计算量, 但是该过程可以在离线情况下完成, 属于训练和学习过程, 实时匹配时并不增加在线计算量。

2.3 特征点置信度

在模板的特征点融合过程中, 同时进行特征点置信度的评估。模板 T_1 的特征点 k 的置信度初始值设为 1; 如果, 在融合时 T_i 中该特征点 k 与模板 $T_i (i=2, \dots, n)$ 中特征点 j 匹配, 则把该特征点置信度加 1; 如此累加直至特征本征模板 T_0 形成。特征点置信度最大值是 n , 最小值是 1。根据特征点置信度的评估方法可以得出, 特征点的置信度反映了特征点是该指纹真实特征点的可信度。特征点的置信度值越大, 该特征点是真实特征点的可信度越高。

3 匹配

设 (b_1, b_2) 是待识图像 I 与融合后本征模板 T_0 的参考基准点对, 分别以 b_1, b_2 为中心, 将各自坐标转换到极坐标下。

$$F_{g,k} = \begin{pmatrix} r_{kb} \\ \varphi_{kb} \\ \theta_{kb} \end{pmatrix} = \begin{pmatrix} \sqrt{(x_k - x_b)^2 + (y_k - y_b)^2} \\ \tan^{-1} \left(\frac{y_k - y_b}{x_k - x_b} \right) \\ D(\theta_k, \theta_b) \end{pmatrix} \quad (17)$$

比较两个特征点:

$$m^{(l)}(i, j) = \begin{cases} \text{如果 } |F_{g,i}(I) - F_{g,j}(T)| < B_g \\ \frac{0.5 + 0.5 C_j^T S^{(l)}(i, j)}{n} \\ \text{其他} \\ 0 \end{cases} \quad (18)$$

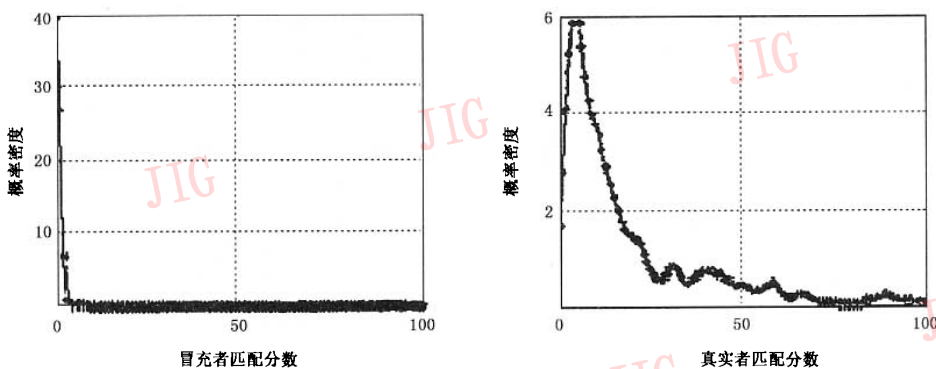
式中, C_j^T 是融合后模板中特征点 j 的置信度, n 是待识图像 B_g 盒中的特征点个数。在以 B_g 盒为单位的匹配过程中, 为了避免一个特征点被多次选中配对, 特别是对于在 B_g 盒的相邻区域中的特征点, 如果待识别图像相邻的 B_g 盒中存在特征点 k 使得 $m^{(l)}(k, j) > m^{(l)}(i, j)$, 则令 $m^{(l)}(i, j) = 0$, 同时标记点 k 。下式比较两个图像模板的特征点集合, 并计算出归一化后的相似度 $M_s, M_s \in [0, 100]$ 。

$$M_s = \frac{100 \sum_{i,j} m^{(l)}(i, j)}{\max(M, N)} \quad (19)$$

这里, M, N 分别是待识图像特征点集和融合模板特征点集的点数。如果 $M_s \geq \lambda$ (λ 是给定的判定阈值), 则认为两个指纹匹配, 否则认为不匹配。

4 实验结果

在实验中选择了包含 80 个指纹图像的 FVC2000 DB2_B 指纹数据库来测试这个算法的性能。FVC2000 指纹数据库中每个手指都有 8 幅指纹图像。每个手指的前面 3 个图像用于构造融合本征模板, 其余 5 个用于匹配。每个融合本征模板与不是同一个手指的图像一一配对组成虚假配对组合(冒充者)。这样, 真实者组合共有 $g = 10 \times 5 = 50$ 种, 冒充者共有 $f = 5 \times 9 \times 10 = 450$ 种。实验结果表明, 相似度 M_s 分布如图 4 所示。

图 4 冒充者和真实者的 M_s 分布Fig. 4 The M_s distribution

根据实验结果统计出来的 ROC 曲线如图 5 所示。

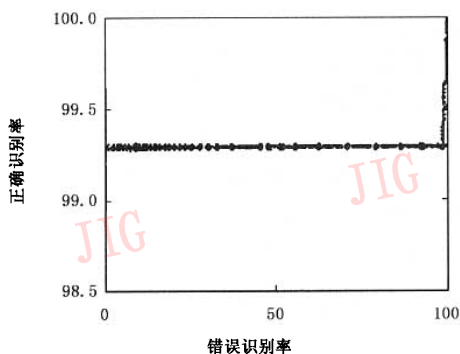


图 5 ROC 曲线

Fig. 5 The receiver operating characteristic curve of the test

5 结论

从图 4 和图 5 可以看出,多模板特征点融合的指纹识别算法使冒充者的匹配分数很低,而真实者匹配分数较高。所以系统即使在零误识率的时候,误拒率也很低(真实接受率很高)。通过把多个同一指纹图像的特征点融合成一个特征点本征模板并且计算各特征点的置信度,然后再与待识图像特征

点进行匹配,能够有效克服由于低质量图像等因素造成的伪特征点的影响。实验结果证明,这种在特征层应用数据融合技术的方法,在不明显增加存储空间和在线计算量的条件下,提高了系统的识别性能。

参考文献 (Reference)

- Xiao Q, Raafat H. Fingerprint image post processing: A combined statistical and structural approach[J]. Pattern Recognition, 1991, 24(10):985 ~ 992.
- Zhang Yun-chao, Zhuang Zhen-quan, Ye Xue-yi. Multi-impression data fusion for fingerprint verification[J]. Computer Engineering and Application, 2003, 39(31):52 ~ 55. [张云超,庄镇泉,简兵等.多印痕指纹识别的数据融合技术[J].计算机工程与应用,2003,39(31):52 ~ 55.]
- Li X R, Zhu Yunmin. Unified optimal linear estimation fusion, [J]. IEEE Transactions on Information Fusion, 2003, 49(9):2192 ~ 2208.
- Jain A K, Hong L, Bolle R. On-line fingerprint verification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(4):302 ~ 314.
- Luo X, Tian J, Wu Y. A minutia matching algorithm in fingerprint verification[A]. In: International Conference on Pattern Recognition (ICPR '00)[C], Barcelona, Spain. 2000, IV: 833 ~ 836.
- Bedworth M, O'Brien J. The Omnibus model: a new model of data fusion? [J]. Aerospace and Electronic Systems Magazine, 2000, 15(4):30 ~ 36.