

基于多尺度序列谱核半定规划优化的 签名认证方法

邱德红,潘昕昕,陈传波
(华中科技大学软件学院,湖北武汉 430074)

摘 要: 变化尺度进行签名序列的相似性描述有利于获得更准确的相似性描述结果. 本文定义了签名序列的变化尺度的谱核矩阵,在多个尺度的核变换空间上进行序列的相似性描述,并利用半定规划对多尺度谱核矩阵进行优化,结合支持向量机建立起一种能够自动优化签名序列多尺度相似性描述的认证方法. 该方法能够适应不同个人的签名特点,克服统一尺度下相似性描述的缺陷,提高签名序列相似性描述的准确性,在相同签名数据集上的实验结果显示该方法可以获得更高的认证准确率.

关键词: 半定规划; 序列谱核; 签名认证

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2008) 12A-044-06

Multi-Scale Sequence Spectrum Kernel Optimization Through SDP for Signature Verification

QIU De-hong, PAN Xin-xin, CHEN Chuan-bo

(School of Software Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China)

Abstract: Using multi-scale to measure the similarities between signature sequences is much helpful to obtain high-qualified similarity measures. This paper puts forward a new sequence similarity kernel, called multi-scale sequence spectrum kernel, to measure signature sequence similarity based on shared occurrences of different-scale continuous subsequences. The multi-scale sequence spectrum kernel is optimized through semidefinite program, and used with support vector machine to verify signature directory sequences. The experiments on the benchmark database show the signature verification accuracy has been enhanced, as this approach could automatically optimize the similarity measures with multi-scales, depending on the signature characteristics.

Key words: semidefinite program; sequence spectrum kernel; signature verification

1 引言

手写签名认证是生物测定学一项重要的研究内容^[1,2],它利用签名过程的动态信息进行身份认证,具有伪造困难、安全性高、认证方式易于接受和使用方便等优点,是一种比较理想的个人身份认证方法. 然而,由于各人签名熟练程度的差异,以及签名时可能受到不同环境因素的影响,如何准确的描述签名动态信息序列之间的相似性,一直是影响认证准确率的一个关键因素. 相关文献对于签名动态信息的稳定性^[3]、时间弯曲现象^[4,5]、局部特征和全局特征均衡融合^[6]、签名认证方法^[7,8]等做了不少的工作,测试结果也比较良好. 但重复使用这些方法认证不同个人的签名时,可以发现同样的认证方法对于不同个人的签名认证效果有较大的波

动. 原因之一是在描述不同个人的签名动态信息序列的相似性时采用了固定的统一的尺度,不能适应由于签名熟练程度差异和环境因素不同的影响,采用变化尺度进行签名动态信息序列的相似性描述,有望获得更好的签名认证结果.

签名序列的相似性可以通过某种距离来反映,常见的有以下三种距离:原始空间距离^[3,4]、特征空间距离^[5,8]和模型距离^[9,10]. 原始空间距离指的是签名序列在原始输入空间上的距离 $S_i - S_j$; 特征空间上的距离指的是签名序列经过某个函数 $f(\cdot)$ 变换后在特征空间上的距离 $f(S_i) - f(S_j)$; 模型距离指的是签名序列的生成模型之间的距离,比如签名序列概率生成模型 $S_i - P_i$ 之间的距离 $P_i - P_j$. 由于签名序列的长度不一致,生成机理尚不清楚,第一和第三种距离在描

述签名序列相似性时有较大的难度和不确定性,特征空间距离 $(S_i) - (S_j)$ 得到了广泛的关注,核变换技术便是其中最受关注的技术之一,它通过核函数

() 将样本从原空间变换到某个内积特征空间,然后在该内积特征空间上来描述样本的像之间的距离. 它的优越之处在于计算时无需显式定义核变换函数 (), 只需要保证特征空间上样本的像之间的距离形成正定对称矩阵即可,具有很强的灵活性和计算效率.

如何获得最有效的核相似性描述引起了人们浓厚的研究兴趣^[11~13]. 本文通过对核变换矩阵的学习来克服统一尺度对签名序列相似描述的不利影响,首先在多种尺度的核变换空间上描述签名序列的相似性,然后利用半定规划方法^[14]对多种尺度下签名序列的谱核矩阵进行优化,建立起因人而异的签名序列相似性描述方法,相似性优化结果与支持向量机相结合生成的签名认证方法能够自觉适应不同签名者的书写特征,在相同数据集上的签名认证实验获得了更高的准确率.

2 签名序列及其多尺度谱核矩阵

2.1 签名序列

签名的位移序列是签名认证的重要动态信息,基于此可以湮生出多种所需要的动态信息序列,其中签名趋势序列记录了签名轨迹的方向变化信息,能够很好的反映签名者的书写习惯和熟练程度,是有效的签名认证信息. 数据集上的原始签名,经过去噪、时间等距插值、正规化等预处理之后,按照其轨迹的运动方向进行离散,便可以获得签名的趋势序列. 这里将签名趋势在 8 个方向进行离散化,分别用数字 0~7 来表示,从而生成字符集合 $S = \{0, 1, 2, 3, 4, 5, 6, 7\}$ 上的签名趋势序列. 具体的签名趋势序列生成方法可以参考文献 [8], 图 1 所示的是 Mario's 签名数据集^[4]上两个签名的位移序列及其对应的符号趋势序列.

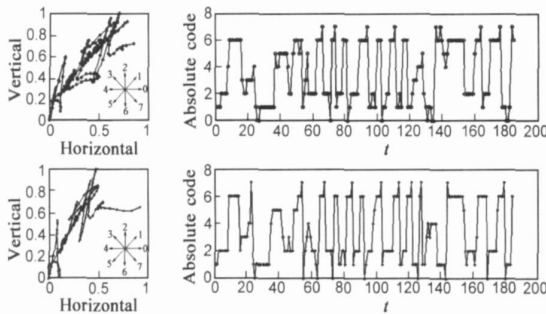


图1 签名位移和对应的签名趋势序列(来自Mario's签名数据库)

2.2 多尺度谱核矩阵

签名序列 S_i 的尺度为 l 的谱定义为它所包含的长度为 l 的连续子序列的集合. 变换函数定义为^[15]:

$${}_l(S_i) = (\phi_a(S_i))_{S^l}$$

其中 a 是由字符集 S 中的字符组成的长度为 l 的连续子序列, S^l 表示长度为 l 的所有连续子序列集合, $\phi_a(S_i)$ 表示签名序列 S_i 中连续子序列 a 出现的次数. 变换函数 ${}_l(S_i)$ 的特征空间维数为 $|S^l|$, S 表示字符集 S 中的字符数目. 尺度为 l 的签名序列谱核矩阵定义为:

$$K_l(S_i, S_j) = {}_l(S_i), {}_l(S_j)$$

由于各人签名的熟练程度不同,不同时刻签名受到的影响因素也可能不一致,不宜采用固定的尺度来描述签名序列的相似性. 变化尺度进行相似性描述有利于兼顾不同稳定程度的签名动态特征. 不同尺度 l 下生成的签名序列谱及其相应的谱核矩阵 K_l 反映出签名序列在不同尺度下的相似性信息,通过对它们的优化可以客观真实的反映签名序列之间的相似程度. 图 2 所示的是尺度 $l=3, 4, 5$ 时, 图 1 中的两个签名趋势序列的多尺度谱. 从中可见,不同尺度下趋势序列谱线的相似性有所差异,很难判断哪个尺度下的描述更接近真实相似程度,因此将多个尺度下的相似性描述综合起来,定义多尺度谱核矩阵为:

$$K_L: K_L(S_i, S_j) = \sum_{m=1}^{|L|} \mu_m K_l(S_i, S_j)$$

其中, L 是尺度的集合, $|L|$ 表示尺度的种数, K_L 是不同尺度下谱核矩阵 K_l 的线性组合,由核变换性质可知 K_L 依然是一个核矩阵,它包含了多个尺度下的相似信息,利用半定规划方法可以确定多尺度融合的最优核矩阵 K_L^* .

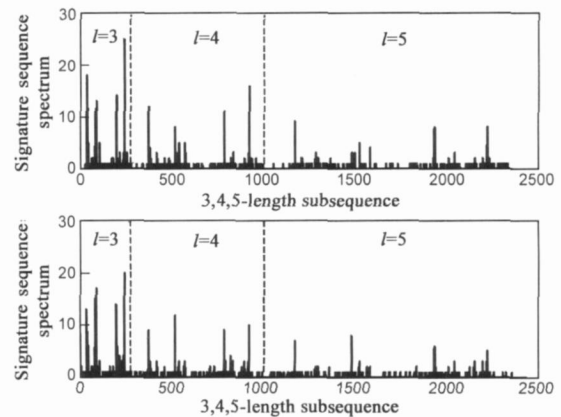


图2 签名序列的多尺度L=(3,4,5)序列谱

3 多尺度谱核半定规划优化

3.1 半定规划

定义 1 设集合 $Z \subseteq R^n$, 如果 $\forall z_1, z_2 \in Z$ 和 $\lambda \in [0, 1]$, 有 $\lambda z_1 + (1 - \lambda) z_2 \in Z$, 则称 Z 为凸集.

定义 2 设集合 $Z \subseteq R^n$, 如果 $\forall z_1, z_2 \in Z$ 和 $\lambda_1, \lambda_2 \in [0, 1]$, 有 $\lambda_1 z_1 + \lambda_2 z_2 \in Z$, 则称 Z 为凸锥.

定义 3 设集合 $Z \subseteq R^n$ 凸集, 函数 $f: Z \rightarrow R$, 如果

$\forall z_1, z_2 \in Z$ 和 $\lambda \in [0, 1]$, 有 $f(\lambda z_1 + (1 - \lambda) z_2) \leq \lambda f(z_1) + (1 - \lambda) f(z_2)$, 则称 $f(z)$ 为凸集 Z 上的凸函数.

半定规划的标准数学形式如下^[14]:

$$\begin{aligned} \min \quad & \text{trace}(C X) \\ \text{s.t.} \quad & \text{trace}(A_i X) = b_i, i = 1, \dots, q \\ & X \succeq 0 \end{aligned}$$

其中, $\text{trace}(C X) = \sum_{i,j=1}^n C_{i,j} X_{i,j}$, 变量 $X \in \mathbf{P}, \mathbf{P} = \{X \in \mathbf{R}^{n \times n} | X = X^T, X \succeq 0\}$, $C, A_1, \dots, A_q \in \mathbf{P}, b_i \in \mathbf{R}$, 变量 X 受线性等式约束和半正定约束. 由半定规划标准数学形式和定义 1-3 可知, 半定规划实质上是一个凸锥上的凸函数的优化问题.

3.2 多尺度谱核的优化组合

序列多尺度谱核矩阵 K_L 属于集合 $\mathbf{K}: \mathbf{K} = \{K \in \mathbf{R}^{n \times n} | K = K^T, K \succeq 0\}$, 可以证明, \mathbf{K} 是一个凸锥, 有定理如下:

定理 1 $\mathbf{K} = \{K \in \mathbf{R}^{n \times n} | K = K^T, K \succeq 0\}$ 是一个凸锥.

将 K_L 引入到支持向量机, 核化之后支持向量机的向量形式为^[16]:

$$\begin{aligned} \max \quad & 2 \gamma - \gamma^T G(K_L) \\ \text{s.t.} \quad & C \succeq 0, \gamma^T y = 0 \end{aligned}$$

其中, $\mathbf{e} \in \mathbf{R}^n$ 为 n 维单位向量, $G(K_L)$ 定义为 $G_{i,j}(K_L) = K_L(x_i, x_j) y_i y_j$, $C \succeq 0$ 表示 $C_{ii} > 0$. 考虑以 K_L 为变量的函数 $c(K_L)$:

$$\begin{aligned} c(K_L) &= \max_{\gamma} 2 \gamma - \gamma^T G(K_L) : C \succeq 0, \gamma^T y = 0 \\ &= \max_{\gamma} 2 \gamma - \gamma^T G(K_L) : C \succeq 0, \gamma^T y = 0 \end{aligned}$$

可以证明 $c(K_L)$ 是 K_L 的凸函数, 有定理如下:

定理 2 对于变量 $K_L \in \mathbf{K}$, 函数 $c(K_L) = \max_{\gamma} 2 \gamma - \gamma^T G(K_L) : C \succeq 0, \gamma^T y = 0$ 是凸函数.

由定理 1 和定理 2, 且约束 $C \succeq 0, \gamma^T y = 0$ 是凸的, 可以推断, 在 \mathbf{K} 中求最优核矩阵 K_L^* 使得 $c(K_L)$ 最小的问题可以看成是凸锥上的凸函数的优化问题(由于支持向量机的边界距离 $\gamma = \frac{1}{W_*}$, 求 $\min_{K \in \mathbf{K}} c(K)$ 等价为使分类边界距离最大且分类误差之和 $\sum_{i=1}^n C_{ii} x_i$ 最小), 因此可以表达成半定规划标准数学形式, 有定理如下:

定理 3 给定一组训练样本 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, 其中 $x_i, i = 1, \dots, n$ 是第 i 个样本向量, $y_i \in \{-1, +1\}$ 表示该样本的类别, $y \in \mathbf{R}^n$, 优化问题 $\min_{K \in \mathbf{K}} c(K)$ s.t. $\text{trace}(K) = c, K_L: K_L(S_i, S_j) = \sum_{m=1}^{L_i} \mu_m K_m(S_i, S_j) \in \mathbf{K}, c$ 为常数, 可表达成以下半定规划问题:

$$\begin{aligned} \min_{\mu_1, \dots, \mu_L} \quad & \text{trace}(\sum_{m=1}^{L_i} \mu_m K_m) = c, \mu_m \in \mathbf{K} \\ \text{s.t.} \quad & \begin{pmatrix} G(\sum_{m=1}^{L_i} \mu_m K_m) & \mathbf{e} + \mathbf{v} - \mathbf{y} \\ (\mathbf{e} + \mathbf{v} - \mathbf{y})^T & t - 2C^T \mathbf{e} \end{pmatrix} \succeq 0 \\ & \mathbf{v} \succeq 0, t \geq 0, \mu \succeq 0 \end{aligned}$$

其中 $\mu = (\mu_1, \dots, \mu_{L_i})$. 由此可知, 半定规划可以求解多尺度序列谱核矩阵 K_L 的优化问题, 优化结果 K_L^* 引入到支持向量机中, 可以生成基于签名序列多尺度相似性优化描述的签名认证器.

图 3 是三种尺度 $L = (3, 4, 5)$ 下的序列谱核矩阵及其优化结果 $K_L^* = \mu_1 K_3 + \mu_2 K_4 + \mu_3 K_5$, 从图上可见, 小尺度 $l=3$ 下的描述结果 K_3 夸大了训练样本之间的相似性(前 15 个为真签名, 后 5 个为伪造签名), 大尺度 $l=4, 5$ 下的描述结果 K_4, K_5 模糊了训练样本之间的相似性, 而优化结果 K_L^* 比较真实地描述了样本之间相似性.

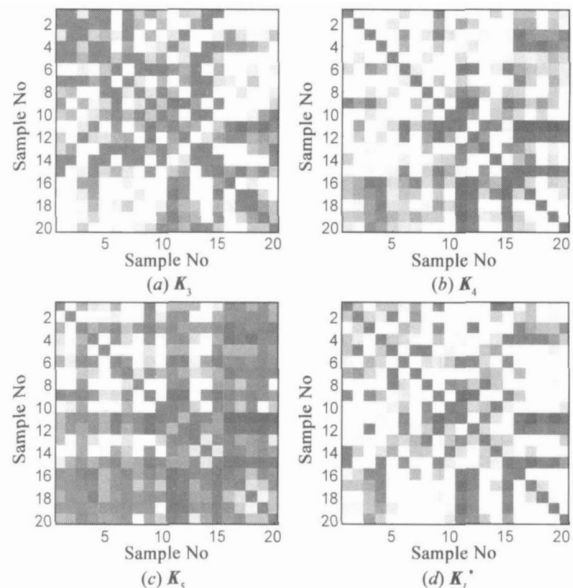


图 3 多尺度 $L=(3,4,5)$ 序列谱核矩阵及其优化结果, 对应 $K_L^* = \mu_1 K_3 + \mu_2 K_4 + \mu_3 K_5$ 中的参数为 $(\mu_1^*, \mu_2^*, \mu_3^*) = (0.66265, 0.087352, 6.9008e-012)$

4 签名认证实验与分析

4.1 认证实验

为了方便比较结果, 签名认证实验在 Mario's 签名数据集^[4]上进行, 该数据集共收集了 50 人的签名, 每人的签名数据中包含有 25 个真签名和 10 个伪造签名. 我们选择前 15 个真签名和前 5 个伪造签名为训练样本, 其余的为测试样本. 采用三种尺度 $L = (3, 4, 5)$ 生成训练样本趋势序列的谱核矩阵并进行半定规划优化, 优化结果 K_L^* 用于核化支持向量机, 生成签名认证分类

器.实验记录真签名判断为伪造签名 ($T = F$) 和伪造签名判断为真签名 ($F = T$) 的数目,对应定义误拒率(FRR)和误纳率(FAR)两类误差,分别为上述两类错误的签名数目与总测试签名数目之比.

实验中序列谱的计算效率十分关键,因为随着尺度 l 的增大,序列谱的空间维数 S^l 递增十分迅速.如果采用蛮力法计算序列谱,即每计算一个子序列的出现次数,搜索一次长度为 l 的签名序列,那么穷尽 S^l 中所有的子序列得到序列谱的计算效率为 $O(|S^l|)$.显然,在尺度 l 和字符集 S 较大时,计算效率将严重下降.文献[15]设计了前缀树这种较为复杂的数据结构来存储序列,通过牺牲存储空间提高序列谱的计算效率至 $O(l|S|)$.本文设计了一种更为简单有效的序列谱计算方法,且不需要特殊的数据结构.其思想是将每个子序列看成是 l 位的 S 进制的整数,该整数值即是对应的子序列在序列谱中的位置,比如在 $S = 8$,尺度 $l = 3$ 时子序列“111”在序列谱中的位置为 73,直接递增该位置的谱值即可,从而将计算效率提高到 $O(|S|)$,并且避免了尺度变化时文献[15]中的方法面临的数据结构空间的变化分配问题.

半定规划问题有比较有效的求解算法,实验中我们用开源的基于 Matlab 的半定规划求解软件 SeDuMi 及其友好应用界面 SeDuMi Interface 来求解^[17,18],SeDuMi 求解的计算复杂度为 $O(q^2 n^{2.5})$,其中 q 为半定规划数学标准形中等式约束的数目, n 为方阵的维数^[19].

4.2 结果分析

表 1 给出了数据集中 50 位签名者多尺度谱核矩阵半定规划优化结果和测试签名的认证结果.从优化结果 μ_1^* 、 μ_2^* 和 μ_3^* 可以看出,大多数情况下,某个尺度下的签名序列相似性描述最合适,说明这些签名者形成了相对稳定的签名动态特征.但对于不同签名者,最适合的相似描述尺度可能有所不同,比如对于 1 号签名者最为合适的尺度是 $l = 3$,2 号签名者最为合适的尺度是 $l = 5$,20 号签名者最为合适的尺度是 $l = 4$.也有一些签名者的签名序列,需要综合不同尺度来描述其相似性,比如 22 号签名者需要综合 $l = 3$ 和 $l = 4$ 两种尺度,而 48 号签名者需要综合 $L = (3, 4, 5)$ 三种尺度,反映出这些签名者的动态特征有所波动.由此可见,本方法可以根据不同的签名者的签名动态特征自动决定相似性的描述尺度,为获得好的认证结果提供了前提条件.

表 1 50 位签名者数据的认证结果

No	c	2^l	μ_1^*	μ_2^*	μ_3^*	T	F	F	T	Sum	
1	10	0.1	0.5	3.00E-13	1.85E-13	0	0	0	0		
2	10	2	5.09E-14	8.48E-14	0.5	0	0	0	0		
3	10	0.1	0.52	5.4E-13	1.65E-13	0	0	0	0		
4	10	0.2	0.417	0.082999	1.23E-12	0	0	0	0		
5	10	0.1	0.5	2.08E-13	1.45E-13	0	0	0	0		
6	10	5.9	4.22E-15	3.85E-13	0.5	0	0	0	0		
7	10	0.1	0.5	1.72E-12	9.31E-13	0	0	0	0		
8	10	0.3	0.5	3.71E-12	2.36E-12	0	1	1	1		
9	10	0.1	0.5	1.64E-13	1.05E-13	0	0	0	0		
10	10	0.1	0.5	4.13E-13	2.26E-13	0	0	0	0		
11	10	0.3	0.39275	6.91E-12	0.10725	0	0	0	0		
12	10	0.1	0.5	1.18E-13	7.97E-14	0	0	0	0		
13	10	0.1	0.5	1.15E-12	7.44E-13	0	0	0	0		
14	10	0.1	0.5	1.37E-13	1.07E-13	0	0	0	0		
15	10	0.1	0.5	2.69E-13	2.24E-13	0	0	0	0		
16	10	0.1	0.5	4.47E-13	2.23E-13	0	0	0	0		
17	10	0.1	0.5	9.90E-13	6.66E-13	0	0	0	0		
18	10	0.1	0.5	2.94E-13	1.92E-13	0	0	0	0		
19	10	0.1	0.5	9.69E-13	5.36E-13	0	0	0	0		
20	10	0.6	1.30E-12	0.5	5.60E-11	0	0	0	0		
21	10	0.1	0.5	1.72E-12	1.11E-12	0	1	1	1		
22	10	0.4	0.39939	0.10061	6.21E-12	0	0	0	0		
23	10	0.1	0.5	1.33E-12	9.03E-13	0	1	1	1		
24	10	0.6	0.029903	0.4701	5.36E-10	0	0	0	0		
25	10	0.1	0.5	4.33E-13	2.20E-13	0	2	2	2		
26	10	0.1	0.5	6.87E-13	2.31E-13	0	0	0	0		
27	10	0.1	0.5	8.93E-13	4.18E-13	0	0	0	0		
28	10	0.6	5.58E-13	1.14E-12	0.5	1	1	2	2		
29	10	0.1	0.5	1.03E-12	5.82E-13	0	0	0	0		
30	10	0.1	0.5	3.60E-13	2.29E-13	0	1	1	1		
31	10	0.1	0.5	2.23E-13	1.44E-13	0	0	0	0		
32	10	0.2	0.5	3.87E-13	3.03E-13	0	1	1	1		
33	10	0.1	0.5	1.52E-12	9.66E-13	0	0	0	0		
34	10	0.1	0.5	4.51E-13	2.87E-13	0	0	0	0		
35	10	0.1	0.5	2.60E-13	1.29E-13	0	0	0	0		
36	10	0.1	0.5	1.97E-13	1.15E-13	0	0	0	0		
37	10	0.3	8.92E-12	0.36265	0.13735	0	0	0	0		
38	10	0.3	0.34825	8.54E-12	0.15175	0	0	0	0		
39	10	0.1	0.5	4.70E-13	3.30E-13	0	0	0	0		
40	10	0.1	0.5	1.07E-12	9.35E-13	0	0	0	0		
41	10	0.5	0.037118	0.46288	1.61E-11	0	0	0	0		
42	10	0.1	0.5	2.15E-12	9.18E-13	0	2	2	2		
43	10	0.1	0.5	1.43E-13	9.74E-14	0	2	2	2		
44	10	0.1	0.5	1.51E-12	9.11E-13	0	1	1	1		
45	10	0.1	0.5	3.91E-13	1.86E-13	0	0	0	0		
46	10	0.1	0.5	9.60E-13	6.01E-13	0	1	1	1		
47	10	0.3	0.5	4.28E-12	2.17E-12	0	3	3	3		
48	10	6.5	0.1511	0.17042	0.17849	0	0	0	0		
49	10	4.1	0.29862	0.20138	1.20E-12	0	0	0	0		
50	10	0.1	0.5	2.59E-13	1.23E-13	0	1	1	1		
Error Rate			19/ (15 * 50) = 2.5333 %			Sum	1	18	19		

表 1 最后 3 栏所示的是对每个签名者的测试样本的认证结果,综合认证误差为 2.53%,其中误纳率 FAR 比重较大(2.4%),这与该组数据说明文件中指

出的该数据集集中的伪造签名是由熟练的签名伪造者伪造的事实相符合.在表 2 的结果对比中,本方法获得的认证结果有了提高.

利用序列多尺度谱核矩阵的优化结果进行序列识别,不仅对签名认证有作用,对于其它序列识别问题同样具有借鉴意义.但是,在解决问题时需要判断问题中相似性描述是否有波动,并确定相适应的尺度范围,否则不仅会增加计算量,还会导致结果恶化.由于 SeDuMi 求解半定规划问题的计算复杂度为 $O(q^2 n^{2.5})$, n 为核矩阵的维数,因此本方法适合训练样本数量不多,样本特征向量维数较大的识别问题求解,大训练样本使得核矩阵维数 n 过大导致计算效率下降.在进行尺度选择时,还需要考虑序列字符集的大小,并兼顾考虑特征的不稳定程度来确定尺度范围.本方法适合解决小幅度特征稳定性波动问题,尺度集合中尺度的数目和大小差异不宜过大.这些约束导致本方法在解决其他问题时效果可能有限,有需要研究更灵活和推广性更好的相似性优化方法.

5 结语

签名动态信息序列之间的相似性描述准确与否深刻影响签名认证效果,签名序列的相似性描述可以通过核函数变换后在特征空间上进行.签名序列谱核函数变换是一种有效的序列相似性描述方法,它通过签名序列在某个尺度下的子序列的特征来刻画签名序列之间的相似性,适当的尺度选择十分关键.半定规划可以很好的解决多尺度下签名序列谱核矩阵的优化问题,结合支持向量机可以生成能够自主适应不同签名者书写特征的签名认证方法,本方法在签名认证实验中取得了良好的效果.

参考文献:

- [1] Prabhakar S, Kittler J, et al. Special issue on biometrics: progress and directions [J]. IEEE Trans on PAMI, 2007, 29(4): 513 - 516.
- [2] Impedovo D, Pirlo G. Automatic signature verification: the state of art [J]. IEEE Trans on SMC-Part C: Application and Reviews, 2008, 38(5): 609 - 635.
- [3] 邱德红,陈传波,等.非线性局部寻优时间弯曲校正及签名特征空间稳定性研究[J].计算机研究与发展,2002,39(10):1227-1232.

表 2 同组数据不同方法的认证结果比较

结果来源	Mario E M,参考文献[4]						参考文献[8]		本文结果	
	time		Euclidean arc-length		affine arc-length		FAR	FRR	FAR	FRR
数据	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR
SET2	4.13 %	6.15 %	4.16 %	3.85 %	3.86 %	4.16 %	3.467 %	0	2.4 %	0.133 %

Qiu De-hong, Chen Chuan-bo, et al. A study of the stability of signature features by nonlinear local optimal time warping [J]. Journal of Computer Research and Development, 2002, 39(10): 1227 - 1232. (in Chinese)

- [4] Mario E M, Pietro P, Visual identification by signature tracking [J]. IEEE Trans on PAMI, 2003, 25(2): 200 - 217.
- [5] 邱德红,陈传波,等.局域动态特征弯曲校正和作用均衡签名认证方法[J].小型微型计算机系统,2006,27(10): 1957 - 1960.
- Qiu De-hong, Chen Chuan-bo, et al. Signature verification based on characteristic parts of signing signal through warp rectifying and effect balancing [J], Mini-Micro Systems, 2006, 27(10): 1957 - 1960. (in Chinese)
- [6] Hu L, Wang Y H. On-line signature verification based on fusion of global and local information [A]. in Proc ICWAPR '07 [C]. Beijing, China, 2007. 3: 1192 - 1196.
- [7] Quan Z H, Huang D S, et al. A hybrid HMM/ANN based approach for online signature verification [A]. in Proc IJCNN [C]. Orlando, FL, 2007, 402 - 405.
- [8] Qiu D H, Chen C B, et al. Machine learning approaches for robust and high performance signature verification [J]. Journal of Computational Information Systems, 2006, 2(3): 933 - 938.
- [9] Wan L, Wan B, et al. On-line signature verification with two-stage statistical models [A]. in Proc 8th ICDAR [C]. Seoul, South Korea, 2005. 1: 282 - 286.
- [10] Fierrez J, Ortega-Garcia J, et al. HMM-based on-line signature verification: feature extraction and signature modeling [J]. Pattern. Recognition Letters, 2007, 28(12): 2325 - 2334.
- [11] Lanckriet G R G, Cristianini N, et al. Learning the kernel matrix with semidefinite programming [J]. Journal of Machine Learning Research, 2004, 5: 27 - 72.
- [12] Zhang D Q, Chen S C, et al. Learning the kernel parameters in kernel minimum distance classifier [J]. Pattern Recognition, 2006, 39(1): 133 - 135.
- [13] Ye J P, Ji S W, et al. Learning the kernel matrix in discriminant analysis via quadratically constrained quadratic programming [A]. in Proc 13th ACM SIGKDD KDD [C]. San Jose, California, USA, 2007. 854 - 863.
- [14] Stephen B, Lieven V. Convex Optimization [M], Cambridge University Press, 2004.
- [15] Christina S, Leslie C, et al. Fast string kernels using inexact matching for protein sequences [J]. Journal of Machine Learn-

- ing Research. 2004, 5(11):1435 - 1455.
- [16] Schölkopf B, Smola A J. Learning with Kernels-Support Vector Machines: Regularization, Optimization and Beyond[M]. Cambridge, Massachusetts: MIT Press, 2002.
- [17] Polik I, Terlaky T, et al. SeDuMi: a package for conic optimization[EB/OL]. <http://imre.polik.googlepages.com/SeDuMi-IMA-poster.pdf>
- [18] Labit Y, Peaucelle D, et al. SeDuMi Interface 1.02: a tool for solving LMI problems with SeDuMi [A]. in Proc IEEE CACSD[C]. Glasgow, U.K., 2002. 272 - 277.
- [19] Antoniou A, Lu W S. Practical optimization: algorithms and engineering applications[M]. Springer, 2007.

作者简介:



邱德红 男, 1971 年生于湖南永州. 博士, 副教授, 机器学习专业委员会委员. 研究方向为机器学习与数据挖掘、软件智能、新媒体技术.
E-mail: qiudehong@163.com



潘昕昕 男, 1981 年生于湖北省沙市. 硕士研究生, 研究方向为数据挖掘与软件智能.

(上接第 59 页)

- [8] A Smola, B Schölkopf. Sparse greedy matrix approximation for machine learning [A]. Proceedings of the 17th International Conference on Machine Learning [C]. San Francisco, USA: Morgan Kaufmann, 2000. 911 - 918.
- [9] R E Fan, P H. Chen, C J Lin. Working set selection using the second order information for training SVM[J]. Journal of Machine Learning Research, 2005, 6:1889 - 1918.
- [10] C C Chang, C J Lin. LIBSVM -- A Library for Support Vector Machines (version 2.85) [OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2007-12-26.
- [11] I W Tsang. LibCVM Toolkit (version 2.02 alpha) [OL]. <http://www.cs.ust.hk/ivor/cvm.html>, 2008-03-05.
- [12] F Rosenblatt. The Perceptron: a probabilistic model for information storage and organization in the brain[J]. Psychological Review, 1958, 65(6):386 - 408.
- [13] C Cortes, V Vapnik. Support-vector networks [J]. Machine Learning, 1995, 20(3):273 - 297.
- [14] P Lingras, C Butz. Interval set classifiers using support vector machines[A]. Proceedings of International Conference of the North American Fuzzy Information Processing Society [C]. Banff, Canada: IEEE, 2004, No. 23, 707 - 710.