

基于并行特征组合与广义 K-L 变换的字符识别*

杨 健, 杨静宇⁺, 高建贞

(南京理工大学 计算机科学系, 江苏 南京 210094)

Handwritten Character Recognition Based on Parallel Feature Combination and Generalized K-L Expansion

YANG Jian, YANG Jing-Yu⁺, GAO Jian-Zhen

(Department of Computer Science, Nanjing University of Science and Technology, Nanjing 210094, China)

+ Corresponding author: E-mail: yangjy@public1.ptt.js.cn

Received 2001-08-22; Accepted 2001-12-17

Yang J, Yang JY, Gao JZ. Handwritten character recognition based on parallel feature combination and generalized K-L expansion. *Journal of Software*, 2003,14(3):490-495.

Abstract: Considering the weaknesses of traditional serial feature fusion technique, a novel parallel features fusion method is proposed in this paper. The main idea of this method can be described as follows. First of all, two sets of feature vectors corresponding to a same sample space are combined together via complex vectors, which are used to construct a complex feature vector space. Then, the classical K-L transform and K-L expansion methods are developed theoretically to suit for feature extraction in the complex feature space. Moreover, the symmetric property of parallel feature fusion is revealed, and, how to combine features effectively is discussed in detail. Finally, the proposed method is used to solve the handwritten character feature extraction and recognition problems. Experiments are performed on NUST603 handwritten Chinese character database built in Nanjing University of Science and Technology as well as the well-known CENPARMI handwritten digit database of Concordia University. The experimental results indicate that the recognition rates are improved significantly after parallel feature fusion, and the proposed parallel features fusion method is superior to the traditional serial feature fusion one.

Key words: feature fusion; feature combination; generalized K-L transform; feature extraction; handwritten character recognition

摘 要: 针对传统的串行特征融合方法的弱点,提出了一种新的并行特征融合方法.该方法的基本思路是:首先,利用复向量将样本空间上的两组特征集组合起来,构成复特征向量空间;然后,从理论上推广了经典的 K-L 变换方法与 3 种基本的 K-L 展开方法,使其适用于复特征向量空间内的特征抽取.此外,还揭示了并行特征融合的对称性质,并详细讨论了并行特征组合的策略问题.最后,用所提出的方法来解决手写体字符的特征抽取与识别问题.在南京理工大学 NUST603HW 手写体汉字库以及 Concordia 大学的 CENPARMI 手写体阿拉伯数字数据库上的实验结果表明,所提出的特征融合方法不仅较大幅度地提高了识别率,而且识别结果优于传统的串行特征融合方法.

关键词: 特征融合;特征组合;广义 K-L 变换;特征抽取;手写体字符识别

* 第一作者简介: 杨健(1973—),男,江苏徐州人,博士,讲师,主要研究领域为字符识别,人脸识别.

中图法分类号: TP391 文献标识码: A

特征级融合在信息融合过程中占有十分重要的地位.特征级融合既保留了参与融合的多特征的有效鉴别信息,又在很大程度上消除了信息的冗余,实现了可观的信息压缩,从而有利于信息的实时处理.目前,有关特征级融合的传统方法,是将两组特征向量首尾相连生成一个联合向量作为新的特征向量^[1],然后再进行特征抽取,我们称其为串行融合方法.尽管这种方法在不少情况下可有效地提高识别率,但其缺点也是明显的.由于特征合并后的维数是两个原始特征的维数之和,这就导致了合并后新特征的维数急剧增加,从而使组合后的特征抽取或识别的速度大幅度地降低.

针对传统的串行特征融合方法的缺点,本文提出了一种新的并行特征融合方法.该方法的基本思想是:利用复向量将样本空间上的两组特征集组合起来,构成复特征向量空间.那么,现在的问题是,如何在复向量特征空间内进行特征抽取?众所周知,K-L 变换是特征抽取的最有效的方法之一^[2],但经典的 K-L 变换方法只适用于实特征空间内的情形.由于利用复向量组合后的特征空间为复向量空间,本文从理论上拓展了经典的 K-L 变换方法,使其完全适用于复特征空间内的特征抽取.本文所提出的方法在手写体字符的特征抽取与识别问题中得到了成功的应用.

1 特征并行组合与广义 K-L 变换方法

设 A, B 为模式样本空间 Ω 上的两组特征.任意模式样本 $\xi \in \Omega$, 其对应的两个特征向量分别为 $\alpha \in A$ 和 $\beta \in B$, 我们用复向量 $\gamma = \alpha + i\beta$ (i 为虚数单位)来表示 ξ 的并行组合特征.注意,若两组特征 α 与 β 的维数不等,低维的特征向量用零补足.例如, $\alpha = (a_1, a_2, a_3)^T$, $\beta = (b_1, b_2)^T$, 则组合特征为 $\gamma = (a_1 + ib_1, a_2 + ib_2, a_3 + i0)^T$.

样本空间 Ω 上的组合特征空间定义为 $C = \{\alpha + i\beta \mid \alpha \in A, \beta \in B\}$. 明显地,该空间为 n 维复向量空间,其中, $n = \max\{\dim A, \dim B\}$. 我们定义如下内积:

$$(X, Y) = X^H Y, \text{ 其中 } X, Y \in C, H \text{ 为共轭转置符号.} \tag{1}$$

定义了以上内积的复空间称为酉空间.相应地,酉空间内的类间散布矩阵、类内散布矩阵和总体散布矩阵分别定义为

$$S_b = \sum_{i=1}^L P(\omega_i) (m_i - m_0)(m_i - m_0)^H, S_w = \sum_{i=1}^L P(\omega_i) E\{(X - m_i)(X - m_i)^H / \omega_i\}, S_t = S_b + S_w. \tag{2}$$

其中 $P(\omega_i)$ 为第 i 类训练样本的先验概率, $m_i = E\{X / \omega_i\}$ 为第 i 类训练样本的均值, $m_0 = E\{X\} = \sum_{i=1}^m P(w_i) m_i$ 为全体训练样本的均值.由式(2)的定义可知, S_w, S_b, S_t 均为 Hermite 阵,且非负定.当 S_w 可逆时,易知 S_w, S_t 均为正定矩阵.

引理 1^[3]. Hermite 阵的特征值均为实数.

由于 S_w, S_b, S_t 均为 Hermite 阵,且非负定,故易得出以下结论:

推论 1. S_w, S_b, S_t 的特征值均为非负实数.

下面我们以 S_t 为例来讨论酉空间内的 K-L 变换方法.不妨设 S_t 的特征向量为 ξ_1, \dots, ξ_n (彼此正交),所对应的特征值为 $\lambda_1, \dots, \lambda_n$, 且满足 $\lambda_1 \geq \dots \geq \lambda_n$. 选择前 d 个最大特征值所对应的特征向量 ξ_1, \dots, ξ_d 作为投影轴.由于酉空间内 S_t 的特征值均为非负实数,组合特征(复)向量在轴 ξ_j 上投影后的物理意义十分明确,即总体散布量为 λ_j . 在酉空间内,以 S_t 为产生矩阵的广义 K-L 变换即为

$$Y = \Phi^H X, \text{ 其中 } \Phi = (\xi_1, \dots, \xi_d). \tag{3}$$

该变换用于组合特征抽取.

与文献[2]中的方法比较可见,以往讨论的实向量空间内的 K-L 变换方法只是本文方法的一个特例,也就是说,本文在复空间内建立的 K-L 变换方法更具一般意义,它完全适用于实空间内的情形.

2 组合的对称性问题

同一样本的两特征 α, β 组合可以有两种不同的方式: $\alpha + i\beta$ 或 $\beta + i\alpha$, 按照这两种不同的方式组合, 再利用广义 K-L 变换进行特征抽取后, 在同一分类器下进行分类, 分类结果是否一致? 若结果一致, 我们认为两特征组合具有对称性, 即并行特征融合后的分类结果与组合次序无关, 这也是我们所期望的. 幸运的是, 我们可以从理论上证明, 酉空间内的并行特征组合满足对称性.

设样本空间 Ω 上的两组合特征空间分别定义为 $C_1 = \{\alpha + i\beta \mid \alpha \in A, \beta \in B\}$, $C_2 = \{\beta + i\alpha \mid \alpha \in A, \beta \in B\}$. 易证明以下引理成立.

引理 2. 设矩阵 $H(\alpha, \beta) = (\alpha + i\beta)(\alpha + i\beta)^H$, $H(\beta, \alpha) = (\beta + i\alpha)(\beta + i\alpha)^H$, 则 $H(\beta, \alpha) = \overline{H(\alpha, \beta)}$. 其中 α, β 为 n 维实向量.

设组合特征空间 C_i 内的类内、类间和总体散布阵分别表示为 $S_w^i, S_b^i, S_t^i (i=1, 2)$, 由引理 2 易证明它们满足以下性质:

性质 1. $S_w^2 = \overline{S_w^1}, S_b^2 = \overline{S_b^1}, S_t^2 = \overline{S_t^1}$.

性质 2. 设 ξ 是 S_t^2 (S_w^2 或 S_b^2) 的属于特征值 λ 的特征向量, 则 $\bar{\xi}$ 是 S_t^1 (S_w^1 或 S_b^1) 的属于特征值 λ 的特征向量.

证明: $S_t^1 \xi = \lambda \xi \Rightarrow \overline{S_t^1 \xi} = \overline{\lambda \xi} \Rightarrow \overline{S_t^1} \bar{\xi} = \overline{\lambda} \bar{\xi}$, 由性质 1 和推论 1 可知, $S_t^2 = \overline{S_t^1}, \bar{\lambda} = \lambda$, 故 $S_t^2 \bar{\xi} = \lambda \bar{\xi}$. \square

性质 2 说明, 两种不同的特征组合方式所对应的广义 K-L 变换的投影轴互为共轭关系. 即若 ξ_1, \dots, ξ_d 为组合特征空间 C_1 中 K-L 变换的投影轴, 则 $\bar{\xi}_1, \dots, \bar{\xi}_d$ 为组合特征空间 C_2 中 K-L 变换的投影轴.

引理 3. 设在组合特征空间 C_1 中, 样本 $x + iy$ 在投影方向 ξ 上的投影分量为 $p + iq$, 则在组合特征空间 C_2 中, 样本 $y + ix$ 在投影方向 $\bar{\xi}$ 上的投影分量为 $q + ip$.

证明: 设 $\xi = (a_1 + ib_1, \dots, a_n + ib_n)^T$, 则

$$\xi^H (x + iy) = (a_1 - ib_1, \dots, a_n - ib_n)(x_1 + iy_1, \dots, x_n + iy_n)^T = \sum_{j=1}^n (a_j x_j + b_j y_j) + i \sum_{j=1}^n (a_j y_j - b_j x_j) = p + iq,$$

$$\bar{\xi}^H (y + ix) = (a_1 + ib_1, \dots, a_n + ib_n)(y_1 + ix_1, \dots, y_n + ix_n)^T = \sum_{j=1}^n (a_j y_j - b_j x_j) + i \sum_{j=1}^n (a_j x_j + b_j y_j) = q + ip.$$

命题得证. \square

由引理 3, 不难得出以下结论:

性质 3. 在组合特征空间 C_1 内, $x + iy \xrightarrow{\text{K-L变换}} u + iv$; 在组合特征空间 C_2 内, $y + ix \xrightarrow{\text{K-L变换}} v + iu$.

在酉空间内, 由式(1)定义的内积引入以下度量(范数):

$$\|Z\| = \sqrt{Z^H Z} = \sqrt{\sum_{j=1}^n (a_j^2 + b_j^2)}, \text{ 其中 } Z = (a_1 + ib_1, \dots, a_n + ib_n)^T. \quad (4)$$

在该度量意义下, 复向量的实部和虚部具有对称性, 即 $\|u + iv\| = \|v + iu\|$.

相应地, 在酉空间内, 两复向量间的距离定义为

$$\|Z_1 - Z_2\| = \sqrt{(Z_1 - Z_2)^H (Z_1 - Z_2)}. \quad (5)$$

由此定义, 两复向量之间的距离只与实部和虚部的取值有关, 而与实部和虚部的次序无关. 更具体地讲, 即设 $Z_1^1 = u_1 + iv_1, Z_2^1 = u_2 + iv_2; Z_1^2 = v_1 + iu_1, Z_2^2 = v_2 + iu_2$, 则 $\|Z_1^1 - Z_2^1\| = \|Z_1^2 - Z_2^2\|$.

综上所述, 我们可以得出以下结论:

定理 1. 在酉空间内, 两特征的组合具有对称性.

即两特征组合经广义 K-L 变换进行特征抽取后, 在酉空间的内积所确定的距离意义下, 识别结果与组合次序无关.

3 几种典型的 K-L 展开方法的推广

3 种经典的 K-L 展开方法^[2]为:(1) 从类平均向量中提取判别信息;(2) 包含在类平均向量中判别信息的最优压缩;(3) 包含在类中心化特征向量中判别信息的提取.这 3 种方法均可推广到复向量空间用于组合特征的抽取.下面,我们分别进行讨论.

方法 1. 从类平均向量中提取判别信息.我们采用以下准则函数:

$$J(\xi_j) = \frac{\xi_j^H S_b \xi_j}{\lambda_j}, \tag{6}$$

其中 ξ_j 为 S_w 的特征向量, λ_j 为所对应的特征值, $j=1, \dots, n$. 由 S_b 的非负定性可知, $\xi_j^H S_b \xi_j \geq 0$; 当 S_w 非奇异时, 特征值 λ_j 恒大于 0, 故准则函数(6)可用来表征样本在轴 ξ_j 上投影后的分类性能, 其物理意义与在实空间的情形完全相同. 若满足 $J(\xi_1) \geq \dots \geq J(\xi_d) \geq \dots \geq J(\xi_n)$, 取 ξ_1, \dots, ξ_d 作为投影轴即可.

方法 2. 包含在类平均向量中判别信息的最优压缩. 在酉空间内, 当 S_w 非奇异时, 由于 S_w 为正定阵, 故存在矩阵 B 使得 $B^H S_w B = I$. 事实上, $B = U A^{-\frac{1}{2}}$, 其中, $U = (\xi_1, \dots, \xi_n)$, A 为所对应的 S_w 的特征值对角阵. 令 $\tilde{S}_b = B^H S_b B$, 易证明, \tilde{S}_b 为酉空间内的非负定阵, 取其 d 个非零特征值所对应的特征向量 v_1, \dots, v_d , 构成变换矩阵 $W = BV$, 其中, $V = (v_1, \dots, v_d)$.

方法 3. 包含在类中心化特征向量中判别信息的提取. 设第 i 类样本的类内散布阵为 $S_w^i (i=1, \dots, c)$, 则第 i 类样本在方向 $\xi_j (S_w$ 的特征向量)上投影后的方差为

$$\gamma_{ij} = \xi_j^H S_w^i \xi_j, \quad i=1, \dots, c; \quad j=1, \dots, n.$$

由于各类的类内散布阵 $S_w^i (i=1, \dots, c)$ 均为非负定阵, 故方差 γ_{ij} 均为实数, 从而, 该方法雷同于实空间内的情形.

4 组合的策略问题

由于特征抽取方法与量纲选择的不同, 参与组合的同一模式样本的两组特征 α 与 β 之间在数值上可能存在较大的差异. 比如, $\alpha = (10, 11, 9)^T$, $\beta = (0.1, 0.9)^T$. 若直接以 $\gamma = \alpha + i\beta$ 的方式进行组合, 组合后两特征的比重将明显失调. 为了消除参与组合的两特征在数值上的非均衡性所造成的不利影响, 我们首先将原始特征 α 与 β 分别进行如下标准化处理:

$$Y = \frac{X - \mu}{\sigma}. \tag{7}$$

其中, μ 为训练样本的均值向量, σ 为训练样本的标准差向量在各个分量上的均值.

另外, 当参与组合的两特征 α 与 β 的维数不等时, 经过标准化以后, 维数较高的特征在组合后的优势依然明显, 这是因为在组合后生成的散布矩阵中, 维数较高的特征在数值上占据较大比重. 为了克服这种影响, 当两特征维数不等时, 我们对标准化后的特征采取加权组合的方法. 设加权组合的形式为 $\gamma = \alpha + i\theta\beta$, 其中, 权值 θ 称为组合系数. 如何精确地估计组合系数 θ , 使得组合后的效果达到最好是一个有待研究的问题. 在此, 我们根据经验给出 θ 的选择区间. 令 δ 表示特征 α 与 β 的维数之比, 则 θ 在区间 (δ, δ^2) 内选取.

5 实验结果与分析

5.1 实验1

本实验的目的在于验证第 1 节建立的广义 K-L 变换方法的有效性.

实验采用南京理工大学 NUST603HW 手写体有限汉字样本库. 该汉字库包含银行货币金额大写常用的“零、壹...”等 19 个汉字的样本. 每类样本数为 400 个(共 7 600 个). 其中 200 个样本作为训练样本, 另外 200 个样本作为测试样本, 这样, 训练样本和测试样本总数均为 3 800 个. 我们采用文献[4,5]的方法分别抽取汉字的以下两种特征:

Cross_feature: 128 维的交叉数特征,该特征反映汉字的结构特征.

Peripheral_feature: 128 维的周边特征,该特征反映汉字的字形特征.

二者的并行组合采用 Cross + $i\theta$ Peripheral 的形式.由于两特征的维数相同,取组合系数 $\theta=1$ 即可.用本文第 1 节给出的广义 K-L 变换方法(以 S_l 作为产生矩阵),将原始特征及其组合特征压缩到低维的鉴别特征空间内,采用二次 Bayesian 分类器,由于该实验中各模式类的先验概率相同,我们采用如下定义的 Bayes 鉴别函数:

$$g_l(x) = \frac{1}{2} \ln |\Sigma_l| + \frac{1}{2} (x - \mu_l)^H \Sigma_l^{-1} (x - \mu_l), \quad (8)$$

其中, μ_l 和 Σ_l 分别表示第 l 类的均值与协方差矩阵.若模式样本 x 满足 $|g_k(x)| = \min_l |g_l(x)|$, 则 $x \in \omega_k$. 分类结果见表 1.

另外,我们采用文献[1]提供的方法,分别将 Cross 与 Peripheral 两组特征向量规一化后,合并为一个 256 维的特征向量,然后采用 K-L 变换进行二次特征抽取,所得到的串行组合特征在二次 Bayesian 分类器下的识别结果见表 1.

Table 1 Comparison of error rates based on four sets of features in experiment 1

表 1 实验 1 中基于 4 类特征的分类错误率对照

Dimension	Cross feature	Peripheral feature	Parallel combination	Serial combination
20	0.058	0.053	0.017	0.027
22	0.056	0.043	0.016	0.025
24	0.055	0.041	0.013	0.022
26	0.056	0.037	0.012	0.020
28	0.057	0.037	0.012	0.021
30	0.060	0.034	0.013	0.021
32	0.059	0.034	0.013	0.018
34	0.063	0.036	0.012	0.016
36	0.067	0.037	0.013	0.015
38	0.067	0.036	0.012	0.013
40	0.068	0.038	0.013	0.014

由表 1 可见,两类特征组合经过广义 K-L 变换后,在各个维度上,识别错误率均较大幅度地得以降低.交叉数特征在变换后的最优识别率(维数压缩为 24)为 94.5%,周边特征在变换后的最优识别率(维数压缩为 30)为 96.6%,并行组合特征在变换后的最优识别率为 98.8%,而且,当维数在 24 与 40 之间发生变化时,识别结果始终稳定在 98.7%之上.由表 1 亦可见,并行组合后的分类结果优于串行组合后的识别结果.

5.2 实验 2

本实验针对参与组合的两组特征维数不等的情况下,验证第 3 节所推广的 K-L 展开方法的有效性.

采用国际上广泛使用的 Concordia University CENPARMI 手写体阿拉伯数字(0~9)数据库,其中有 4 000 个训练样本和 2 000 个测试样本.利用文献[6]已提取出的如下两个图像特征:

G^f : 256 维 Gabor 变换特征; L^f : 121 维 Legendre 矩特征;

二者的组合采用 $C^f = G^f + i\theta L^f$ 的形式.由于两特征的维数不等,组合系数取为 $\theta=2.8$.

利用推广的 K-L 展开方法 1(从类平均向量中提取判别信息)将原始特征及其并行组合特征压缩到低维的鉴别特征空间内,采用二次 Bayes 分类器分类,分类结果见表 2.

另外,我们采用文献[3]提供的方法,分别将 Gabor 与 Legendre 两组特征向量规一化后,合并为一个 377 维的特征向量,然后采用 K-L 展开方法 1 进行二次特征抽取,所得到的串行组合特征在二次 Bayesian 分类器下的识别结果见表 2.

由表 2 可见,在组合前,Gabor 压缩特征的最优识别率为 90.0%,Legendre 压缩特征的最优识别率为 95.3%;并行融合后,最优识别率达到 96.4%.此外,融合后的识别结果较融合前更为稳定,当维数在 40 与 50 之间发生变化时,识别率稳定在 96.2%之上.由表 2 亦可看出,并行特征融合的分类错误率大大低于串行特征融合的分类错误率.

Table 2 Comparison of error rates based on four sets of features in experiment 2**表 2** 实验 2 中基于 4 类特征的分类错误率对照

Dimension	Gabor feature	Legendre feature	Parallel combination	Serial combination
30	0.107	0.055	0.046	0.088
32	0.108	0.056	0.045	0.081
34	0.108	0.051	0.043	0.082
36	0.109	0.047	0.041	0.080
38	0.111	0.051	0.040	0.079
40	0.111	0.052	0.037	0.081
42	0.106	0.056	0.038	0.083
44	0.105	0.057	0.037	0.084
46	0.107	0.059	0.036	0.085
48	0.100	0.057	0.038	0.083
50	0.103	0.059	0.037	0.083

6 结 论

本文提出了一种利用复向量进行特征级融合的新方法,从理论上推广了经典的 K-L 变换技术,使其完全适用于组合(复)特征空间内的特征抽取.事实上,经典的 K-L 变换只是本文方法的一个特例.在 NUST603HW 手写体汉字库和 CENPARMI 阿拉伯数字库上的实验结果证实了基于并行特征组合与广义 K-L 变换方法的有效性.本文提出的方法具有一定的实际意义和广泛的应用前景,它为特征级信息融合提供了一条行之有效的途径.今后的工作一方面是,从理论上进一步完善该方法,比如给出组合系数的精确估计方法等;另一方面是,进一步拓宽该方法的应用范围,将其用于图像的融合及识别问题之中.

References:

- [1] Liu CJ, Wechsler H. A shape- and texture-based enhanced Fisher classifier for face recognition. *IEEE Transactions on Image Processing*, 2001,10(4):598~608.
- [2] Young T, Fu K-S. *Handbook of Pattern Recognition and Image Processing*. New York: Academic Press, 1986. 78~81.
- [3] Ding XR, Cai MK. *Matrix Theory in Engineering*. Tianjin: Tianjin University Press, 1995. 115~118 (in Chinese).
- [4] Tang YT. Offline recognition of Chinese handwriting by multifeature and multilevel classification. *IEEE Transactions on PAMI*, 1998,20(5):556~561.
- [5] Tseng YH, Kuo CC, Lee HJ. Speeding up Chinese character recognition in an automatic document reading system. *Pattern Recognition*, 1998,31(11):1601~1612.
- [6] Hu ZS, Hu Z, Lou, JY, Liu K, Sun JY. Handwritten digit recognition based on multi-classifier combination. *Chinese Journal of Computers*, 1999,22(4):369~374 (in Chinese with English Abstract).

附中文参考文献:

- [3] 丁学仁,蔡庙可.工程中的矩阵理论.天津:天津大学出版社,1995.115~118.
- [6] 胡钟山,娄震,杨静宇,刘克,孙靖夷.基于多分类器组合的手写体数字识别.计算机学报,1999,22(4):369~374.