

一种基于学习的自动图像配准检验方法

李晓明^{1,2} 张继福³

摘要 图像配准是众多具体应用的共性核心技术, 如图像融合, 变化检测等. 然而, 当参考图像经过变换后, 如何自动地确定变换后的图像是否与目标图像真正达到了配准仍然是目前文献中一个尚未很好解决的问题. 究其原因, 主要是很难找到一种图像相似性的度量方法来有效地对配准后的图像进行评价. 不同于传统的方法, 本文提出了一种基于学习的相似性度量方法, 即将图像配准的度量问题转化为模式分类问题, 由基于机器学习设计的分类器自动检验图像是否配准. 本文对 400 组图像进行了配准检验, 实验结果显示了该方法的可行性和可靠性. 尽管本文方法的具体实现是针对基于 Fourier-Mellin 变换的配准算法, 但这种基于学习的图像配准检验思想同样可以应用到其他配准方法中.

关键词 图像配准, 检验准则, 基于学习的分类, Fourier-Mellin 变换
中图分类号 TP391

An Evaluation Method for Image Registration by Machine Learning

LI Xiao-Ming^{1,2} ZHANG Ji-Fu³

Abstract Image registration is a key step in many real applications, such as image fusion and change detection. However, after transforming the reference image, it is usually difficult to assess whether the transformed image is indeed registered with the target image. The underlying reason is that a suitable measure is unavailable currently in literature to adequately evaluate the similarity of two images. In contrast to the traditional evaluation methods, in this paper, the registration evaluation problem is converted into a classification problem through machine learning. Experiments based on 400 pairs of images demonstrate the validity and reliability of our proposed method. Although the proposed method is specifically designed for Fourier-Mellin transform based registration technique, its basic principles could also be applied to other registration techniques.

Key words Image registration, evaluation measure, classification by machine learning, Fourier-Mellin transform

图像配准是计算机视觉及其相关应用领域中的一个基本问题. 如何度量二幅图像是否真正配准, 即如何合理度量配准变换后二幅图像之间的相似性, 是图像配准的一个核心问题, 直接关系到是否可以进行图像的全自动配准. 关于图像的相似性度量, 经典的方法是基于灰度的相关性度量^[1]; 对于不同模态的图像配准问题, 基于互信息的相似性度量^[2] 成为近几年的主流方法; Hausdorff 距离常用于二值图像的相似性度量^[3] 等. 对图像配准而言, 上述这些相似性度量方法的共同不足是仅仅考虑了图像间灰度的相似性, 而没有考虑图像内容间的相似性, 特别是图像几何要素之间的相似性. 而在实际应用中, 图

像几何要素之间的相似性却是至关重要的. 事实上, 如何评价图像的相似性, 涉及到了人类的视觉感知问题. 从目前人们对视觉感知模型的研究进展来看, 还很难给出一种合适的方式来定量描述视觉感知模型, 这也是进行图像配准自动检验的难点. 不同于上述这些直接对图像灰度进行相似性度量的方法, 本文提出了一种基于机器学习的自动图像配准检验方法. 该方法将配准检验问题转换为模式分类问题, 通过机器学习的方法对图像是否配准做出决断. 尽管这种方法的具体实现是针对基于 Fourier-Mellin 变换的配准算法的, 但这种基于学习的相似性度量思想同样可以应用到其他配准方法中.

基于 Fourier-Mellin 变换的图像配准算法^[4-5] 是一种经典的基于非特征的图像配准方法, 可以对两幅近似满足相似变换的图像进行配准. 由于 Fourier-Mellin 配准算法在配准检验中存在一些问题 (将在第 2 节详细论述), 一些常用的相似性度量方法在这里不适用, 所以本文探讨了基于机器学习和分类的配准检验方法. 具体地讲, 本文将图像的配准度量问题转化为参数空间的模式分类问题, 通过机器学习的方法, 较好地解决了图像配准中的自动检验问题.

收稿日期 2007-01-24 收修改稿日期 2007-04-29
Received January 24, 2007; in revised form April 29, 2007
国家自然科学基金 (60773132) 资助
Supported by National Natural Science Foundation of China (60773132)

1. 山西大学数学科学学院 太原 030006 2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100080 3. 太原科技大学计算机科学与技术学院 太原 030024
1. School of Mathematical Sciences, Shanxi University, Taiyuan 030006 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080 3. School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024
DOI: 10.3724/SP.J.1004.2008.00007

本文的组织结构如下: 第 1 节简单回顾了基于 Fourier-Mellin 变换的基本配准算法; 该方法在配准检验中存在的问题将在第 2 节进行深入的剖析; 第 3 节阐述了如何将检验图像的配准问题转化为模式分类问题以及解决这个问题的总体思路; 第 4 节对分类器的设计、不同类型特征数据的组织以及对应的分类性能等进行了研究和比较; 最后是全文的结论.

1 基于 Fourier-Mellin 变换的两步相位相关图像配准算法

1.1 傅里叶 (Fourier) 变换位移理论

设 $f_2(x, y)$ 为 $f_1(x, y)$ 在 x 和 y 方向分别平移 x_0 和 y_0 后的图像, 即

$$f_2(x, y) = f_1(x - x_0, y - y_0) \quad (1)$$

若 f_1 和 f_2 对应的傅里叶变换分别为 $F_1(u, v)$ 和 $F_2(u, v)$, 则它们之间有如下关系

$$F_2(u, v) = F_1(u, v)e^{-j(ux_0 + vy_0)} \quad (2)$$

则 $f_1(x, y)$ 和 $f_2(x, y)$ 的互功率谱为

$$\frac{F_1(u, v)F_2^*(u, v)}{|F_1(u, v)F_2^*(u, v)|} = e^{j(ux_0 + vy_0)} \quad (3)$$

其中 F_2^* 表示 F_2 的复共轭. 通过对式 (3) 进行傅里叶逆变换, 在 (x, y) 空间的 (x_0, y_0) 处将形成一个脉冲函数. 在具体应用中, 由于两幅图像间只有部分重叠及其他噪声和误差, 实际的情况是, 在 (x_0, y_0) 处有一个突出的峰值, 而其他位置近似为 0. 当然, 如果两幅图像之间不满足平移变换关系, 那么式 (3) 傅里叶逆变换后的函数没有明显的峰值, 且呈现出不规则分布. 式 (3) 还表明, 互功率谱的相位等价于图像间的相位差, 故该方法也称作相位相关法.

1.2 基于 Fourier-Mellin 变换的两步相位相关图像配准基本算法

考虑被配准的两幅图像 $s(x, y)$ 和 $r(x, y)$, 其中 $s(x, y)$ 是 $r(x, y)$ 经过平移、旋转和一致尺度缩放 (即两个方向的尺度变换因子相等) 变换后的图像, 即

$$s(x, y) = r[(\sigma(x \cos \alpha + y \sin \alpha) - x_0, \\ -x \sin \alpha + y \cos \alpha) - y_0] \quad (4)$$

那么 $s(x, y)$ 和 $r(x, y)$ 对应的傅里叶变换 $S(u, v)$ 和 $R(u, v)$ 之间满足

$$|S(u, v)| = \sigma^{-2} |R[\sigma^{-1}(u \cos \alpha + v \sin \alpha), \\ \sigma^{-1}(-u \sin \alpha + v \cos \alpha)]| \quad (5)$$

其中 $|\cdot|$ 表示频谱幅度. 从式 (5) 可看出, 旋转角度 α 和缩放因子 σ 可以与平移量 x_0 和 y_0 进行分离计算. 频谱幅度仅与旋转角度 α 和缩放因子 σ 有关, 而与平移量 (x_0, y_0) 无关, 因此, 相似变换参数可分两步分别求出, 第一步通过图像幅度谱求出旋转角度 α 和缩放因子 σ . 第二步求出平移参数 x_0 和 y_0 .

1.2.1 第一步: 用相位相关法求旋转角度 α 和缩放因子 σ

定义

$$r_{p1}(\theta, \log \rho) = r_p(\theta, \rho) \quad (6)$$

$$s_{p1}(\theta, \log \rho) = s_p(\theta, \rho) \quad (7)$$

其中 r_p 和 s_p 分别是图像 r 和 s 在极坐标系 (θ, ρ) 中的幅度谱. 那么很容易得出

$$s_{p1}(\theta, \log \rho) = r_{p1}(\theta - \alpha, \log \rho - \log \sigma) \quad (8)$$

或者

$$s_{p1}(\theta, \lambda) = r_{p1}(\theta - \alpha, \lambda - \kappa) \quad (9)$$

其中 $\lambda = \log \rho$, $\kappa = \log \sigma$.

可以看出, 通过上述变换, 式 (9) 变换为与式 (1) 相同的形式, 这样就可以在对数-极 (Log-polar) 空间对 s_{p1} 和 r_{p1} 依次根据式 (2) 和 (3) 进行计算, 然后, 对式 (3) 进行傅里叶逆变换求得相关值的分布, 根据其对应的峰值位置即可求得 α 和 κ . 本文称这一步求得的相关值为相关值 I, 其最大相关值称为峰值 I.

如果对数的底取为 e , 那么

$$\sigma = e^\kappa \quad (10)$$

这样就求出了旋转角度 α 和比例因子 σ .

1.2.2 第二步: 用相位相关法求平移参数 x_0 和 y_0

根据第一步求出的 α 和 σ 对图像 $s(x, y)$ 进行反变换得到图像 $s_1(x, y)$, 通过在图像空间对 $s_1(x, y)$ 和 $r(x, y)$ 依次计算式 (2) 和式 (3), 然后对式 (3) 进行傅里叶逆变换求得相关值的分布, 根据其对应的峰值位置即可求得两个方向的平移参数 x_0 和 y_0 . 本文称这一步求得的相关值为相关值 II, 其最大相关值称为峰值 II.

式 (9) 称为 Fourier-Mellin 变换, 更详细的内容可参见文献 [4-5].

2 基于 Fourier-Mellin 变换的图像配准算法的性能分析

2.1 第一步相位相关法分析

从第 1.2 节中的式 (4) 和式 (5) 可以看出, 该算法有一个基本前提, 即一幅图像旋转一定角度后的傅里叶变换频谱等价于未经旋转的图像的傅里叶变

换频谱旋转相同的角度. Stone 等^[6] 对由于图像旋转产生的频谱混叠所引起的图像配准噪声进行了详细分析, 并且推导出, 虽然对无限长连续图像而言, 这个假设是正确的, 但是对于有限长离散图像数据而言, 这个基本前提是不正确的. 其主要原因是由于依赖于旋转的频谱混叠 (即对不同的旋转角度其频谱混叠程度是不同的) 引起的 artifacts 和旋转变换中插值误差产生的 artifacts 造成的. Stone 等人得出的结论是: 1) 第一步相位相关法求得的峰值 I 并不能可靠地对应正确的旋转角度位置; 2) 应用相位相关法求得的峰值 I 大大小于理论值 1.0, 一般来说降到了 0.03 以下. 文献 [5-7] 等都针对这种由于旋转引起的问题进行了详细研究, 并采取了一些诸如加窗和滤波等措施来提高峰值、减少频谱混叠、增加鲁棒性, 但仍然未能从根本上解决这个问题.

需要说明的是, 在文献 [5-7] 中, 主要研究在两幅图像内容完全相同时, 由于图像旋转带来的图像配准问题. 然而, 在一般的应用中, 被匹配的两幅图像之间仅有部分重叠, 同时还可能存在平移和尺度缩放等变形, 还有其他多种噪声. 对这类情况, 无法象文献 [6] 中那样给出解析分析结果, 亦即目前理论上还不能定量地推导出这些因素造成的影响.

实验发现, 根据峰值 I 的绝对大小也无法判断匹配结果的正确与否. 文献 [4] 中提到, 一般情况下, 峰值 I 小于 0.03 时结果就不可靠了, 但实验发现, 即使峰值 I 大于 0.1, 结果也未必可靠, 而两幅图像真正匹配时, 峰值 I 小于 0.03 的情况在噪声较大的情况下是很常见的. 因此可以认为, 这个取值 0.03 是和具体应用有关的. 在文献 [5] 中, 对该算法的性能定义了 3 种量化评价指标, 即输出信噪比、峰值的尖锐性和 PD (Percent discrimination), 但都不适宜作为确定配准与否的判断标准. 实验还发现, 根据相关值 I 的位置分布来进行配准检验也是没有规律可循的.

2.2 第二步相位相关法分析

将第 1.1 节中介绍的傅里叶变换位移理论应用于两幅图像之间仅有平移变换的配准问题, 当两幅图像之间的平移量是像素的整数倍时, 该方法是一种精确的配准方法. 虽然也存在频谱混叠的问题, 但不同于具有旋转的情况, 因为这时对于被配准的两幅图像而言其影响是相同的^[6].

在基于 Fourier-Mellin 变换的图像配准应用中, 具体情况和傅里叶变换位移理论的假设有所不同, 这主要体现在: 1) 图像本身是离散取样的, 图像间的平移量可能对应于非整数倍的像素值; 2) 一般来说, 被配准的两幅图像并不严格满足相似变换关系;

3) 两幅图像之间多数情况下只有部分重叠; 4) 在基于 Fourier-Mellin 变换的图像配准基本算法中, 第一步运算过程中的 Log-polar 坐标变换和第二步运算过程的图像反变换, 都存在插值误差; 5) 由于变换参数都是离散取值, 所以第一步求得的旋转角度和尺度缩放系数是有误差的, 而且将直接影响第二步的运算; 6) 一些其他的诸如背景噪声等原因也会引入误差.

通过理论分析和大量的实验发现, 在第二步的相位相关计算中, 由于上述多种噪声的存在, 对相关值 II 的大小和相关位置的分布造成了不同程度的影响, 但总的来说, 相关值 II 呈现出如下三类分布形式.

1) 单峰分布

在理想情况下, 即在没有任何误差的情况下, 在正确的匹配位置处的相关值最大, 且等于 1, 而其他位置的相关值为 0, 是标准的脉冲函数; 在具有较小误差的情况下, 仍具有明显的单峰值, 即峰值 II 明显大于其他位置的相关值.

2) 无规律分布

和单峰分布的情况正相反, 相关值 II 在每点的数值都很小, 且呈现无规律分布.

3) 局部凸峰分布

考虑到上述各种噪声的影响, 被配准的两幅图像之间的相关值 II 还常常会呈现出另一种分布, 这种分布既不是明显的单峰值结构, 也不是毫无规律分布. 表现出来的形式是, 相关值 II 中前几个依次最大的相关值相差不大, 且其对应的相关位置全部或大部分都集中分布于最大相关位置的较小邻域 (并非一定是以最大相关位置为中心的对称分布), 形成一个局部“凸峰”. 更具体地, 相关值 II 的这种凸峰分布还可以进一步细化为以下三种形式:

a) 前几个最大相关值相对较大, 但数量较少, 通常只有 2、3 个.

b) 前几个最大相关值数值较小 (甚至和噪声数据相差不多), 但数量较多, 通常可达到 4 个到 10 个以上.

c) 有时, 前几个最大相关值对应的相关位置中会有个别位置偏离这个集中的凸峰区域, 但这些位置往往是孤立的, 不会形成相对集中的区域, 这通常是由随机因素造成的.

上述三种分布情况可以解释为: 在较为理想的情况下, 相关能量主要集中在某一点, 且数值较大; 随着误差的增大, 主要的相关能量分散在一个较小的区域, 而且数值越大, 分布范围越窄, 数值越小, 分布范围越宽; 当两幅图像本身就不配, 或者误差大到无法判断是否配准时, 相关能量呈现出无规律的分

布,且每点的相关值都很小.

3 问题的提出及总体构想

我们所关心的是,如何判定 Fourier-Mellin 配准算法计算出来的匹配参数的正确性,换句话说,如何根据这些参数自动判断两幅图像之间的配准情况.

根据第 2 节的分析,本文的第一个想法是:不依赖于相关值 I 的大小和分布,而主要依靠相关值 II 的大小和分布来判断匹配参数的正确性.

从第 2.2 节的分析可以看出,尽管相关值 II 的大小和分布具有一定的规律,但要定量和准确地描述这些规律却比较困难,尤其是凸峰分布的多样化,使问题描述变得十分复杂.然而,基于 Fourier-Mellin 变换的图像配准方法能够成立的前提是对应于正确配准的参数在对应的参数空间可以聚类,因此,就有可能通过机器学习的方法,采用分类技术来对配准结果做出判定.

这个思路构成了本文的第二个想法:将相关值 II 的大小和分布看作一个向量样本,输入到已经学习好的分类器,由分类器自动分为配准和不配准两种情况,这样,就将图像的匹配度量问题转化成了参数空间的模式分类问题.

向量样本可以这样理解:相关值 II 是对应每个相关位置的函数,为了便于理解,可将这个函数通过一幅假想的图像来描述,在图像中每个像素的位置对应于相应的相关位置,像素点的灰度值由对应该位置的相关值 II 代替,该位置的相关值大,类似于普通图像中该点的灰度值高,同样,该位置的相关值小,类似于该点的灰度值低,这样就可以将相关值 II 的大小和分布看作一幅“相位相关图”,提取每幅相位相关图的特征数据就构成了一个向量样本.

归纳起来,本文的总体思路可用图 1 进行表述.

在本文中,由于相关值 II 的大小和分布特征主要体现在最大相关位置周围较小的邻域内,所以仅取“相位相关图”中最大相关位置周围的一个“子图像”就足够了.本文中后面的分类器设计和实验仅取以最大相关位置为中心的 5×5 大小的子区域作

为研究对象.

4 基于学习的自动图像配准检验方法

4.1 分类器设计和分类决策

从图 1 可以看出,图像配准的检验问题转化成了一个典型的模式分类问题.本文采用有监督的学习方法来设计分类器:首先用一定数量的已知样本(训练集)进行分类器训练,然后用训练好的分类器对被识别的样本进行分类决策.

由于支持向量机(SVM)具有较好的泛化能力,本文采用支持向量机作为分类器模型^[8-9],其中使用的核函数为高斯函数,定义为

$$k(x_1, x_2) = \exp\left(-\frac{|x - x_i|^2}{\sigma^2}\right), \quad \text{取 } \sigma^2 = 0.1$$

相应的分类器决策函数为

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i^* y_i k(x, x_i) + b^*\right) \quad (11)$$

通过机器学习的方法确定了上述判别函数中的 n 个支持向量和其他参数 α_i^* 和 b^* 后,也就完成了分类器的设计.当任一新来的样本数据输入上述分类器后,符号函数 $f(x)$ 的两种不同取值,分别代表了图像匹配和不匹配的情况.

4.2 原始数据的获取

这部分内容对应于图 1 中的模块 1.手持数码相机,旋转不同的角度、调节不同的焦距,选取了 5 种不同的场景类型(如下页图 2 所示),每种场景拍摄了 80 对图片,共计 400 对,拍摄时使这些图像有的近似满足相似变换关系,有的则专门不满足这种关系,分别用于图像配准和不配准的数据.

应用基于 Fourier-Mellin 变换的基本配准算法计算图像间的匹配参数,然后采用图像拼接技术来主观判断两幅图像是否匹配.这样,把每次的实验结果全部记录下来,作为原始数据库.数据库中的每个记录格式为:

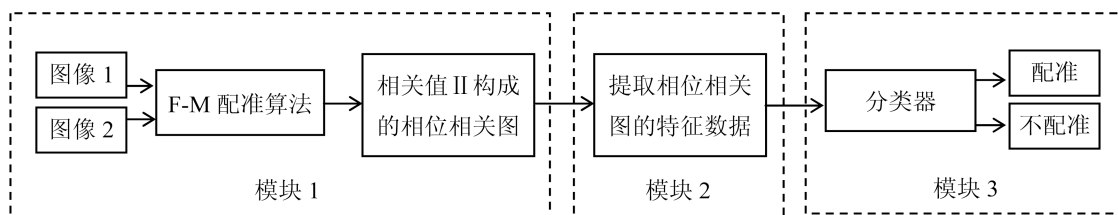


图 1 自动图像配准检验方法流程图

Fig. 1 Flowchart of automatic evaluation method for image registration



图 2 用于机器学习的不同图像类型

Fig. 2 Some samples used for machine learning

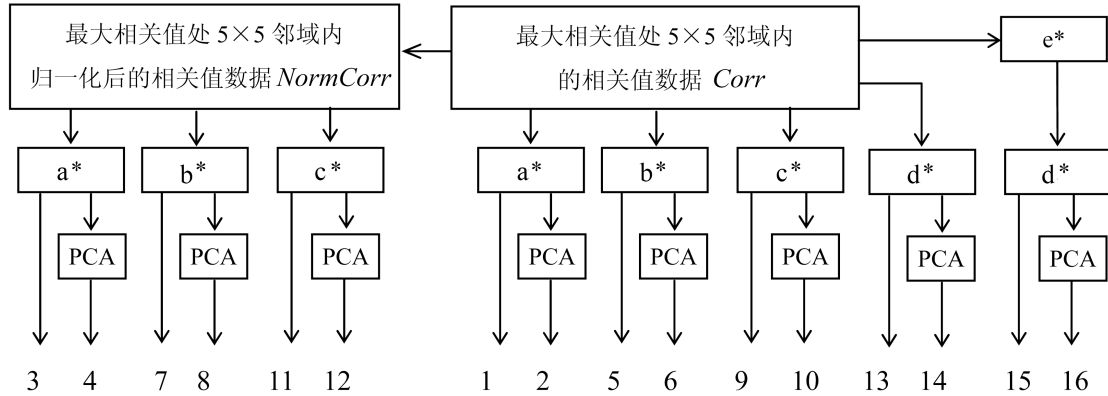


图 3 不同的特征数据构造方式

Fig. 3 Construction forms for various characteristic data

$$Record_i = \{BaseInfo_i, Corr_i, Sort_i\}$$

$$i = 1, 2, \dots, 400$$

其中, $BaseInfo_i$ 、 $Corr_i$ 、 $Sort_i$ 是对应每个记录的数组. $BaseInfo$ 中记录一些基本信息, 如图像名称、图像大小、是否配准等; $Corr$ 是个 5×5 数组, 记录以最大相关位置为中心的 5×5 邻域内每个位置的相关值大小 (每个记录 $Corr_i$ 对应一幅相位相关图); $Sort$ 也是个 5×5 数组, 记录 $Corr$ 数组中对应位置的相关值在图像所有相关值中按从大到小的排列序号.

我们得到的 400 对样本数据中, 共有 256 对匹配数据, 144 对不匹配数据. 其中各取一半组成训练集和测试集.

4.3 不同类型特征数据的分类效果比较

特征数据选取的好坏, 对模式分类系统的性能有重要的影响. 为了比较不同类型的特征数据的分类性能, 我们对原始的相关值数据、归一化后的相关值数据、不考虑相关值大小仅考虑位置分布构成的数据, 以及是否采用主分量分析 (PCA) 先作预处理等不同的组合共构造了 16 种不同的特征数据, 如图 3 所示, 各种特征数据的分类性能比较见表 1 (见下页).

下面给出图 3 中一些术语和符号的解释:

1) 对应于每个记录 $Corr_i$, 归一化的相关值定义为

$$NormCorr(m, n) = \frac{|Corr(m, n)|}{|\max(Corr(m, n))|}$$

其中, $Corr(m, n)$ 代表 5×5 邻域中位于位置 (m, n) 处的相关值, $|\cdot|$ 代表相关值的模, $\max(\cdot)$ 代表该记录中的最大相关值.

2) 符号 “a*”: 数据按先行后列依次排列为 25 维.

3) 符号 “b*”: 数据按先行后列依次排列为 25 维, 如果这些数据是前 n 个最大数之一, 那么相应的数据保留, 否则置 0.

4) 符号 “c*”: 取图像所有相关值中的前 n 个最大数按从大到小顺序组成 n 维特征数据, 如果这些相关值对应的位置位于 5×5 邻域, 则保留这些数据, 否则, 对应的分量置 0.

5) 符号 “d*”: 不考虑相关值大小而仅考虑前 n 个最大值的分布位置构造的数据. 取相关图中前 n 个最大数据组成 n 维特征数据, 如果这些相关值对应的位置位于 5×5 邻域, 那么对应分量置 1, 否则,

表 1 不同特征数据的机器分类效果比较

Table 1 Comparison of classification results for different types of characteristic data

| 特征数据类型 | 数据维数 | SVM 学习结果 | | | 效果评价 |
|--------|------|----------|------------|------------|------|
| | | 支持向量个数 | 训练集正确率 (%) | 测试集正确率 (%) | |
| 1 | 25 | 141 | 65 | 64 | 差 |
| 2 | 4 | 18 | 99.5 | 100 | 较好 |
| 3 | 25 | 116 | 73 | 67.5 | 差 |
| 4 | 19 | 87 | 85.5 | 68.5 | 差 |
| 5 | 25 | 141 | 65 | 64 | 差 |
| 6 | 4 | 19 | 99.5 | 100 | 较好 |
| 7 | 25 | 141 | 65 | 65 | 差 |
| 8 | 12 | 29 | 95 | 95 | 中 |
| 9 | 6 | 17 | 100 | 100 | 好 |
| 10 | 2 | 13 | 100 | 100 | 好 |
| 11 | 6 | 21 | 97 | 96.5 | 中 |
| 12 | 4 | 24 | 98 | 95 | 中 |
| 13 | 6 | 10 | 97 | 97 | 中 |
| 14 | 6 | 29 | 98.5 | 97 | 中 |
| 15 | 6 | 10 | 98 | 99.5 | 中 |
| 16 | 3 | 28 | 99.5 | 100 | 较好 |

对应的分量置 0.

6) 符号“e*”: 基本同情况 5), 不同的是将记录在 *Corr* 中对应于单峰分布的数据剔除 (这时数据库中只剩 378 个记录) 后的数据作为基本数据.

7) “PCA”: 主成分分析. 按前几个最大特征值之和占有所有特征值之和 95% 的原则, 保留这些特征值对应的特征向量, 然后将数据投影到这些特征向量空间, 特征数据维数也相应地被减少.

8) 最后需要说明的是, 上述 3)~6) 中提到的关于“ n 个”的问题. 我们曾对 $n = 3, 4, 6, 8, 10$ 时的情况进行了比较, 结果没有太大的影响, 故在这里不单独列出.

即使对于宽峰分布来说, 组成宽峰的数据一般也仅 10 个左右, 而其他数据大都是噪声数据, 这里取前 n 个数据而忽略其他数据主要是为了尽量避免噪声数据的影响.

4.4 讨论

表 1 中不同种类特征数据的学习和分类的效果比较给出了很多有益的启示和结论.

1) 总体上看, 将相关值归一化后的学习效果不好. 如表中 3、4、7、8、11、12, 和归一化前的数据比较, 结论显而易见. 这就是说, 相关值的绝对大小仍然是数据的很重要的特征, 因此采用归一化数据是不可取的.

2) 总体上看, 将原始数据通过主分量分析 (PCA) 后, 其结果好于原始数据. 表 1 中, 两两依次组成的 8 组数据中, 其中偶数号对应于 PCA 后的结果.

3) 从表 1 中的 1 和 5 可看出, 对最大相关位置 5×5 邻域内的相关值直接按行列顺序排成 25 维的特征数据效果不好, 但是经过 PCA 预处理后, 如 2 和 6, 效果接近最好. 2 和 6 的区别在于是否保留邻域内的噪声数据, 结果发现两者的区别不大. 这主要是由于无规律的噪声数据相比前几个最大相关值来说比较小, 因此起主导作用的仍然是前几个最大的相关值.

4) 观察表 1 中的 13 和 14, 这里忽略了相关值的绝对大小, 仅考虑前 6 个最大相关值对应的位置, 同时考虑了凸峰分布中出现个别位置偏离集中区域的情况, 效果还是不错的, 但仍有一定的错误率. 这种错误判断主要是由于峰值分布的差异造成的. 如单峰分布时, 前 6 个最大相关位置中可能只有一个最大相关位置落入这个 5×5 区域内, 而对不同宽度的凸峰分布来说, 可能会有 2~6 个最大相关位置位于 5×5 区域内.

5) 表 1 中的 15 和 16 是剔除单峰分布的数据后的结果. 可以看出, 结果有所改善, 仍然存在的错误识别源自于不同凸峰分布之间的差异. 同时也说

明, 仅利用峰值的位置来判断, 要想绝对避免判断误差是很困难的.

6) 最好的结果是表 1 中的 9 和 10, 这里既考虑了峰值的大小, 也考虑了出现个别位置偏离集中区域的情况, 100% 的正确率也说明了这种特征数据的构成方式具有鲁棒性. 当然, 在我们的实验中, 没有发现误分类的情况, 但这并不是说在任何情况下均不会出现误分, 这里仅仅说明这种方法有很高的可信度.

7) 在实验中, 我们还对 SVM 学习中的不同核函数, 如线性函数、多项式函数和 S 函数进行了简单的比较, 结果发现本文采用的高斯核函数效果较好, 而且对核函数中的参数并不很敏感.

5 结论

由于基于 Fourier-Mellin 变换的图像配准算法中第一步相位相关计算出来的参数的不确定性, 再加上各种噪声的影响, 使得第二步相位相关计算出来的相关值的大小和分布情况变得十分复杂. 针对这个问题, 本文将图像的配准度量问题转化为参数空间的模式分类问题, 研究探讨了一种基于机器学习的自动图像配准检验方法, 较好地解决了图像配准中的自动检验问题.

文中还对 400 组样本的匹配参数, 采用不同类型的特征数据进行了分类性能比较和讨论. 实验发现, 在分类时, 样本数据不应该归一化, 同时, 用 PCA 进行预处理可以有效提高检验精度.

尽管这种相似性度量方法的具体实现是针对基于 Fourier-Mellin 变换的配准算法的, 但我们认为这种基于学习的相似性度量思想同样可以应用到其他一些方法中.

References

- Zitova B, Flusser J. Image registration methods: a survey. *Imaging and Vision Computing*, 2003, **21**(11): 977–1000
- Viola P, Wells W M. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 1997, **24**(2): 137–154
- Huttenlocher D P, Klanderman G A, Rucklidge W A. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1993, **15**(9): 850–863
- Reddy B S, Chatterji B N. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 1996, **5**(8): 1266–1271
- Chen Q S, Defrise M, Deconinck F. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, **16**(12): 1156–1168
- Stone H S, Tao B, McGuire M. Analysis of image registration noise due to rotationally dependent aliasing. *Journal of Visual Communication and Image Representation*, 2003, **14**(2): 114–135
- Lucchese L, Cortelazzo G M, Monti C. High resolution estimation of planar rotations based on Fourier transform and radial projections. In: *Proceedings of IEEE International Symposium on Circuits and Systems*. Hong Kong, China: IEEE, 1997. 1181–1184
- Vapnik V N. *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995. 314
- Joachims T. Making large-scale SVM learning practical. *Advances in Kernel Methods: Support Vector Learning*. Cambridge, USA: MIT Press, 1999. 169–184



李晓明 博士, 副教授. 主要研究方向为图像配准, 计算机视觉和模式识别. 本文通信作者. E-mail: lixm@sxu.edu.cn
(**LI Xiao-Ming** Ph.D., associate professor. His research interest covers image registration, computer vision, and pattern recognition. Corresponding author of this paper.)



张继福 博士, 教授. 主要研究方向为智能信息处理和模式识别. E-mail: zjf@tyust.edu.cn
(**ZHANG Ji-Fu** Ph.D., professor. His research interest covers intelligent information processing and pattern recognition.)