

多分辨率迭代切距离及其在目标识别中的应用

王国刚^{1,2,3,4}, 史泽林^{1,3,4}, 刘云鹏^{1,3,4}

(1. 中国科学院沈阳自动化研究所, 辽宁 沈阳 110016; 2. 沈阳化工大学 信息学院, 辽宁 沈阳 110142;
3. 中国科学院研究生院, 北京 100049; 4. 中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016)

摘要: 针对在目标识别过程中切距离的线性逼近、易陷入局部最优的局限性, 提出了一种新的仿射变换不变距离度量, 即多分辨率迭代切距离(MITD), 并将其用于仿射变形下的目标识别。MITD 将迭代切距离嵌入多分辨率框架, 计算模板图像和实时图像之间的变换不变距离, 以扩大算法的收敛域, 提高目标识别算法的识别概率。实验结果表明, 所提出的基于 MITD 的目标识别方法较基于欧氏距离(ED)和切距离(TD)的算法具有明显优势, 优于基于迭代切距离(ITD)的算法, 具有较高的鲁棒性。

关键词: 迭代切距离; 仿射变换; 切距离; 多分辨率

中图分类号: TP391 **文献标志码:** A **文章编号:** 1007-2276(2012)05-1369-05

Multi-resolution iterative tangent distance for object recognition

Wang Guogang^{1,2,3,4}, Shi Zelin^{1,3,4}, Liu Yunpeng^{1,3,4}

(1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China;

2. Institute of Information, Shenyang University of Chemical Technology, Shenyang 110142, China;

3. Graduate University of the Chinese Academy of Sciences, Beijing 100049, China;

4. Key Laboratory of Optical-Electronics Information Processing, Chinese Academy of Sciences, Shenyang 110016, China)

Abstract: A new transformation invariant metric that was multi-resolution iterative tangent distance was proposed based on analyzing the limitation of tangent distance, which approximate the real distance between object and template linearly and prone to fall into local optimum limitations in the process of target recognition. The iterative tangent distance was embedded into a multi-resolution framework and the real transform invariant distance was obtained between template and real image through iterations in order to expand the domain of convergence of the algorithm, and improve the probability of identification of the target recognition algorithm. Experiments of object recognition show that the proposed algorithm outperforms existing algorithms based on Euclidean distance, tangent distance and iterative tangent distance, and it has strong robustness.

Key words: iterative tangent distance; affine transform; tangent distance; multi-resolution

收稿日期: 2011-09-05; 修订日期: 2011-10-03

基金项目: 中国科学院国防科技创新基金(CXJJ-65); 辽宁省高等学校杰出青年学者成长计划(LJQ2011039);
辽宁省博士启动基金(2011045)

作者简介: 王国刚(1977-), 男, 博士生, 主要从事自动目标识别方面的研究。Email: ggwang@sia.cn

导师简介: 史泽林(1965-), 男, 研究员, 博士生导师, 博士, 主要从事图像处理与模式识别方面的研究。Email: zlshi@sia.cn

0 引言

利用 2D 图像进行目标识别是计算机视觉的重要组成部分,在城市规划、交通监控、图像检索及工业自动化等领域得到了广泛应用^[1]。目标识别系统是现阶段和未来武器系统的重要组成部分。

在成像制导目标识别系统中,由于基准图像和实时图像之间存在着各种差异,自动目标识别系统面临的问题是需要不同的条件下识别出目标。导弹导引头所获取的图像是在不同位置、不同季节、不同时段实时图像,并且所攻击的目标在导弹飞行中也可能不断地进行机动,因此,基准图和实时图之间往往存在着由于拍摄距离、角度、位置和姿态等变化引起的几何变形。若只存储少量的基准图,则实时图与基准图之间的巨大差异会使匹配精度严重下降。即使不限制基准图的数量,识别系统需要在多维变量空间中对目标类别进行复杂的搜索和匹配,也会使匹配效率下降,并增大匹配的困难。

长期以来,很多学者提出了多种解决几何变形下图像匹配问题的方法。预存目标的多个模板是一种常用的方法,分别用每一个模板在图像中进行相似性度量,寻找所有结果中最相似的一个作为定位结果,这种方法需要较大的计算量和存储量。子空间算法所需存储的特征向量空间数据量相比多模板策略要少很多^[2],但是它对尺度变化和旋转角度变化很敏感。

如果能针对几何变换找到目标的不变特征或度量进行分类就可以通过将测试样本归类为与其最近的目标特征不变点来进行分类,这将有效降低目标识别系统的难度。在目标成像过程中,目标发生的几何变形在更多的情形下可模型化为一个仿射变换。现有的仿射不变特征主要包括简比、交比、角点以及仿射不变矩等^[3],这些不变量只适用于在目标清晰、成像环境好的条件下识别目标,因此其使用范围受到了极大的限制。此外,这些不变量对噪声十分敏感,从而使得特征的不变性并不稳定,因此,大多数的研究思路是寻找模式之间具有仿射不变性的距离度量。切距离(Tangent Distance)及其相关技术^[4-5]是这类方法的典型代表。Haasdonk B 等^[6]将切线距离与支持向量机等分类器集成,应用于数字识别中,取

得了较好的效果。

文中在深入研究了切距离、迭代切距离及其局限性的基础上提出了一种多分辨率迭代切距离,通过迭代找到模板图像和实时图像之间的不变距离度量,并应用于仿射变换下的目标识别中。实验结果验证了所提出的多分辨率迭代切距离的优越性。

1 多分辨率迭代切距离

1.1 切距离

切距离的核心思想是尽量使距离度量保持对图像目标变换的不变性。切距离方法认为一个原型模式经过各种变换产生的模式在空间中扩展为一个流形结构,该流形在原型模式点处可以用一个线性切空间来近似。所谓切距离就是输入模式到该原型模式线性切空间的距离。假设已经知道所需处理的问题会涉及 r 种变换,比如水平平移、垂直平移、剪切、旋转、尺度变换等。对每一个原型样本点 M (模板)都进行所有变换操作 $F_i(M, t_i)$,这样 $F_i(M, t_i)$ 代表模板图像经过变换 t_i 后得到的新图像,然后对每一种操作,都有一个切向量 TV_i :

$$TV_i = F_i(M, t_i) - M \quad (1)$$

这样,对模板图像的每一种变换,只需计算一次。对于模板图像,构造 $r \times d$ 的矩阵 T ,矩阵 T 由 M 的切向量组成, d 表示图像的像素数。在由 r 个通过 M 的切向量张成的子空间中,每一个点代表对所有操作的线性逼近。切距离的概念如图 1 所示,其中 M 代表模板, P 表示变换后的图像目标。

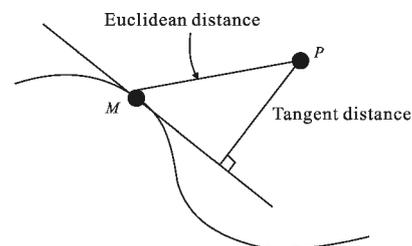


图 1 P 和 M 之间的切距离

Fig.1 Tangent distance of P and M

如果矩阵 T 由 M 处的 r 个切向量组成,则测试样本 P 到 M 之间的切距离为:

$$D_T(M, P) = \min [\| (M + T\alpha) - P \|] \quad (2)$$

公式(2)的计算是一个最小二乘问题,该优化问题的

一个必要条件为公式(2)对参数 α 的偏导数为零。因此有以下命题:

设 P 和 M 为图像流形上的两个点, $D_T(M, P)$ 为公式(2)给出的 P 到 M 之间的切线距离。

如果 T 为图像流形 M 的切向量, 则公式(2)的解 α 应满足:

$$T'(M+T\alpha-P)=0 \quad (3)$$

证明:

$$\begin{aligned} \frac{\partial D(M, P)}{\partial \alpha} &= \frac{\partial}{\partial \alpha} (M+T\alpha-P)'(M+T\alpha-P) = \\ &= \frac{\partial}{\partial \alpha} (M'+\alpha'T'-P')(M+T\alpha-P) = \\ &= T'(M+T\alpha-P)+(M+T\alpha-P)'T = \\ &= 2(M+T\alpha-P, T) = 2T'(M+T\alpha-P) = 0 \end{aligned}$$

由于 M, P 和 α 为实向量, T 为实矩阵, 因此有 $T'(M+T\alpha-P)=0$ 。

切距离本质上是一种可变形模板匹配, 用变化空间在模板处的切平面对变化空间进行近似, 具有对几何变换保持不变的性质。然而, 切空间距离只是对真实变换的一种线性逼近, 易陷入局部最优, 因此切空间距离只会在参数 α 较小的变化范围内有效。

1.2 多分辨率迭代切距离

由公式(2)可知, 切线距离 $D_T(M, P)$ 是用线性化的方法, 以切线向量组矩阵 T 和参数 α 来修正模板 M , 使得变换图像目标与模板的距离最小。模板变化的方向由切向量确定, 模板在该变化方向的大小由切线参数 α 确定。切距离是模板变换的一阶逼近, 因此, 参数 α 的适用范围不会太大。据此, 考虑采用逐步求精的方法对公式(2)进行求解, 就是采用迭代的方法, 每次迭代更新变量 $D_T(M, P)$ 的值, 直到收敛, 得到的 $D_T(M, P)$ 就是迭代切距离。迭代过程为:

$$\alpha^{n+1} = \alpha^n - \lambda [\nabla_{\alpha}^2 D|_{\alpha=\alpha^n}]^{-1} \nabla_{\alpha} D|_{\alpha=\alpha^n} \quad (4)$$

式中: $\nabla_{\alpha} D$ 和 $\nabla_{\alpha}^2 D$ 分别是目标函数(2)对于参数 α 的梯度和赫森矩阵。选择 $\alpha^0=1, \lambda=1$, 切距离对应迭代切距离的第一次迭代结果, 迭代切距离的计算直到公式(4)收敛。与图像配准算法中的高斯牛顿迭代算法类似, 可以证明迭代切距离的收敛性^[7]。迭代切距离消除了切距离的一阶近似, 表征了模板图像和变换目标图像之间的变换不变距离。

然而, 公式(4)的迭代过程的一个明显缺陷是迭代算法的收敛域较小, 容易陷入局部最优。为了克服

这一缺点, 将迭代切距离嵌入多分辨率框架, 以提高算法性能。在物理层面上, 多分辨率的使用相当于扩大了收敛域。为了计算多分辨率迭代切距离, 将模板和要分类的模式通过高斯金字塔进行分解, 在金字塔的每层以上层的结果为初始点计算迭代切距离:

$$\alpha_l^{n+1} = \alpha_l^n - \lambda [\nabla_{\alpha}^2 D|_{\alpha=\alpha_l^n}]^{-1} \nabla_{\alpha} D|_{\alpha=\alpha_l^n} \quad (5)$$

式中: l 代表金字塔分解的第 l 层。相对切距离, 多分辨率迭代切距离保持了切距离的几何变换不变特性。与切距离和迭代切距离相比, 多分辨率迭代切距离对仿射变换具有更好的鲁棒性。

2 仿射不变目标识别

考虑仿射变换下的目标识别, 二维欧氏空间上的仿射变换可以表示为:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r_1 & r_3 \\ r_2 & r_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} r_5 \\ r_6 \end{bmatrix} \quad (6)$$

式中: $(x, y), (x', y')$ 分别是图像平面中相对应两个点的坐标; $\begin{bmatrix} r_1 & r_3 \\ r_2 & r_4 \end{bmatrix}$ 是一个 2×2 非奇异矩阵; $\begin{bmatrix} r_5 \\ r_6 \end{bmatrix}$ 是平移向量。 $\vec{r}=(r_1, r_2, r_3, r_4, r_5, r_6)$ 表示仿射变换的 6 个参数。首先, 根据给定的模板图像计算其切向量, 由于仿射变化有 6 个元素, 因此含有 6 种操作: 平移、旋转、缩放和斜切。在计算切向量的过程中, 要计算图像关于 x 和 y 的偏导数, 这里采用中心差分的方式:

$$Q_x(i, j) = \frac{Q(i+1, j) - Q(i-1, j)}{2} \quad (7)$$

$$Q_y(i, j) = \frac{Q(i, j+1) - Q(i, j-1)}{2} \quad (8)$$

式中: $Q(i, j)$ 为图像矩阵。图 2 给出了模板图像及其对应的 6 个切向量示意图

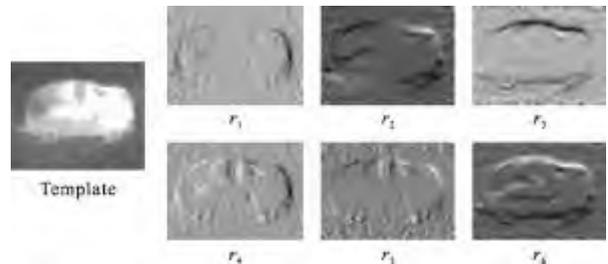


图 2 模板及切向量

Fig.2 Template and tangent vectors

对于给定的测试图像, 根据公式(3)计算参数 α 来修正模板 M , 使得变换图像目标与模板的距离最

小。将 α 代入公式(4),通过高斯牛顿迭代计算新的参数 α ,重复修正,直到 $\|\alpha\|$ 小于一定的阈值,得到的 $D_r(M, P)$ 即为仿射不变距离。根据此距离可构建最近邻分类器对目标进行识别分类,其中计算迭代切距离算法流程如下:

- (1) 计算模板的切向量,参数初始化为: $\alpha^0=1$, $\lambda=1$;
- (2) 计算目标公式(2)对于参数 α 的梯度和赫森矩阵;
- (3) 根据公式(4)计算新的变换参数 α ;
- (4) 若 $\|\alpha\| < \varepsilon$ 则转(5),否则转(2);
- (5) 给出模板与测试样本之间的迭代切距离。

在实际计算过程中,选择金字塔的层数为 4。当前层参数 α 得到后,将该参数传入多分辨率的更低层,作为初始估计值,以最后一层计算的切距离作为最终结果。

3 实验分析

为了验证所提出算法的有效性,将所提出的基于多分辨率迭代切距离(MITD)的识别方法与传统基于欧氏距离(ED)的方法、参考文献[5-6]介绍的基于切距离(TD)及基于迭代切距离(ETD)的识别方法进行了比较,并用 matlab 语言编程实现。

此实验分为两部分,前一部分用已知简单的飞机目标图像,在给定的仿射变换下进行识别实验;后一部分验证实际场景中的目标识别。

实验中首先验证了 5 个不同的飞机模型图像在相同仿射变换下的可区分性。图 3 所示的 5 个飞机目标图像分别为 F-15、F-16、Mig-29、幻影以及空

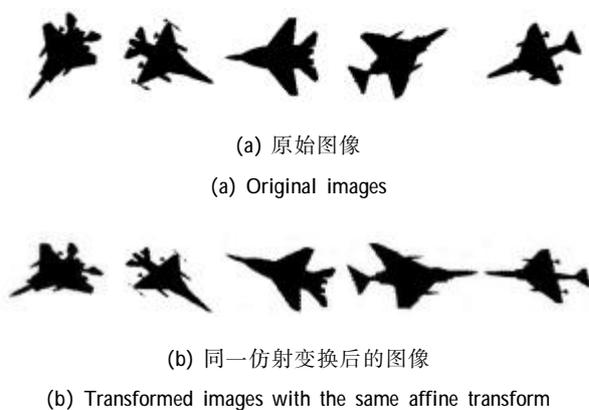


图 3 飞机图像及变形后图像

Fig.3 Plane images and the transformed images

中之鹰。根据参考文献[8]产生仿射变换的方式对 5 个飞机目标进行图像变换,对每一个飞机都采用相同的仿射变换。变换标准差从 2 到 20,步长为 2。对每一个标准差随机生成 100 个仿射变换。分别使用欧氏距离、切距离、迭代切距离和多分辨率迭代切距离,根据最近邻分类器对变换后的飞机目标进行分类。

图 4 给出了识别结果,横坐标表示变换标准差,

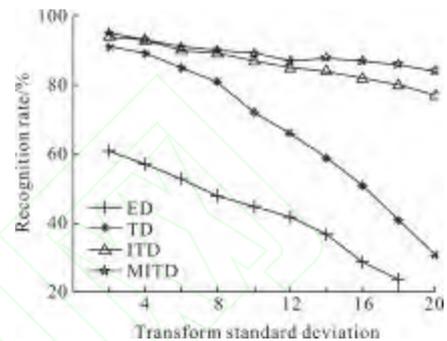


图 4 飞机图像识别率

Fig.4 Recognition rates of plane images

纵坐标表示识别率。从图 4 可以看出,基于 MITD 的识别算法的识别率明显优于基于 ED 和 TD 的识别算法,性能略好于基于 ITD 的算法,当仿射变换的标准差达到 20 时,所提出的基于 MITD 的识别方法的识别率仍然达到 84%,基于 ITD 的识别算法达到 77%,而 ED 和 TD 算法的识别率已经降到 30%左右。实验结果验证了所提出算法的有效性和优越性。

第二部分实验验证了 MITD 对实际场景中的坦克和运输车辆进行分类识别的效果。图片来自于 Google Earth,如图 5 所示,数据是这两类目标在不

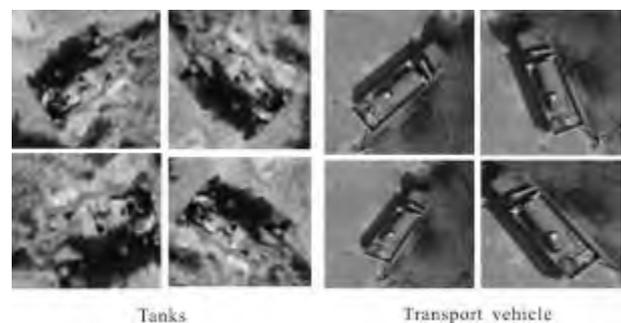


图 5 不同类型目标

Fig.5 Different object types

同方位角和距离下的目标图像,各 50 幅。左上角的

图像为模板图像,根据模板与实际目标之间的距离进行分类,识别结果如表1所示。

表1 坦克和运输车辆识别结果

Tab.1 Recognition results of tank and transport vehicles

Target	Recognition rate/%			
	ED	TD	ITD	MITD
Tanks	44	62	76	82
Transport vehicles	46	70	80	84

从实验结果看,在坦克和运输车辆的识别中,基于MITD的识别算法的识别率优于基于ED、TD和ITD的识别算法,MITD算法可有效地提高目标识别率。

4 结论

文中对成像末制导目标识别中遇到的几何变形问题进行了研究。提出了一种新的仿射变换不变距离度量,即多分辨率迭代切距离,并将其应用于仿射不变目标识别中。迭代切距离更逼近模板和实时图之间的真实距离,多分辨率框架扩大了收敛域。因此所提出的多分辨率迭代切距离与切距离和迭代切距离相比具有较高的鲁棒性。仿射变换下的目标识别实验结果验证了算法的有效性,为几何变形下的目标识别提供了新的研究思路。但是文中使用MITD识别算法的识别过程是简单的最近邻分类过程,没有加入学习训练的环节,还有性能提升的潜力,是下一步工作的主要方向。

参考文献:

- [1] Zhao D M, Chen J. Affine curve moment invariants for shape recognition pattern recognition[J]. 1997, 30(6): 895-901.
- [2] Zhang Zhijia, Zhang Yu, Huang Shabai, et al. Image matching based on region template correlation [J]. Infrared and Laser Engineering, 2004, 33(6): 615-618. (in Chinese) 张志佳, 张瑜, 黄莎白, 等. 基于区域模板相关的图像匹配方法研究[J]. 红外与激光工程, 2004, 33(6): 615-618.
- [3] Wang Songwei, Li Yanjun, Zhang Ke. Seemingly similar 3D target recognition based on local wavelet -moment [J]. Infrared and Laser Engineering, 2008, 37 (6): 1106 - 1110. (in Chinese) 王松伟, 李言俊, 张科. 基于局部小波矩的易混淆三维目标识别[J]. 红外与激光工程, 2008, 37(6): 1106-1110.
- [4] Simard P Y, LeCun Y A, Denker J S, et al. Transformation invariance in pattern recognition tangent distance and tangent propagation [J]. International Journal of Imaging System and Technology, 2001, 11(3): 181-194.
- [5] Effrosyni Kokopoulou, Pascal Frossard. Minimum distance between pattern transformation manifolds[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31(7): 1225-1238.
- [6] Haasdonk B, Keysers D. Tangent distance kernels for support vector machines [C]//Conference on Pattern Recognition, 2002: 864-868.
- [7] Yuan Yaxiang, Sun Wenyu. Optimization Theory and Methods[M]. Beijing: Science Press, 1997. (in Chinese) 袁亚湘, 孙文瑜. 最优化理论与方法 [M]. 北京: 科学出版社, 1997.
- [8] Baker S, Matthews I. Lucas-Kanade 20 years on: a unifying framework [J]. International Journal of Computer Vision, 2004, 56(3): 221-255.