一种目标识别中基于关键点的仿射不变矩

吴高洁 李 超 熊 璋

(北京航空航天大学 计算机学院，北京 100191)

摘 要：针对常用仿射不变矩对轮廓边缘信息敏感的特点，提出了一种基于关键点的局部仿射不变矩，以分割出来的物体灰度图像为基础，首先计算出物体的质心，然后以质心为扩展点向周围引伸出多条射线，寻找每条射线方向上的最近灰度极值点，将所有灰度极值点当作关键点集合，并按照文中提出的计算仿射不变矩的方法提取出多阶不变矩，将其当作神经网络的输入向量，放入已经训练好的神经网络来达到识别物体的目的，将此方法应用在了飞机图像的识别上，实验结果证明此方法在物体轮廓分割不完整和有噪点污染的情况下能保持很好的稳健性，简单有效并具有较广的应用范围.

关 键 词：目标识别；关键点；仿射变换；仿射不变矩

中图分类号：TP 391.4

文献标识码：A **文 章 编 号**：1001-5965(2009)09-1043-05

Affine invariant based on determinant points in object recognition

Wu Gaojie Li Chao Xiong Zhang

(School of Computer Science and Technology, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: The common affine invariants were very sensitive to the edge of the image, so a new method entitled "an affine variant based on the determinant points in image recognition" was proposed. First the centroid of the separated object image was computed, and then many line segments were derived through the centroid, at the end the nearest extreme grayscale points were found on every line. In order to compute the affine invariants, the extreme points were used to build up a collection of determinant points. The affine invariants of the collection could be served as an input vector of the trained neural networks to distinguish whether the source image was the destination image. The method was applied to the plane recognition and was proved to keep highly stable even if the object contour was ill-segmented or noisy. It is much easier and effective compared with the traditional methods and has a very wide scope of application.

Key words: object recognition; determinant points; affine transform; local affine variant

在目标识别、遥感图像几何校正、图像检索等许多智能图像处理领域，需要从多幅不同视角获取的图像中提取相同特征量，并以此特征量作为后续处理的基础. 图像仿射不变特征量[1-12]已经成为图像目标识别的核心问题之一，比较常用的提取仿射不变特征的方法有以下几种.

1) Hu[1]矩. Hu于1962年第1次系统地提出了不变矩理论，构造了7个不变矩，Hu矩对尺度、旋转及平移具有不变性. 后来Reiss证明了Hu矩对一般形式的仿射变换不具有不变性，纠正了原文中的错误，并提出了Hu矩的改进形式.

2) Zernike[2]矩. 文献[2]中提出了Zernike矩，实质上，Zernike矩是一种映射，它将图像函数变换到一组正交基函数上，取代了几何矩定义中的非正交基函数集，因此在理论上Zernike矩比Hu矩具有更好的数学性质.

3) 文献[3,5]提出利用多尺度自卷积(MSA, Multi-Scale Auto-convolution)[3]变换的

收稿日期：2008-08-04

作者简介：吴高洁(1983—)，男，湖北荆州人，硕士生，wugaojie1983@163.com.

方法提取不变特征,具有良好的抗噪声性能,并且对一定范围内的局部遮挡、三维旋转和透视变换具有适应性,其分类能力优于 Hu 矩.

本文提取的方法主要针对飞机图像的识别,针对飞机提取中可能出现的物体轮廓分割不完整和有噪点污染的情况提出了一种基于关键点的局部仿射不变特征的概念.

1 基于关键点的仿射不变矩

1.1 局部仿射不变特征

仿射不变特征[4]可分为全局仿射不变特征和局部仿射不变特征[3-6].本文将引入局部不变矩的概念,提出了基于关键点的仿射不变矩[7],并将其引入神经网络中对物体进行识别,整体流程图如图 1 所示.

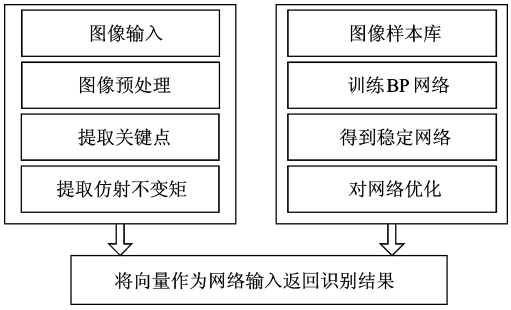

图 1 图像识别整体流程图

1.2 关键点提取

关键点是图像的一种重要局部特征,也可以称之为兴趣点[10],其在保留了图像中物体的重要特征信息的同时有效地减少了数据量,使得对图像处理时运算量大大减少.

J. Matas[8]方法用图像中灰度极值点作为关键点,灰度极值不受仿射变换的影响,因而其关键点在仿射变换前后易检测出.但是图像中灰度极值点的数量不确定,也易受遮障因素的影响.为使关键点能够较好克服上述影响,本文将选择质心作为第一关键点,可以较好地避开小部分轮廓缺失对关键点的影响,根据后面的证明,质心在仿射变换前后是满足仿射不变性的,关于质心[9]的计算公式如下:

$$\left.\begin{aligned} E_x &= \frac{\sum\sum x\boldsymbol{I}(x,y)}{\sum\sum \boldsymbol{I}(x,y)} \\ E_y &= \frac{\sum\sum y\boldsymbol{I}(x,y)}{\sum\sum \boldsymbol{I}(x,y)} \end{aligned}\right\} \tag{1}$$

单纯依赖质心的灰度和位置信息还不足以构造可分性较高的不变特征,因而需要结合质心周围的图像信息构造能够提供稳健的特征的感兴趣区域,称之为仿射不变区域,从质心点出发,寻找当前方向上与质心灰度差异最大且离质心最近的点,这样的点称为最近灰度极值点.用公式表示如下:

$$f(t) = \frac{|\ I(t) - I_0\ |}{\max\left(\dfrac{\int_0^t |\ I(t) - I_0\ |\,\mathrm{d}t}{t}, d\right)} \tag{2}$$

其中,t 为射线的长度;I_0 为质心的灰度值;d 为一个较小的整数防止分母为 0,由此得到构造仿射不变区域的方法:以质心为中心向各个方向作射线,在每条射线上沿发散方向寻找式(2)的最近灰度极值点 $f_{\max}(t_i)$.

将 $f_{\max}(t_i)$ 点依次连接得到闭合区域,然后在对应的仿射变换后的图像中,按照上述方法得到闭合区域,在经过质心的任意一条直线上 $f(t)$ 上都有的两个最近灰度极值点,文中 $f_{\max}(t_i)$ 的集合就是要找的关键点的集合,进而由对整幅图像的识别转而对关键点集合进行识别.关键点提取示意图如图 2 所示.

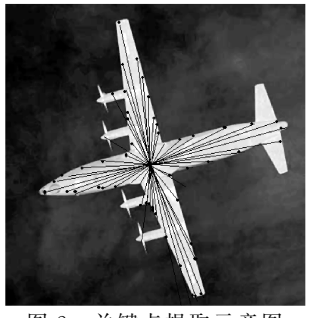

图 2 关键点提取示意图

1.3 基于关键点的仿射不变特征

下面证明"关键点"是仿射不变的.

命题 设图像 $\boldsymbol{I}$,其仿射变换后的图像 $\boldsymbol{I}'$,图像 $\boldsymbol{I}$ 中的点的坐标用 $\boldsymbol{F}(x,y)$ 表示,关键点的坐标用 $\boldsymbol{T}(x,y)$ 表示,仿射变换后图像 $\boldsymbol{I}'$ 中的点的坐标用 $\boldsymbol{F}'(x,y)$ 表示,关键点用 $\boldsymbol{T}'(x,y)$,已知 $\boldsymbol{F}'(x,y)=\boldsymbol{A}\cdot\boldsymbol{F}(x,y)$

证明 $\boldsymbol{T}'(x,y)=\boldsymbol{A}\cdot\boldsymbol{T}(x,y)$

证明过程:设点 $\boldsymbol{F}(x,y)$ 是图像中的任意一个点的坐标,点 $\boldsymbol{F}'(x,y)$ 是仿射变换后图像中对应点的坐标,它们之间的关系满足 $\boldsymbol{F}'(x,y)=\boldsymbol{A}\cdot\boldsymbol{F}(x,y)+\boldsymbol{b}$,$\boldsymbol{A}$ 为非奇异仿射变换矩阵,$\boldsymbol{b}$ 是仿射变换的偏移量.则对于任意一关键点 $\boldsymbol{T}(x,y)$ 必有 $\boldsymbol{T}(x,y)\in\boldsymbol{F}(x,y)$,由于仿射变换是一一对应的 $\boldsymbol{T}'(x,y)=\boldsymbol{A}\cdot\boldsymbol{T}(x,y)+\boldsymbol{b}$,所以有 $\boldsymbol{T}'(x,y)\in\boldsymbol{F}'(x,y)$,故上述命题成立. **证毕**

对于任意非负整数 p,q,二维图像灰度分布函数 $f(x,y)$在平面 R^2 上的 $p+q$ 阶原点矩定义为

$$m_{p,q}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^p x^q f(x,y)\mathrm{d}x\mathrm{d}y \tag{3}$$

中心矩定义为

$$u_{p,q}=\sum_{1}^{M}\sum_{1}^{N}(x-\overline{x})^p(y-\overline{y})^q f(x,y) \tag{4}$$

上述表达式对于数字图像的离散化形式如下:

$$m_{p,q}=\sum_{1}^{M}\sum_{1}^{N}x^p x^q f(x,y) \tag{5}$$

$$u_{p,q}=\sum_{1}^{M}\sum_{1}^{N}(x-\overline{x})^p(y-\overline{y})^q f(x,y)\mathrm{d}x\mathrm{d}y \tag{6}$$

如果一幅图像上的点的坐标向量用$[x,y]$来表示,对应的仿射变换后图像上的点的坐标向量用$[x',y']$表示,于是有

$$\begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}a,b\\c,d\end{bmatrix}\cdot\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}e\\f\end{bmatrix} \tag{7}$$

上述不等式通过坐标平移可以简化为

$$\begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}a,b\\c,d\end{bmatrix}\cdot\begin{bmatrix}x\\y\end{bmatrix} \tag{8}$$

对于图像上的任意两点$[x_i,y_i]$,$[x_j,y_j]$有

$$\left.\begin{aligned}\begin{bmatrix}x'_i\\y'_i\end{bmatrix}&=\boldsymbol{A}\begin{bmatrix}x_i\\y_i\end{bmatrix}\\\begin{bmatrix}x'_j\\y'_j\end{bmatrix}&=\boldsymbol{A}\begin{bmatrix}x_j\\y_j\end{bmatrix}\end{aligned}\right\} \tag{9}$$

于是有

$$\begin{bmatrix}x'_i,y'_i\\x'_j,y'_j\end{bmatrix}=\boldsymbol{A}\begin{bmatrix}x_i,y_i\\x_j,y_j\end{bmatrix}$$

对两边取行列式值,于是有

$$\det\begin{bmatrix}x'_i,y'_i\\x'_j,y'_j\end{bmatrix}=\det\boldsymbol{A}\cdot\det\begin{bmatrix}x_i,y_i\\x_j,y_j\end{bmatrix} \tag{10}$$

假设 P 是图像上点的坐标向量的集合,P'是对应的仿射图像上的点的坐标向量的集合,另假设一个有限集合 Q,Q 集合里面的元素是 P 集合里面的所有元素(x_i,y_i),$(x_j,y_j)\in P$ 经过如下公式计算出来的:

$$Q=\{d(i,j)\mid \mathrm{abs}(x_i y_i-x_j y_i)\} \tag{11}$$

由于 $P(x,y)$和 $P'(x',y')$的一一对应关系,所以对于仿射变换后的 P'也有集合

$$Q'=\{d'(i,j)\mid \mathrm{abs}(x'_i y'_j-x'_j y'_i)\} \tag{12}$$

且 Q 与 Q'是一一对应的关系,Q 与 Q'的元素之间相差一个 $\det\boldsymbol{A}$ 常数的关系

对两边取数学期望,有:

$$E(Q)=\det\boldsymbol{A}\cdot E(Q') \tag{13}$$

根据 K 阶中心矩的概念,中心矩 U_k 为

$$U_k=E\{[X-E(X)]^k\} \tag{14}$$

设 U_n 为 Q 的中心 N 阶矩,于是有

$$U'_n=\det\boldsymbol{A}\cdot{}^nU_n$$

两边开 n 次根号

$$\sqrt[n]{U'_n}=\det\boldsymbol{A}\cdot\sqrt[n]{U_n} \tag{15}$$

式(15)除以式(13)

$$\frac{\sqrt[n]{U'_n}}{E(Q)}=\frac{\sqrt[n]{U_n}}{E(Q')} \tag{16}$$

所以,如果假设

$$\frac{\sqrt[i]{U'_i}}{E(Q)}=M(i)\qquad i=1,2,3,4,5,6,\cdots \tag{17}$$

$M(i)$对于图像仿射变换前后都是一个常量,把 $M(i)$称之为 i 阶仿射不变矩,将式(14)带入式(17)有

$$\frac{\sqrt[i]{E\{[Q-E(Q)]^i\}}}{E(Q)}=M(i) \tag{18}$$

根据公式的推断,可以知道 $M(i)$对于给定图像集合 Q,仿射变换前后值保持不变.

2　实验结果及分析

为验证基于关键点的仿射不变矩的实用性,下面将进行多组实验(实验采用 Matlab7.0 平台)

1) 第 1 组实验

首先进行的实验是对同一物体波音 737 飞机(图 3)在 4 组不同仿射变换下(仿射变换的矩阵如表 1 所示)的 1～7 阶矩值情况.

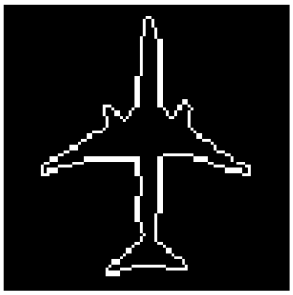

图 3　波音 737 飞机原图

表 1　4 组仿射变换矩阵

仿射变换	第 1 组	第 2 组	第 3 组	第 4 组
矩值	$\begin{pmatrix}1&0&0\\-0.5&1.5&0\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0.8&0.8&0\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\-0.1&0.8&0\\0&0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0.5&-1&0\\0&0&1\end{pmatrix}$

图 4 是波音 737 飞机在 4 组不同仿射变换下的图像.

表2列出了4组图像在1～7阶矩的值，可以看到在每一阶矩上4组图像矩值差异都很小，说明了本文提出的仿射不变矩在不同仿射变换下保持着很好的仿射不变性，这样对于任何视角下拍摄的目标计算得到的各阶不变量，都可以作为目标的不变特征，这也验证了此方法的实用性.

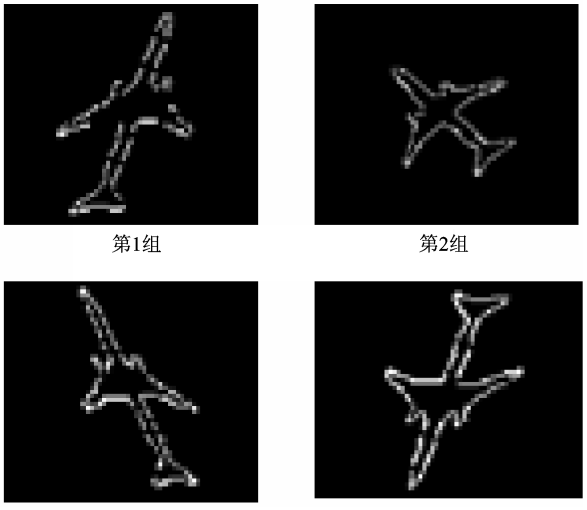

图4 波音737飞机4组仿射变换图像

表2 4组图像的1～7阶矩值

图像	K_1	K_2	K_3	K_4	K_5	K_6	K_7
原图	0.46	1.03	3.41	6.03	13.3	6.58	11.7
组1	0.74	0.87	1.86	4.89	10.6	6.01	9.89
组2	0.62	0.74	3.04	5.81	12.7	5.98	10.6
组3	0.93	1.42	2.85	5.60	11.9	5.99	9.85
组4	0.63	1.66	3.16	8.59	15.8	7.14	10.9

2）第2组实验

在下面的实验中，将采用2种不同飞机图像作为对比来验证文中提出的仿射不变矩对不同物体具有很好的区分性，选取的2种不同飞机图像如图5所示.

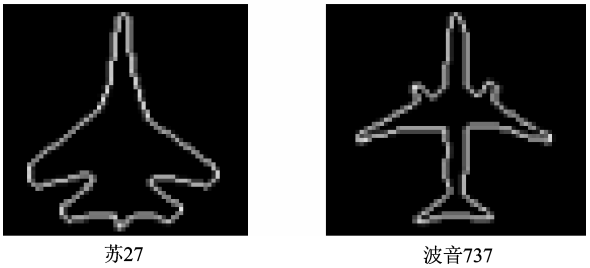

图5 2种不同飞机的图像

表3列出了2种飞机的仿射不变矩的值，最后一行表示仿射不变量从1～7阶矩值的区分度，关于区分度 I 的定义是

$$I=\frac{\sigma_X(X)}{E(X)} \tag{19}$$

其中，$\sigma_X(X)$是样本的标准差；$E(X)$是样本均值，标准差可以描述样本的波动性的大小，I 描述的是每一阶矩对样本的区分度，I 值越大，代表区分度越好.

表3 2种飞机的1～7阶矩值

图像	K_1	K_2	K_3	K_4	K_5	K_6	K_7
苏27	0.56	4.00	1.10	3.14	8.91	6.04	6.11
波音737	0.46	1.03	3.41	6.03	13.3	6.58	11.7
区分度	0.07	0.14	0.15	0.17	0.18	0.18	0.18

从表3中可以看出随着阶数的增大，区分度越来越好，但每一阶矩对于不同物体依然具有很好的区分度.

2）第3组实验

现在将它应用于真实的飞机图像识别中，本实验中将选择3类飞机建立飞机样本图像数据库，3类飞机俯视图模板图像如图6所示.

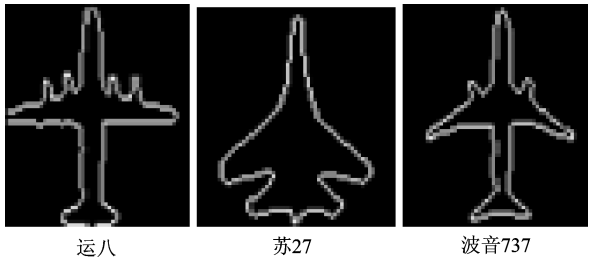

图6 3类飞机模板图像

训练图像的几何变化包括：

① 旋转变换：旋转30°，45°，60°，90°，135°，180°，270°.

② 大小变换：缩小0.5，0.75，0.45，放大1.5，2，4倍.

③ 简单仿射变换：通过仿射变换矩阵实现，包括7种仿射变换.

测试图像的几何变化包括：

① 旋转变换：旋转30°，60°，90°.

② 大小变换：缩小0.5，放大1.2，2.5倍.

③ 简单的仿射变换：通过仿射变换矩阵实现，包括4种随机生成的仿射变换矩阵.

训练的图像一共包括60幅（每种飞机20幅），用于测试的图像包括30幅（每种飞机10幅），然后将对应的30张幅测试图像进行污染（去除部分轮廓、增加噪声、添加遮挡）等，所以测试图片一共包括60张（30张未污染的，30张污染的）.

实验结果：采用BP神经网络对图像样本库进行训练，关于BP网络的具体参数如下：

① 特征值分别采用了Zernike矩和本文提出的仿射不变矩的方法（取1～7阶矩）.

② 输入层神经元个数为7，隐含层神经元个数为15，输出神经元个数为2.

③ 隐含层的传递函数采用 S 型的正切函数 Transig.

训练次数为 1 000,训练目标为 0.01,学习速率 0.1,训练到 729 次网络达到稳定.

根据训练的结果输入测试数据,得到的对比实验数据如表 4 所示(随机从 60 幅图像中抽取的情况).

表 4　输入未污染图像的识别结果

识别方法	输入飞机数量		正确识别	
	未污染	混合	未污染	混合
Zernike 矩	30	60	30	60
本方法	30	34	30	52
正确识别率	100	57	100	86.67

实验分析:由表 4 的实验结果可以看出,在理想情况下 Zernike 矩和本文提出方法正确率达到了 100%,但当物体轮廓出现缺失或者有阴影遮挡物情况下 Zernike 矩的检测正确率大大下降,而本文提出的方法依然具有比较高的准确率.

3　结 束 语

本文提出了一种新的计算统计仿射不变量的方法,并在此基础上引申出基于关键点的局部仿射不变矩,利用基于图像提取出的关键点,摒弃无关信息,适用于复杂背景和有局部遮挡情况下的特征提取.实验数据表明,它具有描述不同目标特征的能力,对仿射变换稳定,对于数字轮廓的噪声和断裂有比较好的稳健性,并在实际的飞机图像识别中发挥了比较好的性能.

参考文献 (References)

[1] Hu M K. Visual pattern recognition by moment invariants [J]. IRE transforms Theory, 1962: 179－187

[2] Khotanzad A, Hong Y H. Invariant image recognition by Zernike moments[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(5):489－497

[3] Esa R, Mikko S, Janne H. Affine invariant pattern recognition using multiscale auto-convolution[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2005, 27(6): 908－918

[4] Resis T H. The revised fundamental theorem of moment invariants[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1991, 13(8): 830－834

[5] Kadyrov K, Petrou M. The trace transform and its application[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(8):811－828

[6] Tomas S, Jan F. Combined blur and affine moments invariant and their use in pattern recognition[J]. Pattern Recognition, 2003, 36: 2895－2907

[7] David L. Distinctive image features from scale-invariant key points[J]. International Journal of Computer Vision, 2004, 60(2): 99－110

[8] ObdrZalek S, Matas J. Object recognition using local affine frames on distinguished regions [J]. Proceedings of the British Machine Vision Conference,2002:113－122

[9] Lazebnik S, Schmid C, Ponce J. A sparse texture representation using local affine regions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(8): 1265－1278

[10] Lazebnik S,Schmid C,Ponce J. Sparse texture representation using affine-invariant neighborhoods[J]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR03), 2003:319－342

[11] Zhang H Y, Yang X Q, Guo HT. Aircraft image recognition based on affine transformation[J]. Acta Aeronautics and Astronautics, 2003:15－20

[12] Cai H P, Lei L, Su Y. An affine invariant region detector using the 4th differential invariant[C]//Proceedings of 19th IEEE International Conference on Tools with Artificial Intelligence. Washington,DC:IEEE Computer Siciety,2007, 1:540－543

(责任编辑:赵海容)