

基于图像外形的多尺度相关性检索算法

黄宝娟¹, 柯芬蓉², 庄健¹, 徐仁鹏¹, 于德弘¹

(1. 西安交通大学现代设计及转子轴承系统教育部重点实验室, 710049, 西安;

2. 西安理工大学理学院, 710048, 西安)

摘要: 采用数学分析的方法比较了图像检索算法中常用的2种坐标系,发现极坐标更加有利于图像的多尺度处理. 设计的边缘序列点的插值算法,既保证了2个比较序列的长度相同,又保留了序列中所有的边缘特征点. 根据极坐标下边缘序列点的插值算法和相似轮廓在空间距离上相关性最大的特点,给出了多尺度相关性的检索算法(MSRA),该算法具有对图像尺度变化不敏感,而对图像轮廓变化敏感的特性. 通过对自然类图像库和人工类图像库200幅图像的检索,表明该算法的性能高于常用的图像检索算法,与笛卡儿坐标下普通边缘点傅里叶搜索算法(CBPFD)相比,MSRA的搜索识别率在自然类图像库中高于CBPFD的29%,而在人工类图像库中几乎是CBPFD的2.5倍.

关键词: 检索算法;相关性;多尺度

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 0253-987X(2007)01-0069-04

Multiscale Relativity Algorithm for Shape-Based Image Retrieval

Huang Baojuan¹, Ke Fenrong², Zhuang Jian¹, Xu Renpeng¹, Yu Dehong¹

(1. Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing Systems, Xi'an Jiaotong University, Xi'an 710049, China; 2. School of Science, Xi'an University of Technology, Xi'an 710048, China)

Abstract: Two kinds of image retrieval reference frames are compared mathematically, and it is found that the polar coordinates facilitate processing multi-scale images. The interpolation method for edge sequence points is designed, where two compared clusters edge points are endowed with the same length, and all original characteristic boundary points are reserved. By means of the characters of similar image boundaries with the maximum relativity in the space distance, the multi-scale relativity algorithm for shape-based image retrieval (MSRA) is proposed, which is sensitive to the changed image contours, but insensitive to the image size. Compared with common boundary points Fourier descriptor algorithm (CBPFD), the identification rate of MSRA gets higher 29% than that of CBPFD in the natural image collections and the identification rate of MRASB is almost 2.5 times of that of CBPFD in the artificial image collections.

Keywords: retrieval algorithm; relativity; multiscale

图像检索是图像处理中的关键技术. 目前,图像检索的处理方法可以大致分为以下两类:一类是基于区域的方法,该方法从待识别目标的整体区域所有像素点的特性入手,其处理方法复杂、计算量大;另一类则是基于轮廓的方法,此方法以外形作为特

征空间^[1],目标的外形轮廓是待识别目标的重要特征,与区域识别算法相比,所处理的数据点要少得多. 在轮廓检索法中常用的算法有:统计检索算法^[2]、傅里叶变换检索算法^[3]、CSS检索算法^[4-5]、神经网络检索算法^[6]等.

1 检索坐标系分析

不失一般性,假定数字图像目标轮廓集合为闭集,在平面笛卡儿坐标系下,该目标轮廓集合 P 由若干个具有横、纵坐标值的像素点构成,可以写为

$$P = \{P_0, P_1, \dots, P_n, P_0\}, P_i = \{x_i, y_i\} \quad (1)$$

式中: P_i 为像素点; x_i, y_i 为 P_i 在 x, y 轴的坐标值。

平移坐标系的原点,消除检索图像轮廓的位置影响,式(1)可以改写为

$$\left. \begin{aligned} P &= \{P'_0, P'_1, \dots, P'_n, P'_0\} \\ P'_i &= \{x_i - x_c, y_i - y_c\} \\ x_c &= \frac{1}{n+1} \sum_{i=0}^n x_i, y_c = \frac{1}{n+1} \sum_{i=0}^n y_i \end{aligned} \right\} \quad (2)$$

式中: x_c, y_c 为 P 检索图像轮廓中心点在 x, y 轴的坐标值。

平面笛卡儿坐标系不利于尺度空间变换和不同旋转角度图像的识别,因为这 2 种变换关系是非线性的。假定 P'_0 是轮廓集合中的一点, P'_1 是原轮廓放大 k 倍后与 P'_0 相对应的点, P'_2 是原轮廓旋转 α 角度后与 P'_0 相对应的点,它们的坐标值存在如下关系

$$k(x_0'^2 + y_0'^2)^{1/2} = (x_1'^2 + y_1'^2)^{1/2} \quad (3)$$

$$\tan^{-1}(y_0'/x_0') = \tan^{-1}(y_2'/x_2') - \alpha \quad (4)$$

很显然,式(3)、式(4)的非线性关系必然会导致在笛卡儿坐标系下检索图像时,多尺度归一化处理的困难、对旋转角度敏感等问题。有研究者发现,在眼睛成像中,视网膜神经末端到大脑视觉皮层形成的是一种类似于极坐标的映射方式^[7],受此启发,对于上述图像轮廓集合中的点,在极坐标下可以改写为

$$\left. \begin{aligned} P'_i &= \{\rho'_i, \theta'_i\} \\ \rho'_i &= ((x_i - x_c)^2 + (y_i - y_c)^2)^{1/2} \\ \theta_i &= \tan^{-1} \frac{x_i - x_c}{y_i - y_c} \end{aligned} \right\} \quad (5)$$

P'_0, P'_1, P'_2 的对应关系为

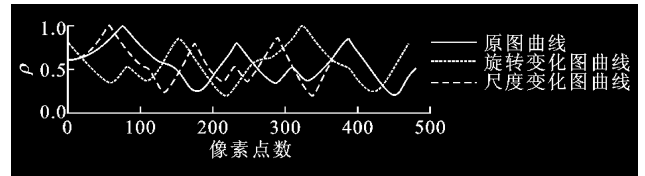
$$k\rho'_0 = \rho'_1, \quad \theta'_0 = \theta'_1 \quad (6)$$

$$\rho'_0 = \rho'_2, \quad \theta'_0 = \theta'_2 - \alpha \quad (7)$$

由式(6)、式(7)可见,在极坐标系下,图像的尺度变化或旋转变换是线性关系,有利于目标的识别。无论是图像做尺度变化还是旋转变换,图像的轮廓特征不变,即轮廓线上特征点的相互空间关系不变。当图像发生尺度变化时,可以认为是所有特征向量的模(ρ 值)乘以相同的倍数,而图像旋转变换时,是特征向量的幅角同时旋转了相同的角度,并且存在

特征向量之间的夹角不变(特征向量是中心点与特征点构成的向量)。

从图 1 中可以看出,当图像的尺度变化时,通过归一化处理后的 ρ 值相等。对于图像的旋转变换,封闭轮廓的 ρ 值变化具有周期性。对于图像的检索问题,显然极坐标系要优于笛卡儿坐标系,极坐标系将图像的尺度和旋转变换转换为线性关系,便于后续的处理。



(d) ρ 值变化曲线

图 1 图像的尺度旋转变换

2 ρ 值序列的插值算法

由于数字图像的离散特性,所以当数字图像旋转或缩放后,坐标值会产生一定的采样误差,该误差转化到极坐标下,对 θ 值的影响远远大于对 ρ 值的影响,因此在图像检索中不考虑 θ 值的变化序列,而只考虑 θ 值变化量的大小。当图像尺度变化或旋转变换后,往往会导致边缘点的总数发生变化,从而导致目标图像与待识别图像的 ρ 值与 θ 值分布不相同。通常,解决此问题的方法是对图像的离散数据点重采样,但不合适的采样会导致轮廓特征点的丢失,为了避免这个问题,在图像检索过程中,对 2 个待辨识的 ρ 值序列做短序列向长序列的插值处理,插值时采用两点间线性插值,并保留原序列端点的处理算法。该算法的具体描述如下。

假定 M_1 为目标图像轮廓数据点的 ρ 值序列, l_1 是 M_1 序列的长度, M_2 是待检索图像轮廓数据点的 ρ 值序列, l_2 是其序列长度。如果 $l_1 < l_2$, 那么待插值序列为 $M' = M_1$, 插值后的序列长度 $l_m = l_2$, 被插值序列长度为 $l_n = l_1$, 插值间距 $\Delta l = l_1 / l_2$ 。否则, $M' = M_2$, 插值后的序列长度 $l_m = l_1$, 被插值序列长度为 $l_n = l_2$, 插值间距 $\Delta l = l_2 / l_1$ 。

For $i=1$ to l_m do $M'[i] = -1$; End For
While $(i < l_m)$ do

```

 $x = i * \Delta l;$ 
Down=floor( $x$ ); Up=ceil( $x$ );
If(Dwon==Up) then Up=Down+1;
End If
If(Up> $l_n$ ) then Up=Up- $l_n$ ;
End If
 $f_1 = M'(Down); f_2 = M'(Up);$ 
If( $M'(Dwon) = -1$ )
then  $M'(i) = M'(Dwon);$ 
else  $M'(i) = f_1 + (f_1 - f_2) * (Down - x);$ 
End If
 $i++;$ 
End While

```

3 极坐标下的多尺度相关性图像检索算法

3.1 检索代价函数的选取

从图1中可看出,图像轮廓的 ρ 值变化具有外形的相似性,因此定义 ρ 值序列的相关性函数为

$$R(M, M', k) = \sum_{i=0}^n |M(i) - M'(i+k)| \quad (8)$$

相关性函数的极小值能够反映 ρ 值序列形状的相似性,2个序列之间的极小值越小,表明2个序列的形状越接近.结合图像极坐标幅角变化量的大小,定义图像搜索算法的代价函数为

$$\text{cost}(M_1, M_2, \theta_1, \theta_2) = \min(R(M, M', k)) e^{t(\theta_2 - \theta_1)^2} \quad (9)$$

式中: M 与 M' 为插值后 ρ 值的等长序列; θ_1 为目标图像轮廓幅角的变化值; θ_2 为待检索图像轮廓幅角的变化值; t 为调节系数.

3.2 检索算法的描述

设目标图像为fig,期望输出检索图像个数为 n ,把被检索图像库所含有的 x 幅图像存放在图像数组map'[x]中,此时极坐标下的多尺度相关性图像检索算法可描述为:

```

For  $i=1$  to  $n$ 
    bestcost[ $i$ ]= $1e+20$ ;
    BestPosition[ $i$ ]=-1;
End For
计算 fig 在极坐标系下的  $M_1$  与  $\theta_1$ ;
For  $i=1$  to  $x$ 
    Position= $i$ 
    计算 map'[ $i$ ]在极坐标系下的  $M_2$  与  $\theta_2$ ;

```

按照文章第2部分,短序列做端点保留插值,获得插值后的 M 与 M' ;

按照式(9)计算代价函数值 cost_now;

```

For  $j=1$  to  $n$ 
    If (bestcost[ $k$ ]<cost_now) then
        For  $k=j$  to  $n-1$ 
            bestcost[ $k$ ]=bestcost[ $k+1$ ];
        end for
        bestcost[ $j$ ]=cost_now;
        BestPosition[ $j$ ]=Position; break;
    End If
End For

```

End For

For $i=0$ to n

显示 map'[BestPosition[i]];

End For

4 实验结果与分析

实验图像分为2大类,共200幅.一类选自 Temple University 的鱼类外形轮廓数据库(<http://knight.cis.temple.edu/~latecki/TestData/fish.tar.gz>),从数据库中挑选100幅图像,构成20组,每组有5幅图像基本类似,称之为自然类图像(简称类1);另一类图像是手工绘制的100幅图像,也分为20组,称它们为人工类图像(简称类2).

实验的计算机CPU为PIV 1.6 GB,1 GB内存,编程语言为C++.在每大类中,100幅图像依次作为检索图像,并由算法检索出5幅最相似的图像.实验中共采用了2种检索算法,分别为笛卡儿坐标下普通边缘点的FT算法(CBPFD)和本文所提出的极坐标下多尺度相关性图像检索算法(MSRA).

图2为2类算法分别在上述2大类图像集中的检索结果示例,从检索结果中可以看出,极坐标下多尺度相关性搜索算法无论在自然类和人工类图像集合中都具有比较好的性能特性.

为了进一步对比2种检索算法的特性,设计的实验为:将2类图像集中的所有图像依次作为检索图像,统计由算法搜索所得到的正确图像的数目,并定义性能指标识别率为算法搜索到的正确图像数目除以理论上正确图像的总数目.如图3所示,极坐标下多尺度相关性图像检索算法在自然和人工图像集合的平均识别率分别为78.2%和97.2%;笛卡儿坐标下普通边缘点FT图像的检索算法的识别率分别为49.2%和39.0%.

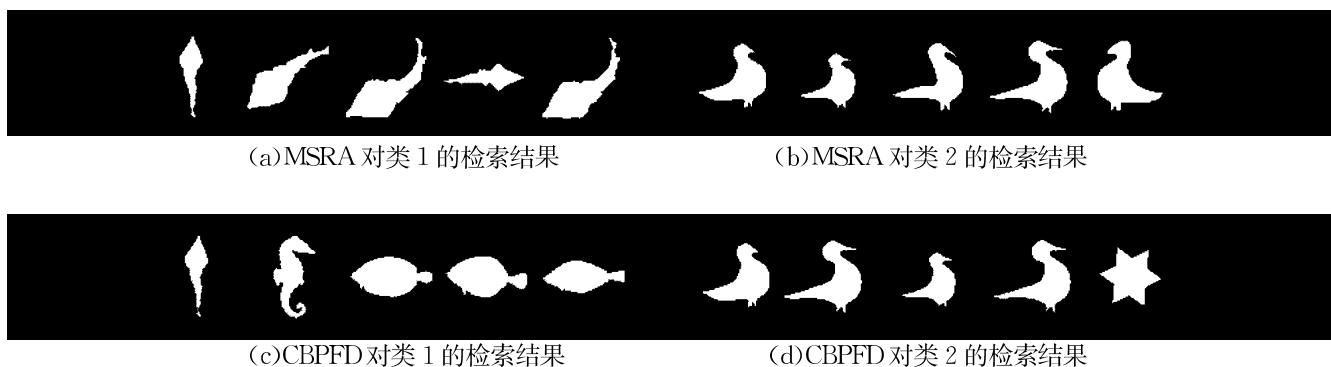
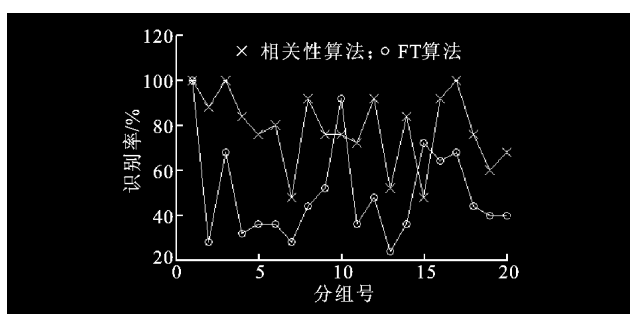
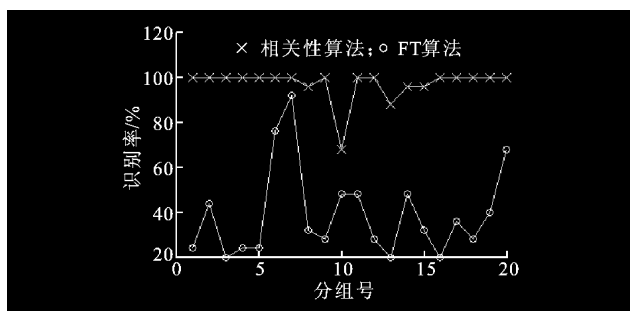


图2 2种算法的检索结果



(a)自然类图像



(b)人工类图像

图3 2种算法的识别率分布图

由图3可见,普通边缘点FT图像的检索算法在自然类图像集合的搜索性能要高于在人工类图像集合中的搜索性能,这是因为普通边缘点FT图像的检索算法未对图像轮廓做归一化处理,对图像的缩放敏感,实验中自然类图像集合中的图像轮廓缩放情况不多,而人工类图像情况则相反.将极坐标下多尺度相关性图像检索算法与普通边缘点FT图像的检索算法相比,前者的搜索识别率在自然类图像集合中要高于后者的29%,在人工类图像集合中前者几乎是后者的2.5倍,主要原因是前者不仅对图像做了归一化处理,而且还忽略了边缘点的角度变化序列.由于FT图像检索算法的坐标体系是笛卡儿坐标,其边缘点的坐标值 x 、 y 包含有边缘点的角度变化分量,因此该分量容易受到数字图像中采样误差的影响.实验表明,在极坐标下多尺度相关性图

像检索算法对图像的缩放、旋转不敏感,能够辨析图像边缘点的细节变化,其性能要好于一般的图像检索算法,因此可以较好地完成图像的检索工作.

5 结论

本文通过分析发现,极坐标空间更加有利于图像多尺度、旋转的处理,此类转换在极坐标空间下转换为线性关系,因此非常适合做归一化处理.但是,由于数字图像的离散特性,当图像发生尺度或旋转变化时,由普通笛卡儿坐标系向极坐标系转化时会存在着误差,并且幅角序列误差要大于轮廓边缘点的序列误差,因此本文选择了轮廓的边缘点序列作为图像搜索的主要特征量.本文还设计了基于轮廓相关性的搜索代价函数,给出了极坐标下的多尺度相关性图像检索算法.实验表明,本文的算法性能要高于常用的图像检索算法,与笛卡儿坐标下边缘点FT检索算法相比,MSRA的搜索识别率在自然类图像集合中要高于CBPFD的29%,在人工类图像集合中几乎是CBPFD的2.5倍.文章的程序和测试图像库可以在<http://202.117.58.58/files/ImageRetrieval.rar>下载.

参考文献:

- [1] Hung Kingchu. The generalized uniqueness wavelet descriptor for planar closed curves [J]. IEEE Transactions on Image Processing, 2000, 9(5): 834-845.
- [2] Hu Mingkuei. Visual pattern recognition by moment invariants [J]. IRE Transactions on Information Theory, 1962, 8(2): 179-187.
- [3] Bartolini I, Ciaccia P, Patella M. WARP: accurate retrieval of shapes using phase of Fourier descriptors and time warping distance [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(1): 142-147.
- [4] Mokhtarian F, Abbasi S. Shape similarity retrieval un-

- der affine transforms [J]. Pattern Recognition, 2002, 35(1): 31-41.
- [5] Mokhtarian F, Mackworth A K. A theory of multi-scale, curvature-based shape representation for planar curves [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(8): 789-805.
- [6] 黄金,梁彦,程咏梅,等. 基于序列图像的自动目标识别算法[J]. 航空学报, 2006, 27(1): 87-93.
Huang Jin, Liang Yan, Cheng Yongmei, et al. Automatic target recognition method based on sequential images[J]. ACTA Aeronautica ET Astronautica SINICA, 2006, 27(1): 87-93.
- [7] Kauppinen H, Seppanen T, Pietikainen M. An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(2): 201-207.
- (编辑 管咏梅)