

# 利用不变量进行基于内容的图像检索

王文惠<sup>1</sup>, 孟 兵<sup>2</sup>, 万建伟<sup>1</sup>, 周良柱<sup>1</sup>

(1. 国防科学技术大学电子科学与工程学院, 湖南长沙 410073; 2. 国防科学技术大学人文与管理学院, 湖南长沙 410073)

**摘 要:** 本文利用一种新的图像特征描述子进行基于内容的图像检索. 该描述子不仅实现了对图像内容的综合概括, 即图像的形状、纹理和颜色内容, 而且具有平移、旋转和尺度变换不变性, 实验证明了该描述子优良的检索性能.

**关键词:** 图像检索; 不变性; 离散傅氏变换; log 极坐标变换

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112 (2002) 07-0949-03

## Using Invariant Feature in Content-Based Image Retrieval

WANG Wen-hui<sup>1</sup>, MENG Bing<sup>2</sup>, WAN Jian-wei<sup>1</sup>, ZHOU Liang-zhu<sup>1</sup>

(1. School of Electronics Science and Engineering, NUDT, Changsha, Hunan 410073, China;

2. School of Humanities and Management, NUDT, Changsha, Hunan 410073, China)

**Abstract:** In the paper, a new image signature is computed for content-based image retrieval. The signature provides a compact description of all image aspects, including shape, color and texture. Also the signature is invariant to 2D rigid transformation, such as rotation, scaling and translation. By experiment, the retrieval result is satisfactory.

**Key words:** image retrieval; DFT; invariant; log-polar transformation

### 1 引言

数字图书馆和图像处理的发展使基于内容的图像检索技术成为一个重要的研究课题. 目前研究的重点集中在利用计算机自动抽取图像的视觉特征, 通过进行视觉特征的相似性测量, 得到图像检索的结果. 成功的检索系统, 如 IBM 公司的 QBIC 系统, Virage 公司的 VIR 图像工程系统, Excalibur 公司的 Excalibur 视觉检索产品等<sup>[2]</sup>. 它们的主要原理是抽取图像的多个描述子, 描述图像视觉特征的变化. 这些描述子包括描述颜色特征的颜色直方图, 描述纹理特征的互相关矩阵. 描述形状特征的滤波器输出. 但是如何组合这些特征, 对系统和用户来说, 都是一个巨大的难题. 不同的加权值, 可能带来完全不同的结果. 为了解决这个问题, 又提出了基于图像正交变换的描述子, 比如小波分解<sup>[3,4]</sup>. 小波分解成功地解决了多特征组合的问题, 但是由于小波分解的多分辨率和小波分解的方向性, 使得该描述子不具有旋转和尺度变换的不变性. 如果检索的图像内容包括产生运动的物体, 则不能得到优良的检索结果.

本文提出了一种基于傅氏变换和 log 极坐标变换的描述子, 该描述子不仅具有平移、旋转和尺度变换不变性, 而且实现了对图像内容的综合概括, 包括图像的形状、纹理和颜色内容, 从而克服了上述的缺陷, 得到了优良的检索结果.

### 2 提取特征描述子进行图像检索

设灰度图像  $I$ , 空域表示  $f(x, y)$  为:

$$\{f(x, y) | x=0, \dots, M-1; y=0, \dots, N-1\} \quad (1)$$

它的离散傅里叶变换  $F(m, n)$  为<sup>[5]</sup>:

$$F(m, n) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j(2\pi x/M)m} e^{-j(2\pi y/N)n} \quad (2)$$

由离散傅里叶变换得到图像的相位和能量谱  $(m, n)$ 、 $P(m, n)$ :

$$\phi(m, n) = \arctan \frac{\text{Im}[F(m, n)]}{\text{Re}[F(m, n)]} \quad (3)$$

$$P(m, n) = \sqrt{\text{Re}[F(m, n)]^2 + \text{Im}[F(m, n)]^2}$$

保留图像的能量谱  $P(m, n)$ , 对它进行高斯滤波得到  $P(m, n)$ :

$$P(m, n) = \frac{1}{2^2(m, n)} P(k, l) \exp \left[ -\frac{(m-k)^2 + (n-l)^2}{2^2(m, n)} \right] \quad (4)$$

$$^2(m, n) = \max(a, b \log \sqrt{(m)^2 + (n)^2}, (m, n) = (m - m_0, n - n_0) \quad (5)$$

其中  $(m_0, n_0)$  是为将要进行 log 极坐标变换所选定的极坐标空间的原点, 通常取  $(0, 0)$ . 对图像进行高斯滤波, 主要是因为从笛卡尔坐标变换到极坐标, 图像的网格点发生了变化, 为了减少误差, 进行平滑处理. 其中,  $a=2$  为常数, 确保低通滤波器在  $(m, n)$  靠近  $(m_0, n_0)$  仍有作用.  $b=0.07$  为尺度系数, 使滤波器发挥最佳性能. 对固定尺寸的图像,  $a$  和  $b$  的值是确定的, 可由实验得到最佳数值.

收稿日期: 2000-11-15; 修回日期: 2001-04-12

基金项目: 高等学校骨干教师资助基金

对  $P(m, n)$  进行极坐标变换得到  $L(\rho, \vartheta)$ , 对  $L(\rho, \vartheta)$  进行  $\log$  采样得到  $L(\rho, \vartheta)$  即:

$$= \log \sqrt{(m - m_0)^2 + (n - n_0)^2}, \vartheta = \arctan \frac{n - n_0}{m - m_0} \quad (6)$$

$$L(\rho, \vartheta) = \log(L(\rho, \vartheta)) \quad (7)$$

对  $L(\rho, \vartheta)$  求边缘函数得到  $L(s), L_\theta(t)$ :

$$L(s) = \int_0^{2\pi} L(s, \vartheta) d\vartheta, L_\theta(t) = \int_0^{\infty} L(\rho, t) d\rho \quad (8)$$

对  $L(s), L_\theta(t)$  进行傅里叶变换, 得到它们的能量谱  $\ell(s), \ell_\theta(t)$ , 则二者就是我们得到的图像特征描述子. 它们具有平移、旋转和尺度变换不变性.

当比较两幅图像之间的相似性时, 利用下式计算图像  $I$  和  $I'$  图像之间的距离即:

$$d(i, j) = \frac{1}{2} \left( \sqrt{\int_s (\ell(s) - \ell'(s))^2 ds} + \sqrt{\int_s (\ell_\theta(s) - \ell'_\theta(s))^2 ds} \right) \quad (9)$$

进行基于内容的图像检索时, 对数据库内的每幅图像事先抽取  $\ell, \ell_\theta$  特征描述子, 并加以存储. 当查询者给出一幅模板图像, 首先抽取检索模板图像的特征描述子, 并且计算数据库内的每幅图像与模板图像之间的距离. 根据计算结果, 把距离最小的前  $K$  幅图像作为检索结果.  $K$  的大小可由检索者自行确定.

### 3 特征描述子的不变性

下面主要证明第一部分给出的图像特征描述子的不变性.

#### 3.1 平移不变性

设图像  $I$  在空间经过平移变换得到:

$$f(x, y) \Rightarrow f_1(x, y) = f(x - a, y - b) \quad (10)$$

经过傅氏变换和取绝对值计算的能量谱:

$$f(x, y) \Rightarrow P(m, n), f_1(x, y) \Rightarrow P_1(m, n) \quad (11)$$

$$\text{则: } P(m, n) = \left| \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j(2\pi mx/M)} e^{-j(2\pi ny/N)} \right|$$

$$P_1(m, n) = \left| \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f_1(x, y) e^{-j(2\pi mx/M)} e^{-j(2\pi ny/N)} \right|$$

$$= \left| \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j(2\pi mx/M)} e^{-j(2\pi ny/N)} \right|$$

$$\text{即: } P(m, n) = P_1(m, n) \quad (12)$$

上述证明假设图像具有周期性. 图像的周期性在后面的实验中表述. 所以在经过傅氏变换和计算能量谱后, 图像的特征具有平移不变性. 从而, 最后的特征  $\ell, \ell_\theta$  也具有平移不变性.

#### 3.2 尺度不变性

设图像尺度变换的系数是  $a$ . 则尺度变换可以表示成:

$$f(x, y) \Rightarrow f_2(x, y) = f(ax, ay) \quad (13)$$

由傅氏变换公式可以得到:

$$f(x, y) \Rightarrow P(m, n)$$

$$f_2(x, y) \Rightarrow P_2(m, n) = \frac{1}{|a|^2} P\left(\frac{1}{a}m, \frac{1}{a}n\right) \quad (14)$$

对式(13)、(14)进行上述定义的  $\log$  极坐标变换, 并且原点选

在(0,0):

$$P(m, n) \Rightarrow L(\rho, \vartheta), P_2(m, n) \Rightarrow L_2(\rho_2, \vartheta_2) \quad (15)$$

则:

$$= \log \sqrt{m^2 + n^2}, \vartheta = \arctan \frac{n}{m}$$

$$2 = \log \sqrt{(a^{-1}m)^2 + (a^{-1}n)^2} \quad (16)$$

$$= -\log|a| + \log \sqrt{m^2 + n^2} = -\log|a| +$$

$$\vartheta_2 = \arctan \frac{an}{am} = \arctan \frac{n}{m} = \vartheta$$

从而得到:  $L(\rho, \vartheta) \Rightarrow L_2(\rho_2, \vartheta_2) = \frac{1}{|a|^2} L(\rho - \log|a|, \vartheta)$

经过  $\log$  采样:  $L(\rho, \vartheta) \Rightarrow L(\rho, \vartheta), L_2(\rho_2, \vartheta_2) \Rightarrow L_2(\rho, \vartheta)$

$$L(\rho, \vartheta) \Rightarrow L_2(\rho, \vartheta) = \log\left(\frac{1}{|a|^2} L(\rho - \log|a|, \vartheta)\right) = -2\log|a| + L(\rho - \log|a|, \vartheta)$$

$$\text{即: } L_2(\rho, \vartheta) = -2\log|a| + L(\rho - \log|a|, \vartheta) \quad (17)$$

即尺度变换使得自变量和函数的取值都产生了平移. 其中  $-2\log|a|$  只增加了直流分量, 极坐标中  $\rho > 0$ , 克服了它的影响. 即:

$$L_2(\rho, \vartheta) = L(\rho - \log|a|, \vartheta) \quad (18)$$

经过式(8)计算, 再求能量谱, 消除了平移  $-\log|a|$  的影响. 从而  $\ell, \ell_\theta$  也具有尺度不变性.

#### 3.3 旋转不变性

由于时域空间到傅氏空间的变换是一一映射的, 并且两者具有同样的周期, 则图像旋转, 傅氏系数同样旋转, 在  $\log$  极坐标空间表现为  $L(\rho, \vartheta)$  平移了  $\phi$ , 该结论在文[6]中进行了证明.

则原图像和旋转后的图像对应为:

$$L(\rho, \vartheta) \Rightarrow L_3(\rho, \vartheta) = L(\rho, \vartheta + \phi) \quad (19)$$

同样经过傅氏变换, 并求能量谱, 消除了平移了  $\phi$  的影响. 从而  $\ell, \ell_\theta$  也具有旋转不变性.

以上的证明说明了提取  $\ell, \ell_\theta$  的过程完全消除了平移、旋转、尺度不变性. 在提取过程中的高斯滤波主要是平滑图像, 消除笛卡尔坐标到极坐标的采样计算误差. 通过式(10)、(11)求边缘函数, 减少了特征描述子的数目, 并且对平移、旋转、尺度不变性无影响. 本文提出的算法完全是非线性、不可逆计算过程, 它高度概括了图像的基本属性.

### 4 实验仿真

实验一 首先验证图像特征描述子的平移、旋转、尺度不变性. 对  $256 \times 256$  的图像, 基于数字图像的特征,  $\log$  极坐标变换采用如下形式:

$$= \frac{1}{5}r, \frac{1}{5}r^2, \dots, \frac{64}{5}r^{R-1}, \vartheta = 0, \frac{\pi}{S}, \frac{2\pi}{2S}, \dots, \frac{2(S-1)\pi}{S} \quad (20)$$

其中:  $r = 10^{1/64}, S = 32$ . 经过计算这种选取是完全合适的.

图1是利用本文所述的方法计算特征描述子的结果, 图像(a)和(b)是经过平移和旋转变换的 lena 图像. 它们的特征描述子  $\ell$  和  $\ell_\theta$  完全相同, 证明了方法的正确性. 实验图像说明, 图像的平移也是建立在周期性基础上的.

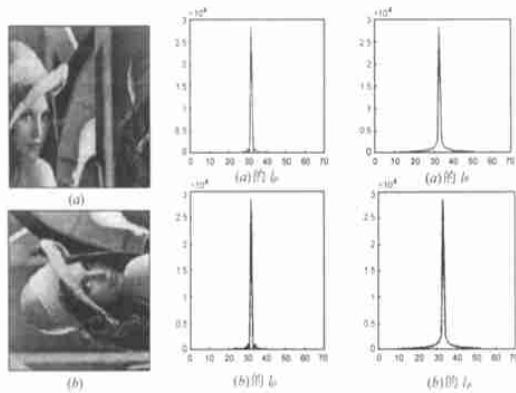


图 1 图像特征不变性实验结果

利用描述子的对称性,只保留特征值的一半,在进行基于内容的图像检索时,会大大节约存储空间。

实验二 利用图像特征不变量进行检索. 图像库由 2000 幅灰度图像组成,包括 5 个目录:纹理、建筑、文字、动物、花草. 每个目录包括 400 幅图像. 图 2 所示,上边是模板图像,下边依次是按距离排列的检索结果. 对图像库中的 10 幅图像进行检索的结果,得到图像返回的正确率为 70.3%(返回的正确图像/图像库中的正确图像)。



图 2 基于内容的图像检索

## 5 结束语

通过理论分析和实验仿真,证明了本文提出的图像特征描述子具有平移、旋转和尺度不变性. 应用到基于内容的图像检索中具有优良的性能. 本文的算法以灰度图像为例,对彩色图像将分别计算 RGB 三个空间的特征描述子。

## 参考文献:

- [ 1 ] John P Eakins , Margaret E Graham. Content-based Image Retrieval [A]. A report to the JISC Technology Applications Programme. Institute for Image Data Research , University of Northumbria at Newcastle January 1999.
- [ 2 ] Lawrence D Bergman , Vittorio Castelli , Chung-Sheng Li. Progressive Content-Based Retrieval from Satellite Image Archives [J]. D-Lib Magazine ,1997.
- [ 3 ] F Idris , S Panchanathan. Image indexing using wavelet vector quantization [J]. Proc. SPIE Storage and retrieval of image and video databases ,1995 ,2420 :373 - 380.
- [ 4 ] C E Jacobs , A Finkelstein , D H Salesin. Fast Multiresolution Image Querying [A]. Proc. Siggraph '95 , ACM [C]. New York : SIGGRAPH , 1995. 200 - 205.
- [ 5 ] 孙即祥. 数字图像处理 [M]. 石家庄 :河北教育出版社 ,1993.
- [ 6 ] Kenneth R Castleman. DIGITAL IMAGE PROCESSING [M]. New York : PRENTICE HALL ,1996.

## 作者简介:



王文惠 女,1971年9月生于山东,1998年考入国防科技大学攻读博士学位,主要研究方向为数字图像处理,计算机视觉,数字信号处理等。



孟兵 男,1972年8月生于四川,博士,主要研究方向为指挥控制自动化系统,信号处理技术等。