

# 基于多特征融合和相关反馈的图像检索

党长青, 宋凤娟, 刘树明

DANG Chang-qing, SONG Feng-juan, LIU Shu-ming

唐山学院 信息工程二系, 河北 唐山 063020

Department two of Information Engineering, Tangshan College, Tangshan, Hebei 063020, China

E-mail: Dangchangqing2007@126.com

**DANG Chang-qing, SONG Feng-juan, LIU Shu-ming. Image retrieval based on multi-feature integration and relevance feedback. Computer Engineering and Applications, 2008, 44(3): 186-188.**

**Abstract:** Current image retrieval methods use only one kind of image features, which can not describe image content completely. In this paper, a image retrieval method based on multi-feature integration is proposed. In order to improve the precision of image retrieval and reduce the semantic gap between the visual feature and semantic feature, a relevance feedback mechanism, based on support vector machines is invoked. The experiment results show that the proposed method have an excellent retrieval performance.

**Key words:** image retrieval; color feature; shape feature; texture feature; feature integration; relevance feedback

**摘要:** 鉴于单一视觉特征不能很好地表达图像内容, 提出一种融合图像颜色、形状、纹理特征的图像检索方法。最后采用支持向量机(SVM)的相关反馈算法提高图像检索的准确度, 缩小低层特征和高层语义之间的语义鸿沟。实验结果说明提出的方法具有良好的检索性能。

**关键词:** 图像检索; 颜色特征; 形状特征; 纹理特征; 特征融合; 相关反馈

**文章编号:** 1002-8331(2008)03-0186-03 **文献标识码:** A **中图分类号:** TP391

## 1 引言

基于内容的图像检索近年来一直是研究的一个热点问题, 其主要方法是以抽取的图像特征作为图像的索引, 计算查询图像和目标图像的相似距离, 按相似度匹配进行检索<sup>[1]</sup>。

颜色特征是描述图像内容的最直接的视觉特征, 因为在图像检索中颜色往往和图像中所包含的物体或场景十分相关。另外, 颜色特征对图像本身的尺寸、方向、视角的依赖性较小, 从而具有较高的鲁棒性<sup>[1,2]</sup>。纹理特征是图像的基本特征之一, 是一种全局特征, 它描述了图像或图像区域所对应的景物的表面性质。包括表面结构组织及其与周围环境关系的许多重要信息, 纹理是一种统计特征, 具有旋转不变性, 并具有较强的抗噪音能力<sup>[3]</sup>。形状是图像目标的显著特征之一, 许多物体具有不同的颜色, 但其形状总是类似的。形状的概念非常明确, 而且由形状的特征来区分物体也非常直观<sup>[4,5]</sup>。

相关反馈技术是在检索过程中允许用户对检索结果进行评价和标记, 然后将用户标记的信息作为训练样本反馈给系统进行学习, 以指导下一轮检索<sup>[6,7]</sup>。本文提出基于颜色、纹理和形状特征融合的图像检索方法, 应用支持向量机(SVM)的相关反馈方法提高图像检索的精确度。实验结果表明, 本文提出的方法具有良好的检索性能。

## 2 图像特征提取

### 2.1 颜色特征提取

颜色特征在图像检索中占有重要位置, 通常采用 RGB 三色表示。但为使表示法更符合人眼视觉特征, 把 RGB 空间模型转换为 HSV 空间模型进行颜色特征的统计。

人眼对  $H$  值较敏感, 对  $S, V$  值的感知较弱, 下面就将 HSV 三个分量按照人的颜色感知进行非等间隔的量化:

(1) 按照人的视觉分辨能力, 把色调  $H$  空间分成 8 份, 饱和度  $S$  和亮度  $V$  空间分别分成 3 份。

(2) 根据色彩的不同范围和主观颜色感知进行量化:

$$H = \begin{cases} 0 & h \in [316, 359] \\ 1 & h \in [21, 40] \\ 2 & h \in [41, 75] \\ 3 & h \in [76, 155] \\ 4 & h \in [156, 190] \\ 5 & h \in [191, 270] \\ 6 & h \in [270, 295] \\ 7 & h \in [296, 315] \end{cases} \quad S = \begin{cases} 0 & s \in [0, 0.2] \\ 1 & s \in [0.2, 0.7] \\ 2 & s \in [0.7, 1] \end{cases} \quad V = \begin{cases} 0 & v \in [0, 0.2] \\ 1 & v \in [0.2, 0.7] \\ 2 & v \in [0.7, 1] \end{cases} \quad (1)$$

(3) 构造一维特征矢量。按照以上的量化级, 把 3 个颜色分量合成为一维特征矢量:

$$G = HQ_i + SQ_i + V \quad (2)$$

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.50127402)。

作者简介: 党长青(1966-), 女, 副教授, 主要研究方向为图像检索、模式识别; 宋凤娟(1961-), 女, 副教授, 主要研究方向为图像处理、模式识别、自动控制; 刘树明(1963-), 男, 副教授, 主要研究方向模式识别。

其中,  $Q_s$  和  $Q_v$  分别是分量  $s$  和  $v$  的量化级数, 在实验中  $Q_s=3$ ,  $Q_v=3$ , 因此:

$$G=9H+3S+V \quad (3)$$

这样,  $H, S, V$  三个分量在一维矢量上分布开来.  $G$  的取值范围为  $[0, 1, \dots, 71]$ , 计算  $G$  获得 72 柄(bin)的一维直方图.

### 2.2 纹理特征提取

本文采用 Gabor 小波变换提取纹理特征, 一个 2-D Gabor 函数为:

$$\Psi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2})] \cdot \exp(j2\pi\omega x) \quad (4)$$

式中  $\omega$  为调制频率,  $\Psi(x, y)$  是经过复数正弦函数调制的 Gaussian 函数. 对小波函数  $\Psi(x, y)$  经过伸缩及旋转可得一组自相似 Gabor 小波函数:

$$\psi_{mn}(x, y) = a^{-m} \psi(\tilde{x}, \tilde{y}) \quad (5)$$

其中  $m, n$  分别是指定小波的尺度和方向,  $m=0, 1, \dots, M-1; n=0, 1, \dots, N-1$ .  $M, N$  分别表示尺度数和方向数. 式(5)中:  $\tilde{x} = a^{-m}(x\cos\theta + y\sin\theta)$ ,  $\tilde{y} = a^{-m}(-x\sin\theta + y\cos\theta)$ , 其中  $a$  为伸缩因子,  $a > 1$ ,  $\theta = n\pi/N$ .

假设对一给定尺寸为  $P \times Q$  的图像  $I(x, y)$ , 它的离散 Gabor 小波变换为:

$$G_{mn}(x, y) = \sum_s \sum_t I(x-s, y-t) \Psi_{mn}^*(s, t) \quad (6)$$

式中:  $\Psi_{mn}^*(s, t)$  是自相似函数  $\Psi_{mn}(s, t)$  的共轭复数,  $s, t$  是滤波器模板大小的变量.

在不同的方向和尺度上计算变换后系数幅度序列  $E(m, n) =$

$\sum_x \sum_y |G_{mn}(x, y)|$ , 式中  $m=0, 1, \dots, M-1; n=0, 1, \dots, N-1$ . 这些系数幅度值在不同方向和尺度上表征了图像的能量. 以系数幅度序列的均值  $\mu_{mn}$  和标准方差  $\sigma_{mn}$  作为图像的纹理特征:

$$\mu_{mn} = E(m, n) / P \times Q$$

$$\sigma_{mn} = \sqrt{\sum_x \sum_y (|G_{mn}(x, y)| - \mu_{mn})^2 / P \times Q} \quad (7)$$

以  $\mu_{mn}$  和  $\sigma_{mn}$  为分量构成特征向量  $f$  来描述图像的纹理, 通常取尺度  $M$  为 5, 方向  $N$  为 6, 即特征向量表示为  $f = (\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \dots, \mu_{45}, \sigma_{45})$ .

### 2.3 形状特征提取

小波分析作为多尺度分析方法, 本身不具有平移、尺度和旋转不变性, 矩特征具有平移、尺度和旋转不变性和良好的抗噪性. 因此在本文中结合小波多尺度分析和矩不变性的优点, 提出一种基于小波多尺度分析的矩不变量特征, 即小波矩不变量.

#### 2.3.1 矩

对于二维平面上灰度图像  $f(x, y)$ , 它的  $(p+q)$  阶混合原点矩定义:

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy \quad p, q=0, 1, 2, \dots \quad (8)$$

$(p+q)$  阶混合中心矩定义为:

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\bar{x}-\bar{x})^p (\bar{y}-\bar{y})^q f(x, y) dx dy \quad p, q=0, 1, 2, \dots \quad (9)$$

其中,  $\bar{x} = \frac{m_{10}}{m_{00}}$ ,  $\bar{y} = \frac{m_{01}}{m_{00}}$ . 规格化的中心矩定义为:

$$\eta_{pq} = \frac{\mu_{pq}}{r^p} \quad (10)$$

式中:

$$r = \frac{p+q+2}{2} \quad p, q=0, 1, 2, \dots \quad (11)$$

由上面这些定义可以得到 Hu 的 7 个低阶不变矩  $M_1 \sim M_7$ . 这 7 个矩对于图像的平移、旋转和尺度变化都具有不变性.

#### 2.3.2 极坐标下的图像小波矩

在式(8)中令  $x=r\cos(\theta)$ ,  $y=r\sin(\theta)$  从直角坐标系转换得到极坐标系, 简单演算后, 可以得到极坐标下矩特征的一般表示式:

$$\|F_{m,q}\| = \left\| \int \int f(r, \theta) g_p(r) e^{j\theta q} r dr d\theta \right\| \quad (12)$$

其中  $g_p(r)$  为变换核的径向分量, 而  $e^{j\theta q}$  是变换核的角度分量. 以小波函数作为式中的  $g_p(r)$ , 就可以得到小波不变矩的表示式.

小波函数沿着径向可以定义为  $\psi_{m,n}(r) = 2^{\frac{m}{2}} \psi(2^m r - n)$ , 令

$S_p(r) = \int f(r, \theta) e^{j\theta q} d\theta$ , 由式(12)得:

$$\|F_{m,n,q}^r\| = \left\| \int S_q(r) \psi_{m,n}(r) r dr \right\| \quad (13)$$

可以证明, 当图像发生旋转后特征值的模  $\|F_{m,n,q}^r\|$  保持不变, 即小波矩  $\|F_{m,n,q}^r\|$  具有旋转不变性. 为了使它具有平移不变性和尺度不变性, 在计算小波不变矩前先要对图像进行归一化处理, 使各个图像的中心位于坐标原点, 且各图像的尺寸要保持一致. 当把图像尺寸归一化为  $r \leq 1$  的范围后, 利用不同的尺度因子  $m$ 、位移因子  $n$ ,  $\psi_{m,n}(r)$  可以遍及整个径向空间  $[0, 1]$ , 小波矩  $\|F_{m,n,q}^r\|$  提供了图像  $f(r, \theta)$  在不同尺度水平上的特征, 得到图像的全局特征和局部特征.

### 3 图像相似度计算

颜色、纹理和形状信息都只反映了图像某方面的特征, 因此在本文中融合多个特征进行图像检索. 合并前先对单个特征的相似性计算进行特征间归一化.

(1) 计算图像库中两个不同图像  $x, y$  间距离  $d(x, y)$  ( $x \neq y$ ), 并计算所有距离的均值  $m_d$  和  $\sigma_d$  方差;

(2) 对查询图像  $q$ , 计算其与图像库中各图像的相似距离, 记为  $d(1, q), d(2, q), \dots, d(m, q)$  ( $m$  为图像库图像数目);

(3) 对  $d(1, q), d(2, q), \dots, d(m, q)$ , 按下式进行高斯归一化得到  $d'(1, q), d'(2, q), \dots, d'(m, q)$ ;

$$d'(x, q) = \frac{d(x, q) - m_d + 1}{2 \cdot 3\sigma_d} \quad (14)$$

其中  $x \in [1, m]$ , 得到归一化后的相似度  $d'(x, q)$  落于  $[0, 1]$  区间的概率大于 99%.

根据上述方法分别求出  $d'_{color}(x, q), d'_{texture}(x, q), d'_{shape}(x, q)$  分别表示图像  $x$  和  $q$  之间的归一化的相似性测度,  $w_{color}, w_{texture}, w_{shape}$  是对应特征的权值, 若用  $d_{mix}(x, q)$  表示了图像  $X$  和  $q$  之间的特征融合相似度, 则有:

$$d_{mix}(x, q) = w_{color} \times d'_{color}(x, q) + w_{texture} \times d'_{texture}(x, q) + w_{shape} \times d'_{shape}(x, q) \quad (15)$$

其中  $w_{color} + w_{texture} + w_{shape} = 1$ .

### 4 相关反馈

由于图像内在语义与系统自动提取的图像视觉特征(颜色、形状、纹理)之间难以匹配,所以检索效果难以满足实用要求,为了提高检索精度,引入相关反馈机制,用户对查询结果进行评价,标记出相关的图像,将标记的图像作为训练样本对支持向量机进行学习,使用户反馈的信息将图像底层特征和语义层内容联系起来,从而实现更好的查询效果<sup>[6,7]</sup>。

#### 4.1 支持向量机

支持向量机是 Vladimir N.Vapnik 等人提出的一种基于统计学习理论的机器学习算法<sup>[8]</sup>。其在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势,并能够推广应用到函数拟合等其他机器学习问题中支持向量机是一种基于结构风险最小化原理的学习算法,它比基于经验风险原理的神经网络学习算法具有更强的理论依据和更好的泛化性能。

SVM 是从线性可分情况下的最优分类面发展而来的,给定线性可分样本 $(x_i, y_i), i=1, \dots, N, y_i \in \{-1, 1\}, x_i \in R_{d_0}$  假定某个分类面  $w \cdot x + b = 0$  使得样本集 $(x_i, y_i)$ 满足:

$$y_i[(w \cdot x_i + b) - 1] \geq 0 \quad i=1, 2, \dots, n \quad (16)$$

此时分类间隔等于  $2 / \|w\|$ , 使间隔最大等价于使  $\|w\|^2$  最小。满足式(16)且使  $\frac{1}{2} \|w\|^2$  最小的分类面就叫最优分类面。

利用 Lagrange 优化方法可以把上述最优分类面问题转化为其对偶问题,即在约束条件:

$$\sum_{i=1}^n y_i \alpha_i = 0 \text{ 和 } \alpha_i \geq 0, i=1, 2, \dots, n \quad (17)$$

下面对  $\alpha$  求解下列函数的最大值:

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (18)$$

这是一个不等式约束下二次函数寻优的问题,存在唯一解。解上述问题后得到的最优分类函数是:

$$f(x) = \text{sign}(\sum_i \alpha_i y_i x_i \cdot x + b) \quad (19)$$

对于非线性问题可以通过一个核函数  $K(x, y)$  将输入空间映射到一个高维空间求解。上述分类函数式(19)相应的变为:

$$f(x) = \text{sign}(\sum_i \alpha_i y_i K(x_i, x) + b) \quad (20)$$

获得分类器后,将分类函数式(20)转化形式,用来计算图像库中所有图像与分类超平面的距离,从而得到各个图像与检索目标的相似度并进行排序。转化形式之后的分类函数为:

$$f(x) = \sum_i \alpha_i y_i K(x_i, x) + b \quad (21)$$

#### 4.2 基于 SVM 的相关反馈算法

图像检索的相关反馈过程是一个二分类问题,因此也可以将 SVM 算法用于相关反馈的学习和检索过程。将检索结果图像作为训练样本,由用户标记出正例样本(其它为负例样本),作为有类别标号的训练样本由 SVM 进行学习,详细的学习和检索算法如下:

- (1)使用多特征融合的方法进行检索,在初始检索结果中选择  $n$  幅图像提供给用户,初始时,所有图像权值  $w(i)=0$ ,正例集合  $I_p$  和  $I_n$  反例集合为空集。
- (2)用户标注与检索目标相关的图像,得到当前反馈结果

中正例集合  $I_p^1$  和反例集合  $I_n^1$ , 用来更新集合  $I_p$  和反例集合  $I_n$ :

$$I_p = (I_p \cup I_p^1) - I_n^1, I_n = (I_n \cup I_n^1) - I_p^1 \quad (22)$$

(3) 分类器准备训练样本  $(x_i, y_i), x_i \in I_p \cup I_n, y_i =$

$$\begin{cases} +1, & x_i \in I_p \\ -1, & x_i \in I_n \end{cases}$$

(4)利用 SVM 算法构造分类函数:

$$f(x) = \sum_i \alpha_i y_i K(x_i, x) + b \quad (23)$$

(5)利用分类函数  $f(x)$  的输出更新每一幅图像的权值:

$$w(i) = (1 - \beta)w(i) + \beta f(x_i), \beta \text{ 为衰减系数且 } 0 \leq \beta \leq 1 \quad (24)$$

(6)根据权值对图像库进行排序,返回检索图像。

### 5 实验结果与分析

本文在 P4 Windows XP 操作系统, 内存为 1 G 上实现一个图像检索系统, 实验中共有 1 000 幅图像, 本文设计了四个实验:(1)分别作基于颜色、纹理、形状的单一特征检索;(2)结合颜色、纹理、形状特征进行特征融合检索;(3)根据用户的反馈图像,通过支持向量机相关反馈,获得更接近检索者求的检索结果;(4)不同特征融合方法之间的对比实验。对检索算法的性能评价的重要指标是查准率(Precision),查全率(Recall)。查准率定义为检索出的图像中相关图像的数目占的比例,查全率定义为检索出的相关图像的数目占数据库中所有相关的图像数目的比例。表 1 给出了基于单一特征和特征融合的查准率和查全率。图 1 给出了相关反馈的次数对系统性能影响的查准率和查全率曲线。图 2 是本文的部分实验结果,在经过 3 次相关反馈后,返回的 12 幅图像。通过实验比较可知,本文的算法具有较好的检索性能。

表 1 单一特征和综合特征的查准率和查全率比较

	Color	Texture	Shape	Proposed
平均查准率	0.287	0.405	0.434	0.715
平均查全率	0.726	0.659	0.492	0.786

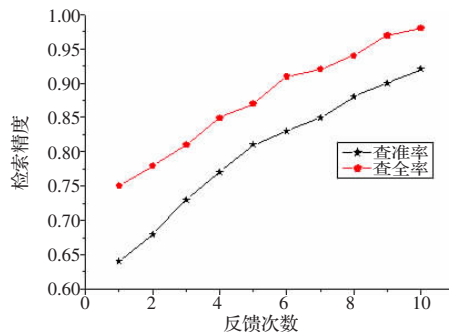


图 1 相关反馈的次数对系统性能影响



图 2 本文方法的检索结果