

感兴趣区域特征提取的图像检索

刘 恒, 马 涛, 高金辉

(河南师范大学 物理与信息工程学院, 河南 新乡 453007)

摘 要: 在基于内容的图像检索系统(CBIR)中, 特征的提取和应用是检索的关键技术。在CBIR中综合颜色、纹理和形状的提取方法是关注的重点。在传统的CBIR方法上, 利用Gabor滤波器获取感兴趣区域(ROIS)数目, 实现快速高效的特征提取。当感兴趣区域被确定后, 提取基于Gabor特征阈值的纹理特征, 基于直方图的颜色特征, 在YUV空间中的颜色矩量以及基于Zernike矩量的形状特征。通过在Corel图像库上进行测试, 证实在判定图像间的相似度上非常有效, 可以在CBIR中得到广泛应用。

关键词: 基于内容的图像检索; 感兴趣区域; Zernike矩量; 纹理分析; 特征提取

中图分类号: TP391

文献标识码: A

Image retrieval based on feature extraction of interest-based regions

LIU Heng, MA Tao, GAO Jin Hui

(Physics & Information Engineering College, Henan Normal University, Xinxiang 453007, China)

Abstract: Feature extraction and the use of the features as query terms are crucial problems in content-based image retrieval (CBIR) system. The main focus in this paper is on integrated color, texture and shape extraction methods for CBIR. We have developed original CBIR method that uses Gabor filtration for determining the number of ROIs, in which fast and effective feature extraction is performed. In the ROIs extracted, texture features based on thresholded Gabor features, color features based on histograms, color moments in YUV space, and shape features based on Zernike moments are then calculated. The features presented proved to be efficient in determining similarity between images. Our system was tested on Corel photo libraries and can be widely used in CBIR applications.

Key words: CBIR; ROIs; Zernike moment; texture analysis; feature extraction

传统的图像数据库是文本注释的, 因此图像检索是基于关键词查询的过程。这种方法存在很多弊端: (1)没有一套合适的词集能准确描述图像内容, 因此基于关键词的图像检索具有不确定性; (2)关键词注释具有很强的主观性。近几年研究的一些针对图像和视频的CBIR系统逐步取代了基于文本的检索。CBIR旨在从数据库中根据视觉内容检索相关图像。当用户提交一个查询图像, 检索系统将根据相似度自动显示检索结果。

在CBIR系统中, 从图像内容中提取的底层特征构成特征库。颜色、纹理、局部形状和空间信息等一系列特征在CBIR中使用得最为广泛^[1-3]。QBIC系统^[4]允许基于颜色、形状、纹理和示例图像的查询。Virage^[5]应用的是颜色、纹理和对象轮廓信息。Photobook^[6]由三个子系统组成, 分别是提取形状、纹理和人脸特征。NeTra^[7]、IRIS^[8]和CORE^[9]提供了基于全局颜色、形状、纹理和语

义内容的查询。VisualSEEK^[10]是基于颜色、纹理和空间布局的查询。

目前, CBIR主要采取两种特征提取技术: 全局和局部。在传统的CBIR中, 图像匹配通常依赖于全局方法。

在局部方法中, 常需用户标示出感兴趣区域或者事先将感兴趣区域定位于固定区域。本文提出一种自动的局部特征提取方法, 无需用户提供感兴趣区域或兴趣点。

该系统实现过程如图1所示。描述如下:

(1)利用Gabor滤波器作为兴趣点POIs(Points of Interest)筛选器选出一定数目的兴趣点;

(2)根据每个兴趣点, 选定围绕它的感兴趣区域ROIs;

(3)对于每一个ROI, 提取其特征(颜色、纹理和形状)的描述符;

(4)检索图像时, 依据测量出来的特征向量之间的

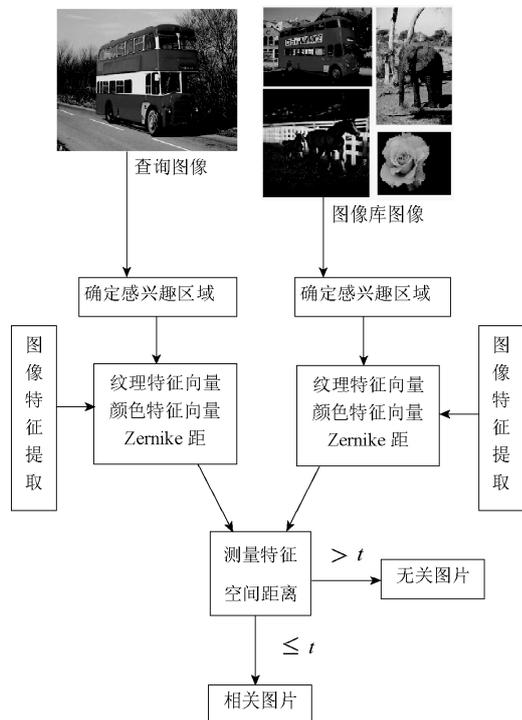


图1 图像检索过程

距离(相似度)提供检索结果。

从数学含义上,检索过程可描述如下:对于一个查询图像 Q , 当与图像库中图像 D_i 的特征向量之间的距离小于一个确定的阈值 t 时, 即 $\text{Dist}(\text{Feature}(Q), \text{Feature}(D_i)) \leq t$; 图像库中的图像 D_i 被确定为查询图像 Q 的相关图像。

本文提出了一种综合 Gabor 滤波器和局部特征提取方法, 提取基于 Gabor 阈值的纹理特征。基于直方图的颜色特征和 YUV 空间的颜色矩量、基于 Zernike 矩量的形状特征。为了提高系统效率和鲁棒性, 只提取感兴趣区域的颜色、形状和纹理特征。本文在现有的提取兴趣点方法上做了改进, 再根据兴趣点创建感兴趣区域。

1 Gabor 滤波器和 POIs 的提取

兴趣点的提取是计算机视觉中的 1 项关键技术, 很多研究者提出了兴趣点提取的代表性方法。例如, 基于小波分析的方法提取兴趣点^[11]、利用彩色 Harris 角点检测算法提取角点作为兴趣点^[12], 以及尺度不变特征变换(SIFT)法提取兴趣点^[13]等。

SIFT 变换基于窗口中围绕兴趣点的像素的图像梯度, 并不考虑颜色信息。而且, 一个标准的 SIFT 特征向量的建立较复杂, 它的特殊设计使得选择不具备直接性。Wolf 等人^[14]用 Gabor 滤波器组的响应作为 CBIR 系统中兴趣点的特征, 仅仅提取了纹理信息。

本文提出了一种与图像类别无关的 POIs 检测方法, 适用于包含多种分类的图像数据库的 CBIR 系统。该方法基于 Gabor 小波变换的分解, 并利用 Gabor 滤波器组提取纹理特征。

常用的二维 Gabor 滤波器组函数可以用复杂的正弦信号调制表示成高斯函数。Gabor 滤波器可以用公式(1)表示:

$$\psi(x, y; \sigma, \lambda, \theta_k) = \exp\left(-\frac{x_{\theta_k}^2 + \gamma^2 y_{\theta_k}^2}{2\sigma^2}\right) \exp\left(\frac{2\pi x_{\theta_k}}{\lambda} i\right) \quad (1)$$

其中 $x_{\theta_k} = x \cos \theta_k + y \sin \theta_k$; $y_{\theta_k} = -x \sin \theta_k + y \cos \theta_k$; σ 是沿着 x 轴和 y 轴的 Gaussian 包络的标准差; γ 是空间展弦比; λ 和 θ_k 分别为波长和方位。

本文仅采用该信号的奇分量。因为 γ 是常量, 不作为 Gabor 滤波器的参数^[7]。在 xy 平面上旋转 θ_k 将会产生一个以 θ_k 为参数的 Gabor 滤波器; 方位角取 $\theta_k = \frac{\pi}{n}(k-1)$, 其中 $k=1, 2, 3, \dots, n$; $n \in N$, n 表示所取方位个数。本文系统取 $n=4$ 。

参数 λ 为波长, $\frac{1}{\lambda}$ 为调和因子 $\cos(2\pi x_{\theta_k}/\lambda)$ 或 $\sin(2\pi x_{\theta_k}/\lambda)$ 的空间频率。 $\frac{\sigma}{\lambda}$ 决定了 Gabor 滤波器空间频率的频宽。库中所有滤波器的比率 $\frac{\sigma}{\lambda}$ 均取不变值 0.56, 相当于一个 8 倍频的半响应空间频宽。

输入图像 $f(x, y)$ 通过二维 Gabor 单滤波器 $\psi_0(x, y; \sigma, \lambda, \theta_k)$ 卷积处理得到 1 个 Gabor 图像响应函数:

$$\phi_0(x, y; \sigma, \lambda, \theta_k) = \sum_{\eta} \sum_{\zeta} f(\eta, \zeta) \psi(x-\eta, y-\zeta; \sigma, \lambda, \theta_k) d\eta d\zeta \quad (2)$$

为了实现查询的高效性, 只提取围绕 POIs 的 ROIs 的图像特征。

POIs 提取算法步骤如下:

- (1): 图像被 $\phi(x, y; \sigma, \lambda, \theta_k)$ 分为规模为 $a \times a$ 的互不重叠的块, $b^k(i, j)$, $k=1, 2, 3, 4$;
- (2): 计算出每块方差 $V^k(i, j)$;
- (3): 搜索 p 块 $b_p^k(i, j)$ 的最大方差。这些包含重要的亮度变化信息, 通常与图像中对象的轮廓相一致;
- (4): 在每个块 $b_p^k(i, j)$ 中, 搜索具有最大滤波响应的点。在 POI 图片中将这坐标点呈现出来。

POIs 提取算法结果如图 2 所示。

2 基于 Gabor 滤波器的纹理特征

首先利用一组对方向和尺度敏感的滤波器对图像进行滤波处理, 计算出图像的纹理描述信息, 检测 POIs, 确定 ROIs, 并计算在频域中滤出的 ROIs 的能量和熵参数。利用滤波器组所产生的滤波器响应可以直接被用作纹理特征, 而本文应用了 Gabor 滤波器结果的阈值。阈值计算如下:

$$T_0(x, y; \sigma, \lambda, \theta_k) = \chi(\Phi_0(x, y; \sigma, \lambda, \theta_k)) \quad (3)$$

$$\text{其中, } \chi(z) = \begin{cases} 0 & \text{当 } z < 0 \\ z & \text{当 } z \geq 0 \end{cases} \quad (4)$$



(a)

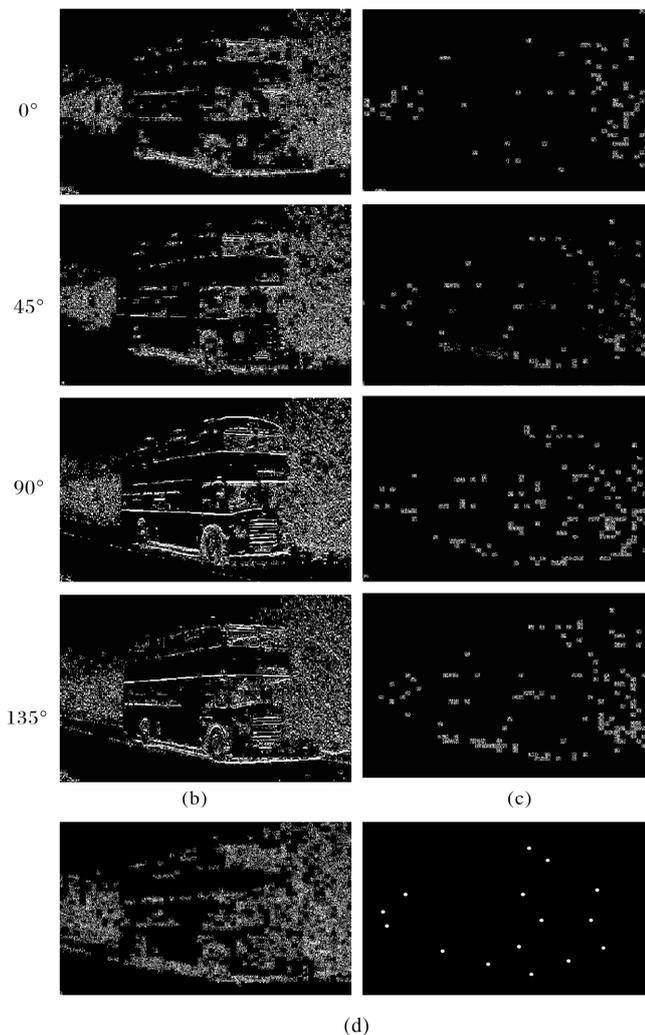


图2 POIS提取算法结果

(a)原始图像 (b)Gabor滤波器在0°、45°、90°、135°方向上的响应
(c)根据响应提取出 POIs 的图像 (d)提取出的兴趣点和选出的用来决定感兴趣区域的 p 点(具有最大滤波响应的点)

$\phi_0(x, y; \sigma, \lambda, \theta_k)$ 是 Gabor 滤波器响应的奇分量。

特征提取的方法包括分割图像和计算每个 ROI 区域的熵和能。在一个 $A \times B$ 的 ROI 区域中，归一化熵和能量计算如下：

$$E_r = \frac{\sum_x \sum_y T_0^2(x, y; \sigma, \lambda, \theta_k)}{(A \cdot B)^2} \quad (5)$$

$$Entropy = \frac{1}{(A \cdot B)} \times \sum_x \sum_y \{ [T_0^2(x, y; \sigma, \lambda, \theta_k)] \times \log_2 [T_0^2(x, y; \sigma, \lambda, \theta_k)] \} \quad (6)$$

最终，获得纹理特征向量。

$$F_{texture} = \{ E_{r1}, E_{r2}, \dots, E_{rp}, Entropy_1, Entropy_2, \dots, Entropy_p \} \quad (7)$$

其中， P 为 ROI 区域的数目。

查询图像 Q 和库中图像 D 纹理相似度定义如下：

$$d^{(Q)(D)}(E_r, Entropy) = \sum_p d_p^{(Q)(D)} \quad (8)$$

其中，

$$d_p^{(Q)(D)} = |E_{rp}^{(Q)} - E_{rp}^{(D)}| + |Entropy_p^{(Q)} - Entropy_p^{(D)}| \quad (9)$$

3 颜色特征的提取

颜色是人类视觉能识别分辨的最直接的特征之一。提取图像中感兴趣区域 p 中颜色特征，创建一个有效向量，以供检索。 YUV 空间广泛应用于图像压缩和图像应用。 Y 代表颜色亮度， U 和 V 代表色度。视觉系统对于 Y 成分相对于其他两个色度分量更为敏感，典型的 Y 分量有很高的密度(例如，在 JPEG 和 MPEG 标准)。

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (10)$$

因此，颜色特征的提取分为两步。首先，将图像中像素的 RGB 值转换为 YUV 值。其次，建立 Y (亮度)分量直方图，如图 3 所示。

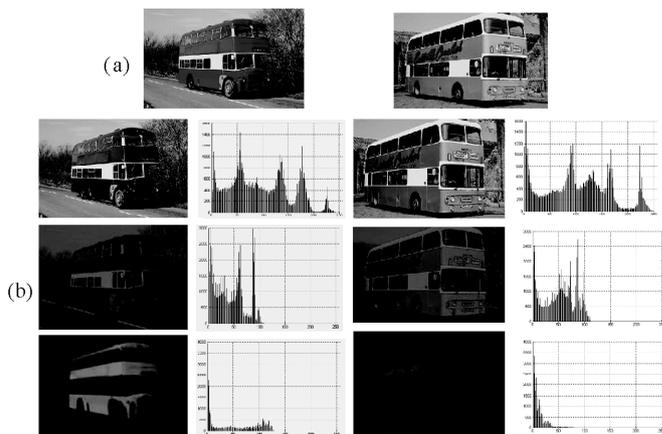


图3 Y (亮度)分量直方图

(a)原始图像 (b)直方图体现的 YUV 分量

为每一个感兴趣区域(包括 $A \cdot B$ 个像素)创建直方图，给出 $(A \cdot B)$ 柄值，如图 4 所示。

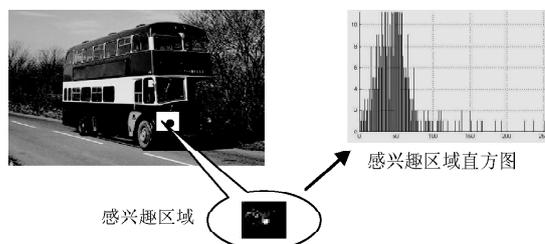


图4 提取出的 ROI 的直方图

$$H_x(j) = \frac{\text{亮度为}j\text{的像素数}}{x\text{中的像素总数}} \quad (11)$$

另外，颜色矩也作为颜色特征体现出来。

$$E_p = \frac{1}{A \cdot B} \sum_{s=1}^{A \cdot B} Y_{ps} \quad (12)$$

$$\sigma_p = \sqrt{\frac{1}{A \cdot B} \sum_{s=1}^{A \cdot B} (Y_{ps} - E_p)^2} \quad (13)$$

最终，我们获得颜色特征向量：

$$F_{color} = \{H_{Q1}, H_{Q2}, \dots, H_{QP}, E_1, E_2, \dots, E_p, \sigma_1, \sigma_2, \dots, \sigma_p\} \quad (14)$$

其中， P 是 ROI 的数目，查询图像 Q 和数据库中的图像 D 颜色相似度定义如下：

$$d^{(Q)(D)}(D_p, E_p, \sigma_p) = \sum_p d_p^{(Q)(D)} \quad (15)$$

此处，

$$d^{(Q)(D)} = |II_p^{(Q)} - II_p^{(D)}| + |E_p^{(Q)} - E_p^{(D)}| + |\sigma_p^{(Q)} - \sigma_p^{(D)}| \quad (16)$$

4 形状特征

基于形状的图像检索为两形状之间特征的相似度的测量。包括形状特征提取和特征的相似度比较。

Zernike 矩 (ZM) Z_{nm} 由图像 $f(x, y)$ 唯一决定，相反， $f(x, y)$ 被 Z_{nm} 唯一描述。ZM 的这种正交性允许描述对冗余的删减，有助于提高计算效率。

为了计算图片 $f(x, y)$ 的 Zernike 矩，图像首先被投影在极坐标单位圆中，图片中心定位为圆心。落在单位圆外的像素忽略不计，如图 5 所示。

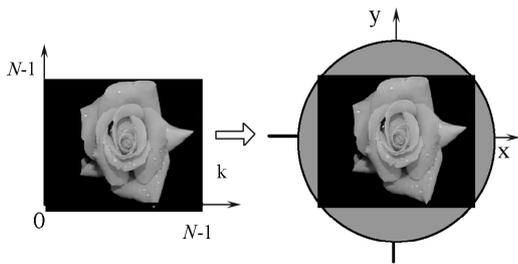


图 5 直角坐标向极坐标变换

ZM 是一系列定义在极坐标单位圆内正交的 Zernike 多项式，从感兴趣区域 p 中提取对象并计算 Zernike 矩。

Zernike 矩做为描述符。ZM 的阶由 2 个参数决定： n 和 m 。每一个 Zernike 矩通过径向多项式与相位值相乘获得。阶为 n 、重复度为 m 的 Zernike 矩定义如下：

$$Z_{nm} = \frac{n+1}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [V_{nm}(x, y)]^* f(x, y) dx dy \quad (17)$$

其中， $f(x, y)$ 表示图像笛卡尔坐标 (x, y) 处的亮度； $[V_{nm}(x, y)]^*$ 是下面函数的复共轭函数，

$$V_{nm}(x, y) = V_{nm}(\rho \cos(\theta)), \rho \sin(\theta) = R_{nm}(\rho) e^{im\theta} \quad (18)$$

其中， (ρ, θ) 表示极坐标， $i = \sqrt{-1}$ 。图像域中极坐标

(ρ, θ) 与笛卡尔坐标 (x, y) 的关系为： $x = \rho \cos(\theta), y = \rho \sin(\theta)$ ； $n \geq 0, n - |m|$ 是正偶数。

Khotanzad 和 Hong^[15] 定义了径向函数：

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} \frac{(-1)^s [(n-s)!] \rho^{n-2s}}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \quad (19)$$

因为 Zernike 基函数取单位圆作为定义域，在实现时，一幅 $N \times N$ 矩形图被变换并归一化到一个单位圆 $x^2 + y^2 \leq 1$ 内，如图 7 所示，单位圆作为边界。

公式(17)的离散逼近如下：

$$Z_{nm} = \frac{4(n+1)}{(N-1)^2 \pi} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f(k, l) R_{nm}(\rho_{k,l}) e^{-jm\theta_{k,l}} \quad 0 \leq \rho_{k,l} \leq 1 \quad (20)$$

离散极坐标如下：

$$\rho_{k,l} = \sqrt{x_k^2 + y_l^2} \quad \theta_{kl} = \arctan\left(\frac{y_l}{x_k}\right) \quad (21)$$

变换关系为：

$$x_k = \frac{\sqrt{2}}{N-1} k + \frac{-1}{\sqrt{2}}, y_l = \frac{\sqrt{2}}{N-1} l + \frac{-1}{\sqrt{2}} \quad (22)$$

再建一个离散函数

$$\tilde{f}(x, y) = \sum_{n=0}^{n_{\max}} \frac{C_{n0}}{2} R_{n0}(\rho) + \sum_{n=1}^{n_{\max}} (C_{nm} \cos(m\theta) + S_{nm} \sin(m\theta)) R_{nm}(\rho) \quad (23)$$

其中， n_{\max} 是 ZM 的最大阶； C_{nm} 和 S_{nm} 分别是 Z_{nm} 的实部和虚部。在极坐标 (ρ, θ) 中，

$$C_{nm} = 2\Re(Z_{nm}) = \frac{2n-2}{\pi} \iint_{x^2+y^2 \leq 1} F(x, y) R_{nm}(\rho) \cos(m\theta) dx dy \quad (24)$$

$$S_{nm} = 2\Im(Z_{nm}) = \frac{-2n-2}{\pi} \iint_{x^2+y^2 \leq 1} F(x, y) R_{nm}(\rho) \sin(m\theta) dx dy$$

因为， Z_{nm} 是复数，把模 $|Z_{nm}|$ 作为形状特征。

由于 ZM 值具有旋转不变的特性。就均方差而言，利用一组 ZM 作为图像描述符优于其他描述符：几何矩、Legendre 矩、旋转矩和复数矩。

用以下包含 Zernike 矩的特征向量描述形状特征：

$$SFV = (Z_{1m}, Z_{2m}, \dots, Z_{nm}) \quad (25)$$

使用 SFV 建立数据库中每个形状的索引。

5 实验与分析

在 Corel 图片库中检验了本系统。其中包括 1000 张图片。这些图片选自 10 个类别：人物、美食、海边、公交车、大象、恐龙、花、建筑、马、雪山。采用 2 个众所周知的指标评价检索效率：查准率和查全率。通常随着被检索图像数目的增加，查全率增加，查准率减小。

为了证明采用算法的有效性。本系统执行了一系列查询，利用 Corel 数据库中的 50 张图片作为查询图片，覆盖了数据库中图像 5%，计算每次查询，得到相应查准率和查全率。同时，考察了在 CBIR 系统执行中的一系列特征对检索的影响，实验证明基于综合特征的查精率高于基于单一或两种特征。

图 6 为利用不同特征查询 50 张图片的平均 precision recall 曲线。

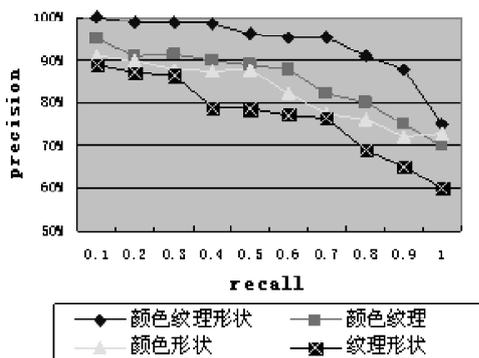


图 6 平均查准 - 查全率图

对数据库中的原始图片和变化后的图片进行测试。

为了研究变化后图像的检索效果，对一些查询图片进行尺度变化和旋转处理。图片分别旋转 10、45 和 90 度。尺度变化因子为 0.9 和 1.1。试验结果查准率减小约 7%。在进一步的实验中，将查询图片边缘部分随机剪裁 20%，平均查准率降低约 10%。实验结果说明对于局部图像的查询，在识别图像时，并不是所有感兴趣区域都是必须的。针对修改后图像的 precision-recall 曲线如图 7 所示。

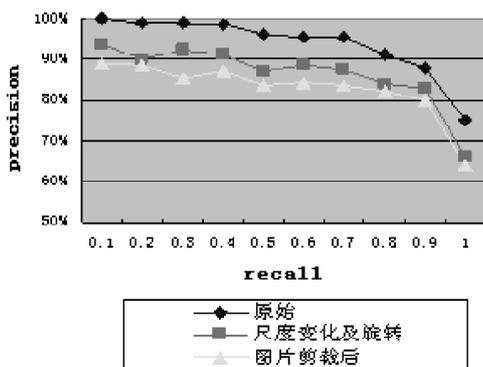


图 7 修改后图像的平均查准 - 查全率图

本文提出了一种新的检索方法，综合了颜色、纹理和形状信息。基于所有特征分量的图片相似度有所提升。不同于现有 CBIR 系统，仅计算提取出的有限感兴趣区域的特征。也有其他方法确定 ROI 区域(比如，SIFT, Harris)，但是本文的方法是基于兴趣点检测和 Gabor 滤波器，这有利于利用滤波器响应描述纹理特征。在 Corel 库中对该方法进行了测试，取得了满意的结果。

参考文献

- [1] SMEULDERS AWM, WORRING M, GUPTA A, et al. Content-based image retrieval at the end of the early years[J]. IEEE Trans Pattern Anal Machine Intell, 2000, 22: 1349-1380.
- [2] CONNERS R, HARLOW C. A theoretical comparison of texture algorithms[J]. IEEE Trans Pattern Anal Machine Intell, 1980, 2: 204-222.
- [3] HOWARTH P, RUGUR S. Evaluation of texture features for content-based image retrieval[J]. ENSENER P (eds) Image and video retrieval. Springer LNCS, 2004, 3115: 326-334.
- [4] FLICKER M, SAWHNEY H, NIBLACK W, et al. Query by image and video content: the QBIC System[J]. IEEE Comput Mag, 1995, 28: 23-32.
- [5] BACH J R, FULLER C, GUPTA A, et al. An open framework for image management[J]. SPIE Storage Retr Still Image Video Database, 1996: 76-87.
- [6] PENTLAND A, PICARD R, SCLAROFF S. Photobook: Content-based manipulated of image databases[J]. Int J Comput Vis 1996, 18: 233-254.
- [7] MA W Y, MANJUNATH B S. Netra: A toolbox for navigating large image databases[M]. Proceedings of ICIIP'97. Santa Barbara, CA, 1997: 568-571.
- [8] ALSHUTH P, TERMES P, KLAUCK C, et al. IRIS image retrieval for images and video[M]. Proceedings of the First International Workshop on Image Database and Multimedia Search, Amsterdam, The Netherlands, 1996, 170-179.
- [9] WU J K, NARASHIHALU A D, MEHTRE B M, et al. CORE: A content-based retrieval engine for multimedia information systems [J]. Multimed Syst, 1995, 3: 25-41.
- [10] SMITH J R, CHANG S F. VisualSEEK: A fully automated content-based image query system[M]. Proceedings of the ACM international conference on multimedia, Boston, MA, 1997: 87-98.
- [11] 张佳佳, 徐东平. 基于兴趣点特征对互信息准则的图像配准算法[J]. 计算机与数字工程, 2006, 36(10): 162-191.
- [12] 刘毅, 张明. 基于兴趣点的图像检索[J]. 计算机工程与应用, 2006, (12): 102-106.
- [13] 高健, 黄心汉, 彭刚, 等. 一种简化的 SIFT 图像特征点提取算法 [J]. 计算机应用研究, 2008, 25(7): 2213-2222.
- [14] WOLF C, JOLION J M, KROPATSCHE W, et al. Content based image retrieval using interest points and texture features[J]. Proceedings of the IEEE International Conference on Pattern Recognition, 2000: 1511-1514.
- [15] KHOTANZAD A, HONG Y H. Invariant image recognition by Zernike moments[J]. IEEE Trans Pattern Anal Machine Intell 1990, 12: 489-498.

(收稿日期: 2009-02-24)