

基于感兴趣区的小波域彩色图像检索新方法

王向阳^{1,2)} 杨红颖¹⁾ 胡峰丽¹⁾

¹⁾(辽宁师范大学计算机与信息技术学院,大连 116029) ²⁾(南京大学计算机软件新技术国家重点实验室,南京 210093)

摘要 提出了一种基于感兴趣区的小波域彩色图像检索新方法。该方法首先结合人眼视觉感知特性,在小波变换域内利用 K-均值聚类提取出感兴趣区域,然后以感兴趣区的小波系数局部能量作为纹理特征,颜色均值和均方差作为颜色特征,重心坐标作为位置特征,计算图像间内容的相似度并进行检索。仿真实验结果表明,当图像中有明显的感兴趣区域时(特别是背景简单的图像),该方法能够更加准确地查找出用户所需内容的图像,明显地提高了检索精度。

关键词 图像检索 感兴趣区 小波变换

中图法分类号: TP391 **文献标识码:** A **文章编号:** 1006-8961(2006)02-0175-05

A Novel Color Image Retrieval Scheme Based on Regions-of-Interest

WANG Xiang-yang^{1,2)}, YANG Hong-ying¹⁾, HU Feng-li¹⁾

¹⁾(School of Computer and Information Technique, Liaoning Normal University, Dalian 116029)

²⁾(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract In this paper, a new ROI based color image retrieval using DWT is proposed. Firstly, the ROI are extracted in DWT domain by using the human visual characteristic and K-mean clustering. Secondly, the local energy of wavelet coefficients of ROI is used for the texture feature, the mean values and standard deviations are used for the color feature, and the weight center coordinate of ROI is used for the location feature. Finally, the similarity between images is computed according to the ROI's above features. Experimental results show that our image retrieval is more accurate in retrieving the user-interested images when there are ROI in the image(especially for the image with simple background).

Keywords image retrieval, regions-of-interest, DWT

1 引言

基于内容的图像检索方法主要分为两大类^[1,2]: 全局特征方法和感兴趣区特征方法。其中,基于感兴趣区的检索技术不仅利用了图像的局部特征,而且消除了次要信息的干扰,从而在一定程度上实现了物体层次的图像检索,改善了检索效率。然而,理论分析和实验结果表明,现有基于感兴趣区的图像检索方法均不同程度存在以下不足:(1)感兴趣区的提取多在空间域内进行,且仅仅考虑了图像的灰

度信息,未能充分利用人眼视觉系统的感知特性;(2)没有准确描述并利用图像中感兴趣区域的位置等空间特征;(3)相似度计算模型不够合理,与人的感知特性存在差异。

为此提出了一种基于感兴趣区的小波域图像检索新方法,该方法具有 3 个特点:(1)结合人眼视觉感知特性,采纳了基于小波变换的感兴趣区提取策略;(2)不仅比较准确地描述了感兴趣区的纹理特征(以小波系数局部能量作为感兴趣区的纹理特征),同时考虑了感兴趣区的位置特征;(3)完善了相似度计算模型。

基金项目:辽宁省自然科学基金项目(20032100);计算机软件新技术国家重点实验室(南京大学)开放基金项目(03-02)

收稿日期:2004-12-03;**改回日期:**2005-04-18

第一作者简介:王向阳(1965~),男,教授。1995 年于吉林大学获计算机专业硕士学位。主要研究领域为网络信息安全技术、多媒体信息处理技术。E-mail: wxy37@126.com

2 感兴趣区的提取

由小波分析理论知:在小波变换域内,感兴趣区的局部小波系数方差要明显大于背景区域的局部小波系数方差。因此,完全可以通过定义图像局部区域的小波系数方差来区分感兴趣区和背景,即提取出感兴趣区。

设原始彩色图像为 $I = \{I(i, j), 0 \leq i < M, 0 \leq j < N\}$, $I(i, j)$ 代表原始彩色图像的第 i 行、第 j 列像素的颜色值。

选取适合人眼视觉感知特性的 HSV 颜色空间进行彩色图像的分割与检索,而且在彩色图像的亮度分量上提取感兴趣区域。则感兴趣区提取过程(关键步骤)可描述如下:

(1) 对彩色图像 I 的亮度分量 V 进行 1 级小波变换,得到小波系数矩阵 U 。然后,将亮度分量 V 划分成互不重叠的图像子块,令空间域图像子块的大小为 $L \times W$ 。

(2) 在小波系数矩阵 U 的 3 个高频子带中,计算对应于空间域图像块的 3 个小波系数块的块内系数方差,并求出其平均值,以此作为图像分割的特征。设在亮度分量 V 中,位于 (i, j) ($0 \leq i < M/L, 0 \leq j < N/W$) 处的空间域图像块为

$$V_{i,j} = \{v(mi, nj), 0 \leq m < L, 0 \leq n < W\}$$

则 U 的 3 个高频子带 LH、HL 和 HH 内相应小波系数块分别为

$$U_{i,j}^\lambda = \{u^\lambda(mi, nj), 0 \leq m < L/2, 0 \leq n < W/2\} \quad (\lambda = \text{LH, HL, HH})$$

而 3 个小波系数块的块内系数方差为

$$\sigma_{i,j}^\lambda = \frac{1}{L/2 \times W/2} \sum_{m=0}^{L/2-1} \sum_{n=0}^{W/2-1} (|u^\lambda(mi, nj)| - \mu_{i,j}^\lambda)^2 \quad (\lambda = \text{LH, HL, HH})$$

其中

$$\mu_{i,j}^\lambda = \frac{1}{L/2 \times W/2} \sum_{m=0}^{L/2-1} \sum_{n=0}^{W/2-1} |u^\lambda(mi, nj)| \quad (\lambda = \text{LH, HL, HH})$$

则 3 个小波系数块的块内系数方差平均值(即图像分割特征)为

$$\sigma_{i,j} = \frac{\sigma_{i,j}^{\text{LH}} + \sigma_{i,j}^{\text{HL}} + \sigma_{i,j}^{\text{HH}}}{3}$$

(3) 利用 K-均值聚类算法对上述图像分割特征进行聚类(分成 2 类),聚类中心分别代表感兴趣

区和背景,即图像分割特征 $\sigma_{i,j}$ 大,表示对应的空间域图像块属于感兴趣图像块。

(4) 将所有感兴趣图像块作为要处理的目标集合 S ,并将集合 S 按照 8 连通准则划分为满足如下条件的感兴趣图像块连通集 S_0, S_1, \dots, S_{P-1} ,其中, P 为连通集个数。

① $\bigcup_{k=0}^{P-1} S_k = S$, 即所有的感兴趣图像块连通集组成了整个目标集合;

② $S_i \cap S_j = \emptyset$ ($i \neq j, 0 \leq i, j \leq P-1$), 其中 \emptyset 为空集;

③ S_k ($0 \leq k \leq P-1$) 是一个感兴趣图像块连通集;

④ S_i 与 S_j 非连通,其中 $i \neq j, 0 \leq i, j \leq P-1$ 。

假设感兴趣图像块连通集 S_k 中共包含有 M_k 个感兴趣图像块,且 $p_i^{(k)}$ ($0 \leq i < M_k(L \times W)$) 为连通集 S_k 中第 i 个像素,并记像素 $p_i^{(k)}$ 的坐标为 $(x_i^{(k)}, y_i^{(k)})$ 。取变量

$$l_k = \min \{x_0^{(k)}, x_1^{(k)}, \dots, x_{M_k(L \times W)-2}^{(k)}, x_{M_k(L \times W)-1}^{(k)}\}$$

$$r_k = \max \{x_0^{(k)}, x_1^{(k)}, \dots, x_{M_k(L \times W)-2}^{(k)}, x_{M_k(L \times W)-1}^{(k)}\}$$

$$t_k = \min \{y_0^{(k)}, y_1^{(k)}, \dots, y_{M_k(L \times W)-2}^{(k)}, y_{M_k(L \times W)-1}^{(k)}\}$$

$$b_k = \max \{y_0^{(k)}, y_1^{(k)}, \dots, y_{M_k(L \times W)-2}^{(k)}, y_{M_k(L \times W)-1}^{(k)}\}$$

则可以定义一个矩形 $R = \{(l_k, t_k), (r_k, b_k)\}$, 并称矩形 R 为感兴趣图像块连通集 S_k 的连通矩形,记为 $R(S_k)$ 。再取变量 $A = |r_k - l_k| \cdot |b_k - t_k|$, 并称 A 为感兴趣图像块连通集 S_k 的连通面积,并记为 $A(S_k)$ 。

不难看出,对所有感兴趣图像块集 S 进行连通集划分,将得到一个连通集集合 $\{S_k\}$ ($0 \leq k \leq P-1$),而只要在感兴趣图像块连通集集合 $\{S_k\}$ ($0 \leq k \leq P-1$) 中选取连通面积较大的连通集 Ω (本文选取了连通面积大于 2% 的连通集),则相应的连通矩形 $R(\Omega)$ 即为所要提取的感兴趣区。

3 感兴趣区的特征提取

经过以上图像分割,得到了一系列反映图像内容的感兴趣区域,而图像的主要内容就由这些感兴趣区的颜色、纹理和位置等特征来表示。

3.1 颜色特征

感兴趣区的颜色特征用该区域的颜色均值和均方差表示。设感兴趣区为 R (大小为 $m \times n$), 则其颜

色均值与颜色均方差分别为

$$\mu_i = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} I_i(x,y)}{m \times n}$$

$$\sigma_i = \sqrt{\frac{\sum_{x=0}^{m-1} \sum_{y=0}^{n-1} (I_i(x,y) - \mu_i)^2}{m \times n}}$$

其中, $I_i(x,y)$ 为感兴趣区 R 内像素点 (x,y) 的第 i 个颜色分量。选用能够较好符合人眼视觉感知特性的 HSV 颜色空间, 于是得到感兴趣区 R 的颜色特征 $C = (\mu_H, \sigma_H, \mu_S, \sigma_S, \mu_V, \sigma_V)$ 。

3.2 纹理特征

纹理通常定义为图像的某种局部性质, 是相邻像素的灰度(颜色)空间相关性或灰度(颜色)空间位置变化的视觉表现。利用纹理特征进行图像检索是一种非常有效的手段。

在小波变换域内提取感兴趣区的纹理特征, 以期充分利用小波系数与图像的平移、旋转及缩放均无关等特性。基于小波变换的感兴趣区纹理特征提取方法如下:

(1) 从感兴趣区 R 中提取出其亮度分量 R^V , 并对亮度分量 R^V 进行 2 级小波变换。

(2) 在小波变换域内, 计算出感兴趣区亮度分量 R^V 的小波系数局部能量^[3]。设坐标位置为 (x,y) 的小波系数 $\{d_{r,s}^{l,i}\}$ 的局部能量为 $E_l^o(x,y)$ (其中, $l \in \{1,2\}$ 表示小波变换层数, $\theta \in \{LL, LH, HL, HH\}$ 表示小波域内的子带图像方向, r,s 表示移动因子), 则

$$E_l^o(x,y) = \sum_{r,s \in Z} (d_{r,s}^{l,i})^2 K(2^{-j}x - r, 2^{-j}y - s)$$

这里, $K(x,y)$ 选取为 Gaussian 内核函数^[3]。

(3) 计算出亮度分量 R^V 各个子带图像的能量, 第 l 变换层、 θ 方向的子带图像能量为

$$E_l^o = \sum_x \sum_y E_l^o(x,y)$$

(4) 为了减少计算复杂度, 选取亮度分量 R^V 各子带图像的下列能量特征(局部能量的对数值)作为感兴趣区 R 的纹理特征 T , 即

$$T = \{\log E_1^{LL}, \log E_1^{LH}, \log E_1^{HL}, \log E_1^{HH}, \log E_2^{LL}, \log E_2^{LH}, \log E_2^{HL}, \log E_2^{HH}\}$$

3.3 位置特征

考虑到计算方便, 感兴趣区 R 的位置特征 O 用其重心的归一化坐标表示, 即

$$O = \left(\frac{\bar{x}}{M}, \frac{\bar{y}}{N} \right)$$

这里, \bar{x}, \bar{y} 表示重心坐标 M 和 N 是图像的宽和高。

最终, 第 i 个感兴趣区 R_i 的特征用向量表示为 $F_i = (C_i, T_i, O_i)$, 而含有 m 个感兴趣区的图像的特征向量表示为

$$\{F_0, F_1, F_2, \dots, F_{m-2}, F_{m-1}\}$$

4 图像相似性的测度

采纳高斯函数计算任意感兴趣区 R_i, R_j 之间各特征的相似程度, 并通过调整各特征的方差 σ_i 将其归一化为具有不同量纲的特征变量, 使得相似度 s 总在 $[0,1]$ 之间。当 $s=1$ 时, 两个感兴趣区相似; 当 $s=0$ 时, 两个感兴趣区不相似。相似度的计算公式如下(特征值已经过归一化)。

(1) 颜色相似度

$$s_C = \exp \left[- \frac{\sum_{g \in \{R, S, V\}} [(\mu_g(R_i) - \mu_g(R_j))^2 + (\sigma_g(R_i) - \sigma_g(R_j))^2]}{6\sigma_C^2} \right]$$

(2) 纹理相似度

$$s_T = \exp \left[- \frac{\sum_{l \in \{1,2\}, \theta \in \{LL, LH, HL, HH\}} [\log E_l^o(R_i) - \log E_l^o(R_j)]^2}{7\sigma_T^2} \right]$$

(3) 位置相似度

$$s_O = \exp \left[- \frac{(\bar{x}(R_i) - \bar{x}(R_j))^2 + (\bar{y}(R_i) - \bar{y}(R_j))^2}{2\sigma_O^2} \right]$$

于是, 图像中任意两个感兴趣区 R_i, R_j 之间的相似度是各特征相似度的加权平均, 即

$$s_R(R_i, R_j) = w_C s_C + w_T s_T + w_O s_O, \text{ 且 } w_C + w_T + w_O = 1$$

假设示例图像 Q 有 m 个感兴趣区, 数据库中任意图像 I 有 n 个感兴趣区, 则图像 Q 相对于 I 的相似度为

$$s(Q, I) = \sum_{i=0}^{m-1} W_i s_R(R_i, P_i(R_i)), \text{ 且 } \sum_{i=0}^{m-1} W_i = 1$$

其中, W_i 为感兴趣区 R_i 的权重。由于图像中各感兴趣区的面积不同, 因此重要程度各不相同。本文将 W_i 初始化为感兴趣区 R_i 占整个图像面积的百分比, 从而使权重 W_i 与该感兴趣区 R_i 的面积成正比。 $P_i(R_i)$ ($i=0,1,\dots,m-1$) 表示图像 Q 的感兴趣区 R_i 与图像 I 中感兴趣区之间的映射关系, 即返回在图像 I 中与图像 Q 的第 i 个感兴趣区 R_i 最为相似的感兴趣区。

5 实验结果

为了验证本文方法的性能,以下给出了基于内容的图像检索实验结果,并与 SIMPLIcity 方法^[4]进行了对比。实验中,选用了著名的 Corel 图像库,该图像数据库包括飞机、人物、动物、建筑、海滩、餐具、花卉、体育、汽车、村庄等各类图片 10 000 多张(每类图像大约包含 100 幅)。空间域图像子块的大小为 $L \times W = 4 \times 4$,相关权值参数分别设置为

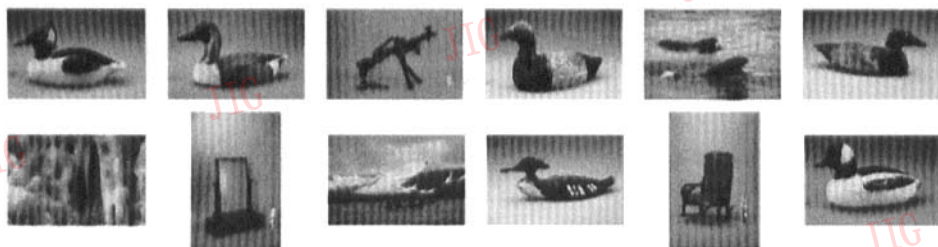
$$\sigma_c = \frac{1}{6}, \sigma_r = \frac{1}{7}, \sigma_o = \frac{1}{2}, w_c = w_r = w_o = \frac{1}{3}$$

另外,为评价图像检索算法的效果,同时考虑到

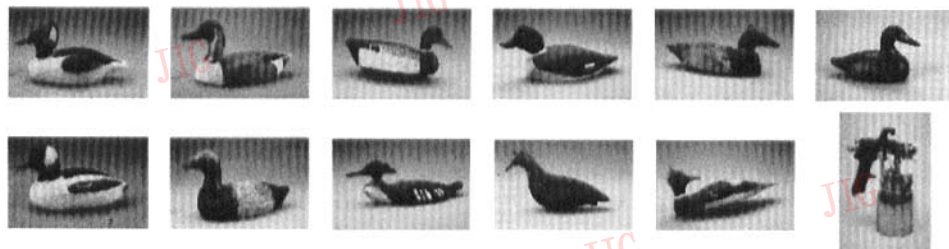
检索系统返回的相似图像数目(不是所有)不适宜采用常规的查准率(Precision)和查全率(Recall)评价指标,故采用了“标准查准率(Normal Precision)”和“标准查全率(Normal Recall)”^[5]作为相似检索的评价准则。

5.1 实验 1

图 1 给出了应用本文方法和 SIMPLIcity 方法进行图像检索的结果。其中,显示图像中的左上角是示例图像(水鸟),其余 11 幅图像是查询结果。不难看出,在 SIMPLIcity 方法(图 1(a))的查询结果中,正确的图像有 5 幅,而在本文方法(图 1(b))的查询结果中,正确的图像有 10 幅。



(a) SIMPLIcity 方法



(b) 本文方法

图 1 示例图像和检索图像队列

Fig.1 Query image and retrieval results

5.2 实验 2

在 Corel 图像库中选取 5 类图像(飞机、动物、海滩、建筑、汽车),每类随机抽取 5 幅图像作为示例图像,总共构成 25 次查询,每次查询选取前 30 幅最相似的图像作为检索结果。对于每类图像而言,计算其 5 次查询结果的标准查准率和标准查全率的平均值,并将其作为该类图像的平均标准查准率和平均标准查全率,如表 1 所示。

表 1 两种图像检索方法的性能对照

Tab.1 Performance comparison between two methods (retrieval correct ratio)

图像类别	平均标准查准率(%)		平均标准查全率(%)	
	SIMPLIcity 方法	本文方法	SIMPLIcity 方法	本文方法
飞机	57.7	80.2	51.2	68.2
动物	60.2	74.4	55.1	67.1
海滩	53.5	63.9	47.6	58.4
建筑	54.1	62.1	43.8	53.9
汽车	60.9	70.4	53.4	62.5

6 结 论

本文提出了一种新的基于感兴趣区的小波域彩色图像检索方法。实验结果表明,该算法不仅具有较好的检索性能,而且其检索结果能较好地接近人的视觉感知效果。进一步的研究工作包括:(1)探索更为有效的感兴趣区提取方法(特别是对于背景复杂图像);(2)结合相关反馈和机器学习技术,进一步提高查询效率,以便在查询过程中更好地实现自适应和自学习等。

参考文献(References)

- 1 Smeulders A W M, Worring M, Santini S, *et al.* Content-based image retrieval at the end of the early years[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22** (12): 1349 ~ 1380.
- 2 Veltkamp R C, Tanase M. Content-based Image Retrieval Systems; A Survey [R]. UU-CS-2000-34, Utrecht, The Kingdom of the Netherlands; Department of Computer Science, Utrecht University, 2002.
- 3 Randen T, Husoy J H. Filtering for texture classification: A comparative study [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, **21**(4):291 ~ 310.
- 4 Wang J Z, Li J, Wiederhold G. SIMPLcity: Semantics-sensitive integrated matching for picture libraries [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23** (9): 947 ~ 963.
- 5 Tan Kian-Lee, Ooi Beng Chin, Yee Chia Yeow. An evaluation of color-spatial retrieval technique for large image database [J]. *Multimedia Tools and Applications*, 2002, **14**(1):55 ~ 78.

1 Smeulders A W M, Worring M, Santini S, *et al.* Content-based