

基于颜色特征的图像检索技术^①

王娟, 孔兵, 贾巧丽

(河南大学 计算机与信息工程学院, 开封 475004)

摘要: 在基于内容的图像检索(CBIR)中, 颜色特征作为图像的一种重要视觉信息, 由于计算简单, 鲁棒性强, 已经得到了广泛的应用。从图像的颜色特征出发, 针对颜色空间(颜色空间的转换)、颜色的量化(HSV颜色空间)、颜色特征提取及表示和颜色特征的相似性度量做了系统讨论和研究。接着通过经典的分块颜色直方图算法来介绍基于颜色特征的图像检索全过程。最后总结了基于颜色特征的图像检索研究存在的问题以及进一步的研究方向。

关键词: 图像检索; 颜色特征; HSV颜色空间; 分块颜色直方图算法

Color-Based Image Retrieval

WANG Juan, KONG Bing, JIA Qiao Li

(School of Computer and Information Engineering, Henan University, Kaifeng 475004, China)

Abstract: The color features, as an important visual information of image, have been widely used in the Content-Based Image Retrieval(CBIR) because of the simple calculation and better robustness. According to the color features of the image, this paper makes systematic discussion and research on color space(color space transformation), color quantification(HSV color space), color features extraction and color characteristic similarity measurement. Then, through classic color segment histogram algorithm, the entire process of Color-based Image Retrieval is proposed. Finally, the existent problems and further research direction of the color-based image retrieval is summarized.

Key words: image retrieval; color feature; HSV color space; color segment histogram algorithm

1 引言

近年来, 随着多媒体技术及网络技术的迅速发展, 图像资源越来越丰富, 传统的基于文本的图像检索方法由于对文字标注的主观性和繁琐性, 已经不能满足检索需要, 在这种情况下, 基于内容的图像检索(Content-Based image Retrieval, CBIR)^[1]技术应运而生。基于内容的图像检索技术的基本思路是: 首先对图像内容进行分析, 自动或半自动的从中抽取形状、颜色、纹理等特征, 然后利用相似性度量函数计算或评价特征之间的相似性, 最后将最相似的图像作为检索结果返回给用户。

在图像检索中, 图像的内容通过图像的特征反映。图像的特征分为两大类, 即低层视觉特征和高层语义

特征。在这两大类特征中, 虽然基于语义特征的检索比较符合人的使用要求, 但是由于受当前计算机视觉和图像理解发展水平的限制, 且还涉及到心理学、生物视觉等其它一些学科, 使得对于这种方式的检索还难以实现。因此目前基于内容的图像检索的研究热点主要集中在基于对低层视觉特征的检索。

在基于低层视觉特征的检索中, 颜色特征又是描述一幅图像最简便而有效的特征, 它不仅具有旋转不变性和尺度不变性, 而且比其它特征更容易获取, 因此在文中主要讨论基于颜色特征的图像检索。包括以下几个方面: 颜色空间的选择、颜色的量化、颜色特征的提取以及颜色的相似性度量。

① 基金项目: 国家自然科学基金(60873039)

收稿时间: 2010-10-22; 收到修改稿时间: 2010-11-15

2 颜色空间

2.1 RGB 颜色空间

RGB 颜色空间是最常用的颜色空间。数字图像一般都采用 RGB 颜色空间来表示,其三维空间包括 RGB 三个坐标轴,如图 1 所示。原点对应黑色,离原点最远的顶点对应白色,立方体与三个坐标轴的交点对应三基色:红色、绿色和蓝色,剩余的三个顶点对应于 3 补色:品红、蓝绿和黄。

在这个模型中,从黑到白的灰度值分布在从原点到离原点最远顶点间的连线上,而立方体内其余各点对应的不同颜色可以用该点到原点的向量来表示。为了方便,可将立方体归一化为单位立方体,即 R、G、B 的值都在区间[0, 1]中。

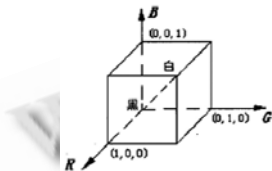


图 1 RGB 彩色立方体

2.2 HSV 颜色空间

HSV 颜色空间作为一种面向视觉感知的颜色模型,主要包括三个要素:色调、饱和度和亮度^[2]。色调是指光的颜色,与混合光谱中主要光波长相联系,如红、橙、黄分别表示不同的色调;饱和度是指彩色的深浅程度,即与一定色调的纯度相关。亮度是指人眼感受到的光的明暗程度。

HSV 颜色模型可以用一个倒置的圆锥体来表示。长轴表示亮度,离开长轴的距离表示饱和度,围绕长轴的角度表示为色调,其模型如图 2 所示。

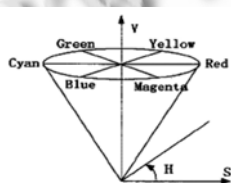


图 2 HSV 颜色空间

2.3 L*a*b 颜色空间

L*a*b 色彩空间模型是国际照明委员会(CIE)于 1976 年定义的一种近似的均匀彩色空间。它由亮度分量(L)和两个色度分量组成:即 a 分量(从绿到红)和 b

分量(从蓝到黄)。L*a*b 颜色设计与设备无关,不管使用什么设备创建或输出图像,这种颜色模型产生的颜色都保持一致。

2.4 颜色空间转换

2.4.1 RGB 颜色空间到 HSV 颜色空间的转换

设 R,G,B 颜色空间值为(R,G,B),其中 R,G,B ∈ [0,1,..., 255],由公式(1)得到的 HSV 值(H,S,V),其中 H ∈ [0,1,..., 360],S, V ∈ [0... 1]。由 RGB 颜色空间模型到 HSV 颜色空间模型的转化如下:

$$H = \begin{cases} \arccos \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B \leq G \\ 2\pi - \arccos \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B > G \end{cases}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)}$$

$$V = \frac{\max(R, G, B)}{255}$$
(1)

2.4.2 RGB 颜色空间到 L*a*b 颜色空间的转换

设 RGB 空间的色分量分别为 R、G、B,则由 RGB 颜色空间到 L*a*b 颜色空间的转换由公式(2)给出:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.430 & 0.342 & 0.178 \\ 0.300 & 0.590 & 0.110 \\ 0.020 & 0.130 & 0.939 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} = \begin{bmatrix} 0.430 & 0.342 & 0.178 \\ 0.300 & 0.590 & 0.110 \\ 0.020 & 0.130 & 0.939 \end{bmatrix} \begin{bmatrix} 255 \\ 255 \\ 255 \end{bmatrix}$$
(2)

X、Y、Z 分别代表设想的红绿蓝原色,则

$$L^* = 116(Y/Y_0)^{1/3} - 16, Y/Y_0 > 0.01$$

$$a^* = 500 \left[(X/X_0)^{1/3} - (Y/Y_0)^{1/3} \right]$$

$$b^* = 200 \left[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3} \right]$$
(3)

其中, L*为米制明度, a*、b*为米制色度, X₀、Y₀、Z₀是 CIE 标准明体照射在完全反射的漫射体上,反射到观察者眼中的白色刺激的三个刺激值。

3 颜色量化

所谓颜色量化就是将颜色空间映射到一个给定大小的子集中,使其总体误差最小。颜色量化方法有三种:均匀量化,非均匀量化和静态量化。其中非均匀量化是常用的量化方法。而 HSV 颜色空间作为适合人眼感知的颜色空间^[3],得到了广泛应用。下面以 HSV 为例进行非均匀量化:

1) 根据人的视觉分辨能力,把色调 H 空间分成 8 份,饱和度 S 空间分成 3 份,亮度 V 空间分成 3 份。

2) 根据色彩的不同范围和主观颜色感知进行量化:

$$H = \begin{cases} 0, H \in [316, 20] & 4, H \in [156, 190] \\ 1H \in [21, 40] & 5, H \in [191, 270] \\ 2H \in [41, 75] & 6, H \in [271, 295] \\ 3H \in [76, 155] & 7, H \in [296, 315] \end{cases} \quad S = \begin{cases} 0, S \in [0, 0.2] \\ 1, S \in [0.2, 0.7] \\ 2, S \in [0.7, 1] \end{cases} \quad V = \begin{cases} 0, V \in [0, 0.2] \\ 1, V \in [0.2, 0.7] \\ 2, V \in [0.7, 1] \end{cases}$$

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

$$W = HQ_S Q_V + SQ_V + V \quad (4)$$

其中, Q_S 和 Q_V 分别是 S 、 V 的量化级数, $Q_S=3$, $Q_V=3$ 。

$$\text{公式(4)实际上为: } W=9H+3S+V \quad (5)$$

这样, H, S, V 三个分量在一维矢量上分布开来。

W 的取值范围同 $[0, 1, \dots, 71]$ 。计算 W 获得 72 bin 一维直方图。

4 颜色特征提取及表示

在图像检索中, 基于颜色特征的图像检索算法主要集中在两个方面: 全局颜色特征和空间颜色特征。

4.1 全局颜色特征

4.1.1 颜色直方图

在确定颜色空间的基础上, 对一幅数字图像, 统计每一种颜色出现的象素点数, 然后把颜色值作为横坐标, 颜色出现的象素点数作为纵坐标, 以此绘出的图形就称为该图像的颜色直方图。图像的归一化颜色直方图 H 可以定义^[4]为:

$$H = \left\{ h[c_1], h[c_2], \dots, h[c_k], \dots, h[c_n] \mid \sum_{k=1}^n h[c_k] = 1, 0 \leq h[c_k] \leq 1 \right\} \quad (6)$$

$h[c_k]$ 表示第 K 种颜色在该图像中出现的象素点频数:

$$h[c_k] = \frac{\sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} \begin{cases} 1 & \text{if } I[i, j] = c_k \\ 0 & \text{otherwise} \end{cases}}{N_1 \times N_2} \quad (7)$$

其中, N_1 和 N_2 为图像的宽和高。虽然颜色直方图具有平移和旋转不变性, 也得到了广泛的应用。但是, 它丢失了颜色的空间分布信息, 有可能造成不同的图像有相同的颜色直方图, 从而降低检索效率。若颜色集中的某些颜色, 图像中可能没有该颜色, 颜色直方图中就会出现若干零值, 这样会造成稀疏颜色直方图, 而稀疏直方图对噪声是很敏感的。

4.1.2 累积直方图

在全局直方图的基础上, Stricker 和 Orengo 进一

步提出了使用累积颜色直方图^[5]的概念。累积直方图是以颜色值作为横坐标, 以颜色累积出现的频数为纵坐标。因此累积直方图的定义如下:

$$H(K) = \frac{\sum_{i=1}^k n_i}{N} \quad k = 0, 1, \dots, L-1 \quad (8)$$

其中, k 是图像的特征取值, L 是特征可取值的个数, n_i 是图像中特征值为 k 的像素的个数, N 是像素的总数。在累积直方图中, 相邻颜色在频数上是相关的。虽然累积直方图的存储量和计算量有很小的增加, 但是累积直方图消除了直方图中常见的零值问题, 经过累积处理, 可以保持两种颜色在特征轴上的距离与它们之间匹配的相似程度成正比关系。

4.1.3 颜色矩

颜色矩^[6]是一种简单有效的颜色特征表示方法。分为一阶矩(均值), 二阶矩(方差)和三阶矩(斜度)。三个颜色矩的数学定义如下:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij} \quad \sigma_i = \sqrt{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2} \quad \xi_i = \sqrt[3]{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3} \quad (9)$$

其中, f_{ij} 表示像素 j 的颜色值为 i 的概率, N 为图像中像素点的个数。若 Q, I 表示两幅图像, 则其矩的距离公式为:

$$D[Q, I] = \sum_{i=1}^3 \left(\omega_{i1} |\mu_i^Q - \mu_i^I| + \omega_{i2} |\sigma_i^Q - \sigma_i^I| + \omega_{i3} |\xi_i^Q - \xi_i^I| \right) \quad (10)$$

其中, $\omega_{ij} \geq 0 (1 \leq i, j \leq 3)$ 是用户指定的加权系数, i 表示颜色通道。公式中的加权系数值可以根据具体的应用调整。

4.1.4 颜色熵

根据颜色直方图特性和信息论中信息熵的概念, John 提出采用图像颜色的信息熵来表示图像的颜色特征, 从而将图像的颜色直方图由多维降低到一维。设含有 N 种颜色值的数字图像 I , 其归一化颜色直方图可表示为 (P_1, P_2, \dots, P_N) , P_i 为第 i 种颜色值在图像中出现的概率。将图像的颜色直方图看作是图像中不同颜色的像素在图像空间中出现的概率密度函数, 且各颜色值概率相互独立, $\sum_{i=1}^N P_i = 1, P_i \geq 0$ 。图像的信息熵可表示为:

$$H(T) = - \sum_{i=1}^N P_i \log_2 P_i \quad (11)$$

其中, P_i 表示第 i 种颜色值像素在整幅图像中所占的百分比, 即出现的频率。虽然采用熵的方法有效的降低了图像直方图特征的维数, 但在利用熵进行图像检索时, 其分辨能力是较低的。

4.2 空间颜色特征

4.2.1 颜色聚合向量

颜色聚合矢量是图像直方图的一种演变，其核心思想是当图像颜色相似的像素所占据的连续区域的面积大于一定的阈值时，该区域中的像素为聚合像素，否则为非聚合像素，这样统计图像所包含的每种颜色的聚合像素和非聚合像素的比率称为该图像的颜色聚合矢量，在图像检索过程中匹配目标图像的颜色聚合矢量和检索图像的颜色聚合矢量，聚合矢量中的聚合信息在某种程度上保留了图像颜色的空间信息。如令定义为第 i 个聚合像素，为第 j 个非聚合像素，那么颜色聚合矢量定义：

$$\langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_N, \beta_N) \rangle \quad (12)$$

则 $\langle (\alpha_1 + \beta_1), (\alpha_2 + \beta_2), \dots, (\alpha_N + \beta_N) \rangle$ 为图像的颜色直方图。由于加入了空间信息，采用颜色聚合矢量比采用颜色直方图检索的效果要好，特别是对大块的均匀区域和大部分为纹理的图像检索效果更好，但是增大了计算量。

4.2.2 颜色相关图

颜色相关图是指对于颜色值为 C_i 的像素点 P_i ，与像素点 P_i 相距距离为 k 的另一个像素点 P_2 的颜色为 C_j 的概率的大小。其数学定义为：设一个距离 $d \in R$ ， $(i, j) \in \{1, 2, \dots, N\}$ ，并且 $k \in [d]$ ，定义图像 I 的颜色相关图为：

$$\gamma_{C_i, C_j}^{(k)}(I) = \text{Pr}[p_2 \in I_{C_j}, |p_1 - p_2| = k, p_1 \in I_{C_i}] \quad (13)$$

由于颜色相关图的计算量很大 ($O(N^2 d)$)，为了减小计算量，一般只考虑相同颜色之间的相关性，即颜色自相关图为 $\alpha_c^{(k)} = \gamma_{c,c}^{(k)}(I)$ ；相对于颜色直方图和颜色聚合矢量，颜色相关图大大提高了检索效率，但是由于高维数而使得计算量同样很大。

5 颜色特征的相似性度量

5.1 直方图相交距离

设图像 Q 与 I 在某色彩通道 C 上的颜色直方图分别为 $Q_c(i)$ 和 $I_c(i)$ ，其中 $i \in [1, n]$ ， n 为亮度的水平数，则在通道 C 上的直方图距离 $D_c(Q, I)$ 定义为：

$$D_c(Q, I) = \sum \min \left\{ \bar{Q}_c(i), \bar{I}_c(i) \right\} \quad (14)$$

式中， $\bar{Q}_c(i) = \frac{Q_c(i)}{\sum_{i=1}^n Q_c(i)}$ ， $\bar{I}_c(i) = \frac{I_c(i)}{\sum_{i=1}^n I_c(i)}$

综合 R,G,B 三个通道，则图像 Q 和 I 之间的距离为：

$$D(Q, I) = \sum_{C=R,G,B} \left(\sum_{i=1}^n \min \left\{ \bar{Q}_c(i), \bar{I}_c(i) \right\} \right) \quad (15)$$

5.2 Minkowsky 距离

Minkowsky 距离是基于 L_p 范式定义的：

$$D(x, y) = \left(\sum_{i=1}^N |x_i - y_i|^n \right)^{1/n} \quad (16)$$

当 $n=1$ 时， $D(x, y) = \sum_{i=1}^N |x_i - y_i|$ 称为绝对值距离。

当 $n=2$ 时， $D(x, y) = \sqrt{\sum_{i=1}^N |x_i - y_i|^2}$ 称为欧氏距离。

其中， $D(x, y)$ 表示向量 x, y 之间的距离， N 表示向量的维数， x_i 表示向量 x 的第 i 维分量， y_i 表示向量 y 的第 i 维分量， n 为一常数。如果图像特征的分量之间是正交无关，且各维度的重要程度相同，则两者之间的距离可用欧式距离。

5.3 二次式距离

对于基于颜色直方图的图像检索来说，二次式距离比欧氏距离及直方图相交法更为有效，因为这种距离考虑到不同颜色之间存在的相似度。二次式距离可以表示为：

$$d_{quad}(A, B) = \sqrt{(A - B)^T M (A - B)} \quad (17)$$

其中， A 和 B 表示图像的两个颜色直方图， $M = [m_{ij}]$ ， m_{ij} 表示直方图中下标为 i 和 j 颜色之间的相似度。这种方法通过引入颜色相似性矩阵 M ，使其能够考虑到相似但不相同的颜色间的相似性因素。

6 基于分块颜色直方图算法的检索过程

6.1 颜色分块

基于分块的颜色直方图算法采用固定分块的图像划分方法，人为地将图像划分为适当的子块，提取相应的局部特征。根据人对图像的不同理解，赋予不同分块不同的权值。在本例中将图像分成 4×4 个子块，根据人对不同图像的理解，对不同的子块赋予不同的权值。

6.2 图像相似度度量

相似度计算采用本文提到的欧氏距离计算方法。假设示例图像的直方图特征矢量为 (x_0, x_1, \dots, x_n) ，图像数据库中的图像直方图特征矢量为 (y_0, y_1, \dots, y_n) ，然后根据直方图欧氏距离算法公式 $D(x, y) = \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$ 进行图像相似性度量。

6.3 实验流程

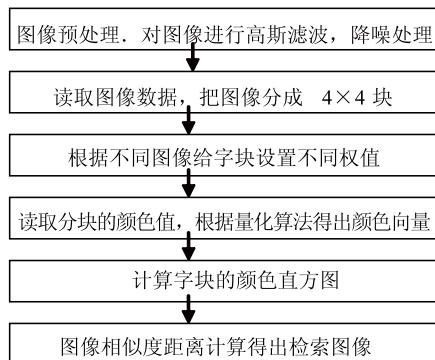


图3 图像检索算法流程图

7 总结

本文以图像检索中的颜色特征为角度,介绍了颜色模型的种类、颜色的量化、颜色特征的提取以及度量方法,扫描了算法检索流程。从中可知图像的颜色特征提取与匹配技术已经相对成熟,但依然存在一些问题,如颜色特征作为图像底层特征如何符合高层语义以及如何提高检索结果的用户主观满意度等,在以后的研究工作中,将重点研究和解决这些问题。通过

本文的系统讨论和研究可以预见基于颜色特征的图像检索技术将会有着更广阔的发展前景。

参考文献

- 1 Aslandogan YA, Yu CT. Techniques and systems for image and video retrieval. *IEEE Trans. on Knowledge and Data Engineering*, 1999,11(1):56-60.
- 2 章毓晋.图像工程(上册:图像处理与分析).北京:清华大学出版社,1999.
- 3 Malljunath BS, Ohm JR, VasudenVan VV, et al. Color and texture descriptors. *IEEE Trans. on CSVT*, 2001,11(6):703-715.
- 4 薛向阳,罗航哉.一种新的颜色相似度定义及其计算方法. *计算机学报*,1999,22(9):918-922.
- 5 Stricker M, Orengo M. Similarity of color images. *IS&T/SPIEC for Storage and Retrieval for Image and Video Databases III*, 1995.
- 6 Stricker M, Orengo M. Similarity of color images. *SPm Storage and Retrieval for Image and Video Databases III*, 21 85:381-392.

(上接第34页)

- 2001.
- 2 3GPP TS23.207. End-to-End QoS Concept and Architecture, Version 9.0.0, Release 9, 2009.
- 3 Calhoun P, Loughney J, Guttman E, Zorn GJ. Arkko. Diameter Base Protocol. Internet Engineering Task Force, Request for Comments(Informational) RFC 3588, September 2003.
- 4 邱锡鹏,刘海鹏. Diameter 协议研究. *计算机科学*, 2003,30(2): 75-78.
- 5 Rosenberg J. The Extensible Markup Language (XML)

- Configuration Access Protocol (XCAP). Internet Engineering Task Force, Request for Comments (Informational) RFC 4825, May 2007.
- 6 Ozianyi VG, Good R, Carrilho N, Ventura N. XML-Driven Framework for Policy-Based QoS Management of IMS Networks. *IEEE Global Telecommu-nications Conference*, doi:10.1109/GLOCOM. 2008.327.
- 7 3GPP TS29.213. Policy and charging control signaling flows and QoS, Version 7.10.0, Release 7, 2010.