

基于颜色直方图的图像检索技术

窦建军, 文俊, 刘重庆

(上海交通大学 图像处理与模式识别研究所, 上海 200030)

摘要: 讨论了基于颜色直方图的图像检索中颜色空间和直方图距离的选择问题, 提出了利用颜色直方图来检索图像的一般框架, 建立了基于两种颜色空间 (RGB 和 HSV) 和四种直方图距离共八种图像检索算法, 并且采用了两种客观评价标准对所建立算法的有效性进行评价。结果证明, 在基于颜色直方图的应用中, HSV 空间比 RGB 空间更有效, EMD 距离在所有四种距离量度中性能较好, 但计算量稍大, 与 EMD 距离相比, 虽交集距离性能稍有不足, 但所需计算量较小。

关键词: 图像检索; 直方图; 颜色空间; 直方图距离; EMD

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1007-2276(2005)01-0084-04

Histogram-based color image retrieval

DOU Jian-jun, WEN Jun, LIU Chong-qing

(Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: The problem of color space and histogram distance in histogram-based image retrieval is discussed, and the common frame of histogram-based image retrieval is proposed. Then eight kinds of image retrieval algorithms based on two color spaces and four histogram distance algorithms are created and also two objective metrics for evaluating their effectiveness are adopted. The results show that HSV color space is more effective than RGB color space in common. For histogram distance, EMD has showed better performance than other distance, but it has a relative greater computational burden. Comparing with EMD, histogram intersection distance needs less computational burden and only has slightly weak performance.

Key words: Image retrieval; Histogram; Color space; Histogram distance; EMD

0 引言

在过去的 10 年里, 已经开发了很多图像检索系统, 如 IBM Almaden 研究中心的 IBM QBIC 系统、Virage 公司开发的 VIRAGE 系统、MIT 媒体实验室开发的 Photobook 系统和加州伯克利分校的 Blobworld

系统等^[1]。所有这些系统的潜在原理都是利用图像的像素值生成图像的信号, 进而采用一定的规则对所生成的图像信号进行比较, 信号可以由图像的颜色、纹理、图像中目标的形状以及其他可以用来在图像间进行比较的特征表示。其中, 颜色作为目标的一种重要特征, 常常和目标的身份有密切的关系, 如天空是蓝

收稿日期: 2004-02-08; 修订日期: 2004-06-23

作者简介: 窦建军 (1978-), 男, 山东博兴人, 硕士, 主要从事模式识别、图像检索方面的研究。

色,树木和森林是绿色等等,同时,颜色是日常人们认识一个物体的重要信息来源,因此,利用颜色信息进行图像检索^[1]也就成了必然的选择。然而,就认识一个物体的完整性来看,单纯的颜色信息是远远不够的,还需要利用纹理以及形状等其他信息,很多已开发的检索技术往往都是结合几种特征来实现的。但是,与其他方法相比,颜色检索仍是一种计算既简单又快速的方法,对那些有实时性要求的系统尤为重要。

利用图像的颜色检索是指给定一个查询图像,需要从图像库中找到与之颜色分布类似的图像,通常,这种颜色分布由一定颜色空间的图像颜色直方图表示,目前存在的颜色空间主要有 RGB、HSV、HSL、YUV 以及 CIE-LAB 等,其中 HSV 颜色模型与人的主观感觉最接近。为了对两幅图像的颜色分布即颜色直方图的相似性进行度量,经常采用直方图距离来定义颜色直方图之间的相似性,其中包括欧氏距离、交集距离、二次方距离以及最近提出的 EMD (Earth Move Distance) 等。本文详细讨论了几种距离公式在表示颜色相似性时的特点,并且利用不同的颜色模型来构造图像的颜色直方图,根据以上几种距离公式提出了利用颜色直方图进行图像检索的一般框架,最后采用两个客观评价标准对文中讨论的九种检索方案进行了比较。

1 颜色空间

颜色空间可以定义为一种用亮度值来表示颜色的模型^[1],一般是一个三维空间,空间每一维代表一个颜色分量或一个颜色通道。常见的颜色模型有 RGB、HSV、HSL、YUV 以及 CIE-LAB 等,其中 RGB 是在早期图像检索系统中比较常用的颜色模型,通常用于彩色阴极射线管和彩色光栅图形显示器中,由红、绿、蓝三基色组成,这三个基色是一种“加性原色”,——各个原色的光能叠加在一起产生复合色。RGB 颜色模型通常可以用一个单位立方体表示,其最大缺点是与人的眼对颜色的感觉并不相符。其他颜色空间如 CIE-LAB 和 CIE-LUV 等比较符合人的感觉,此类模式把颜色空间分成色度、亮度和饱和度,这种模式的优点是:独立地对颜色通道进行处理时不会引入错误的颜色信号(hue)^[2],而且比较容易补偿颜色的扭曲及其他人为误差。但是由于在 RGB 到 CIE-LAB

及 CIE-LUV 的正反变换中涉及到非线性操作,显然是极其不方便的。HSV 模型对应于画家配色的方法,画家通过改变色浓和色深从某种纯色中获得不同色调的颜色。此颜色空间与人眼对颜色的视觉特性紧密相关,这种视觉特性使得图像处理算法建立在生理学基础上。HSV 颜色模式具有 CIE-LAB 颜色模式的优点和缺点,但是在与 RGB 进行逆变换时却比较容易实现,本文采用参考文献[2]中建议的方法,把 RGB 空间上的向量 (r,g,b) 转化为 HSV 空间上的向量 (h,s,v) ,其中 $r,g,b \in [0,1]; s,v \in [0,1]; h \in [0,6]$ 。

$$v = \max\{r,g,b\} \quad s = \frac{v - \min\{r,g,b\}}{v}$$

$$h = \begin{cases} 5+\bar{b} & \text{如果 } r = \max\{r,g,b\} \text{ 和 } g = \min\{r,g,b\} \quad \bar{r} = \frac{v-r}{v-\min\{r,g,b\}} \\ 1-\bar{g} & \text{如果 } r = \max\{r,g,b\} \text{ 和 } g \neq \min\{r,g,b\} \\ 1+\bar{r} & \text{如果 } g = \max\{r,g,b\} \text{ 和 } b = \min\{r,g,b\} \quad \bar{g} = \frac{v-g}{v-\min\{r,g,b\}} \\ 3-\bar{b} & \text{如果 } g = \max\{r,g,b\} \text{ 和 } b \neq \min\{r,g,b\} \\ 3+\bar{g} & \text{如果 } b = \max\{r,g,b\} \text{ 和 } r = \min\{r,g,b\} \quad \bar{b} = \frac{v-b}{v-\min\{r,g,b\}} \\ 5-\bar{r} & \text{其他} \end{cases} \quad (1)$$

2 颜色直方图的定义和量化

一幅图像的灰度直方图是指图像亮度的概率密度函数,从而颜色直方图可以定义为图像中三个颜色通道的联合概率密度函数:

$$h_{A,B,C}(a,b,c) = N \cdot P(A=a, B=b, C=c) \quad (2)$$

式中 A, B, C 表示图像的颜色通道 (R,G,B 或 H,S,V); N 表示图像的大小,即整幅图像包含的像素数目; P 表示概率密度函数; h 为直方图函数,可以用一个四维向量表示,前三维分别为三个颜色通道,最后一维表示该颜色在整幅图像中所占的比例。显然,不同于灰度直方图,颜色直方图是很难用图形来描述的,一般只用二维直方图图形来图示其三个分量。

当利用颜色直方图进行图像检索时,一个重要的问题就是颜色空间的选择和直方图的量化。由于颜色直方图是高维向量数组,就 RGB 空间而言,直接表示需要 $256 \times 256 \times 256 = 16777216$ 个向量,显然,如果再考

考虑直方图上的相关运算,则需要极大的计算量,因此,一般的解决方法是对直方图进行量化来减小其所包含的向量个数。一个合适的量化级要求能够在性能和计算量上取得平衡:一方面,所选择的量化级必须保证差别较大的颜色不能被分在同一量化级,从而要求量化级不能过大;另一方面,过大容量的颜色直方图必然导致过大的计算量,特别是给后面进行的直方图比较操作带来繁重的计算负担。最后,量化还要求根据所选择的颜色空间三个通道是否一致来决定各个通道的量化级是否一致。

3 直方图相似性的比较量度

为了能够反映两幅图像之间的相似程度,必须有一个表示图像之间相似性的客观度量。对于基于直方图的图像检索来说,经常利用直方图之间的距离公式来表示相似性,显然,距离和相似性具有相反的属性,距离越小,相似性越大。表示直方图之间相似性的距离公式有很多,一般来说,那些用来比较直方图之间的概率分布的距离并不适合于图像检索,因为图像检索是基于视觉感觉而不是概率分布的相似性。目前存在的直方图距离主要有欧拉距离、交集距离、二次方距离以及 EMD 等。

3.1 欧拉距离

设 h 和 g 分别表示两幅图像的颜色直方图,则 h 和 g 之间的欧拉距离表示为:

$$d^2(h, g) = \sum_A \sum_B \sum_C (h(a, b, c) - g(a, b, c))^2 \quad (3)$$

此公式的特点是:仅仅在两个直方图中的同一个量化级间进行比较,而没有针对处在不同量化级感觉上却相似的直方图分量进行交叉比较,所有量化级的直方图分量在整个距离公式中有相等的贡献。另外,欧拉距离有一种变化形式,公式如下:

$$d^2(h, g) = \sum_A \sum_B \sum_C \frac{(h(a, b, c) - g(a, b, c))^2}{h(a, b, c) + g(a, b, c)} \quad (4)$$

此变形的距离公式对不同的量化级分别加上不同的权值,显然比较符合人眼观察场景时的非精确性。

3.2 交集距离

交集距离最早由 M. J. Swain^[3]引入图像检索中,

两个直方图 h 和 g 之间的交集距离定义如下:

$$d(h, g) = 1 - \frac{\sum_A \sum_B \sum_C \min(h(a, b, c), g(a, b, c))}{\min(|h|, |g|)} \quad (5)$$

式中 $|h|$ 和 $|g|$ 分别为两个直方图的幅度,即图像中像素的总数。此公式的特点是:用户查询图像中不存在的颜色对交集距离总和的贡献为零,这必然降低了背景颜色的贡献,同样,与欧拉距离类似,交集距离也没有考虑不同量化级上感觉却相似的直方图分量。

3.3 二次方距离

二次方距离又称为交叉距离,顾名思义,它把颜色之间的相关性结合到距离公式中,其最早应用在 QBIC 系统中^[4]。公式定义如下:

$$d(h, g) = (h - g)^T \mathbf{A} (h, g) \quad (6)$$

式中 \mathbf{A} 称为相似矩阵,由直方图各个分量间的交叉相关值组成。设 a_{ij} 为相似矩阵 \mathbf{A} 的第 (i, j) 分量,对于 RGB 空间:

$$a_{ij} = 1 - d_{ij} / \max(d_{ij}) \quad (7)$$

式中 d_{ij} 表示 RGB 空间中颜色 i 和 j 的 L_2 距离(欧拉距离)。对于 HSV 空间, a_{ij} 可按下式得到^[5]:

$$a_{ij} = 1 - \frac{1}{\sqrt{5}} \left[(v_i - v_j)^2 + (s_i \cosh_i - s_j \cosh_j)^2 + (s_i \cosh_i - s_j \cosh_j)^2 \right]^{\frac{1}{2}} \quad (8)$$

此公式相应于 HSV 颜色空间的一个近似。

3.4 EMD 距离

EMD 距离最早是来自著名的运输问题的求解^[6], 后由 Rubner^[7]把 EMD 距离引入了图像检索中,它可以求解某个特征空间中两个多维分布之间的不相似性,通常单个特征之间的距离是给定的,并称此距离为地面距离。

给定两个直方图分布,一个分布 h 可以看成是一堆土的集合,另一个分布 g 可以看成是很多洞的集合,不同的土与洞之间的地面距离是不同的,从而 EMD 可以计算出用土来填满所有洞的最小工作量或最小代价,此处单位工作量定义为运输单位土时经过单位地面距离的工作量,这样 EMD 可以定义为:

$$EMD(h,g)=\frac{\min imum\ cost}{\min\{area(h),area(g)\}} \quad (9)$$

式中 *area* 表示直方图的面积,也是与该直方图相关的图像包含的像素数目。当然,为了求解 EMD,最重要的步骤就是求解最小代价,本文采用如下算法:

设 *h* 和 *g* 是两个直方图,它们的每项分别为 *h*(*i*)、*g*(*i*),且 *i*=1,⋯,*n*。设 *c*_{*ij*}=|*i*-*j*|,表示地面距离,*x*_{*ij*}表示 *h*(*i*)转换成 *g*(*i*)的数量,则需要最小化 $\sum_i \sum_j c_{ij}x_{ij}$:

(1) 求解 $w_i = \max_j c_{ij}, y_j = \max_i c_{ij}$;

(2) 找到一个 *i, j*,使得 $w_i + y_j - c_{ij} = \max_{(k,l)} [(w_k + y_l - c_{kl}) > 0]$;

(3) $x_{ij} = \min\{h(i), g(j)\}$;

(4) 从 *h*(*i*),*g*(*j*) 中减去 *x*_{*ij*},除去 *h* 和 *g* 中等于零的项,如果两个直方图中所有项均等于零,则退出;否则,回到步骤(1)。

4 直方图颜色检索的实现及实验结果

为了评价图像检索算法的有效性,需要一定的指标来进行衡量,由于图像之间的相似性非常依赖于人的主观感觉,一个完全符合人主观感觉的评价标准还有待于研究,本文采用了目前比较流行的 Recall 和 Precision 两个指标。当对一幅图像进行查询时,Recall 指检索到的相关图像在所有与查询图像相关图像中所占的比率,Precision 指检索到的相关图像在所有检索到的图像中所占的比率,如图 1 所示。

$$\begin{aligned} \text{Recall} &= P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{a}{a+c} \\ \text{Precision} &= P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{a}{a+b} \end{aligned} \quad (10)$$

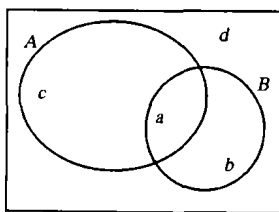


图 1 与检索效率有关的集合图示

Fig.1 Diagram about retrieval efficiency

图 1 中 *A* 为与查询图像相关的图像集合;*B* 为检索到的图像集合;*a* 为检索到的相关图像子集;*b* 为检索到的不相关图像子集;*c* 为未检索到的相关图像子集;*d* 为未检索到的不相关图像子集;Recall 和 Precision 可以用条件概率密度来表示。

4.1 算法实现

为了进一步对本文所考虑的两个颜色空间和四种直方图距离进行分析,利用 visual C++ 建立了一个图像检索的 Demon 系统,实现了共八种颜色图像检索算法,并且利用两个指标来比较八种算法的结果,文中采用的图像库来自 SIMPLcity,为了简单起见,只采用了前 500 张图片。

一般而言,对于一个图像检索系统,都需要经过颜色空间变换和量化。对于基于 RGB 颜色空间的应用,也就不需要颜色变换,由于 RGB 空间中三个分量是等距离的,需要对三个分量采用相同的量化级,此处,采用的量化级是 8,从而把 256×256×256 个颜色量化成 8×8×8=512 个,显然这只是在性能和计算量上取平衡。对于 HSV 颜色空间的应用,首先把 RGB 转换为 HSV 空间,与 RGB 颜色空间不同的是,色度分量是包含 R、G、B 三个颜色的一个圆圈,每个基色占 120°空间,一般把色度分量量化成 18 级时,足够可以分开其中的红、绿、蓝、青、黄等颜色,亮度分量和饱和度分量量化成 3 级时就可符合人的感觉,从而将 HSV 空间直方图量化成 18×3×3。

4.2 结果比较

图 2~图 5 所示的曲线分别是利用八种方法对所

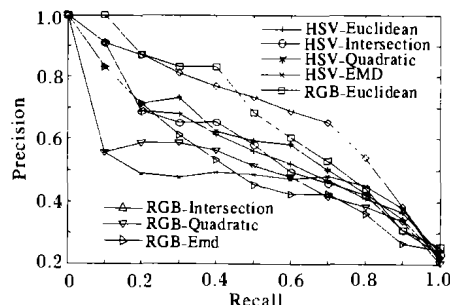


图 2 直方图检索效率(Africa)

Fig.2 Histogram retrieval efficiency(Africa)

用图片库中四个情景进行检索得到的 Precision-Recall 相对图。其中,八种方法分别是基于 HSV 空间的四种距离应用:HSV-Euclidean、HSV-Intersection、HSV-Quadratic、HSV-EMD 以及基于 RGB 空间的四种距离应用:

RGB -Euclidean、RGB -Intersection、RGB -Quadratic、RGB -EMD, 四个场景的内容分别是非洲(Africa)、海滩(Beach)、建筑(Building)以及公交车(Bus), 为了保证结果的相对精确性, 在每个场景中任意抽取十幅查询

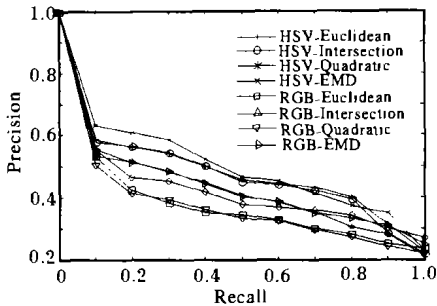


图 3 直方图检索效率(Beach)

Fig.3 Histogram retrieval efficiency (Beach)

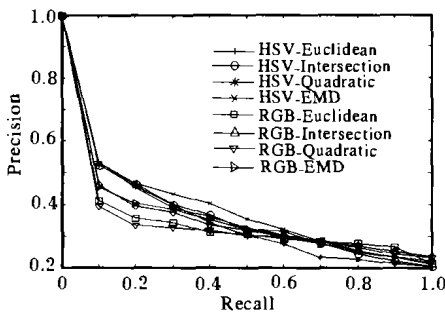


图 4 直方图检索效率(Bus)

Fig.4 Histogram retrieval efficiency(Bus)

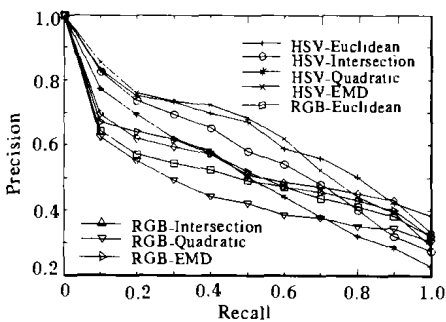


图 5 直方图检索效率(Building)

Fig.5 Histogram retrieval efficiency(Building)

图像, 对所有查询图像检索得到的结果进行平均, 平均值作为最终数据绘制在上面的四幅图中。通过四个场景的 Precision-Recall 相对图可以发现: 除 Africa 场景外, 在相等的 Recall 值之处, 基于 HSV 空间的算法有较高的 Precision 值, 显然, HSV 空间的检索方法优于基于 RGB 空间的方法。对于不同的直方图距离, 图中反映了 EMD 距离在其中三个场景中有着较好的性能, 但是 EMD 距离的求解相对比较复杂, 一般可以采

用一定的快速算法来解决计算量问题。欧拉距离与交集距离的性能比 EMD 稍差, 但是在计算量上, 即使 EMD 距离采用了快速算法, 欧拉距离和交集距离仍然占有优势, 特别是, 交集距离的运算速度比较快。最后一种距离是二次方距离, 其中计算相似矩阵需要相当大的计算量, 但是算法的性能却并没有很大的改善。

5 结束语

本文在两个颜色空间中采用了四种直方图距离建立了八种不同的颜色检索算法, 测试结果证明: HSV 空间比 RGB 空间更有效, 四个直方图距离中 EMD 是比较有效的一个距离量度, 但计算量较大, 交集距离虽然性能上稍差, 但其计算量在四个距离中最小, 因此, 交集距离是较令人满意的直方图距离。从测试结果还可看出, 全局直方图有其固有缺点, 它不能包含形状、位置及纹理等信息, 从而往往检索到颜色组成类似, 内容却截然不同的图像, 因此可以进一步把这类其他信息包含进本文算法中, 从而可以得到效果更好的检索结果。

参考文献:

- [1] James Z Wang. Integrated Region-Based Image Retrieval[M]. Boston: Kluwer Academic Publishers, 2001.
- [2] Smith J R, Chang S F. Tools and techniques for color image retrieval[A]. In: Sethi I K, Jain, R C, eds. Storage & Retrieval for Image and Video Databases IV, IS&T/SPIE Proceedings[C]. 1996, 2670. 426-437.
- [3] Swain M J, Ballard D H. Color indexing [J]. In International Journal of Computer Vision, 1991, 7(1): 11-13.
- [4] James Hafner, Harpreet S Sawhney, Will Equits. Efficient color histogram indexing for quadratic form distance functions[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1995, 17(7): 729-736.
- [5] Smith J R, Chang S F. VisualSEEK: a fully automated content-based image query system[A]. ACM Multimedia[C]. Boston MA: MCM Press, 1996. 87-98.
- [6] Hitchcock F L. The distribution of a product from several sources to numerous localities[J]. J Math Phys, 1941, 20: 224-230.
- [7] Rubner Y, Tomasi C, Guibas L J. A metric for distributions with applications to image databases[A]. IEEE International Conference on Computer Vision[C]. 1998. 59-66.