第 25 卷第 12 期 电子与信息学报 Vol.25No.12
2003 年 12 月 Journal of Electronics & Information Technology Dec. 2003

基于特征元素的图像检索方法 [1]

李 勍 章毓晋
(清华大学电子工程系 北京 100084)

摘 要: 该文提出了一种基于特征元素的新的图像检索算法. 特征元素与特征向量相比更注重根据人的主观感知来表达图像的视觉特征. 在特征元素的基础上, 该文先定义了图像间的相似性度量, 即特征元素间的距离, 又分别实现了对不同特征元素类别的距离计算. 检索实验表明, 基于特征元素的图像检索算法能够取得更符合人们视觉感知的结果.

关键词: 图像检索, 特征元素, 特征向量, 图像相似度
中图分类号: TN911.73 **文献标识码:** A **文章编号:** 1009-5896(2003)12-1591-07

Image Retrieval Based on Feature Element

Li Qing Zhang Yu-jin
(*Department of Electronic Engineering, Tsinghua University, Beijing* 100084, *China*)

Abstract In this paper, a novel image retrieval algorithm based on feature element is proposed. Compared with feature vectors, feature elements tend to represent visual meanings of the image according to subjective perception of human beings. Based on feature elements, the similarity between images is defined which is measured by the distant between feature elements. Retrieval experiments show that image retrieval based on feature elements can retrieve images which are more accordant with visual perception of human beings.

Key words Image retrieval, Feature element, Feature vector, Image similarity

1 引言

近年来随着网络应用的普及和存储能力的增长, 图像已经成为 WWW 上极为普遍的信息载体. 越来越多的人希望能够在 Internet 上检索需要的图像. 一些 WWW 上著名的文字搜索引擎如 Alta Vista 等已经开始提供图像的检索服务, 但大部分图像搜索引擎仍是基于关键字的检索, 前提是描述图像的文本已经存在. 虽然用文字对图像归类对于 WWW 上的图像来说很常见, 但是由于基于文本的搜索引擎在图像信息的索引和检索方面的能力非常有限, 对于更加有效的图像搜索引擎的研究显得日益重要.

基于内容的图像检索 (CBIR) 研究近年来一直十分活跃, 并出现了一些 WWW 上运用该技术的图像搜索引擎 [1,2] . CBIR 的目的是为了弥补传统的基于文本的图像检索技术的不足, 通过提取图像数据中的各种视觉特征作为图像内容的索引, 为用户提供基于这些视觉特征相似性的内容检索. CBIR 的主要思想是从图像中提取出诸如颜色、纹理、形状等视觉特征并将这些特征用向量表示, 图像之间的相似度就定义为特征向量之间的距离. 与基于文本的图像检索相比, 它能够自动提取特征向量, 并且在一定程度上反映了图像的视觉内容. 现在 CBIR 研究的主流趋向于基于语义的检索, 希望通过各种特征的结合来获取人对图像的语义认知, 虽然取得了一定的进步, 但结果都不令人满意. CBIR 是建立在特征向量的基础上, 而在特征向量的基

[1] 2002-07-01 收到, 2002-12-13 改回
国家自然科学基金 (60172025) 资助项目

础上提取图像的语义十分困难，因为特征向量并不能真正地反映人对图像的理解．例如当一个人看一幅彩色图像时，很难意识到该图像的颜色直方图是什么样的，而更可能关心的是这幅图像包含了几个特定的颜色．视觉心理学和神经科学的最新实验结果更倾向于这样一种假设：人们对于图像内容的视觉认知是离散的而不是连续的．特征元素的理论就是在这个基础上提出来的 [3]．与向量空间中连续的特征向量不同，特征元素采用了离散的形式．特征元素比特征向量更为直观，并具有一定的视觉意义，它可以看作是处于底层特征向量与高层语义之间的一个中间层．

本文以 WWW 上的图像搜索引擎开发为背景，对其中的图像检索技术进行了研究，提出了一种基于特征元素的图像检索算法，取得了较好的检索效果．下面第 2 节介绍了特征元素的提取方法，第 3 节对基于特征元素的图像检索算法进行了详细说明，第 4 节列举了一些实验的结果，最后是结论以及对未来工作的展望．

2 特征元素的提取

图像特征的提取是图像检索算法的基础．本文所提取的视觉特征不是用一般 CBIR 系统中所采用的特征向量来表达，而是用特征元素来表达 [4]．本文采用了两种特征元素：颜色特征元素和形状特征元素．其中颜色特征元素的提取已在文献 [5] 中进行了详细介绍．

本文所采用的形状特征提取方法基于小波模极大值和不变矩 [6]．对图像进行了 6 层小波分解，并对每个尺度的边界图像都计算 7 个不变矩，就得到了 42 维的形状特征向量，由 6 层，每层 7 个不变矩构成．特征向量提取完成后，将特征向量聚类，将对应的聚类区作为特征元素的组件．具体将整个特征向量分裂为 6 个子向量，每个 7 维，表示一幅边界图像的 7 个不变矩．另外再将整个特征向量分裂为 7 个子向量，每个 6 维，表示 6 幅边界图像的同一个不变矩．形状特征向量的分裂过程如图 1 所示．

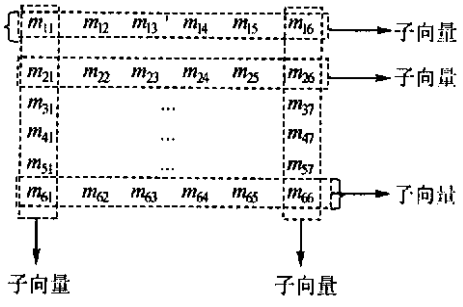


图 1 形状特征向量分裂过程

每个特征向量分裂结束后均得到 13 个子向量，加上整个特征向量，将所有的向量都进行无监督聚类，我们使用 K-均值聚类算法．聚类完成后，每个向量都在各自的向量空间对应一个聚类区．如图 2 所示，将这些对应的聚类区作为形状特征元素的组件．这样形状特征元素总共包含 14 个组件．

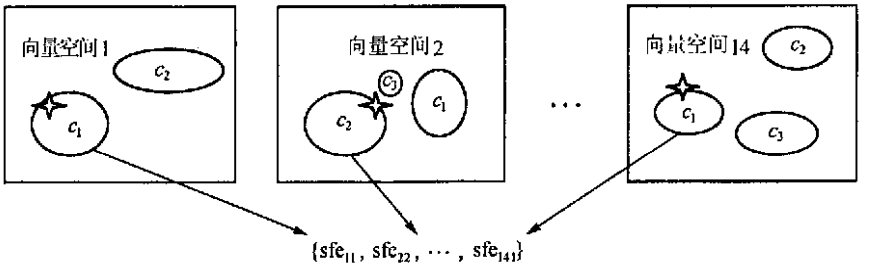


图 2　形状特征元素构建示意图
(* 为向量，c 为聚类区)

3 基于特征元素的图像检索

从检索的角度来说，基于特征元素的图像检索与基于特征向量的图像检索的出发点是相同的，都以特征元素之间的距离作为度量相似性的标准．因此在特征元素提取的基础上，接下来的问题就是如何定义特征元素之间的距离．对于第 2 节中所述的两种特征元素，本文使用了不同的方法定义它们之间的距离．首先对于形状特征元素，采用了非常简单的距离定义方法，即将两个形状特征元素包含的不同的组件个数作为它们之间的距离，如下式：

$$\mathrm{dis}(\mathrm{SFE}_q, \mathrm{SFE}_r) = \sum_{i=1}^{14} d(\mathrm{sfe}_{qi}, \mathrm{sfe}_{ri}) \tag{1}$$

其中

$$d(\mathrm{sfe}_{qi}, \mathrm{sfe}_{ri}) = \begin{cases} 1, & \mathrm{sfe}_{qi} \neq \mathrm{sfe}_{ri} \\ 0, & \mathrm{sfe}_{qi} = \mathrm{sfe}_{ri} \end{cases} \tag{2}$$

如在文献 [5] 中所述，颜色特征元素的组件为图像颜色分割后得到的颜色区域的参数，这些参数反映了颜色区域的特性．为了计算颜色特征元素之间的距离，首先需要构成这样一些颜色区域对 $(\mathrm{CR}_{qi}, \mathrm{CR}_{rj})$，其中 CR_{qi} 为查询图像的第 i 个颜色区域，CR_{rj} 为目标图像的第 j 个颜色区域．颜色区域对的构成按照以下原则：

(1) 可以构成颜色区域对的颜色区域其几何大小应大于阈值，非常小的颜色区域对人的视觉影响很小，可以忽略不计；

(2) 颜色区域对的两个颜色区域的色度差至少对于其中一个颜色区域是最小的，即应满足以下两个条件之一：

$$|\mathrm{Hue}_{qi} - \mathrm{Hue}_{rj}| = \min_{1 \le l \le m} |\mathrm{Hue}_{ql} - \mathrm{Hue}_{rj}| \tag{3}$$

$$|\mathrm{Hue}_{qi} - \mathrm{Hue}_{rj}| = \min_{1 \le k \le n} |\mathrm{Hue}_{qi} - \mathrm{Hue}_{rk}| \tag{4}$$

(3) 查询图像和目标图像中所有可以构成颜色区域对的颜色区域至少被包含在一个颜色区域对中．

在颜色区域对的基础上，颜色特征元素之间的距离定义为所有颜色区域对中两个颜色区域差异的均值，如下式：

$$\mathrm{dis}(\mathrm{CFE}_q, \mathrm{CFE}_r) = \sum_{i,j} \lambda D(\mathrm{CR}_{qi}, \mathrm{CR}_{rj}) / N \tag{5}$$

其中 N 为颜色区域对的个数， $D(\mathrm{CR}_{qi},\mathrm{CR}_{rj})$ 为颜色区域对 $(\mathrm{CR}_{qi},\mathrm{CR}_{rj})$ 中两个颜色区域 CR_{qi} 和 CR_{rj} 的差异，定义如下：

$$D(\mathrm{CR}_{qi},\mathrm{CR}_{rj})=|\mathrm{Hue}_{qi}-\mathrm{Hue}_{rj}|\sum_{t}|\mathrm{cfe}_{qi}^{t}-\mathrm{cfe}_{rj}^{t}| \tag{6}$$

其中色度 Hue 为颜色特征元素中的一个组件，表示了颜色区域的颜色信息， cfe^t 为颜色区域的其它参数，表示了颜色区域的大小、空间分布等信息． λ 为一个常数，根据颜色区域对的构成不同取不同的值，具体取值如下式：

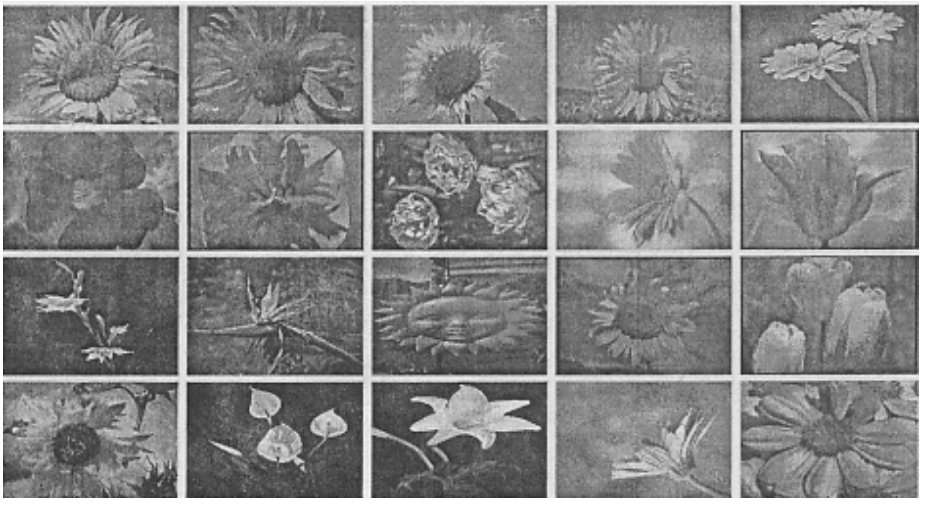
$$\lambda=\begin{cases}1, & |\mathrm{Hue}_{qi}-\mathrm{Hue}_{rj}|=\min\limits_{1\le l\le m,1\le k\le n}|\mathrm{Hue}_{ql}-\mathrm{Hue}_{rk}|\\ 0.5, & |\mathrm{Hue}_{qi}-\mathrm{Hue}_{rj}|=\min\limits_{1\le l\le m}|\mathrm{Hue}_{ql}-\mathrm{Hue}_{rj}|\\ 0.75, & |\mathrm{Hue}_{qi}-\mathrm{Hue}_{rj}|=\min\limits_{1\le k\le n}|\mathrm{Hue}_{qi}-\mathrm{Hue}_{rk}|\end{cases} \tag{7}$$

其中下标为 q 的表示查询图像颜色特征元素的组件，下标为 r 的表示目标图像颜色特征元素的组件．这样取值的原因是为了得到的目标图像含有与查询图像中的颜色区域相似的颜色区域，同时也尽量排除含有与查询图像差异很大的颜色区域的那些目标图像。计算得到两种特征元素的距离之后，将两个距离之和作为查询图像与目标图像之间相似度的度量，即距离越小，相似度越高；距离越大，相似度越低。最后，与一般的 CBIR 系统相同，计算查询图像与所有目标图像之间的距离，选出距离最小的若干幅图像作为检索结果返回给用户。

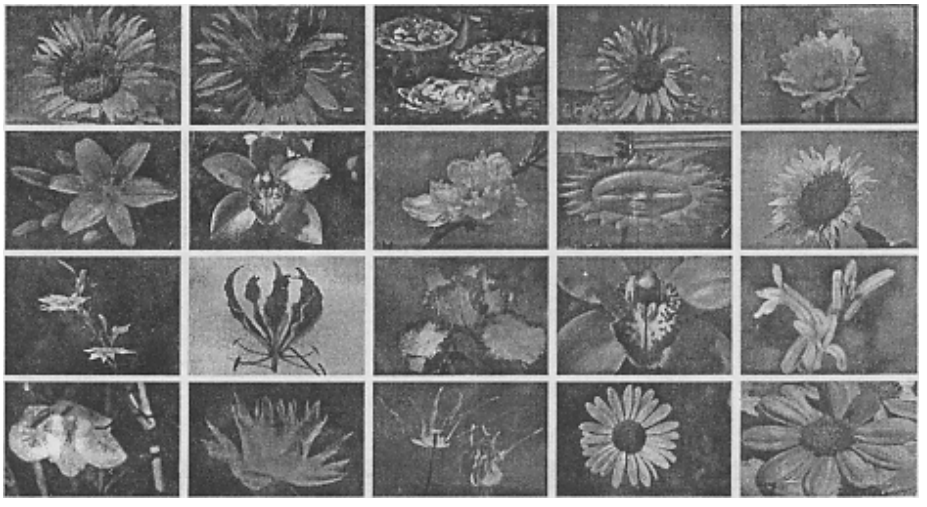
4 实验结果

我们按照上述算法进行了检索实验．由于在图库中图像是按类别存储的，检索时只计算属于同一类别的目标图像与查询图像之间的距离，并且只返回距离最小的 20 个结果．为了与本文提出的基于特征元素的检索算法相比较，我们还实现了一个基于特征向量的检索算法，以特征向量间的欧氏距离作为图像之间的相似度，其中采用的特征向量为颜色特征向量 CCV[7] 、CAC[8] 以及基于小波和矩的形状特征向量 [6] ．我们分别使用两种方法对同一幅查询图像进行检索，检索结果的示例如图 3 和图 4，每个示例的第一幅图像为查询图像，其后的图像为按照相似度排序的检索结果。

图 3 的查询图像为一幅蓝天下的黄色向日葵，属于花朵类．基于特征元素的检索算法得到的结果中，如图 3(a)，有 6 幅图像为蓝天下的向日葵，即与查询图像非常相似，有 9 幅图像或者具有蓝色背景，或者具有黄色花朵，或者为与向日葵形状相似的菊花，即与查询图像部分相似，剩下的 4 幅图像与查询图像很不相似。而基于特征向量的检索算法得到的结果中，如图 3(b)，有 4 幅图像与查询图像非常相似， 7 幅图像与查询图像部分相似，剩下 8 幅图像与查询图像很不相似。



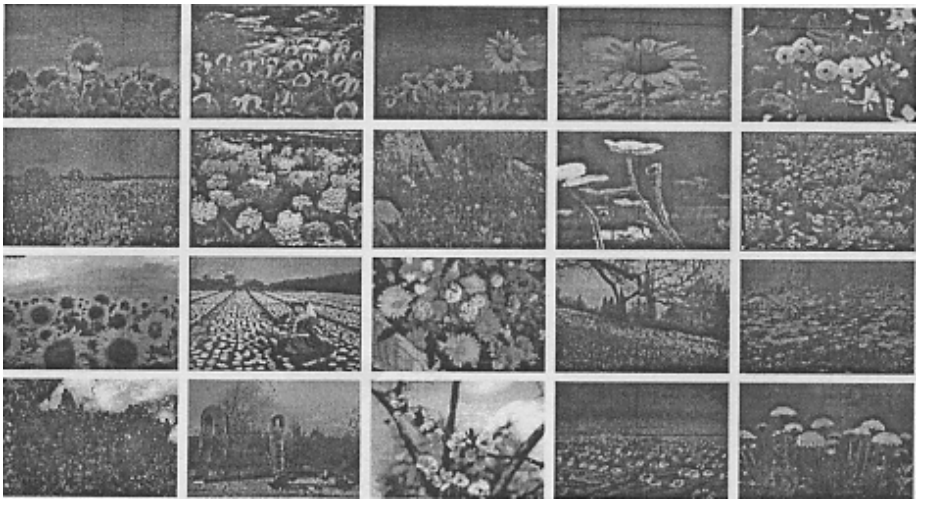
(a) 基于特征元素的检索结果



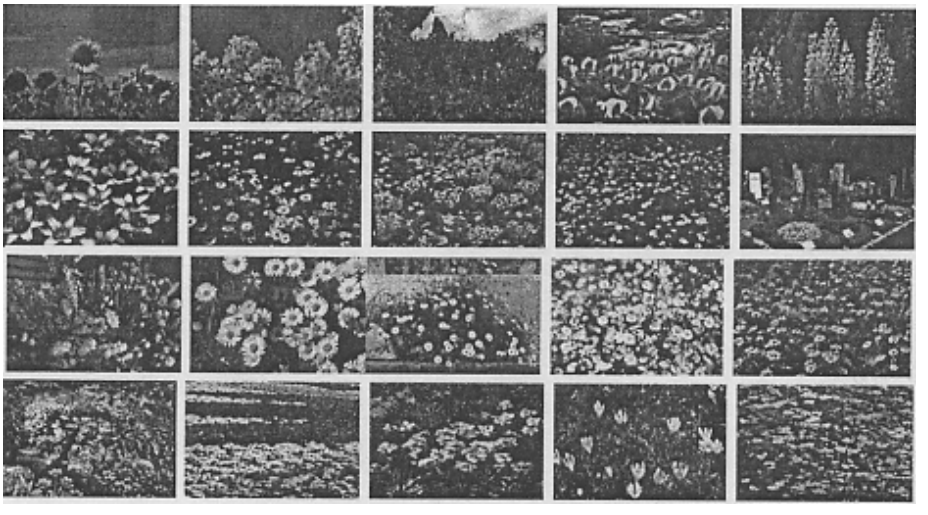
(b) 基于特征向量的检索结果

图 3　花朵类图像检索示例

图 4 的查询图像为蓝天下的大片黄色向日葵，属于花丛类．基于特征元素的检索算法得到的结果中，如图 4(a)，有 4 幅为蓝天下的大片向日葵，与查询图像非常相似．有 6 幅图像或者为蓝天下的大片小黄花，或者为绿草从中的大片小黄花，即与查询图像部分相似，剩下 9 幅图像与查询图像很不相似．基于特征向量的检索结果中，如图 4(b)，没有图像与查询图像非常相似，有 5 幅图像与查询图像部分相似，剩下 14 幅图像与查询图像很不相似．



(a) 基于特征元素的检索结果



(b) 基于特征向量的检索结果

图 4 花丛类图像的检索示例

为了进一步验证算法性能，我们挑出花朵类中所有与图 3 的查询图像最相似的图像，如图 5, 即全部为蓝天下的黄色向日葵，共 7 幅。将这 7 幅图像逐一作为查询图像，在得到的检索结果中，统计除去查询图像本身外，其它 6 幅图像有几幅出现在结果中。表 1 为分别用两种检索算法进行检索得到的统计结果。可以看出，基于特征元素的图像检索算法几乎对于每幅查询图像都能够将其余 6 幅最相似的图像全部检索出来，而基于特征向量的检索算法对于 7 幅查询图像均无法将其余 6 幅图像全部检索出来，有些甚至只能检索出一两幅。

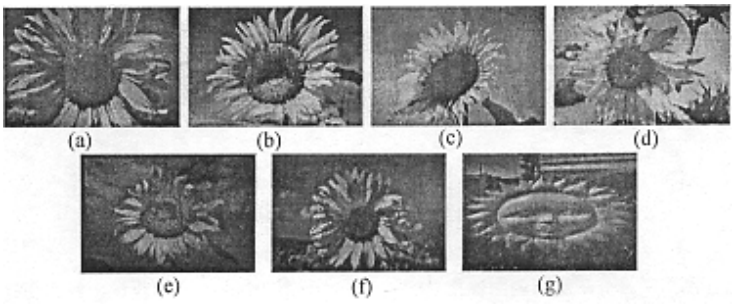


图 5　实验用查询图像

表 1　两种检索算法的统计结果

查询图像	图 5(a)	图 5(b)	图 5(c)	图 5(d)	图 5(e)	图 5(f)	图 5(g)
基于特征元素的检索算法结果	5	6	6	5	6	6	6
基于特征向量的检索算法结果	2	4	5	3	1	2	4

从以上实验的结果可以看出，与基于特征向量的检索算法相比，本文提出的基于特征元素的检索算法更为符合人的视觉感知，取得了较好的检索结果。

5 结论与展望

本文提出了一种基于特征元素的图像检索算法。与特征向量相比，特征元素能够根据人们对图像的主观认知反映图像的视觉内容。在特征元素的基础上，本文定义了图像之间的相似度，即特征元素之间的距离度量。实验结果表明，与基于特征向量的图像检索相比，本文实现的基于特征元素的图像检索更为符合人的视觉感知。

目前本文实现的检索算法仅是对查询图像返回一次性的检索结果，缺乏对于检索结果的优化，因此下一步的工作将着眼于用户反馈的处理，抓住用户的兴趣所在，从而达到更高的检索精度。

参　考　文　献

[1] Mukherjea S, Cho J H. Automatically determining semantics for World Wide Web multimedia information retrieval[J]. *Journal of Visual Languages and Computing*, 1999, 10(6): 585–606.

[2] Smith J R, Chang S-F. Visually searching the Web for content[J]. *IEEE Multimedia Magazine*, 1997, 4(3): 12–20.

[3] Xu Y, Zhang Y-J. Feature element theory for image recognition and retrieval[C]. Proc. Storage and Retrieval for Media Databases, SPIE, 2002, Vol.4676: 126–137.

[4] Xu Y, Zhang Y-J. Image retrieval framework driven by association feedback with feature elements evaluation built in[C]. Proc. Storage and Retrieval for Media Databases, SPIE, 2001, Vol.4315: 118–129.

[5] 李　勍，章毓晋. 基于特征元素的图像分类方法. 电子学报，2002, 30(9): 1262–1265.

[6] Yao Y-R, Zhang Y-J. Shape-based image retrieval using wavelets and moments[C]. Proc. of Workshop on Very Low Bitrate Video'99, Kyoto, Japan, 1999: 71–74.

[7] Pass G, Zabih R. Histogram refinement for content-base image retrieval[C]. Proc. of the IEEE Workshop on Applications of Computer Vision, Sarasota, FL, USA, 1996: 96–102.

[8] Huang J, *et al.*. Spatial color indexing and applications[C]. Proc. of the IEEE International Conference on Computer Vision, Princetor, NJ, USA, 1998: 602–607.

李　勍：　女，1977 年生，硕士，研究方向包括多媒体数据库以及基于网络的图像分类和图像检索.

章毓晋：　男，1954 年生，教授，博士生导师，主要研究领域是图像处理，图像分析，图像理解及其技术应用.