

基于仿生形象思维方法的图像检索算法

王守觉^{1,2}, 孙 华¹, 柳培忠², 廖英豪², 尹 号², 郭东辉²

(1. 中科院半导体研究所, 北京 100083; 2. 厦门大学信息科学与技术学院, 福建厦门 361000)

摘 要: 在人类对图像的理解过程中, 认知结果是由人脑形象思维活动得到的. 模仿人脑从形象思维角度, 本文提出了一种图像特征提取方法, 设计了一种图像检索算法. 把每幅图像都映射为高维特征矢量(每个矢量都被看作是 高维特征空间中的一个点), 计算高维特征空间中点与点的距离判别函数即可得到图像之间的关系. 与其他检索算法比较, 实验结果表明该方法在检索效率和检索速度具有优越性.

关键词: 图像检索; 形象思维; 特征提取; 特征空间

中图分类号: TP391.4 文献标识码: A 文章编号: 0372-2112 (2010) 05-0993-05

An Image Retrieval Algorithm Based on Biomimetic Imaginal Thinking

WANG Shou-ju^{1,2}, SUN Hua¹, LIU Pei-zhong², LIAO Ying-hao², DING Xing-hao², GUO Dong-hui²

(1. Institute of Semiconductors, CAS, Beijing 100083, China;

2. Information Science and Technology College, Xiamen University, Fujian, Xiamen 361000, China)

Abstract: In the process of human beings knowing or understanding an image, they get the cognition result from their imaginal thinking actions. Simulating the human brain and viewing from the imaginal thinking, a method of image feature extraction and an algorithm of image retrieval were proposed. After mapping each image to a high-dimensional vector (each vector can be seen as a point in the high-dimensional feature space), calculating the distance function of a pair of points can indicate the relation of two images. The experimental results show that this image retrieval algorithm acts better than other algorithms in retrieval efficiency and speed.

Key words: image retrieval; imaginal thinking; feature extraction; feature space

1 引言

图像检索^[1]是近年信息科学领域重要研究的热点, 根据研究对象的不同形成多种检索技术: 基于文本、基于内容和基于语义的图像检索. 基于文本的图像检索采用的方法是人工标注图像并用标注进行检索, 其最大的优点: 速度快; 一旦标注完整且适当, 会产生很好的检索结果, 目前互联网的图像检索基本采用这种方式. 同时也存在缺点: 当前数码设备繁多, 可产生大量的图像, 所以人工标注工作量太大^[2].

为克服基于文本的图像检索存在的问题, 提出了基于内容的图像检索技术(Content-Based Image Retrieval, CBIR)^[2], 其研究内容主要是基于视觉特征的提取, 基本方法是从图像中自动提取若干低层视觉特征(如颜色、纹理、形状、轮廓等), 通过比较这些特征的相似度来获得检索结果. 如经典的系统有 Virage^[3]、VisualSEEK^[4]、

QBIC^[5]等, 这类方法在某些应用场合得到很好的效果, 但仅仅从底层特征反映查询者的意图容易产生语义上的歧义, 如检索的内容是含有红色的球时, 得到的结果可能是太阳、球、圆形等, 为避免类似情况产生, 需要检索系统融合图像高层语义功能.

基于语义的图像检索是在图像检索研究中引入了更多的交叉领域学科, 其原因在于图像的高层语义和低层特征之间没有直接联系, 存在巨大的“语义鸿沟”, 系统要获取图像的高层语义是非常困难的. 因此为建立图像从低层特征到高层语义的映射, 基于语义的图像检索融合了模式识别、人工智能等多个知识领域^[6], 从而弥补这个“鸿沟”, 主要的技术手段有^[7-9]:

(1) 相关反馈: 按照最初的查询条件, 系统返回给用户查询结果, 用户选择几个最符合查询意图的返回结果称为正反馈, 反之为负反馈. 反馈信息送入系统更新查询条件, 搜索更符合查询者的意图.

(2) 图像分割: 图像的特征包括全局特征和局部特征. 可依据不同的算法进行图像分割, 划分出不同的分割区域, 增加局部特征的信息量, 实现弥补语义鸿沟.

(3) 建立复杂的非线性分类模型: Li^[10] 等提出的图像自动语义索引系统 (Automatic Linguistic Indexing of Pictures, ALIP), 该系统使用二维多分辨率马尔可夫模型捕获给定语义类别的图像特征之间和内部空间的依赖关系, 该系统对图像检索技术有深远的影响; 采用支持向量机 (Support Vector Machine, SVM)^[11] 技术来弥补语义鸿沟. 这些方法有效地解决语义鸿沟的问题, 但是也存在一些不足:

(1) 低层特征的提取后原有数据的信息量急剧下降.

(2) 目前图像分割算法不成熟, 影响语义的提取.

(3) 目前的语义映射机制存在很大的不足, 其原理是基于先验知识或人类已知的知识进行训练, 需要人机交互.

(4) 基于语义的图像检索技术需要建立能表示现实世界的语义模型是很困难的.

总而言之, 现有的图像检索技术是从逻辑思维的方法来解决形象思维的问题, 以定理、推论等手段的逻辑思维方法^[1], 很难反映图像检索认知过程的本质性, 而形象思维方法是人的一种本质思维方式, 通过人的各种感官来认识图像的信息, 因此形象思维的问题通过形象思维的方法来解决是具有优势的^[16]. 随着各个领域对多媒体检索的需要, 特别是跨媒体检索技术的日渐要求, 这些传统的基于逻辑思维方法的检索技术难以从根本上解决现有算法检索效率低的问题. 因此本文提出了基于形象思维方法的图像检索算法, 结合高维空间仿生信息学^[2]的几何分析方法来提取图像的特征空间, 形成一种关于图像检索的新思想和新方法.

2 仿生形象几何分析图像检索算法

仿生形象思维方法应用于信息科学领域以来, 成功地应用于模式识别领域的技术研究, 并提出了一种新的理论——仿生模式识别理论^[13], 根据仿生模式识别理论研究成果^[13-15], 可以看到对于实物识别、人脸识别等方面的实际应用方面, 其效果都远优于国际上近年发展的 SVM 技术, 在图像复原和清晰化新技术、高维信息大比例压缩新技术等提供很好的解决方法. 总之相对于用逻辑思维方法来解决信息科学领域问题, 以形象思维方法研究为基础的方法, 是目前解决图像处理、语音识别等形象思维问题的有效方法.

本文图像检索算法的理论基础是基于分析了人们对图片视觉反应的以下几个特点发展而得的:

(1) 图片宏观观看时大粒度的模糊概念; 当人们从远处观看一张图片后慢慢走近时, 图片的细节人们并没

有看清, 但却先能看到是什么类型的照片, 是风景照还是人物照等, 因而宏观观看分类, 重点在大粒度信息.

(2) 图片分块参数排序的位置不变性概念: 当人们宏观观看一张图片时, 图片上的山、水、房屋、树木等细节, 如果把它们在水平方向做互相调换位置, 则这个图片对人们的总体观感的变化不大, 也就是说在仿生分类检索下应该检索距离是很近的. 这一用分块参数排序解决位置不变性的方法作者称为“Order-based Block Features Method (OBFM)”, 将在以后的论文介绍与讨论.

(3) 人们在宏观观看一张图片时图片横竖不对称的概念: 当人们在宏观观看一张图片时, 如果把它从反面看, 则变化的感觉与反过面来的旋转轴的方向有关, 旋转轴为横向或纵向的效果是不对称的.

从人们对图片视觉以上三个特点出发, 按照仿生信息学的原理实现了图像信息检索的新算法, 并设计一种能识别图像的特征编码, 如依据本文的算法可将像素为 640×480 的图像文件用 576Byte 的特征数据表示 (程序放在 <ftp://59.77.36.71/ImageRetrievalAlgorithm>, 用户名: eda, 密码: eda401). 在使用本程序把样本图像和待检索图像库中的图像分别压缩为多个 576Byte 的数据后, 每幅图像都是 576 维特征空间中的一个点, 计算两点之间的判别函数距离就可以得到两幅图像的远近关系.

在分析高维度的特征空间中点与点之间和子空间之间的运算时, 用高维空间仿生信息学^[12]的几何分析方法和数学概念将便于清楚地分析的问题, 使运算过程简单. 故以下算法多使用高维空间仿生信息学的数学符号和定义^[12].

主要步骤如下:

第一步: 图像分块.

一个 $m \times n$ 的原图像数组的矩阵形式为:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

在高维仿生信息学中即为 $m \times n$ 维空间中的一个点, 用数学符号 $A_{(m \times n)}$ 来表示.

将原图像分块, 成 $M \times N$ 个子块, 则图像子块形式为:

$$\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1N} \\ B_{21} & B_{22} & \cdots & B_{2N} \\ \vdots & \vdots & & \vdots \\ B_{M1} & B_{M2} & \cdots & B_{MN} \end{bmatrix}$$

其中,

$$B_i = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ik} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{h1} & a_{h2} & \cdots & a_{hk} \end{bmatrix}$$

即每一图像子块在高维仿生信息学中即为 $h \times k$ 维空间中的一个点, 故原图像转化为 $h \times k$ 维空间中的 $M \times N$ 个点, 表示为 $B_i^{(h \times k)}$, ($i = 1, 2, \dots, M \times N$).

第二步: 子块的特征提取.

对 $h \times k$ 维空间中的点 $B_i^{(h \times k)}$, ($i = 1, 2, \dots, M \times N$), 用形象思维方法和高维空间几何分析方法^[12]进行几何映射即特征提取, 找到一个维数为 u 的特征空间记为 $R^u = p_1 p_2 p_3 \dots p_{M \times N}$, 使得 u 维特征空间中的 $p_1, p_2, p_3 \dots p_{M \times N}$ 能与原图像信息融合, $u < (h \times k)$.

第三步: 特征空间中点的排序.

根据“图片分块参数排序的位置不变性概念”, 把 u 维特征空间中的 $p_1, p_2, p_3 \dots p_{M \times N}$ 各点在水平方向按照特定规则排序, 得到了 u 维特征空间中新的点序列 $q_1, q_2, q_3 \dots q_{M \times N}$.

第四步: 将此序列按照水平到垂直的顺序合成 v 字节的特征数据. 至此完成了对一幅图像的特征提取, 将其映射为 v 维特征空间中一个点(对于 640×480 的图像 $v = 576$).

运算公式如下:

$$Q = \bigcup_{j=1}^M \bigcup_{i=1}^N q_{ji}$$

第五步: 图像匹配是图像检索的一个关键步骤, 其基本思想是判断一个样本图像与图像库中图像的距离. 高维空间仿生信息学中, 每幅图像的 v 维特征数据都是 v 维特征空间中的一个点. 判别样本图像和图像库中的图像之间距离就是分析 v 维特征空间中的点与点的关系. 本文用函数 θ 来计算样本图像与图像库中每一幅图像的相似距离, 因而图片类似判别函数:

$$T = \xi_j \theta(Q_1, Q_2), S_j$$

式中:

(1) Q_1, Q_2 为样本图像和要判别的图像的 v 维特征向量.

(2) S 为阈值, 可根据实验设定:

$$T = \begin{cases} 1, & \text{当 } \theta(Q_1, Q_2) < S \\ 0, & \text{其他} \end{cases}$$

$T = 1$ 时样本图像和要判别的图像为类似图像, 反之则不是.

(3) 函数 θ 的计算公式为:

$$\theta = \prod_{i=1}^v \frac{\max(Q_{1,i}, Q_{2,i})}{\min(Q_{1,i}, Q_{2,i})}$$

3 实验与结果分析

3.1 实验环境

硬件环境: CPU 频率 INTEL PENTIUM 1.7GHz+ 内存容量 1G; 软件环境: 操作系统为 WIN XP, 开发工具为 Vc++ 2008 和 open cv workspace, net 2005; 实验使用的

COREL 图像数据库约 900 幅彩色图像, 包恐龙、云、汽车、风景图、大象、房屋、花等 10 类.

3.2 相似图像检索结果分析

相似图像检索是衡量检索算法的检索结果是否符合人们直接意图的一个重要手段, 本实验用纹理和色彩较为相似的云类型图像做比较, 用三种算法 HSV 算法(图 1)、传统图像分割算法(图 2)、仿生形象思维方法算法(图 3). 依据结果可知, 本文算法可实现被检索图像库与样本图像按相似程度的远近排序, 可见本文算法适合于做图像检索.

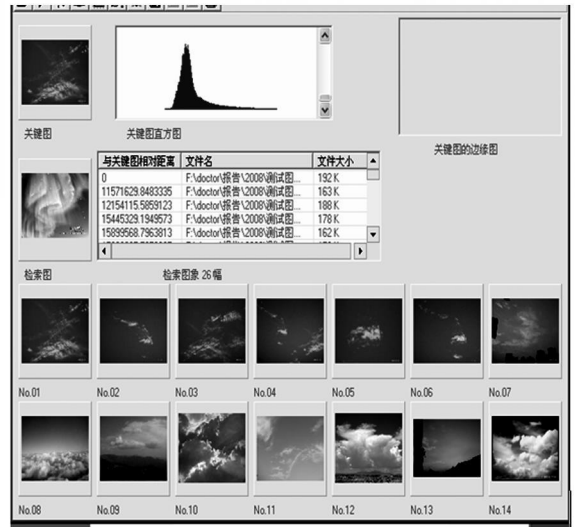


图1 HSV算法检索结果

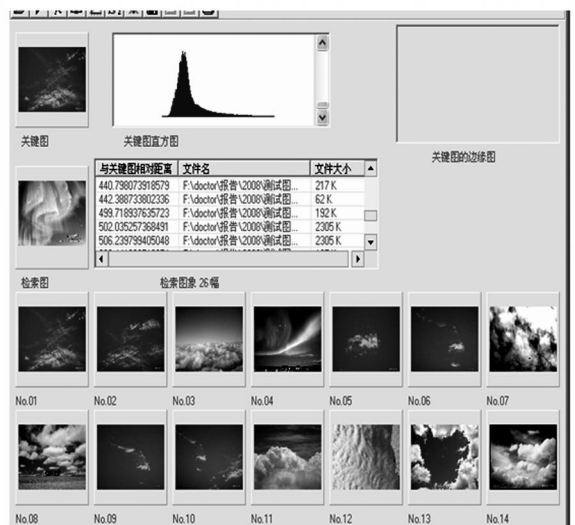


图2 传统图像分割算法检索结果

3.3 查准率分析

用传统的图像分割算法、ALIP 算法和 HSV 算法^[8]计算他们的查准率与用本文方法检索结果对比(见表 1)显而易见, 本文提出的算法总体上优于其他算法.

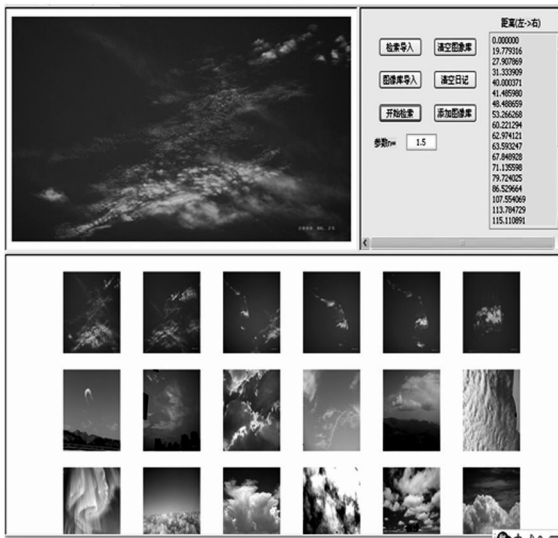


图3 仿生形象思维方法算法检索结果

表 1 各种算法检索结果查准率比较(单位: %)

算法 图片类型	HSV 算法	ALIP 算法	传统分割算法	本文算法
马	40	34	18	56
大象	42	44	14	48
花	70	90	14	75
食物	6	32	12	42
风景图	38	4	12	36
恐龙	88	22	14	90
汽车	18	10	60	50
人	36	16	18	40
房屋	28	18	44	52
大海	44	20	10	32
综合查准率	41	29	21.6	52.1

3.4 检索速度分析

本文算法、图像分割算法、ALIP 算法和 HSV 算法等多种算法的检索速度作比较, 分析不同算法的速度. 检索同一批图像文件(共 26 张, 每张图片大小至少为 640×480) 计算不同算法的检索时间(见表 2). 从实验结果得出基于形象思维方法的方法在检索速度方面高于其他比较的算法.

表 2 各种算法的检索速度比较(单位: s)

算法 时间	HSV 算法	ALIP 算法	图像分割算法	本文算法
	20	8	10	5

4 结论

由上述实验结果分析可以看到:

(1) 基于仿生形象思维方法的检索算法整体效果优于作比较的图像检索算法, 说明基于形象思维方法的图像检索算法是可行的.

(2) 基于仿生形象思维方法的检索算法提出了对

图像进行特征编码的方法, 可以广泛应用嵌入式系统、特定领域的图像检索系统等.

(3) 基于仿生形象思维方法的检索算法的速度快于作比较的算法, 当图像特征编码生成后其速度则更快.

综上所述, 由于仿生形象思维方法的图像检索更接近人类搜索图像信息的本质特征, 使得检索结果有很好的提高. 随着多媒体数据的日益庞大, 对跨媒体的检索技术已经成为当今一个非常重要的研究热点, 但用传统思维方法来解决跨媒体检索的问题显得无能为力, 特别是在跨媒体的特征提取和基于认知角度的搜索引擎的建立. 本文已经实现仿生形象思维方法在图像检索技术上的有益尝试, 经过仿生对形象思维方法应用的进一步研究, 将新的研究成果应用于跨媒体检索技术中, 促进检索技术跨越式的发展.

致谢: 与本文比较的算法程序(图 1、图 2)由云南师大现代教育中心制作, 在此表示感谢.

参考文献:

- [1] R Datta, J Li, J Z Wang. Content-based image retrieval approaches and trends of the new age[A]. In Proceedings of the Seventh International Workshop on Multimedia Information Retrieval[C]. Singapore, 2005. 253-262.
- [2] 黄祥林, 沈兰荪. 基于内容的图像检索技术研究[J]. 电子学报, 2002, 30(07): 1065-1071.
Huang Xiang-lin, Shen Lan-sun. Research on content-based image retrieval techniques[J]. Acta Electronica Sinica, 2002, 30(07): 1065-1071. (in Chinese)
- [3] Gupta A, Jain R. Visual information retrieval[J]. Communications of the ACM, 1997, 40(5): 71-79.
- [4] J Smith, S Chang. VisualSEEK: a fully automated content-based image query system[C]. In ACM Multimedia, Boston, Massachusetts, 1996. 87-98.
- [5] Flickner M, Sawhney H, Niblack W, et al. Query by image and video content: the QBIC system[J]. Computer, 1995, 28(9): 23-32.
- [6] 李志欣, 施智平, 等. 图像检索中语义映射方法综述[J]. 计算机辅助设计与图形学学报, 2008, 20(8): 1085-1096.
Li Zhixin, Shi Zhiping. A survey of semantic mapping in image retrieval[J]. Journal of Computer-Aided Design & Computer Graphics, 2008, 20(8): 1085-1096. (in Chinese)
- [7] WeiJiang, GuihuaEr. Similarity-based online feature selection in content-based image retrieval[J]. IEEE Transactions on Image Processing, 2006, 15(3): 702-712.
- [8] 杨红颖, 等. 一种基于 HSV 空间的彩色边缘图像检索方法[J]. 中国图像图形学报, 2008, 13(10): 2035-2038.
Yang Hong-ying, et al. Content based image retrieval using color edge histogram in HSV color space[J]. Journal of Image and Graphics, 2008, 13(10): 2035-2038. (in Chinese)

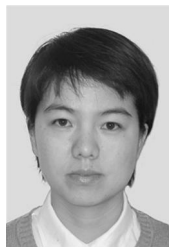
- [9] Deng, Y N . An efficient color representation for image retrieval[J] . IEEE Transactions on Image Processing, 2001, 10 (1) : 140– 147 .
- [10] Li, J Wang, J Z. Real-time computerized annotation of pictures [A] . In Proceedings of the ACM International Conference on Multimedia[C] . October 23– 27, 2006, Santa Barbara, California, USA, 2006. 911– 920 .
- [11] R Min, H D Cheng. Image retrieval based on relevance feedback using fuzzy support vector machines[A] . In Proceedings of the 10th Joint Conference on Information Sciences[C] . Salt Lake City, Utah, USA, 2007. 1009– 1015 .
- [12] 王守觉, 来疆亮. 多维空间仿生信息学入门 [M] . 北京: 国防工业出版社, 2008. 1 .
- [13] 王守觉. 仿生模式识别(拓扑模式识别) —— 一种模式识别新模型的理论与应用[J] . 电子学报, 2002, 30(10) : 1417– 1420 .
Wang Shou-jue. Bionic (topological) pattern recognition —— a new model of pattern recognition theory and its applications [J] . Acta Electronica Sinica, 2002, 30(10) : 1417– 1420 . (in Chinese)
- [14] 王守觉, 王柏南. 人工神经网络的多维空间几何及其理论[J] . 电子学报, 2002, 1(30) : 1– 4 .
Wang Shou-jue, Wang Bai-nan. Analysis and theory of high dimensional space geometry for artificial neural networks[J] . Acta Electronica Sinica, 2002, 1(30) : 1– 4 . (in Chinese)
- [15] Wang Shoujue, Lai Jiangliang. High dimensional imagery geometry and it's applications[J] . Chinese Journal of Electronics, 2006, 15(4): 761– 767 .

- [16] GEVERS, T, SMEULDERS, A. Pictoseek: Combining color and shape invariant features for image retrieval[J] . Image Process, 2000, 9 (1) : 102– 119 .

作者简介:



王守觉 男, 院士, 1925 年生, 中国科学院院士. 我国半导体与信息科学专家、半导体器件与微电子奠基人之一. 近年来致力于半导体神经网络、智能信息处理领域的研究工作. 曾获得过多项国家级科研成果发明奖. 目前兼任《电子学报》和《Chinese Journal of Electronics》主编.



孙 华 女, 博士研究生, 1973 年生. 2007 年获得中国科学院研究生院工学硕士学位, 同年师从王守觉院士攻读博士学位. 研究方向: 图像处理、图像检索、模式识别、人工智能等.

E-mail: shua@semi.ac.cn



柳培忠(通信作者) 男, 博士研究生, 1976 年生. 2007 年获得厦门大学工学硕士学位, 同年攻读博士学位, 师从王守觉院士. 研究方向: 视觉媒体检索, 模式识别等.

E-mail: liupeizhong0921@163.com