

基于内容的图像检索研究与实现

专业：计算机科学与技术 学生：张益强 指导老师：俞 颖

摘要

近年来，大量的图像信息充斥在我们的周围，人们迫切地希望从大量繁杂的图像信息中快速，准确地定位到我们所期望得到的多媒体信息。图像检索技术在物体识别、医疗、军事、遥感分析等领域的应用越来越广泛，逐渐成为专业学者的研究热点问题。早期的图像检索主要是基于文本技术，该方法存在明显的不足之处，图像丰富的内容仅凭寥寥几字的文本是不足以清晰的描述出的，因文本描述都是定性的，而图像中往往含有许多不规则形状及大量难以描述的纹理，而且文本的描述具备描述者一定的主观性。两相比较之下，基于内容的图像检索是直接从图像中去提取所需要的信息线索，不用经过文本描述。基于内容的图像检索中特征提取和检索所需的索引完全是在计算机中自动实现的，检索效率大大提高。

本课题基于 MATLAB 平台，主要研究的是基于内容的图像检索。将待检索的图像集进行图像预处理，图像的特征提取，然后利用欧式距离和曼哈顿距离的图像检索技术将图像进行检索，经过仿真实验表明，基于内容的图像检查技术能有效地克服传统图像检索方法的局限性，更加直观和灵活，检索效果良好，精度较高。

关键词：MATLAB 平台 图像检索 欧式距离 曼哈顿距离

目 录

1 绪论	3
1.1 图像检索及应用	3
1.2 本课题的研究背景及意义	3
1.3 国内外研究现状	4
1.4 基于内容的图像检索的关键技术	5
1.5 基于内容的图像检索优缺点	5
2 开发环境介绍	7
2.1 MATLAB 简介	7
2.2 MATLAB 优势特点	7
2.3 MATLAB 环境	7
2.4 MATLAB 图像处理工具箱	8
3 图像检索原理及流程介绍	9
3.1 基于内容的图像检索原理	9
3.2 基于内容的图像检索过程	11
4 图像特征提取	12
4.1 基于纹理的特征参数提取	12
4.2 基于颜色的特征提取	14
5 图像检索	18
5.1 图像特征的相似性匹配	18
6 仿真及分析	20
6.1 仿真实验	20
结论与展望	28
参考文献	29

1 绪论

1.1 图像检索及应用

图像检索是根据已有的条件,去找寻相同或相近图像的意思,其原理主要包括三个方面:一方面是了解用户的需求,对用户的需求进行分析和转化,使之可以对数据库进行索引检索。另一方面,将图像进行收集,而后对图像进行加工整理,提取出所需要的图像特征,分析并对图像进行标记,完成图像的索引且将索引数据库建成。最后一方面则是按照某种相似度算法,对用户的需求建立索引,与索引数据库中的记录进行相似度比较,提取出满意的结果反馈给用户,将相似度最高的优先反馈。

目前,图像检索技术的应用范围已有了很大的扩展,已在公安系统、数字图书馆、遥感影像等方面使用越来越广泛。在公安系统中,历年的犯罪嫌疑人的有关脸、指纹、照片等的信息都保存在其中,如果有前科的嫌疑人犯罪,工作人员可以运用嫌疑人的信息进行快速检索,加快破案进度。数字图书馆中,主要是数字图书馆存有大量的数字信息,而图像检索就是从大量的信息中得到我们所需的信息。遥感影像则是从影像中得到大量的地理信息,可以通过图像检索得到地质变化的情况。

1.2 本课题的研究背景及意义

近年来,随着计算机网络技术和多媒体技术在网上的快速发展,大数据时代来临,网络中数字图像的数目也成井喷式增长。面对如此繁杂且分布于全世界的大量图像资源,怎样将海量的图像数据中将所需要的信息高速、有效地检索出是当前许多应用领域所需要解决的问题。传统的基于文本的检索技术已难堪大任,主要是因为仅仅使用关键词对图像进行内容描述的局限性;基于内容的图像检索技术(Content-Based Image Retrieval,CBIR)便应运而生。^[1]

基于文本的图像检索(TBIR)就是在整理图像库时对图像的名称、编号、作者等进行文字收集,或者是对图像中的文字或者图像周围的文字一同整理,收集在一起,用做该图像的文本描述。在进行检索时,通过对图像的文本描述检索,将该图像识别出来。但是基于内容的图像检索的不足还是显而易见的。当图像的数量非常大时,基于文本的图像检索所需要的文本描述能力,也就是关键字有限,图像中丰富的内容难以将其描述而出。若文本描述过多,则存储量增大。还有在检索中人们对图像所表达的关键字理解不同,所使用的关键字不同,所需要的同一张图像,经过不同人的检索得到的图像也多不相同。这也就是所说的基于文本的图像检索具备一定的主观性。但基于文本的图案检索若是关键字对了那就能准确、快速的将所需的图像识别出来。

基于内容的图像检索（CBIR）在当前已成为大量学者研究者的热门研究课题。基于内容的图像检索也有人称之为图像内容查询（QBIC）或基于内容的视觉信息检索（CBVIR），是一种基于相似度度量有效地从集合中检索到所需图像的过程，通过手上已有的图像，对图像的底层特征（例如颜色、纹理、形状等）进行提取，再与已整理得到的图像数据库进行检索匹配，得到所符合内容的目标图像以及相似的图像进行选择。让“有图有真相”不再渺茫。基于内容的图像检索是直接从图像内容中提取所需的信息，提取的特征与建立的索引可由计算机实现，不需要进行文本注释，图像库建立所需时间少，检索效率大大提高。因其不需要去识别图像中的对象，只需要关注图像内容，快速地发现信息，在大量的数据中有着较高的检索效率。目前，已在工业和科研等领域具有较大的前景，例如应用于医院的病理图像资料管理、以及公安机关的犯罪数据图像分析、气象局的天气资料照片管理和遥感影像的分类识别等。所以本次课题是基于内容的图像检索研究与实现。

1.3 国内外研究现状

1.3.1 国外研究现状

由于在国外,基于内容的图像检索的出现时间较早,发展比较快,现如今技术水平已达到了一个相当高的水平,已有了一些较为出色的图像检索系统,像 QBIC 检索系统, VisualSeek 检索系统, Photobook 检索系统等。

QBIC（Query By Image Content）系统是由 IBM 公司发布的一款可以实现基于颜色、纹理或形状的组合查询的图像检索系统，它既能当作为一个独立的软件产品进行使用，也可以成为 IBM 公司成立的数字图书馆中的一部分。是一款较为经典的产品，在最新的版本中，QBIC 还结合了基于文本和基于内容为一体的图像查询。^[2]

VisualSeek 检索系统是美国哥伦比亚大学电子工程系研制成功的，是提供人们在 WEB 上检索的一套工具，将基于内容的图像或者视频的检索功能^[3] 在网络上实现。VisualSeek 主要是利用了图像的彩色信息，区域大小以及绝对和相对空间位置的有效索引技术，使系统可以计算出各种复杂的关联颜色和空间的查询。具有高效的 WEB 图像信息检索、用户界面大、操作简单、支持用户直接下载信息等特点。

Photobook 检索系统是由美国麻省理工学院所研发的一套用于图像检索和浏览的交互式系统，它主要是由分别负责提取形状、提取纹理、提取面部特征三个子系统组成。因此，用户可以分别使用这个系统中不同的子系统进行基于形状、基于纹理、基于面部特征的图像检索^[4]。

1.3.2 国内研究现状

在国内，虽然基于内容的图像检索技术研究近期才成为热点课题，目前许多研究人员和研究机构已投身于这一热门课题。有些研究机构也已经推出了自己研发的图像检索系统。其中较为出名清华大学计算机系研制出的 **ImgRetr** 系统^[5]。除了研究领域外，商业领域中，如百度识图、阿里巴巴集团的淘淘搜等也是应用了基于内容的图像检索技术。

ImgRetr 系统主要是利用颜色、颜色直方图、纹理、轮廓和框架等多种特征结合来进行图像检索。百度是国内目前在图像检索领域中研究时间最久的一家公司。百度识图通过对图像底层局部特征的对比，实现了寻找相同或者相似的图像的功能，可以满足用户图片寻源、残图补全、模糊图换清晰图等的需求。而阿里巴巴的淘淘搜更是满足其公司发展需要，在日益火爆的网上购物中，不少人只识图不知其关键字，传统的基于文本的检索方式做不到“有图有真相”，以图搜图的检索方式更易于其商业发展，不需要用户思考输入何种关键字进行查找。

目前，现有的基于内容的图像检索系统基本满足了人们的需求，但是仍然存在不尽如人意的地方，而且由于图像特征描述、图像特征参数提取以及相似量的计算还存在很大的复杂性，因此大肆运用还存在很多问题。

1.4 基于内容的图像检索的关键技术

CBIR 系统的包含以下几个关键技术方面：

(1) 数学模型。基于内容的图像检索系统的核心是数学模型，数学模型的功能决定了基于内容的图像检索系统的查询类型和检索性能。数学模型的建立能够充分地反映图像中对象的内容。

(2) 特征提取。特征提取是图像检索的基石。特征提取的算法简单实用，由计算机自动进行提取，减少人为操作。所选择的特征在不破坏图像的情况下应尽可能详细。

(3) 索引技术。在进行图像检索时，最主要的方式是相似性算法匹配。即在图像数据特征库中寻找与待测图像最为相似的一幅或多幅图像。并将它们反馈给用户。它的过程是将提取出的大量图像特征建立成图像特征库。然而在目前提取的图像特征都是高维的，因此图像检索的问题就变成了对空间数据库中的高维数的近邻检索问题。为了建立的索引能够高效的进行数据特征查找，从而完成图像检索。因此，如何降低索引维数和建立良好的索引方法是目前对索引技术的研究重点。

1.5 基于内容的图像检索优缺点

CBIR 技术相比较于传统的基于文本的图像检索技术，将基于文本的图像检索时的主观性和歧义性克服了，将人机交互，数据库、多媒体等技术融入到 CBIR 技术中，主要具

有以下的优势:

(1) 传统的基于文本的图像检索技术, 是基于关键字匹配进行检索的。系统前期需要人为的将图像库中的每张图像进行有关关键字的描述或总结。随后利用关键字对图像库中的每张图像进行关键字的匹配, 经过索引得到所需的图片, 再将图像反馈给用户。而 CBIR 技术是运用编程, 使计算机自动按照编写好的程序对图像的特征进行提取特征向量。无需人为干预, 避免了对图像库中大量图片的关键字描述。

(2) 基于文本的图像检索技术是用精确匹配实现的, 而基于内容的图像检索技术则是近似匹配, 是将检索相似性中, 最符合待测图像的几张图像结果反馈给用户, 给予用户选择。从而找出最满意的答案。

(3) CBIR 技术都是计算机按照预设好的特征提取算法、检索距离度量公式及索引自动完成图像特征的提取, 使检索效率获得大的提升。

(4) CBIR 系统中的特征向量库中包括图像库、特征库以及知识库等, 可用于众多领域的各自不同的检索案例。

虽然 CBIR 系统经过近几十年的发展, 技术水平愈加成熟, 算法种类更加丰富, 但是依然存在很多需要解决的问题。

(1) 目前特征提取阶段中, 基于颜色特征提取是因为没有正确的表达人眼对色彩特征的感知, 虽然 HSV 颜色空间是与人眼视觉较为符合的, 但是还是不能完全表达人眼对颜色的认知。基于纹理特征提取中是因为无法适用于所有物体, 在某一类特定物体中, 效果卓越, 但在另一类物体中, 检索效果就差强人意了。

(2) CBIR 系统中, 检索算法的实现尤为重要, 目前的 CBIR 系统中, 由于图像特征库的庞大, 运用的算法如何准确、有效得到所需图片也是人们研究的重点。

2 开发环境介绍

2.1 MATLAB 简介

Matlab 是 Matrix LABoratory 的缩写，是矩阵实验室的意思。^[7]是一种集成了计算，可视化编程和数据处理的高性能语言和开发环境。它具备了复杂的数据结构，包括内置的编辑和调试工具，支持面向对象的编程，还能用于创建用户界面以及调用其他语言（例如 C，C++等）编写的程序。这些因素使得 MATLAB 成为了大多数学者教学和研究的工具。

Matlab 在汽车安全系统、行星际宇宙飞船、健康监控设备、智能电网等领域中使用广泛，主要用于机器学习、信号处理、图像处理、计算金融学、控制设计等等。因为 matlab 附加的众多的工具箱（Toolbox），因此，除了用于数值的运算外，也同样使用图像处理、图像检索、通讯等领域。因此，我们选用 matlab 来完成本次课题。

2.2 MATLAB 优势特点

（1）功能强大，适用范围广，Matlab 具有高效的数值计算能力和专业水平的符号计算能力，使用户不必进行繁杂的数学运算。

（2）语言简洁，使用方便，Matlab 拥有大量函数建立的函数库，既包涵常用的基本库函数，还拥有功能多种多样的专用库函数。Matlab 所提供的的库函数都是由专家编写，普通用户不用因函数的不正确而困扰。使用简便，遇到所需要的函数，直接套用函数名即可，可避免使用复杂的子程序编程。

（3）图形界面操作简便，Matlab 具有友好的用户界面，操作简单。如遇到不懂得，还可以使用 help 帮助系统，使学者更容易上手进行操作。

（4）齐全的工具箱，Matlab 为广大学者提供了种类繁多、功能强大的工具箱，其中核心部分由系统自带，另各种可选的工具箱可自行下载使用。

2.3 MATLAB 环境

MATLAB 环境（大多数计算机系统）都是由菜单栏、工具栏、当前目录路径、当前目录窗口、文件详细信息窗口、命令窗口、工作区窗口和命令历史窗口组成。在启动 matlab 时首先看到的是写入区域即为命令窗口、命令窗口是 matlab 工作的主要窗口，在命令窗口中，输入需要执行的命令后，按回车键执行命令，若要中断命令，则键入 ctrl+c 即可。在命令历史窗口中可查看键入的命令，若要执行命令历史窗口中的命令，双击该代码即可。

2.4 MATLAB 图像处理工具箱

Matlab 工具箱中图像处理工具箱是我们此次使用的主要工具箱。该工具箱支持多种图像格式，主要包括 JPG、TIFF、HDF 等。除了基本的图像格式外，matlab 还支持导入/导出 AVI 的格式文件以及其他符合工业标准的数据格式。

图像数据箱提供了大量的函数用于图像处理，使图像数据分析、获取图像信息、图像预处理变得简单易懂。图像处理工具箱支持 4 种图像类型，它们分别是真彩色图像(也就是我们所说的 RGB 图像)、索引色图像、灰度图像、二值图像。丰富的图像处理工具函数为我们进行基于内容的图像检索技术的研究与实现提供了便利和可能性。

3 图像检索原理及流程介绍

3.1 基于内容的图像检索原理

基于内容的图像检索的研究主要是通过对图像的基本底部特征（例如颜色、纹理特征）来进行的，其工作原理是先收集大量的图像，经过对图像内容的分析，提取出图像中的各种特征参数，当做是特征向量，并与图像一起存储在建立的图像特征库中。当开始进行图像检索时，一般是由用户提供一个样例图像或者是一幅草图，提取其图像特征，然后再与图像特征库中的特征向量进行比较，并根据计算结果将得到的相应图像及近似图像反馈给用户。一个完整的基于内容的图像检索（CBIR）系统通常是由四个部分组成，分别为图像数据库系统、对图像进行相似匹配、经过查询方式匹配、输出结果。根据基于内容的图像检索系统的原理，我们可以得到以下信息：

（1）图像特征库系统。其中包括图像库、特征库、以及信息库组成。^[6]图像库中储存着各类图像的信息；特征库中存储的则是从各个图像中提取的各种底层视觉特征；信息库则是包含匹配信息及索引，主要是作为于查询优化和快速匹配的；能够对图像数据库进行有效地管理并建立起所需要的高效索引，将图像的检索速度加快。

（2）图像预处理。图像预处理包括将图像尺寸进行统一、对图像进行增强，去除不需要的噪声，对图像进行平滑等操作，为图像特征提取提供便利。

（3）特征提取。特征提取包括选择出所需的特征范围和对所需特征进行提取，特征选择即根据所需特征选取出正确的图像特征空间形成特征空间，一旦特征选取对了，所对应的特征空间也就是所需的特征空间。而提取特征则是指从所选取的特征空间中将其所需要的特征参数信息从图像中提取表示出来。特征空间的选取可以是整幅图像，也可以是图像中的某个区域或者是图像中具体的某个对象。

（4）相似匹配。相似匹配即如何进行图像匹配，也就是确定所选用的匹配算法。基于内容的图像检索中，使用图像与图像的匹配。这是与基于文本的图像检索中的文本与图像匹配的精确匹配不同，而是采用一种近似的匹配。图像检索是利用特征向量的相似度距离函数来进行相应的检索，将得到的特征参数与图像库中的参数进行距离匹配，再从特征库中寻找匹配的特征。将得到的一张或几张相似的图片反馈给用户。

基于内容的图像检索技术过程原理如图 3-1 所示。

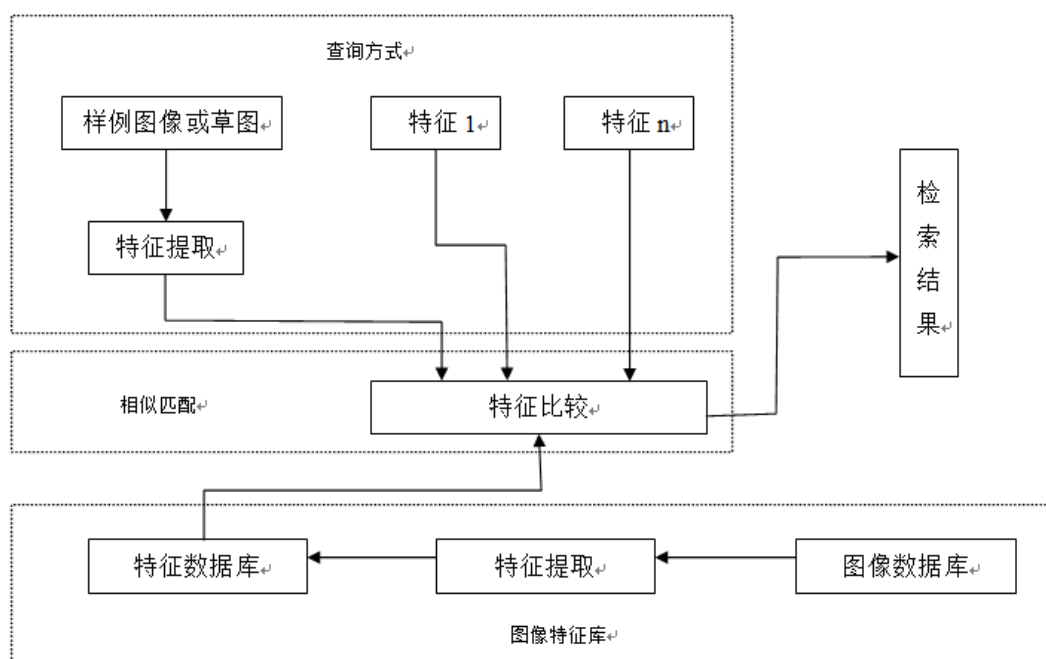


图 3-1 基于内容的图像检索原理图

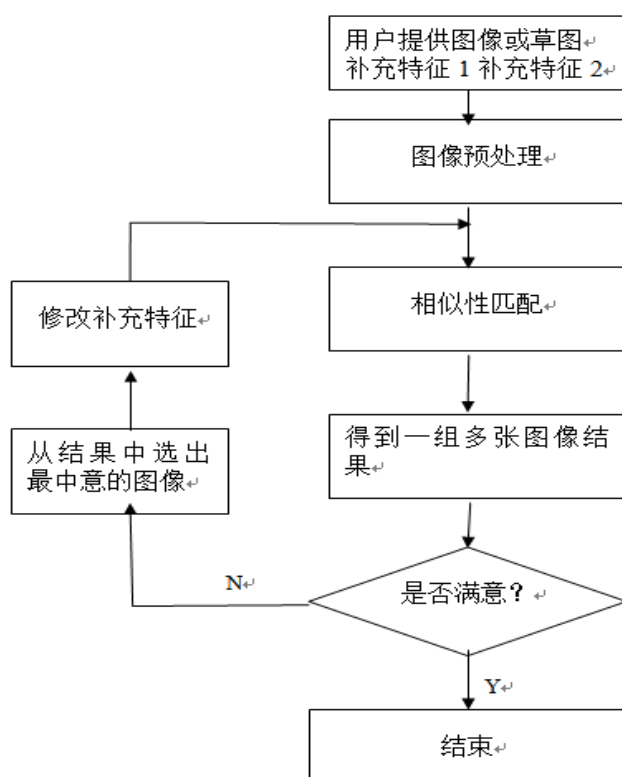


图 3-2 基于内容的图像检索过程

3.2 基于内容的图像检索过程

基于内容的图像检索过程就是一个精益求精的，得到用户所想，所需的图像的过程。其中存在将查询的特征调整，对调整的特征进行重新匹配的循环过程。

（1）查询要求，用户所需要进行图像检索使，需要一个待测图像或者草图，将待测图像作为特征查询的一个条件，还可补充其他特征进行补充检索。通常都先将图像进行预处理。然后传递给特征比较。

（2）相似性匹配。将用户提交的图像经过预处理后，将提取到的特征与数据特征库中的特征按照一定的相似度度量算法进行匹配。

（3）将经过匹配得到的多张图片反馈给用户，若用户不满意得到的图片，则可修改补充的特征，进行更为精确的查找。

基于内容的图像检索过程如图 3-2 所示。

4 图像特征提取

让计算机能够将图像识别出来的,这就是图像特征的目的。也就是将人们所知的低层特征转化为计算机能够识别的技术。图像特征提取作为实现基于内容的图像检索研究与实现的基石,为实验中提供数据基础。本章中主要列出了几种常用的图像提取技术。

4.1 基于纹理的特征参数提取

4.1.1 纹理的概念和特征

纹理是人们任何可见的物体表面凹凸不平结构变化的触感。在此基础上,将触次感与人类视觉所看到的信息所结合起来。纹理成为了当今一种重要的视觉信息,主要描述物体表面具有周期性变化的表面结构属性。纹理是图像对象本身存在的一种属性,它反映了物质的质地,并不依赖于图像本身的颜色深浅和亮度变化,具体的可以从以下三个方面进行描述。

(1) 理通俗来讲是图像表面具有的纹路或者花纹,事实上是图像对象本身的像素点灰度值之间存在的分布规律,表示像素之间的排列关系。

(2) 纹理是图像区域中表达的一个属性,单独一个像素点的纹理是毫无用处的。因此,图像灰度值空间关系的大小取决于纹理的类型,或者说定义纹理基元的大小。

(3) 在不同的尺度以及不同的分辨率下图像的纹理特征仍能被感知。

近年来,计算机图像处理、图像分析、图像检索等研究领域的大热,作为这些热门领域的基础性研究领域之一的纹理特征提取也随之变得热门。

4.1.2 灰度共生矩阵法

在 20 世纪 70 年代,由 Haralick 等人提出的灰度共生矩阵,是目前最常用的描述纹理特征的方法,主要是通过统计图像上像素值的相关性的计算来描述纹理信息,对灰度图像上保持某个特定距离的两个像素,分别对这两个像素所具有的某灰度的状况进行统计所得到的^[8]。

在图像($N \times N$)中,取 (x, y) 作为任意一点,以及偏离 (x, y) 的一点 $(x+a, y+b)$,我们假设灰度值为 (g_1, g_2) ^[9]。假设让 (x, y) 移动在整个画面上,会得出各种 (g_1, g_2) 值,接着假设灰度值的级数为 k ,可得 k^2 种 (g_1, g_2) 的组合。再将所有的 (g_1, g_2) 的值进行统计,记录每种 (g_1, g_2) 出现的总次数,将每种 (g_1, g_2) 出现的总次数结合排列成一个矩阵,再使用 (g_1, g_2) 出现过的总次数将矩阵进行归一化,并将矩阵转化为每种 $(g_1,$

g_2) 出现的概率 $P(g_1, g_2)$, 所得到的对于 $P(g_1, g_2)$ 的矩阵就被称为灰度共生矩阵, 如果取用不同的 (a, b) 组合, 就可以得到不同情况下的灰度共生矩阵。 (a, b) 的选择不能随意选取, 纹理周期分布的特性是 (a, b) 取值选择的根据。若图像纹理较细时, 一般选取的 (a, b) 应当较小, 例如 $(1, 0)$ $(1, 1)$ $(2, 0)$ 等的小的差分值。一般使用灰度共生矩阵时, 会选取 0° 角, 45° 角, 90° 角和 135° 角得到的灰度共生矩阵, 即 (a, b) 取值为 $(0, 1)$ 时, 为水平的像素对, 为 0 度扫描; (a, b) 为 $(1, 1)$ 时, 像素对为右对角线, 为 45 度扫描; (a, b) 为 $(1, 0)$ 时, 是垂直的像素对, 为 90 度扫描; (a, b) 为 $(1, 1)$ 时, 像素对为左对角线, 则是 135 度扫描。

经过 (a, b) 取值不同的处理, 同时出现两个像素灰度级的概率将 (x, y) 的空间坐标进行了转换, 换为对“灰度对” (g_1, g_2) 的描述, 成为了灰度共生矩阵。在实验中, 将灰度共生矩阵归一化, 如公式 (4-1) 所示。

$$p(g_1, g_2) = \frac{p(g_1, g_2)}{R}, R = \begin{cases} N(N-1) & \theta = 0 \text{ 或 } \theta \\ (N-1)^2 & \theta = 45 \text{ 或 } \theta \end{cases} \quad \text{公式 (4-1)}$$

灰度共生

矩阵中, 其他元素值对于主对角线的分布若使用离散化来表示, 则反映出来的通常为纹理的粗细程度。如果离主对角线远的元素经过归一化后的值高, 则元素的离散化程度较大, 即某一位置关系两个像素之间的灰度差的比例高。

如果以位置关系 $|\Delta x|=1$ 或 0 , $|\Delta y|=1$ 或 0 为例子的话, 离散化程度大的意味着相邻像素之间灰度差大的所占的比例高, 则表明图像中与该方向垂直的纹理较薄; 相反, 则表明图像中垂直于该方向上的纹理较厚。当除了主对角线外, 其他的元素的归一化值全都为 0 时, 元素的离散化程度最小, 即纹理不会出现在垂直于该方向的图像上, 主对角线上的元素表示灰度共生矩阵中两像素同灰度组合在一定位置关系下出现的次数。因为沿着纹理方向的相似元素的灰度级基本上相同, 垂直于纹理方向上的相邻像素间的大灰度差的一般规律使得主对角线上元素的大小能够帮助我们确定纹理的方向和粗细, 在纹理分析中起着重要的作用。

灰度共生矩阵主要表现为像素灰度值的相关性, 使用一对像素出现某灰度值的条件概率来表达纹理特征, 还需要计算出相关性、熵值、对比度、能量等来反映整幅图像的特征。但是由于各个特征值的物理意义不一样, 需要把它们相同的权重进行归一化处理^[10]。

纹理能量如公式 (4-2) 所示。

$$Q1 = \sum_{g_1} \sum_{g_2} [p(g_1, g_2)]^2 \quad \text{公式 (4-2)}$$

是表示灰度共生矩阵的平方和，反映了图像灰度分布情况。如果共生矩阵的所有值越接近相同，元素越稀疏，则能量值较小。如果共生矩阵中有的值大而有的值小，则能量值较大，还有若共生矩阵中元素分布较为集中时，能量值也较大。

纹理对比度如公式（4-3）所示。

$$Q2 = \sum_{g1} \sum_{g2} k^2 p(g1, g2)^{2 \cdot k = |g1 - g2|} \quad \text{公式（4-3）}$$

这反映了图像纹理的清晰度以及纹理中沟纹的深浅程度。如果纹理的沟纹越深，则对比度就越大，可视度就越清晰；反之，如果对比度越小，则纹理沟纹越浅，可视度模糊^[11]。如果在这图像中的亮度值变化的很快，则远离矩阵对角线的元素值就越大。

纹理相关性如公式（4-4）所示。

$$Q3 = \frac{\sum_{g1} \sum_{g2} g1 g2 p(g1, g2) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad \text{公式（4-4）}$$

使用纹理的相关性来测量灰度共生矩阵中行方向或列方向上的元素的相似度。如果图像在一定方向上的值大于其他方向，则该方向上的纹理比其他方向更强，因此可以表示纹理的方向。

纹理熵如公式（4-5）所示。

$$Q4 = - \sum \sum p(g1, g2) \lg p(g1, g2) \quad \text{公式（4-5）}$$

如果此时灰度共生矩阵的值分布较均匀，则纹理熵的值也比较大，意味着图像中纹理的非均匀程度较高，即图像复杂程度较高。

4.2 基于颜色的特征提取

颜色空间模型指的是某个三维颜色空间中的一个包含颜色域中所有颜色的可见光子集。主要有 RGB、HSV、HIS 等不同的颜色空间模型，以下介绍 RGB 颜色空间和 HSV 颜色空间。

4.2.1 RGB 颜色空间

RGB 颜色模型中，是由三个分为红色（R）、绿色（G）、蓝色（B）的颜色通道组成的

颜色，R、G、B 的值分别表示每个颜色通道的亮度值，所有的颜色都可由这三种基本色按照不同比例混合而成。

RGB 颜色空间模型在三维立方体进行描述时，三个坐标分量分别表示 R、G、B，原点则表示黑色，与原点立方体中对角线上的另一端则表示白色，灰度等级则是沿着这条体对角线分布的。因 R、G、B 三个分量的图像分别都是一幅 8 比特的图像，所以 RGB 图像都是具有 24 比特的深度。RGB 颜色空间模型如图 4-1 所示。

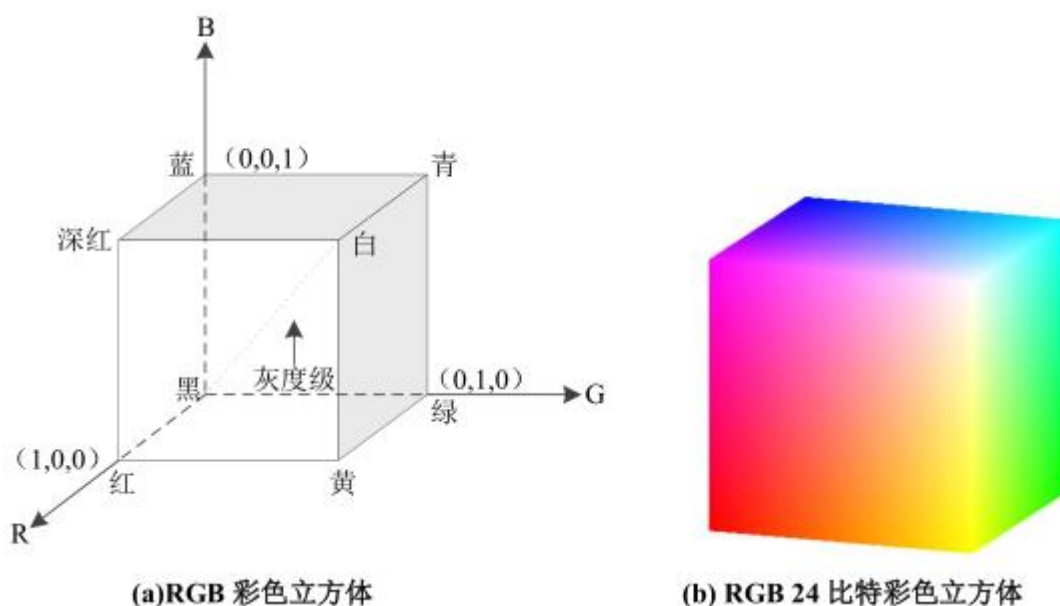


图 4-1 RGB 颜色空间模型

4.2.2 颜色直方图

颜色直方图^[12]做为颜色检索的一种，被广泛应用于各大图像检索系统，颜色直方图只关心图像中各种颜色所占的比例，对于图像中的对象则无动于衷。对于描述难于自动分割的图像使用颜色直方图是最为合适的。颜色直方图最常见于 RGB 颜色空间，因为大多数数字图像都是在 RGB 颜色空间基础上表达出来的。但是 RGB 颜色空间并不能直观的用肉眼直接看出，因此，有人提出了基于 HSV 颜色空间的颜色直方图。

4.2.3 HSV 颜色空间

HSV 空间是一种人类通过视觉进行感知色彩渐变过程的颜色空间。因 RGB 空间颜色变化不均匀，所以 HSV 空间有着 RGB 空间不具备的优点。它将颜色信号分为三种属性来表示，分别为色调 H (Hue)、饱和度 S (Saturation) 和亮度 V (Value)，简称 HSV^[13]。

色调 H 是用颜色的名称来辨别, 从红色开始按逆时针方向来进行计算的, 其中红色为 0° 绿色为 120° , 蓝色为 240° [14]。饱和度 S 可看成由其他颜色与白色混合的成果。饱和度 S 是用百分比来度量的, 取值范围为 $0\%\sim 100\%$, S 值越大, 颜色越饱和, 100% 为完全饱和。亮度 V 通常用百分比表达, 是表示颜色的亮暗程度, 若为 0% 则为最暗的黑色, 而 100% 是最亮的白色。

HSV 颜色空间能够直观的表达出颜色的亮暗, 色彩, 以及鲜艳程度等, 方便对颜色进行对比, 但不能直接不能直接将颜色值传达给电脑, 需要进行转化。

4.2.4 RGB 颜色空间到 HSV 颜色空间的转换

从人类的视觉上来讲, HSV 颜色空间中三个轴互不干扰, 能够直接对应人眼所关注到的色调、饱和度和亮度这三个要素, 因此 HSV 颜色空间相对于 RGB 颜色空间而言, 更有优势, 空间距离更符合人的视觉特征。而且 RGB 颜色空间模型与 HSV 模型之间有着相应的映射关系。因此可以将 RGB 颜色空间中的所有颜色转换到 HSV 颜色空间中, 这是人们所说的颜色转换 (Color Transformation)。也可以将 HSV 颜色模型转换成 RGB 模型上的颜色。但是 HSV 空间是均匀的, 所以由 RGB 空间转换 HSV 空间是非线性的。国外许多 CBIR 产品中都有运用到这种转换。以 RGB 颜色空间转换为 HSV 颜色空间为例, 如公式 (4-6) 所示。

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

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

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

公式 (4-6)

因为, S 和 V 都是用百分比表示, H 是使用角度表示的, $S \in [0, 1]$, $V \in [0, 1]$, 为了使 H 也能使用角度表示, 对 H 进行如公式 (4-7) 所示的处理, 使 $H \in [0^\circ, 360^\circ]$ 。RGB 颜色空间的每一点都能经过此公式进行转换, 得到 HSV 颜色空间。

公式 (4-7)

$$H = \frac{180^\circ \times H}{\pi}$$

5 图像检索

匹配算法作为现在研究基于内容的图像检索的重要研究方面之一，不管是在现有算法下对算法进行优化，还是研究出新算法，其主要目的都是为了实现图像检索及图像检索效果更尽如人意。本章主要介绍了欧氏距离和曼哈顿距离。

5.1 图像特征的相似性匹配

5.1.1 相似性匹配

图像检索的效果好坏在一定程度上是由匹配算法决定的。基于文本的检索方法中是使用文本对图像进行精确的匹配，基于内容的图像检索方法则是通过计算待测图像与图像库中的图像之间有关视觉特征的相似度来完成的，以一定量的数据训练来判断图像内容是否相关。通过对待测图片提取颜色、纹理等的视觉特征，再将其表示为多维的特征向量或者是向量集，就可以将转化后的特征向量之间的相似度来表示待测图像与图像库中图像的相似性，也就是计算两个向量之间的距离。在定义距离测度时应当满足自相似性、最小性、对称性、三角不等性等四个条件。

设现有 x 、 y 、 z 为三个特征向量，以 S 表示距离函数，那么所有的距离函数都能够满足以下条件：

- (1) $S(x, y) = S(y, x) = 0$ (自相似性)
- (2) $S(x, y) - S(x, x) = 0$ (最小性)
- (3) $S(x, y) = S(y, x)$ (对称性)
- (4) $S(x, y) + S(y, z) = S(x, z)$ (三角不等性)

在图像检索系统中，欧式距离、曼哈顿距离、马氏距离等方法经常会被用作距离度量。本文就是采用欧式距离和曼哈顿距离法。

5.1.2 欧式距离法

欧几里得度量 (euclidean metric) 也被称为欧式距离，本是有欧几里得家所发明，被称为欧氏距离，现如今被错误称为欧式距离。是最常用的度量方式^[15]。指的是在 N 维空间中两点之间的距离。定义如公式 (5-1) 所示。

$$L_2(A, B) = \left[\sum_{i=1}^n |a_{i1} - b_{i2}| \right]^{\frac{1}{2}}$$

公式 (5-1)

欧式距离公式不需要考虑各维之间的关系，各维重要性必须是相同的、所以使用范围受到了影响。

在一维中，欧氏距离就是两点之间的距离是其数值差的绝对值，也就是两点间的位移如公式（5-2）所示。

$$\sqrt{(x-y)^2} = |x-y| \quad \text{公式 (5-2)}$$

在二维平面上的两点 a (x1,y1) 和 b (x2,y2) 之间的欧式距离如公式（5-3）所示。

$$d_{12} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad \text{公式 (5-3)}$$

三维中，欧氏距离如公式（5-4）所示。

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \quad \text{公式 (5-4)}$$

欧式距离法是较为准确的相似度度量方法，但也有明显的缺点，它将样品的所有属性等同对待，即使是差异明显的不同属性也等同对待，有时候不能满足实际的要求。

5.1.3 曼哈顿距离法

曼哈顿距离又被称为城市街区距离，定义是在直角坐标系上两点 a, b 所形成的线段 ab 分别对所有轴所产生的投影的距离之和。在对于高维度矢量数据时，曼哈顿距离产生的效果比欧氏距离要好。

(1) 二维平面中假设 a (x1,y1) 与 b (x2,y2)，则两点的曼哈顿距离如公式（5-5）为。

$$d_{12} = |x_1 - x_2| + |y_1 - y_2| \quad \text{公式 (5-5)}$$

(2) 若在 n 维向量 a (x11,x12,...,x1n) 和 b (x21,x22,...,x2n) 中的曼哈顿距离如公式（5-6）所示。

$$d_{12} = \sum_{k=1}^n |x_{1k} - x_{2k}| \quad \text{公式 (5-6)}$$

6 仿真及分析

6.1 仿真实验

6.1.1 系统环境

本实验是在 Windows 7 32 位操作系统下进行的，安装内存 4GB。在 matlab2012a 版本中运用。matlab2012a 是在 2012 年发布的版本，相比较于之前的版本，所包含的函数更多，GUI 用户界面更友好。帮助系统更完善。所拥有的功能更强大。

6.1.2 系统功能描述

本实验中采用平时常见的 60 张图像进行小检索，分别是 20 张婴儿车类，20 张鞋类，20 张糕点类图像。运用 matlab2012a 版本完成基于内容的图像检索系统。

该系统主要实现以下几个功能：

- (1) 导入并生成图像特征库。
- (2) 对待测图像进行基于颜色特征的欧式距离检索。
- (3) 对待测图像进行基于颜色特征的曼哈顿距离检索。
- (4) 对待测图像进行基于纹理特征的欧式距离检索。
- (5) 对待测图像进行基于纹理特征的曼哈顿距离检索。

6.1.3 系统界面效果及实现

设计主界面如图 6-1 所示。包括建立图像库，选择图像，颜色特征提取检索，纹理特征检索，欧式距离，曼哈顿距离几部分。

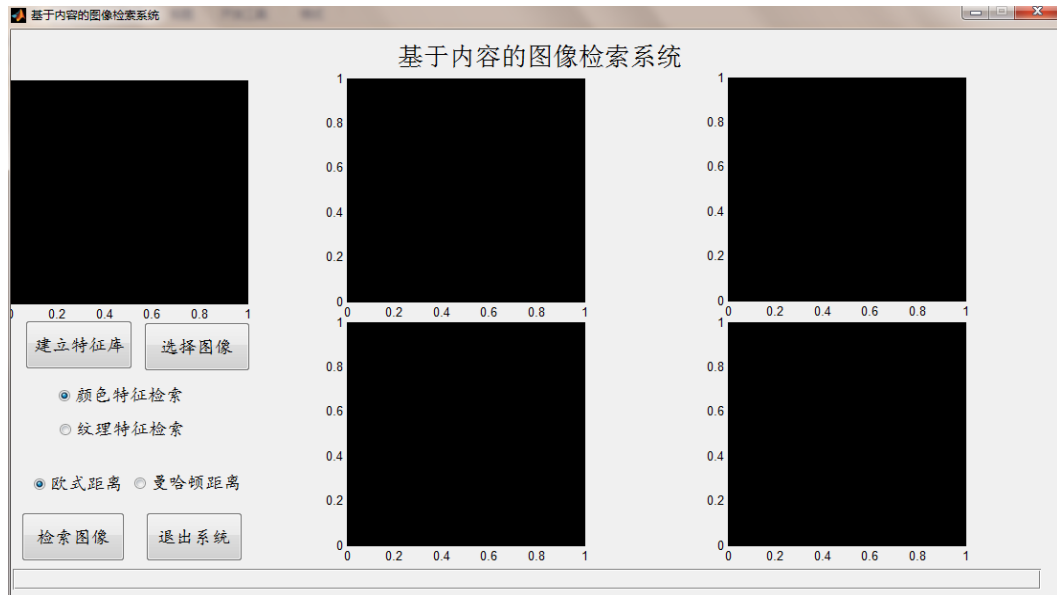


图 6-1 设计主界面

点击建立图像库按钮，执行建立特征库命令，结果如图 6-2 所示。

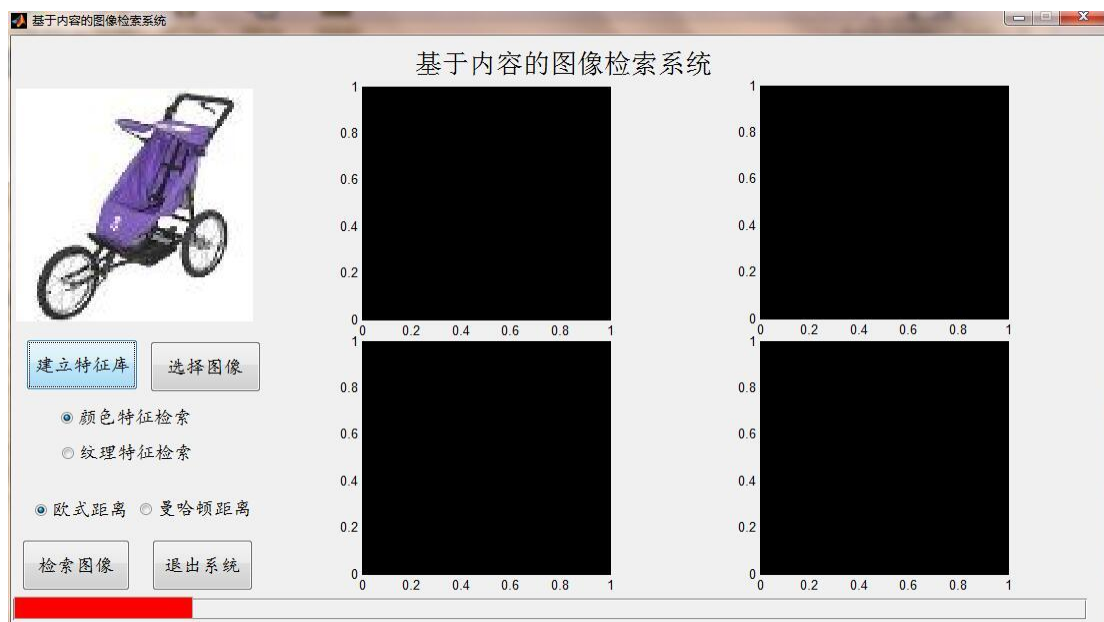


图 6-2 建立特征库

点击选择图像按钮，选取待测图像：鞋子类的白鞋和红鞋。
白鞋基于颜色特征下的欧式距离检索结果如图 6-3 所示。



图 6-3 白鞋图像基于颜色的欧式距离检索结果

红鞋基于颜色特征下的欧式距离检索结果如下图 6-4 所示。



图 6-4 红鞋图像基于颜色的欧式距离检索结果

白鞋图像基于颜色特征下的曼哈顿距离检索结果如图 6-5 所示。



图 6-5 白鞋图像基于颜色的曼哈顿距离检索结果

红鞋基于颜色特征下的曼哈顿距离检索结果如下图 6-6 所示。



图 6-6 红鞋图像基于颜色的曼哈顿距离检索结果

更换待测图像为婴儿车类中的蓝黑车和红灰车，分别进行基于颜色的检索。蓝黑车基于颜色特征下的欧式距离检索结果如下图 6-7 所示。



图 6-7 蓝黑车基于颜色特征的欧氏距离检索结果

红灰车基于颜色特征下的欧式距离检索结果如下图 6-8 所示。



图 6-8 红灰车基于颜色特征的欧氏距离检索结果

蓝黑车基于颜色特征下的曼哈顿距离检索结果如下图 6-9 所示。



图 6-9 蓝黑车基于颜色特征下的曼哈顿距离检索结果

红灰车基于颜色特征下的曼哈顿距离检索结果如下图 6-10 所示



图 6-10 红灰车基于颜色特征下的曼哈顿距离检索结果

经过鞋类与婴儿车类对比，基于颜色特征的图像检索实验已基本实现。

更换图像，验证图像基于纹理特征的图像检索。

图像在基于纹理特征下的欧式距离检索结果如图 6-11 所示。



图 6-11 图像欧式距离检索结果

图像基于纹理特征下的曼哈顿距离检索结果如图 6-12 所示。



图 6-12 图像曼哈顿距离检索结果

经过大量实验后，将结果比对，包括对于鞋，婴儿车，蛋糕等等的检索，比较使用基于颜色检索和基于纹理检索时的效果，发现基于颜色的检索效果要明显优于基于纹理的检索。在与基于纹理的检索比较中，发现基于颜色的检索在 3 个类别的图像中，有 2 个类别的检索性能有较大的优势，但是对于类别蛋糕类中的图像，基于颜色检索的性能明显要低于其它类别，甚至低于基于纹理检索的性能。实验内容基本完成。

结论与展望

本文研究的偏重点是在基于内容的图像检索算法,根据待测图像内容的差别,分别研究了基于颜色特征和纹理特征,分别运用欧氏距离检索算法和曼哈顿距离检索算法。并对其中涉及到的颜色模型、颜色直方图、灰度共生矩阵等进行了较深入的研究。

在对基于内容的图像检索方法的研究基础上,本文采用 matlab 语言编写了一个简单的图像检索演示程序,实验证明,该程序能较为精准的得到所需的图片,实现了图像检索的基本功能。但仍存在不足:

- (1) 仅仅使用单一的特征参数进行检索,得到的相似度结果有时并不尽如人意。
- (2) 使用欧式距离和曼哈顿距离进行检索时,若图像库越大,所需要的时间则越久。
- (3) 今后可以考虑将多种特征参数进行结合,从而提高准确率。

参考文献

- [1] 罗志灶.基于颜色的图像检索[J].福建电脑,2004,12:22-23.
- [2] 王涛.基于内容的图像检索[D].武汉:华中科技大学电子与通信工程专业,2013.
- [3] 师鸣若.基于内容的视频检索系统研究[D]. 西安:西安理工大学信号与信息处理专业,2003.
- [4] 何剑非.基于内容的图像检索系统的研究与设计[D]. 沈阳:沈阳工业大学,2011.
- [5] 姚争儿,李志奎.基于内容的图像检索算法研究[J].商情,2013,6:87.
- [6] 王伟光,解成俊.基于内容的图像检索技术探索[J].产业与科技论坛,2014,13(1):82-83
- [7] 纪元法,孙希延,施浒立.[MATLAB 的两种图形用户界面参数输入方法](#)[J].电脑知识与技术,2008,2(5):728-731.
- [8] 陆光,满庆丽.林木害虫天牛的图像检索方法研究[J].森林工程,2013,5,29(3):71-75.
- [9] 杨浩,罗扬.基于灰度共生矩阵的边缘检测方法[J].智能计算机与应用,2012,12,2(6):43-45.
- [10] 步亚东.图像纹理特征提取的研究[D].济南:山东师范大学计算机软件与理论专业,2012.
- [11] 陈美龙,戴声奎.基于 GLCM 算法的图像纹理特征分析[J].通信技术,2012,45(2):108-111.
- [12] 朱佳丽,李士进,万定生,冯钧.基于特征选择和半监督学习的遥感图像检索[J],中国图象图形学报,2011,16(8):1474-1482.
- [13] 董广宇.基于内容的图像检索研究[D]. 北京:北京信息科技大学计算机应用技术专业,2008.
- [14] 谭阳.基于脑波的图像情感特征映射的研究[D].上海:东华大学计算机应用技术专业,2012.
- [15] 崔瑞苹.数学学习笔记初探[J].数学学习与研究,2012,7.
- [16] Gang Zhang and Zongmin Ma.Texture Feature Extraction and Description Using Gabor Wavelet In Content Medical Image Retrieval[C].Proceedings of Wavelet Analysis and Pattern Recognition ICWAPR International Conference,2007:1:169-173.