

黑白图像彩色化技术

摘要

黑白图像彩色化技术在黑白照片和电影彩色化等领域有着广泛应用，一套优秀的图像彩色化算法的评价标准是能够在尽量少的人工干预下产生自然的彩色化效果。当前的黑白图像彩色化技术需要用户耗费大量时间给出 scribble，并且物体边缘的色彩渗透问题也很难得到解决。本文主要研究自动 scribble 产生技术和基于四元代数的图像结构保持机制。首先我们在空间分布熵极值点上自动产生 scribble。在这些彩色 scribble 的指导下，我们计算四元小波相位来引导色彩在跨越图像结构相似区域的等相位线上蔓延。对于物体边缘，我们建立了尺度空间边缘强度模型来约束色彩在同一边缘上蔓延。最后，我们将颜色作为向量在四元数极坐标色彩表示下进行重构，这样能很好地保持色彩通道之间的内在关系。实验证明使用自动产生的 scribble 和本文提出的彩色化方法能够在极少的人工输入下实现自然的彩色化效果。

关键词：彩色化，自动 scribble 产生，结构保持机制，四元代数

IMAGE COLORIZATION TECHNIQUES

ABSTRACT

Image colorization techniques are widely used in the applications of monochrome photo colorization and movie colorization. The evaluation criterion for an outstanding colorization algorithm is whether it can achieve prospective effect requiring as little manual input as possible. In current colorization techniques, major user intervention is required in the form of tedious, time-consuming scribble drawing. Moreover, color leakage usually occurs across contours and object boundaries. In this paper, we focus on automatic scribble generation and structure-preservation mechanism. Firstly, we generate scribbles automatically along points where the spatial distribution entropy achieves locally extreme value. Given the color scribbles, we compute quaternion wavelet phases to conduct colorization along equal-phase lines. These lines across scribbles and monochrome patches locate textures with similar pattern distribution. Contour 'strength' model is also established in scale space to direct color propagation among similar edge structures. Finally, we reconstruct color image patches as vector elements using polar representation in quaternion algebra, well-preserving interrelationship between color channels. The experimental results demonstrate that the proposed colorization method can achieve natural color transitions between different objects with automatically generated scribbles.

Key words: colorization, automatic scribbles generation, structure-preservation mechanism, quaternion algebra

目 录

第一章 绪论	14
1.1 黑白图像彩色化的概述	2
1.2 黑白图像彩色化算法的研究现状	2
1.2.1 基于 scribble 的彩色化技术	2
1.2.2 基于示例的彩色化技术	33
1.3 本文工作与章节结构安排	33
第二章 自动 scribble 生成算法	55
2.1 空间分布熵理论	55
2.1.1 图像空间划分	55
2.1.2 空间分布熵	55
2.2 scribble 搜索策略	66
2.2.1 算法构思	66
2.2.2 基于图论的图像分割技术	66
2.2.3 算法详细流程	88
2.3 本章小结	99
第三章 基于四元代数的图像彩色化	1040
3.1 四元代数概念	1040
3.1.1 定义	1040
3.1.2 性质	1040
3.1.3 群旋转	1040
3.1.4 以矩阵表示四元数	1144
3.1.5 四元数运算	1144
3.1.6 四元数的极坐标表示	1242
3.2 利用四元 Gabor 相位分析图像结构	1242
3.2.1 四元小波变换演进	1242
3.2.2 使用四元 Gabor 相位衡量结构相似性	1343
3.2.3 相位卷绕问题	1343
3.2.4 PCA 降维处理	1444
3.3 尺度空间结构优先级机制	1545
3.3.1 Canny 边缘检测	1545
3.3.2 尺度空间边缘强度模型	1646
3.4 全局最优色彩蔓延	1747
3.4.1 全局最优化算法介绍	1747
3.4.2 四元数色彩重构	2020
3.5 本章小结	2020
第四章 实验结果及用户界面	2222
4.1 彩色化仿真结果	2222
4.2 交互式选色界面	2424
第五章 结论	2626



参考文献.....	2727
谢辞.....	2929

第一章 绪论

数字图像处理就是利用计算机对图像信息进行加工以满足人的视觉心理或者应用需求的行为。实质上是一段能够被计算机还原显示和输出为一幅图像的数字码。21 世纪是一个充满信息的时代，图像作为人类感知世界的视觉基础，是人类获取信息、表达信息和传递信息的重要手段。数字图像处理，即用计算机对图像进行处理，其发展历史并不长。数字图像处理技术源于 20 世纪 20 年代，当时通过海底电缆从英国伦敦到美国纽约传输了一幅照片，采用了数字压缩技术。首先数字图像处理技术可以帮助人们更客观、准确地认识世界，人的视觉系统可以帮助人类从外界获取 3/4 以上的信息，而图像、图形又是所有视觉信息的载体，尽管人眼的鉴别力很高，可以识别上千种颜色，但很多情况下，图像对于人眼来说是模糊的甚至是不可见的，通过图像增强技术，可以使模糊甚至不可见的图像变得清晰明亮。另一方面，通过数字图像处理中的模式识别技术，可以将人眼无法识别的图像进行分类处理。通过计算机模式识别技术可以快速准确的检索、匹配和识别出各种东西。

数字图像处理技术已经广泛深入地应用于国计民生休戚相关的各个领域。

在计算机中，按照颜色和灰度的多少可以将图像分为二值图像、灰度图像、索引图像和真彩色 RGB 图像四种基本类型。目前，大多数图像处理软件都支持这四种类型的图像。

一幅二值图像的二维矩阵仅由 0、1 两个值构成，“0”代表黑色，“1”代表白色。由于每一像素（矩阵中每一元素）取值仅有 0、1 两种可能，所以计算机中二值图像的数据类型通常为 1 个二进制位。二值图像通常用于文字、线条图的扫描识别（OCR）和掩膜图像的存储。

灰度图像矩阵元素的取值范围通常为[0, 255]。因此其数据类型一般为 8 位无符号整数的（int8），这就是人们经常提到的 256 灰度图像。“0”表示纯黑色，“255”表示纯白色，中间的数字从小到大表示由黑到白的过渡色。在某些软件中，灰度图像也可以用双精度数据类型（double）表示，像素的值为[0, 1]，0 代表黑色，1 代表白色，0 到 1 之间的小数表示不同的灰度等级。二值图像可以看成是灰度图像的一个特例。

索引图像的文件结构比较复杂，除了存放图像的二维矩阵外，还包括一个称之为颜色索引矩阵 MAP 的二维数组。MAP 的大小由存放图像的矩阵元素值域决定，如矩阵元素值域为[0, 255]，则 MAP 矩阵的大小为 256 X 3，用 MAP=[RGB]表示。MAP 中每一行的三个元素分别指定该行对应颜色的红、绿、蓝单色值，MAP 中每一行对应图像矩阵像素的一个灰度值，如某一像素的灰度值为 64，则该像素就与 MAP 中的第 64 行建立了映射关系，该像素在屏幕上的实际颜色由第 64 行的[RGB]组合决定。也就是说，图像在屏幕上显示时，每一像素的颜色由存放在矩阵中该像素的灰度值作为索引通过检索颜色索引矩阵 MAP 得到。索引图像的数据类型一般为 8 位无符号整形（int8），相应索引矩阵 MAP 的大小为 256 X 3，因此一般索引图像只能同时显示 256 种颜色，但通过改变索引矩阵，颜色的类型可以调整。索引图像的数据类型也可采用双精度浮点型（double）。索引图像一般用于存放色彩要求比较简单的图像，如 Windows 中色彩构成比较简单的壁纸多采用索引图像存放，如果图像的色彩比较复杂，就要用到 RGB 真彩色图像。

RGB 图像与索引图像一样都可以用来表示彩色图像。与索引图像一样，它分别用红（R）、绿（G）、蓝（B）三原色的组合来表示每个像素的颜色。但与索引图像不同的是，RGB 图像每一个像素的颜色值（由 RGB 三原色表示）直接存放在图像矩阵中，由于每一像素的颜色需由 R、G、B 三个分量来表示，M、N 分别表示图像的行列数，三个 M x N 的二维矩阵分别表示各个像素的 R、G、B 三个颜色分量。RGB 图像的数据类型一般为 8 位无符号整形，通常用于表示和存放真彩色图像，当然也可以存放灰度图像。

数字化图像数据有两种存储方式：位图存储(Bitmap)和矢量存储(Vector)

我们平常是以图像分辨率（即像素点）和颜色数来描述数字图像的。例如一张分辨率为 640×480 , 16 位色的数字图片，就由 $2^{16}=65536$ 种颜色的 $307200(=640 \times 480)$ 个素点组成。

位图图像：位图方式是将图像的每一个像素点转换为一个数据，当图像是单色（只有黑白二色）时，8 个像素点的数据就占据一个字节（一个字节就是 8 个二进制数，1 个二进制数存放像素点）；16 色（区别于前段“16 位色”）的图像每两个像素点用一个字节存储；256 色图像每一个像素点用一个字节存储。这样就能够精确地描述各种不同颜色模式的图像图面。位图图像弥补了矢量式图像的缺陷，它能够制作出色彩和色调变化丰富的图像，可以逼真地表现自然界的景象，同时也可以很容易地不同软件之间交换文件，这就是位图图像的优点；而其缺点则是它无法制作真正的 3D 图像，并且图像缩放和旋转时会产生失真的现象，同时文件较大，对内存和硬盘空间容量的需求也较高。位图方式就是将图像的每一个像素点转换为一个数据。如果用 1 位数据来记录，那么它只能代表 2 种颜色（ $2^1=2$ ）；如果以 8 位来记录，便可以表现出 256 种颜色或色调（ $2^8=256$ ），因此使用的位元素越多所能表现的色彩也越多。通常我们使用的颜色有 16 色、256 色、增强 16 位和真彩色 24 位。一般所说的真彩色是指 24 位（ 2^{24} ）的位图存储模式适合于内容复杂的图像和真实照片。但随着分辨率以及颜色数的提高，图像所占用的磁盘空间也就相当大；另外由于在放大图像的过程中，其图像势必要变得模糊而失真，放大后的图像像素点实际上变成了像素“方格”。用数码相机和扫描仪获取的图像都属于位图。

1.1 黑白图像彩色化的概述

彩色化这个术语是在 1970 年 Wilson Markle 第一次提出，用来描述他发明的一种计算机辅助的给黑白电影和电视节目着色的技术。

黑白图像彩色化技术是在一些人工提示的帮助下实现在黑白图像上自然地添加色彩的技术。这项技术的基本思想是根据用户提供的色彩提示找到从标量灰度图像到矢量彩色图像的映射。由此可见彩色化问题是一个 ill-posed problem，彩色图像并不存在，而且用户提供的色彩提示中的信息也不足以确定一幅彩色图像。那么黑白图像彩色化算法的目标就是利用有限的色彩信息使彩色化结果自然、逼真。

在计算机辅助彩色化算法问世之前，彩色化工作最大的困难就是昂贵且费时。比如，想要彩色化一个静态图片，一个艺术家需要先把图像进行分割成一个一个的区域然后给每个区域进行着色。不幸的是自动分割算法的结果往往不尽人意，尤其在复杂的图像结构中，比如人的毛发和面部。所以现在的计算机辅助彩色化算法力求在任何复杂的情况下都能避免繁重的人工输入而取得理想的效果。

1.2 黑白图像彩色化算法的研究现状

1.2.1 基于 scribble 的彩色化技术

在基于 scribble 的图像彩色化技术中，大多数研究人员都从 Levin 等人的工作^[1]的启发，假设在空间上临近的像素如果灰度值相近那么就应该具有相近的颜色。这样一来，在灰度值相近但是图像结构有明显变化的地方就容易发生色彩泄露，原因是缺乏结构保持机制。Drew 等人为了得到自然的彩色化效果，假设彩色化后的图像的对比度应该与原始灰度图像的梯度相匹配^[2]。为了产生更稳定的边缘保持彩色化结果，Kim 等人使同一边缘结构内的色彩一致^[3]。Lezora 把彩色化过程建模成一个图论约束问题，取得良好彩色化效果的前提是用户需要提供更加细致复杂的 scribble 提示^[4]。Heu 建立了色彩优先蔓延机制，保证图像关键结构区域优先被着色^[5]。在这些工作中的两个最基本的问题就是可靠的 scribble 指导结构保持机制。

1.2.2 基于示例的彩色化技术

近些年,一些作者提出了基于示例的图像彩色化技术。要彩色化一张黑白目标图像需要用户提供多张彩色参考图像。Welsh^[6]等人通过匹配目标图像和参考图像的光照和纹理信息来把色彩从参考图像上转移到目标图像上。Irony^[7]摒弃了依赖一系列像素级的决策标准,开发了一个新的能够利用像素的到高层次内容的策略,这个策略使用了空间一致性限制。X.P. Liu 专注于利用分离图像的反射分量和光照分量来解决目标图像和参考图像之间存在的光照差^[8]。Sykora 等人^[9]设计了一个特别适合于卡通彩色化的彩色化算法,这个算法能够在彩色化之前很好的分割卡通图像。

在上述方法中,彩色化的效果极大地依赖参考图像与目标图像的相似度。而且参考图像数据库可能很难得到。

1.3 本文工作与章节结构安排

本文依然着眼于基于 scribble 的图像彩色化技术,同时将人工绘制 scribble 的过程自动化。在彩色化过程中,本文很好地保持了图像的结构特征。图 1-1 展示了本文彩色化算法的主要步骤以及彩色化效果,(a)输入待彩色化的灰度图像(b)将图像过分割成若干闭合区域(c)自动产生 scribble 供用户选色(d)用户确定 scribble 的色调(e)彩色化结果(f)彩色原图,用于与彩色化效果进行比较。

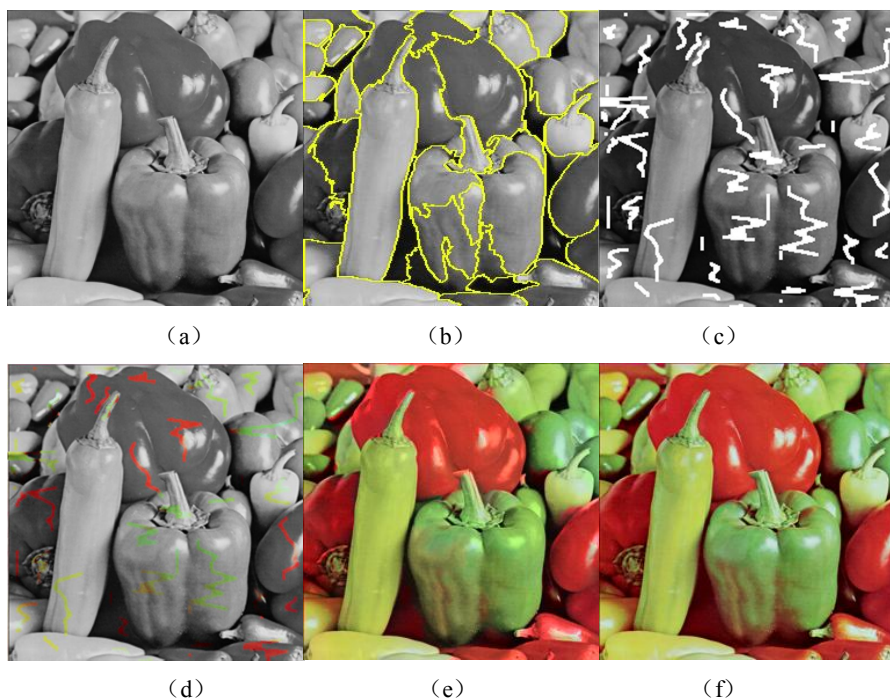


图 1-1 本文彩色化算法操作流程

本文提出的算法能够自动地利用控件分布熵理论在输入的灰度图像上产生可靠精确的 scribble。这样用户只需要为已经产生的 scribble 选择色彩即可。当完成彩色 scribble 之后,我们计算灰度图像的四元小波相位来引导色彩在等相位线上传播。这些等相位线穿过灰度图像中相似的纹理区域。在尺度空间中,我们提出一种基于边缘强度的结构优先机制模型来引导颜色在不同边缘上有不同的蔓延优先级。最后,在四元数色彩表示的极坐标形式下重构整张图片的颜色,这样避免了在彩色化过程中分离各个色彩通道造成的色彩失

真。我们还为彩色化算法设计了交互式用户界面。

文章总体结构安排如下：第二章详细介绍了基于空间分布熵理论和基于空间分布熵理论的自动 scribble 搜索策略；第三章描述了我们的彩色化算法是怎样利用自动产生的彩色 scribble 提示完成彩色化过程的。本章首先介绍了四元代数理论，这是本文分析图像结构和色彩重构使用的数学基础。第二小节介绍了如何使用四元数 Gabor 小波相位描述图像结构来引导色彩蔓延。第三小节介绍了尺度空间下的结构优先级机制。第四小节介绍了如何利用上述所有线索在四元数色彩空间下使用全局最优化算法重构图像色彩。第四章给出了实验结果供视觉评价以及交互式用户界面。第五章对整个工作进行了总结得出了结论和未来努力的方向。

第二章 自动 scribble 生成算法

Scribble 在彩色化技术中非常重要，它包含了彩色化过程所需的所有色彩信息。然而具有复杂结构的图像需要有经验的用户提供细致的 scribble 才能得到令人满意的着色效果。本节我们提出一种基于空间分布熵的自动 scribble 产生算法，将 scribble 产生在具有高信息密度的区域。

2.1 空间分布熵理论

2.1.1 图像空间划分

设 $(P_{xy})_{C \times R}$ 是大小为 $C \times R$ 的一幅图像，这里 P 表示像素 (x, y) 的颜色。假设彩色图像某一颜色通道对应的颜色数为 N 。令 S_q 为颜色值为 B_q 的像素的集合。设 C 为 S_q 的质心。在对图像区间的划分上，采用等距划分的方法，对于给定的一个正整数 M ，把半径分为 M 等级，然后以质心 C 为圆心画圆可得到 M 个圆环。

一般来说，具有相同颜色的像素点围绕其质心的分布密度会随着它们同质心间距离的增大而逐渐减小，也就是说，离质心的距离越远，具有相同颜色的像素的分布越稀疏，因此，文中对图像区间的划分时采用以质心 C 为圆心，以 d 为半径画圆，其中 $d=2, 3, \dots$ 。我们把这样一个规定范围的经分割后的圆形区域称作一个 scope，如图2-1所示。

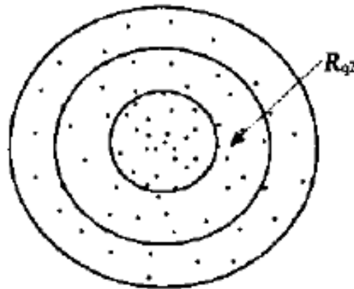


图2-1 scope 模型

设圆 C_q 共被分割为 M 个圆环，以 $|R_{ij}|$ 表示 S_i 的第 j 个圆环中颜色为 B_i 像素的总个数，其中 $1 \leq i \leq N$ ， $1 \leq j \leq M$ 。由此得到的直方图 $(|R_{i1}|, |R_{i2}|, \dots, |R_{iM}|)$ 称为颜色 B_i 的空间颜色直方图（当 $d=1$ 时为传统的直方图）。

记 $P_{ij} = |R_{ij}| / |R_i|$ 表示颜色 B_i 在环 j 中分布的概率密度，其中 $|R_i|$ 表示 S_i 中颜色为 B_i 的所有像素的个数。

2.1.2 空间分布熵

颜色 i 的空间分布熵可表示为

$$E_p = - \sum_{m=1}^M \sum_{v=1}^{V_Q} \frac{|s_{vm}|}{|U_m|} \times \log_2 \frac{|s_{vm}|}{|U_m|} \quad (2-1)$$

即在 P 点所对应的环形统计范围中对灰度值为 V 的像素数量与该区域中的像素总数量的比与其对数的乘积进行求和。

由以上空间分布熵的概念容易看出，引入环形中心空间结构的信息熵的意义在于，对于给定某种颜色，其对视觉的刺激大小很大程度上取决于该颜色在空间分布上的集中或分散程度，分布越集中则对视觉刺激越大，反之则刺激较小，引入上述模型，由(2-1)式可以看出，当空间分布熵值越大，则颜色分布越均匀，越分散，而较小的空间分布熵则表征了该颜色集中分布的区域。根据这种性质我们可以将这种模型引入 scribble 定位所用。

2.2 scribble 搜索策略

2.2.1 算法构思

标色的意义不仅仅在于提供一定的颜色信息，要使有限的标色信息产生尽可能多的彩色效果，我们可以将其理解为找到最关键颜色信息所在的区域。因此 scribble 自动定位的产生标准，也就是探索某种合理的能够表征每个像素点携带信息量大小的度量标准。

由以上空间分布熵的概念，容易联想到，应该对于那些空间分布熵较小的像素点集周围给予更多的颜色信息，使之更精细地上色，并且向灰度平坦的区域——也就是分布熵较大的区域进行色彩蔓延，这样可以优化彩色化的效果。

为了避免一条 scribble 线条横跨不同的同态区域，灰度图像要使用基于图论的图像分割算法^[10]过分割成若干闭合区域。然后我们在每个闭合区域中应用一个搜索策略选取周围信息密度最大的一些点。我们用空间分布熵来计算信息密度。最后，在这些被选取的点上插值形成连续的 scribble。

为保证有限的 scribble 能为彩色化提供足够的色彩信息，我们将 scribble 产生在具有高空间分布熵的区域。图 2-2 直观地展示了这个过程。(a) 搜索过程在一个分割闭合区域（红色区域）内进行，(b) 自适应的 scope 半径是和该点到区域边界的最短距离线（红色箭头）性相关的 (c) 每移动一步都有 8 个候选 scope，搜索路径有 8 个自由度 (d) 8 个方向的空间分布熵直方图 (e) 用光滑的曲线连接被选中的点产生 scribble。

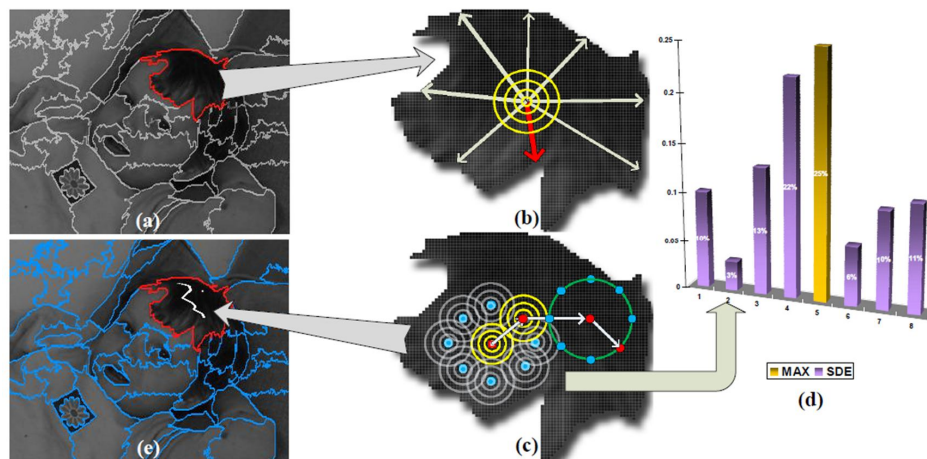


图 2-2 自动 scribble 产生算法细节

2.2.2 基于图论的图像分割技术

图像分割技术在一项重要的图像技术中，在许多的研究领域和实际使用中应用的非常广泛，得到了人们越来越多的重视。图像分割的一个总的原则是将一个图像分割成若干不同的区域，但是分割图像的技术的方法种类却有很多，没有一个唯一的界定和标准，有些算法可以直接对应用图像进行分割，而有些分割技术先要进行一个大致的分割，然后从中提取信息后再进行分割，所以对一个图像要采取何种的分割方法直接影响着分割结果的

效果,它是在图像处理和图像分析之间的关键步骤,往往受到具体场合的影响,也影响着人们对于这幅图像的理解。

在目前有许多图像分割算法都在应用领域获得了长足的发展,比如有阈值法,区域增长法,边缘检测法,模糊方法和聚类法等,是建立在图像相似性上面的;

图论[Graph Theory]是数学的一个分支。它以图为研究对象。图论中的图是由若干给定的点及连接两点的线所构成的图形,这种图形通常用来描述某些事物之间的某种特定关系,用点代表事物,用连接两点的线表示相应两个事物间具有这种关系。基于图论的图像分割算法由于具有对于数据类的性质的不敏感性,同时聚类的效果具有很好的通用性等特点越来越受到广大研究者的青睐,通过应用图论方法来解决图像处理问题,最一般的方法主要是最小生成树算法。

论图像分割算法实现:

首先采用了从仿真角度分析图像在不同 FRFT 域的能量分布特点,通过归一化方法剩余误差因子 P 评估和分析 FRFT 域的能量积聚性和图像所包含的信息,然后采用区域生成法对图像首先进行第一次初始分割,主要的思想如下:

将一些比较相似的像素集合起来放在同一个区域内,然后具体的对每一个需要分割的区域找个统一的像素来作为生长的起点,接着将种子像素周围的领域内的与该种子像素具有相似或者相同的性质的像素一起合并到种子像素所在的区域中,最后将这个最新的像素当成是种子像素,直至没有满足条件的像素为止。对于种子点的选择,主要是根据图像的饱和度来首先确认一组种子像素,采用图论中八邻域联通方案来搜索邻接的像素,给定门限值 Th ,若领域像素与种子的像素相似度小于这个门限值,那么该像素所包括了该种子像素所在的区域,假如该种子相似度大于这个门限值 Th ,那么终止生长。

最小生成树区域分割:

在采用区域生长过后,图像中的所有的特征信息都分布在不同的区域中,下面采用最小生成树算法对区域图像进行分割,最后采用二值化方法对图像进行锐化。

首先把初始分割后的每个区域看成是一个无向图 G 的一个节点,这样方便直接对图像的像素点直接操作。节约了算法的时间。权函数定义的是两个节点之间的相似度,这里考虑到图像的不同特征性质,把性质相同的归为一个区域,设为同一类。

最小生成树图像分割:

无向图 $G=[X, E]$ 具有 N 个节点的网络,设 U 作为生成树顶点的集合,本文采用 Prim 算法来求解最小生成树。

目标函数的建立:

图像的数据集确定后,由生成树得到一个以数据样本作为无向图顶点的最小生成树,该 $D-MST$ 的边权所反映的是数据样本之间的相似程度。根据最小生成树的子树和强大的类间类内的距离比图像质量评价准确,确定一个目标函数来评价每个不同的分类状态,目标函数定义如下:

$$H_t = \frac{\sum_{i=1}^m P_i(m_i - m)^2}{\sum_{i=1}^m \sum_{j=1}^t (x_i - m_i)^2 h_i} \quad (2-2)$$



图 2-3 基于图论的图像过分割

基于图论的图像分割算法能把图像分割成若干闭合的区域，我们将在每个区域中搜索 scribble。

2.2.3 算法详细流程

以下算法流程是在一个闭合的分割区域内实施的。

首先选择 scribble 搜索的起始点，这个起始点的位置将决定 scribble 搜索路径的伸展空间，为了使伸展空间达到最大，我们选择区域的中心作为起始点，起始点坐标值即该区域所有像素点坐标值的加权平均。

整个路径搜索策略参考图 2-4 所示，我们采用回溯的搜索方法来尽量避免引某一步不明确的移动影响找到最佳 scribble 路径。当 8 个方向的候选 scribble 都无法满足以下两个判定条件时，则选择回溯重新定向第二候选 scope 继续搜索。

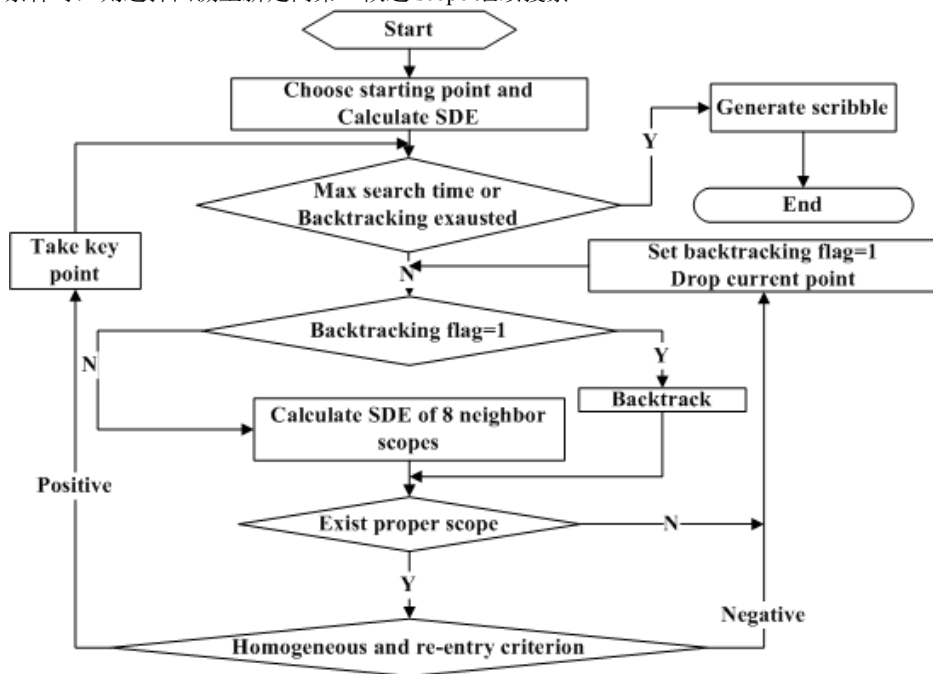


图 2-4 自动 scribble 产生搜索策略流程图

同态性判定条件：我们对当前scope和候选scope的相关性系数与之前路径上所有点的相关性系数均值作比较，当比值小于某个比例时则判定为不满足同态性判定条件。这样做的目的是让scribble产生在一个同态性区域中。设探索路径已经历M+1个scope，其间的相邻scope间的相关性系数为 $R=\{r_1, r_2, \dots, r_M\}$ ，当前的系数为 r_{M+1} ，需要取阈值V，依照(2-3)式判定是否满足同态限定。

$$r_{M+1} \geq \frac{V}{M} \sum_{i=1}^M r_i \quad (2-3)$$

当回溯路径穷尽时搜索结束。最后将搜索路径上的所有点拟合为光滑曲线。二维空间散点与曲线之前没有一一映射的关系因为点可以以不同的顺序连接，在此我们不失一般性地选择沿x方向依次拟合这些点，并对x值重合的点做y方向微调后做三次差值拟合。

2.3 本章小结

本章先介绍了 scribble 搜索算法的设计思路，然后按搜索流程详细介绍了如何实现算法构思。

基于空间分布熵的 scribble 能够定位在图像信息密度大的区域，也就是说能够在需要提供关键色彩信息的地方产生 scribble。这样在方便用户提供色彩信息的同时又能保证后面的色彩蔓延过程能有足够的色彩信息。

同时，由于每条 scribble 产生在一个闭合的分割区域中并且每条 scribble 只能具有一个颜色，所以图像分割时要将图像进行过分割。但是分割的区块面积太小会导致 scribble 数量增加，增加用户负担。如果分割区域面积过大则可能导致一条 scribble 跨越了不同颜色的区域，所以图像分割的精细程度要适中，但分割阈值的设定目前还需要人工给出没有完全做到自动估计。

第三章 基于四元代数的图像彩色化

根据给定的彩色 scribble，彩色化算法需要将 scribble 上的颜色按照图像结构自然地蔓延到图像上的每个区域。这个任务首先需要找到能够描述图像结构的分析方法，这些方法要使色彩在同态区域内蔓延并约束不同色彩被限制在不同的边缘结构两侧。最后要保证正确的色彩附着在对应的位置且色彩的变化清晰自然。

本章首先介绍如何用超解析信号表示分析图像样式和尺度空间下的结构优先级。然后研究怎样利用优化过程使颜色在相似的图像结构中生长。

3.1 四元代数概念

四元数是由爱尔兰数学家威廉·卢云·哈密顿在 1843 年发现的数学概念。四元数的乘法不符合交换律。

明确地说，四元数是复数的不可交换延伸。如把四元数的集合考虑成多维实数空间的话，四元数就代表着一个四维空间，相对于复数为二维空间。

3.1.1 定义

复数是由实数加上元素 i 组成，其中

$$i^2 = -1 \quad (3-1)$$

相似地，四元数都是由实数加上三个元素 i 、 j 、 k 组成，而且它们有如下的关系：

$$i^2 = j^2 = k^2 = ijk = -1 \quad (3-2)$$

每个四元数都是 1 、 i 、 j 和 k 的线性组合，即是四元数一般可表示为 $a+bi+cj+dk$ 。

要把两个四元数相加只需将相类的系数加起来就可以，就像复数一样。至于乘法则可跟随以下的法则：

*	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

图 3-1 四元数运算法则

四元数的单位元的乘法构成了八阶四元群。

3.1.2 性质

四元数不像实数或复数那样，它的乘法是不可交换的。

四元数是除法环的一个例子。除了没有乘法的交换律外，除法环与域是相类的。特别地，乘法的结合律仍旧存在、非零元素仍有唯一的逆元素。

四元数形成一个在实数上的四维结合代数（事实上是除法代数），并包括复数，但不与复数组成结合代数。四元数（以及实数和复数）都只是有限维的实数结合除法代数。

四元数的不可交换性往往导致一些令人意外的结果，例如四元数的 n -阶多项式能有多于 n 个不同的根。

3.1.3 群旋转

如四元数和空间转动条目所释，非零四元数的乘法群在 \mathbb{R}^3 的实部为零的部分上的共轭作用可以实现转动。单位四元数（绝对值为 1 的四元数）若实部为 $\cos(t)$ ，它的共轭作用是

一个角度为 $2t$ 的转动，转轴为虚部的方向。四元数的优点是：

- (1) 表达式无奇点（和例如欧拉角之类的表示相比）
- (2) 比矩阵更简练（也更快速）
- (3) 单位四元数的对可以表示四维空间中的一个转动。

所有单位四元数的集合组成一个三维球 S^3 和在乘法下的一个群（一个李群）。 S^3 是行列式为 1 的实正交 3×3 正交矩阵的群 $SO(3, \mathbb{R})$ 的双重复盖，因为每两个单位四元数通过上述关系对应于一个转动。群 S^3 和 $SU(2)$ 同构， $SU(2)$ 是行列式为 1 的复酉 2×2 矩阵的群。令 A 为形为 $a + bi + cj + dk$ 的四元数的集合，其中 a, b, c 和 d 或者都是整数或者都是分子为奇数分母为 2 的有理数。集合 A 是一个环，并且是一个格。该环中存在 24 个四元数，而它们是施莱夫利符号为 $\{3, 4, 3\}$ 的正二十四胞体的顶点。

3.1.4 以矩阵表示四元数

有两种方法能以矩阵表示四元数，并以矩阵之加法、乘法应用于四元数之加法、乘法。第一种是以二阶复数矩阵表示。若 $h = a + bi + cj + dk$ 则它的复数形式为：

$$\begin{pmatrix} a - di & -b + ci \\ b + ci & a + di \end{pmatrix} \quad (3-3)$$

这种表示法有如下优点：

所有复数 ($c = d = 0$) 就相应于一个实矩阵。

四元数的绝对值的平方就等于矩阵的行列式。

四元数的共轭值就等于矩阵的共轭转置。

对于单位四元数 ($|h| = 1$) 而言，这种表示方式给了四维球体和 $SU(2)$ 之间的一个同型，而后者对于量子力学中的自旋的研究十分重要。

3.1.5 四元数运算

四元数运算在电动力学与广义相对论中有广泛的应用。四元数可以用来取代张量表示。有时候采用带有复数元素之四元数会比较容易，导得结果不为除法代数之形式。然而亦可结合共轭运算以达到相同的运算结果。

此处仅讨论具有实数元素之四元数，并将以两种形式来描述四元数。其中一种是矢量与标量的结合，另一形式两个创建量(constructor)与双矢量(bivector; i, j 与 k)的结合。

定义两个四元数：

$$\begin{aligned} q &= a + \vec{u} = a + bi + cj + dk \\ p &= t + \vec{v} = t + xi + yj + zk \end{aligned} \quad (3-4)$$

四元数加法： $p + q$

跟复数、矢量和矩阵一样，两个四元数之和需要将不同的元素加起来：

$$q + p = a + t + (b + x)i + (c + y)j + (d + z)k \quad (3-5)$$

加法遵循实数和复数的所有交换律和结合律。

四元数乘法： pq

两个四元数之间的非可换乘积通常被格拉斯曼称为积，它的完整型态是：

$$\begin{aligned} pq &= (at - bx - cy - dz) + (bt + ax + cj - dy)i + (ct + ay + dx - bz)j \\ &\quad + (dt + za + by - xc)k \end{aligned} \quad (3-6)$$

由于四元数乘法的非可换性， pq 并不等于 qp 。格拉斯曼积常用在描述许多其他代数函数。

四元数点积： $p \cdot q$

点积也叫做欧几里得内积，四元数的点积等同于一个四维矢量的点积。点积的值是 p 中每个元素的数值与 q 中相应元素的数值的乘积的和。这是四元数之间的可换积，并返回一个标量。

$$p \cdot q = at + bx + cy + dz \quad (3-7)$$

3.1.6 四元数的极坐标表示

像复数一样，四元数也可以在极坐标下进行表示。

$$q = |q|e^{i\theta}e^{j\varphi}e^{k\theta} \quad (3-8)$$

对于一个单位四元数 $q = q_0 + q_x i + q_y j + q_z k$, $|q| = 1$, 它的相位可以首先通过式(2-5)计算估计出来。

$$\psi = -\arcsin(2(q_x q_y - q_0 q_z))/2 \quad (3-9)$$

然后需要对相位的主值进行判断。为了表达的方便，首先定义三个公式：

$$\begin{aligned} T_i(q) &= -i q i = q_0 + q_x i - q_y j - q_z k; \\ T_j(q) &= -j q j = q_0 - q_x j + q_y i - q_z k; \\ T_k(q) &= -k q k = q_0 - q_x i - q_y j + q_z k \end{aligned} \quad (3-10)$$

判断规则为：

当 $\psi \in [-\pi/4, \pi/4]$, $\varphi = \arg(q T_i(q))/2$, $\theta = \arg(T_i(q))/2$

当 $\psi \in [\pi/4, 3\pi/4]$, 则 $\varphi = 0$, $\theta = \arg(q T_k(q))/2$, 或 $\theta = 0$, $\varphi = \arg(q T_k(q))/2$

当 $e^{j\varphi} e^{k\psi} e^{j\theta} = -q$, 且 $\varphi \geq 0$, 则 $\varphi = \varphi - \pi$;

当 $e^{j\varphi} e^{k\psi} e^{j\theta} = q$, 且 $\varphi < 0$, 则 $\varphi = \varphi + \pi$

3.2 利用四元 Gabor 相位分析图像结构

图像中的颜色往往随图像结构的不同而不同。灰度分布确实带有图像结构信息，但是当存在不规则光照变化的图像中，灰度分布无法作为一种有效的工具来表示图像结构。同时，由于复数 Gabor 相位对光照变化和尺度变化不敏感，已经被广泛用于表示图像结构。最近的一些工作中发现四元数 Gabor 相位能更好地表示图像结构，因为它可以实现二维特征分析^[11]。相反，复数 Gabor 相位只能作为分析一维特征分析的有效工具^[12]。

我们可以用傅里叶平移理论来解释：

$$\begin{aligned} F\{f(x - x_0, y - y_0)\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x - x_0, y - y_0) e^{-i2\pi ux} e^{-i2\pi vy} dx dy = \\ F\{f(x, y)\} e^{-i2\pi(u x_0 + v y_0)} &= F\{f(x, y)\} e^{-i\Delta\theta} \end{aligned} \quad (3-11)$$

$$\begin{aligned} F^q\{f(x - x_0, y - y_0)\} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x - x_0, y - y_0) e^{-i2\pi ux} e^{-j2\pi vy} dx dy = \\ F\{f(x, y)\} e^{-i2\pi ux_0 - j2\pi vy_0} &= F\{f(x, y)\} e^{-i\Delta\theta - j\Delta\theta} \end{aligned} \quad (3-12)$$

其中 $F\{\}$ 表示复数傅里叶变换， $F^q\{\}$ 表示四元数傅里叶变换。从上述推导中可以看到四元数 Gabor 可以用两个相位表示二维平移 x_0 和 y_0 ，而复数 Gabor 只能用一个相位表示从而将信息埋没。

3.2.1 四元数小波变换演进

时频分析是信号处理领域的重要工具，其目的是描述信号的幅值和频率含量如何随时间或空间演变，体现信号在空间局部或时间局部的幅值与频率成分的变化。解析信号也即复信号的定义是为了得到合理的实信号的瞬时频率与瞬时幅值。1 维解析信号的定义早已成熟，目前有多种 2 维解析信号的定义，其中，四元数解析信号是 1 维解析信号定义在 2 维上的合理延伸，克服了以往各种 2 维解析信号定义的不足。但是现有的四元数解析信号的相位定义是建立在单位四元数与 3 维空间转动的对应关系上的，不适合分析 2 维信号的瞬时频率。为此，根据四元数解析信号的自身特点，使用四元数解析信号相位定义，该定义能够更合理地体现信号分量在水平与竖直两个正交方向上的瞬时相位，从而得到信号分量更合理的瞬时频率。该定义同时可以检测出信号分量的累加可分离度。

与傅里叶基方程不同，四元数小波基的局部性能够通过将信号能量集中，稀疏地表示多

奇点的信号。

Bulow^[13]定义了四元数 Gabor 理论, 即一个应用在灰度图像上的加高斯窗的四元数傅里叶变换。现在, 四元数 Gabor 是目前最常用的四元数小波并且能够从复数 Gabor 扩展的张量积中获得。

与复数 Gabor 类似, 一个四元数 Gabor 在负频率相界内有 side-slopes。一个典型的 Gabor 图像分析是不可逆的并且由于傅里叶核而计算量大。

从解析信号角度出发, 四元数小波变换是实数小波变换和复数小波变换的扩展, 它使用了四元代数和 2 维希尔伯特变换。解析信号的概念是由 Gabor 在 1946 年提出的。作为复数小波变换的加强扩展, 四元数小波变换可以实现二维特征分析。相反, 复数 Gabor 相位只能作为分析一维特征分析的有效工具, 而实数小波不能成为确切的解析的。有四种基本局部结构能够被 1 维局部相位区分, 而如果使用 2 维局部相位我们可以区分 16 种这样的结构。

3.2.2 使用四元数 Gabor 相位衡量结构相似性

我们用四元数 Gabor 相位来衡量两个相邻区域的结构相似性。首先我们建立倍频程四元数 Gabors,

$$G_{\sigma\alpha}^q(x, u) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} e^{-i2\pi x' u} e^{-j2\pi y' u} \quad (3-13)$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \quad (3-14)$$

G 表示四元数 Gabor 卷积核。我们进一步根据四元代数法则[4]计算三个四元数 Gabor 相位,

$$\Phi_{\sigma\alpha} \{\varphi, \psi, \theta\} = \arg(I * G_{\sigma\alpha}^q) \quad (3-15)$$

Φ 表示四元数相位向量, I 表示灰度图像。

现在定义一个量来估计两个相邻像素点的结构同态性,

$$\begin{aligned} H_{\sigma\alpha}(p_a, p_b) &= \frac{1}{4} \left(\left| \rho(p_a) \varphi_{\sigma\alpha}^q(p_a) - \rho(p_b) \varphi_{\sigma\alpha}^q(p_b) \right|_{2\pi} \right) \\ &+ \frac{1}{2} \left(\left| \rho(p_a) \theta_{\sigma\alpha}^q(p_a) - \rho(p_b) \theta_{\sigma\alpha}^q(p_b) \right|_{\pi} \right) \\ &+ \left(\left| \rho(p_a) \psi_{\sigma\alpha}^q(p_a) - \rho(p_b) \psi_{\sigma\alpha}^q(p_b) \right|_{\frac{\pi}{2}} \right) \end{aligned} \quad (3-16)$$

3.2.3 相位卷绕问题

相位卷绕是信号处理中一个常见的问题, 它源自于相位在主值区间两端的折断。所谓相位折断, 就是某一相频特性曲线在相邻的两个频率, 原本应该连续的相位值由于相位差绝对值大于 π , 造成相位折断。

由于原来连续的相位可以表示为 $\omega = \Phi + 2k\pi$, 其中 Φ 为主值, k 为整数。通过 Gabor 滤波器后的输出图像矩阵为复数矩阵, 相位是通过反正切函数计算的, 因而只能表示在主值范围 $[-\pi, \pi]$, 内造成了原本的相位在主值 $[-\pi, \pi]$ 范围内的卷绕。

相位卷绕是个复杂的问题, 在这个彩色化过程中是个遗留问题。但是我们可以对相位奇点进行适当的处理。

若记经过 Gabor 滤波器输出的图像像素值为 $x = a + bi$, 在非边缘区域的 x 的模值趋近 0, 即 a, b 都趋近于 0。此时的相位值是极其不稳定的, 可能在很大范围内浮动, 我们称之为奇点。



图 3-2 受相位卷绕影响的彩色化效果

但由于奇点都集中在边缘区，我们将幅值也引入特征值判断依据，当 $|x| < V$ ，时就取消相位作为特征值的判据，取消其权重影响。

这样我们在边缘能适当减小奇点相位的影响。采用这种思路，我们进行了实验，图 3-3 是一种参数设置下的仿真结果（ $V=0.12$ ）。由图可见，只有少数模糊边界有渗透产生，当然，这与最初的 scribble 位置也是相关的。



图 3-3 奇点抑制效果图

3.2.4 PCA 降维处理

上述的多方向多尺度四元数 Gabor 滤波处理后提取的相位特征为 $4 \times 8 = 32$ 维特征向量。要通过如此多维的特征量进行特征比对控制颜色的蔓延，一方面计算量较大，另一方面一些不稳定的相位值的影响不容易控制，因此尝试用 PCA 主成分分析法进行降维处理。

基于主成分分析的特征提取来自于模式识别领域，是很重要的一种统计分析方法。对于初始特征的选择大部分是在考虑样本的可分性意义上进行的，很多情况下初始特征集合都会包含大量互相关联的特征，由于其间存在很大的相关性会有一些信息的重叠。而且各个特征对于样本分类的作用和应有贡献都不是均衡的，因此对大的特征向量集合进行特征的选择是很有必要的。

下面对主成分分析的方法做简要的阐述。

主成分分析是把多个特征映射为少数几个综合特征的一种统计分析方法。它采取降维的方法，找出几个综合因子来代表原来众多的特征，是这些综合后的成分尽可能地反映原来更多维的特征量的信息，并且彼此之间互不关联，简化了大规模的特征向量。

假设研究总体中抽取 N 个样品，以两个指标观测样品，此时的主分量其实成了二维空间的几何问题。如图 3-4 所示，如果假设这些样品在 2 个指标下的分布大致为椭圆，记将坐标系正交旋转 θ 角后的椭圆长轴和短轴方向分别为 y_1 和 y_2 ，用矩阵来表达旋转过程即：

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1N} \\ y_{21} & y_{22} & \cdots & y_{2N} \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \end{bmatrix} = UX \quad (3-17)$$

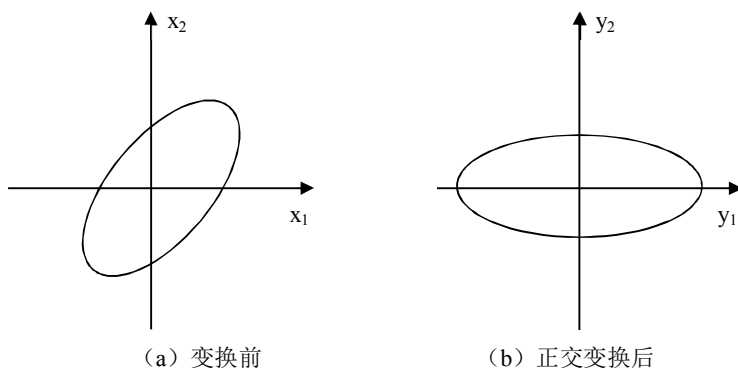


图 3-4 二维特征主成分分析示意图^[24]

其中的 U 作为坐标旋转变换矩阵，其实是一个正交变换矩阵。由图 3-4 可以直观地表现出变换后 y_1 和 y_2 的特征： N 个点的坐标 (y_1, y_2) 几乎没有相关性；二维平面上 N 个点的方差大部分体现在 y_1 的方向上。可以称 y_1 轴为第一主分量，而与其正交且方差较小的 y_2 轴为第二主分量。

上述过程表现了 PCA 的几何意义。将它推广到更一般的情况，可以设 N 个样品的 n 个特征 x_1, x_2, \dots, x_n ，记为列向量 X ，对其进行主成分分析，可以把他们综合为 n 个综合变量，同样用 $y_1 y_2 \dots y_n$ ，记为列向量 Y ，系数矩阵记为 C ，即 $Y=CX$ 。且满足 $c_{k1}^2 + c_{k2}^2 + \dots + c_{kn}^2 = 1$ ($k=1, 2, \dots, n$)，即 C 为单位正交阵。

变化后的 y_1 是 x_1, x_2, \dots, x_n 的所有线性组合中方差最大者， y_2 是与 y_1 不相关的所有 x 向量线性组合中方差次大者，以此类推。通常我们只取前 m 个主分量使得其累计的特征信息的贡献率达到一定量，以此降低特征向量的维数。

用 PCA 对四元数 Gabor 相位产生的 32 维特征向量做降维工作，值得注意的是，并非保留的维数是越多越好的，当前面已有足够的主分量囊括了像素的几乎所有的特征后，后面的次要分量包含的特征分量之间的方差非常小，所以保留变换后过多的分量同样会减小纹理特征的差异度，使得颜色在不同对象间产生渗透，效果适得其反。

3.3 尺度空间结构优先级机制

本节我们介绍一种基于尺度空间边缘强度的结构优先级机制来衡量局部边缘与全局边缘的不同重要性。在本彩色化算法中边缘被提取出来衡量图像边缘的结构作为对四元数 Gabor 相位的补充。另外，全局边缘与局部边缘需要区别对待，因为在大尺度上的边缘更能表示图像的关键结构。

3.3.1 Canny 边缘检测

Canny 边缘检测^[14]的目标是找到一个最优的边缘检测算法，最优边缘检测的含义是：

好的检测 - 算法能够尽可能多地标识出图像中的实际边缘。

好的定位 - 标识出的边缘要尽可能与实际图像中的实际边缘尽可能接近。

最小响应 - 图像中的边缘只能标识一次，并且可能存在的图像噪声不应标识为边缘。

为了满足这些要求 Canny 使用了变分法，这是一种寻找满足特定功能的函数的方法。最优检测使用四个指数函数项的和表示，但是它非常近似于高斯函数的一阶导数。

Canny 算法的步骤

(1) 去噪声

任何边缘检测算法都不可能在未经处理的原始数据上很好地处理，所以第一步是对原始数据与高斯平滑模板作卷积，得到的图像与原始图像相比有些轻微的模糊 (blurred)。这样，单独的一个像素噪声在经过高斯平滑的图像上变得几乎没有影响。

(2) 寻找图像中的亮度梯度

图像中的边缘可能会指向不同的方向，所以 Canny 算法使用 4 个 mask 检测水平、垂直以及对角线方向的边缘。原始图像与每个 mask 所作的卷积都存储起来。对于每个点我们都标识在这个点上的最大值以及生成的边缘的方向。这样我们就从原始图像生成了图像中每个点亮度梯度图以及亮度梯度的方向。

(3) 在图像中跟踪边缘

较高的亮度梯度比较有可能是边缘，但是没有有一个确切的值来限定多大的亮度梯度是边缘多大又不是，所以 Canny 使用了滞后阈值。

滞后阈值需要两个阈值——高阈值与低阈值。假设图像中的重要边缘都是连续的曲线，这样我们就可以跟踪给定曲线中模糊的部分，并且避免将没有组成曲线的噪声像素当成边缘。所以我们要从一个较大的阈值开始，这将标识出我们比较确信的真正边缘，使用前面导出的方向信息，我们从这些真正的边缘开始在图像中跟踪整个的边缘。在跟踪的时候，我们使用一个较小的阈值，这样就可以跟踪曲线的模糊部分直到我们回到起点。一旦这个过程完成，我们就得到了一个二值图像，每点表示是否是一个边缘点。

一个获得亚像素精度边缘的改进实现是在梯度方向检测二阶方向导数的过零点

$$L_x^2 L_{xx} + 2L_x L_y L_{xy} + L_y^2 L_{yy} = 0 \quad (3-18)$$

它在梯度方向的三阶方向导数满足符号条件

$$L_x^3 L_{xxx} + 3L_x^2 L_y L_{xxy} + 3L_x L_y^2 L_{xyy} + L_y^3 L_{yyy} < 0 \quad (3-19)$$

用这种方法得到的边缘片断是连续曲线，这样就不需要另外的边缘跟踪改进。滞后阈值也可以用于亚像素边缘检测。

Canny 算法包含许多可以调整的参数，它们将影响到算法的计算的时间与实效。

高斯滤波器的大小：第一步所用的平滑滤波器将会直接影响 Canny 算法的结果。较小的滤波器产生的模糊效果也较少，这样就可以检测较小、变化明显的细线。较大的滤波器产生的模糊效果也较多，将较大的一块图像区域涂成一个特定点的颜色值。这样带来的结果就是对于检测较大、平滑的边缘更加有用，例如彩虹的边缘。

阈值：使用两个阈值比使用一个阈值更加灵活，但是它还是有阈值存在的共性问题。设置的阈值过高，可能会漏掉重要信息；阈值过低，将会把枝节信息看得很重要。很难给出一个适用于所有图像的通用阈值。目前还没有一个经过验证的实现方法。

Canny 算法适用于不同的场合。它的参数允许根据不同实现的特定要求进行调整以识别不同的边缘特性。对于 PC 上的实时图像处理来说可能慢得无法使用，尤其是在使用大的高斯滤波器的情况下。但是，我们讨论计算能力的时候，也要考虑到随着处理器速度不断提升，有望在未来几年使得这不再成为一个问题。

3.3.2 尺度空间边缘强度模型

我们定义一条边缘的重要程度为边缘强度并提出了一个算法来计算边缘强度的值。本算法是基于多尺度 Canny 边缘检测^[14]的，尺度因子用 σ_i 表示。尺度因子越大，边缘图像就含有更多粗略的结构而少有细节。一条边缘如果在尺度空间下具有很长的生命周期，则需要赋给它较大的强度值，因为这代表着这条边沿是输入图像关键的边缘结构。也就是说它应该优

先被彩色化以保证这个区域的色彩不出现渗透等问题。由于多尺度边缘不一定精确产生在同一位置，我们用形态学算法膨胀来辅助得到更有可靠的边缘强度估计。

该算法在下表中给出。

表3-1 计算图像中边缘强度的算法流程

计算边缘强度算法
1: 初始化: $\text{Strength} = C(a1)$
2: for $i = 2$ to n
3: 以半径 $i-1$ 膨胀 $C(ai)$ 得到图像 di
4: 对于 di 中的每个边缘上的像素 $p(x, y)$
5: if $p(x, y)$ 也是 $C(ai-1)$ 边缘上的像素
6: $\text{Strength}(x, y) = \text{Strength}(x, y) + 1$
7: end if
8: end for
9: 以半径 $i-1$ 膨胀 $C(ai)$ 得到另外一个膨胀图像 di
10: 对于 di 中的每个不在边缘上的像素 $p(x, y)$
11: if $p(x, y)$ 也是 $C(ai-1)$ 边缘上的像素
12: $\text{Strength}(x, y) = \text{Strength}(x, y) + i$
13: end if
14: end for
15: end for

3.4 全局最优色彩蔓延

3.4.1 全局最优化算法介绍

全局最优化问题的形式一般都是在给定的限制条件下寻找一个方程的极值。具体到本文的彩色化问题中去就是在色彩空间各个通道的数值的限制条件下寻找一个图像结构与色彩相关的代价方程的最小值，代价方程的建立在下小节中具体阐述。

对最优化问题求解的理论和算法的研究，是当今数学和计算机科学研究的重要课题之一。所谓优化算法，是一种搜索过程或规则，它是基于某种思想和机制，通过一定的途径或规则得到满足用户要求的问题解。总体上分为局部最优算法和全局最优算法。局部优化的理论比较成熟，具有判断局部极小点的最优性条件(无约束问题的一阶必要条件、二阶充分条件和二阶必要条件；约束问题的 Fritz John 必要条件、Kuhn-Tucker 必要条件和 Kuhn-Tucker 充分条件)。局部优化的算法较多，如，求解线性规划的单纯形方法(Dantzig, 1947)、求解非线性最优化问题的最速下降法、牛顿法、拟牛顿法、共轭梯度法(FR 算法和 PRP 算法等)、罚函数法、可行方向法和信赖域法(Powell, 1975)等；但是，这些传统的数值最优化算法通常只能找到目标函数的平稳点/驻点和局部最优点，而很难找到全局最优点，并且每次只能找到目标函数的一个最优点。

全局最优算法在理论上还没有确定的方法判断一个局部最优点是否为全局最优点。对于大部分实际应用来说，找到全局最优点更好，但并不是必须的。有些问题则必须找到全局最优点，如，机器手的设计、化学中的许多问题及半定规划等。常用的全局最优算法有：遗传算法(Holland, 1975)、蚁群算法(MarcoDorigo, 1992)、模拟退火算法(Kirkpatrick, 1982)、分枝与界定方法、区间方法(Moore, 1962)、Tabu 搜索法(Glover, 1986)、聚类方法、隧道函数法和填充函数法等。

目前有多种方式对全局最优化算法进行分类。2003 年 COCONUT 研究组根据算法找到全局最优解的可信程度，把全局最优算法分成如下 4 类：

(1)不完全方法(Incomplete Method)。它使用一定的启发式搜索技术,不能保证找到全局最优解。(2)渐进完全方法(Asymptotically Complete Method)。它随着运行时间趋向于无穷,将以概率 1 找到全局最优解,但是无法知道何时找到全局最优解。(3)完全方法(Complete Method)。它经过一定时间后可以知道已经找到近似全局最优解(在一定的误差范围内)。(4)严格方法(Rigorous Method)。它在舍入误差存在的情况下一定可以在给定误差范围内找到全局最优解。

不完全方法

最简单的不完全方法是多随机点方法(Multiple Random Start)。首先选取多个随机点,然后对每个点执行局部优化。大多数基于随机技术的方法都可以看成是对这个基本方法的改进,并以不同的方法选取点进行局部优化。主要包括:

(1)局部下降技术(Local Descent Techniques)。多随机点方法(Multiple Random Start)、聚类(Clustering Methods)、隧道法(Tunneling Method)、填充函数法(Filled Function Method)。(2)非单调搜索技术(Non-monotonic Search Techniques)。禁忌搜索(Tabu Search)、模拟退火算法(Simulated Annealing)。(3)种群方法(Ensemble Methods)。遗传算法(Genetic Algorithms)和蚁群算法(Ant Colony)。禁忌搜索(Tabu Search 或 Taboo Search, 简称 TS)是由 Glover 在 20 世纪 80 年代末提出的一套优化理论,是对局部邻域搜索的一种扩展,是一种全局逐步寻优、高效启发式优化技术和对人类智力过程的一种模拟。TS 算法通过引入一个灵活的存储结构和相应的禁忌准则,以避免迂回搜索,并通过藐视准则赦免一些被禁忌的优良状态,进而保证多样化的有效搜索,最终实现全局优化。

禁忌搜索算法在组合优化、生产调度、机器学习、路由选择和电路设计等领域取得了很大的成功。Tan P H 等人^[15, 16]研究了 CDMA 中多用户检测的禁忌搜索算法;Niar S 等人^[17]设计了多维 0-1 背包问题的并行禁忌搜索算法;Watson T P 等人^[18]提出了求解 job shop 的禁忌搜索算法的动态模型;等等。近年来又在函数全局优化方面得到较多的研究,并大有发展的趋势。为了进一步改善禁忌搜索算法的性能,可以对算法本身的操作和参数选择进行改进,也可与模拟退火算法、遗传算法、神经网络等相结合。1982 年, Kirkpatrick S^[19]等将退火思想引入组合优化领域,提出一种解大规模组合优化问题的有效近似算法——模拟退火算法(Simulated Annealing Methods)。此算法的核心在于模仿热力学中液体的冻结与结晶或金属溶液的冷却与退火过程。只要模拟过程足够充分,算法就可以概率 1 收敛到全局最优解。模拟退火算法除了可以接受优化解外,还可用一个随机接受准则(Metropolis 准则)有限度地接受恶化解。

在应用模拟退火算法时,需要确定一组控制算法进程的参数(冷却进度表):初始温度、温度衰减函数、每个 Markov 链的长度和停止准则。

模拟退火算法的原始形式虽然可以从概率的角度证明其收敛性,但是其速度很慢。为了提高其速度已经作了许多改进,如:改进冷却进度表^[20]、加温退火法、有记忆的模拟退火法、带返回搜索的模拟退火法^[21]、多次寻优法、回火退火法、退火演算法、两阶段模拟退火算法^[22,23]和并行模拟退火算法等。

完全方法

几乎所有完全最优化方法的基础都是分枝原则(Branching Principle),其关键在于把原始问题的解空间不断划分成小的易解的子空间,然后进行处理。

分枝与界定(Branching and Bounding)是一种通用的搜索方法。计算目标函数可在某个可行域中的上界、下界和对可行域进行划分^[24]。其基本思想如下:首先在整个可行域中考虑原始问题(根问题),计算根问题的上界和下界,如果找到了问题的最优解则算法结束;否则,把可行域划分(Branching)成 2 个或多个域(每个域都是原来域的真子域并且所有域的并等于原来的域)使得这些子问题是根节点的子节点。其次,对每个子问题递归地应用上面的算法

以形成每个子问题的搜索树。如果找到了某个子问题的最优解,则此解是整个问题的可行解而不一定是全局最优解。最后,可以利用此解对搜索树剩余的节点进行修剪(Bounding)。重复上面的过程,直到满足要求。

严格方法

符号计算在应用数学、物理、工程等领域中起着重要的作用,其特点是可以求得问题的全部解和精确解,主要包括 Gr bner 基方法、Ritt-Wu 方法和结式方法^[25,26]。符号优化算法最具代表性的是代数函数全局优化的符号算法^[27-30],它源于有限核原理,即在多项式约束下,一个多项式的局部最大(最小)值必为有限多个。虽然此时其最优点之集可以是无限的甚至是不可数的,但该算法将求全局最大(最小)值的问题转换为对有限多个不等式的验证。

Hanson 等设计了一种通过计算 Gr bner 基和特征值的方法解多元多项式全局最小值的算法^[31]。当多项式有最小值时,此方法将返回全局最小值并且在每个包含全局最小值的连通区域中至少能找到一个使多项式取最小值的点;当多项式没有最小值时,此方法将计算它的下确界。符号算法虽然可以找到全局最优解,但是由于符号计算所固有的复杂性,所以目前它只能解决一些规模较小的问题。

区间分析是利用区间进行计算的理论与算法的学科。区间分析起源于 1959 年 Moore 的技术报告,并取得了一定的进展^[31]。利用区间技术进行全局优化可以保证找到全局最优解。区间技术主要用来计算目标函数的全局信息、函数的界限、Lipschitz 常量和高阶导数等。一般利用区间技术的全局最优化方法都采用了分枝与界定的策略^[27,32,33]。此方法以下面的伪代码的形式给出:

```
1 Initialize a list of boxes L by placing the initial search region  $X_0$  in L
2 Do While  $L \neq \emptyset$ 
  (a) Remove the first box X from L;
  (b) (Process X) Do one of the following:
    • Reject X;
    • Reduce the size of X;
    • Determine that X contains an unique critical point, then find the critical point to high accuracy;
    • Subdivide X to make it more likely to succeed at rejecting, reducing, or verifying uniqueness.
  (c) Do the following to the box(es) resulting from Step 2(b):
    • If X was rejected, do nothing;
    • If more than one box was derived from X, insert all but one of them into L. Call the remaining box derived from X;
    • If there is a X that has been proven to contain a critical point, insert it into C
    • If there is a X that is small, but has not been proven to contain a feasible point, insert it into U.
End Do
```

其中: X_0 为初始区间; X 为当前正在处理的区间; L 为等待处理的区间的列表; C 为已经得到验证包含最优解的区间列表; U 为区间的宽度满足要求的区间列表。

区间分析技术的优点有: (1)可以很容易地计算出表达式在某个区间上的界限; (2)可利用简单的方法(通过中值形式)处理非线性问题; (3)扩展了传统分析的能力,即提供半局部存在性和最优性条件; 可以给出包含最优点的区间而不只是说明其存在。区间分析方法的缺点

有：(1)运算速度较慢(2)判断某个区间内是否存在局部最优优点比较困难。

3.4.2 四元数色彩重构

我们的四元代数理论不仅应用于图像结构的分析（四元数 Gabor 相位），还作为全局最优色彩重构的色彩空间。关于色彩空间的选择要根据图像处理任务的特点进行选择，但是在大多数情况下研究人员会选择 YUV 空间。因为 RGB 空间的三个色彩通道之间存在很大的相关性，并且地位是相当的，这样就存在大量的冗余信息，而变换到 YUV 空间之后很好地将图像幅值和色度分离开来，这样处理就可以更有针对性地在图像的不同方面进行。但是 YUV 空间对光照变化比较敏感，从这个方面来看，四元数色彩空间在包容光照变化时更稳定，这一点是依赖图像结构的图像彩色化技术需要的。

彩色化的最后一步是根据上面定义的四元数相位和边缘强度进行色彩生长。我们最小化一个关于四元数颜色相位的代价方程来实施色彩生长。在四元数色彩空间，每个像素用一个纯虚四元数表示， $q=a+bi+cj+dk$ ， a 等于 0， b ， c ， d 分别为 RGB 值。在极坐标中，一个四元数可以表示为幅值和三个相位。

色彩空间中的向量运算允许不同的色彩通道被一样对待而不是分开处理。

受 Levin 工作^[1]的启发，我们建立需要最小化的代价方程，

$$\Phi^{opt} = Arg \min_{\Phi} \sum_p \left| \Phi(p) - \frac{\sum_{q \in N(p)} W_{pq} \Phi(q)}{\sum_{q \in N(p)} W_{pq}} \right|, \Phi = \{\varphi, \psi, \theta\} \quad (3-20)$$

这个最优化问题是关于三个相位分别进行的。

$$W_{pq} = W_{pq}^{Strength} \times W_{pq}^{QGabor} \quad (3-21)$$

$$W_{pq}^{Strength} = \exp(-|strength(p) - strength(q)|) \quad (3-22)$$

$$W_{pq}^{QGabor} = \exp(-\sum_{\sigma} \sum_{\alpha} H_{\sigma\alpha}(p, q)) \quad (3-23)$$

$N(p)$ 是一个以像素 p 为中心的 3×3 的像素区域。 W_{pq}^{QGabor} 的设计是为了约束在同一同态区

域内的相邻像素具有同样的颜色； $W_{pq}^{Strength}$ 是为了强调优先彩色化具有高强度的边缘位置，

$W_{pq}^{Strength}$ 在无边缘区域保持常数，这项权重值保证在边缘区域有无误的彩色化效果。

3.5 本章小结

本章内容主要介绍在自动 scribble 产生之后如何根据灰度图像的结构信息进行色彩蔓延，使 scribble 上的色彩自然地分布在整个图像上。图像彩色化问题是一个 ill-posed problem，这个问题不存在所谓的正确的解，因为一个灰度图像不带有任何彩色信息，一切彩色化算法的目的都是让彩色化结果有更自然的视觉表现。

所以本章的重点围绕图像结构的表示方法，给彩色蔓延提供充足的线索并且同时带来一些约束使得蔓延无误。我们用超解析信号分析四元数 Gabor 小波相位来找到图像的同态区域，使相同的色彩在同一同态区域内蔓延。同时我们在尺度空间下引入结构优先级机，使得关键边缘部位在彩色化中占更大的权重以至于被优先彩色化。

最后本文将四元数 Gabor 相位和尺度空间边缘强度作为权重设计了一个代价方程，使用全局最优算法最小化代价方程，重构整幅图像的色彩。色彩重构是在四元数色彩空间中进行

行的，与其他色彩空间相比四元数色彩空间更能包容光照变化，从而减小了光照变化对彩色化的负面影响。

第四章 实验结果及用户界面

本章第一节将展示本次毕业论文的彩色化效果图，给出量化的质量评价并与其他主流彩色化算法进行比较。第二节介绍设计编写的交互式选色系统，这套系统能够让用户方便地使用自动 scribble 产生、选色和彩色化功能。

4.1 彩色化仿真结果

下面给出彩色化实验的参数选择。

首先需要对灰度图像进行分割，需要说明的是我们的彩色化效果不依赖图像分割的边缘是否精细，而且只需要对图像进行粗略的过分割即可，因为图像分割只决定 scribble 的产生。本文在实验中将分割的最小区块大小设置为 $M*N*5\%$ ， $M*N$ 代表图像的分辨率。在 scribble 搜索算法中，对第 M 个 scope 的相关性阈值（见 2.2）取：

$$T_{M+1} = \frac{0.6}{M} \sum_{i=2}^M r_i \quad (4-1)$$

其中 r_i 是第 i 个 scope 和第 $i-1$ 个 scope 的相关性系数。

下图给出了“房屋”和“Lena”两个图像的彩色化结果。白色线条是自动产生的 scribble。从图中可以看出，scribble 都自动产生在了有关键颜色结构的区域，并且没有单独一条 scribble 跨越了两个不同颜色的区域。第三列的原始彩色图像是为了与彩色化效果进行对比，可以看出本文算法能很自然地恢复失去的色彩信息。



图 4-1 彩色化效果与原图对比

在图 5 中，将本文的彩色化效果与依靠灰度相似性和复数 Gabor 相位的算法得到的结果进行比较，证明四元数 Gabor 相位和尺度空间边缘强度模型能够很好地表示图像结构。图 5 中的彩色化质量用 PSNR 进行衡量。

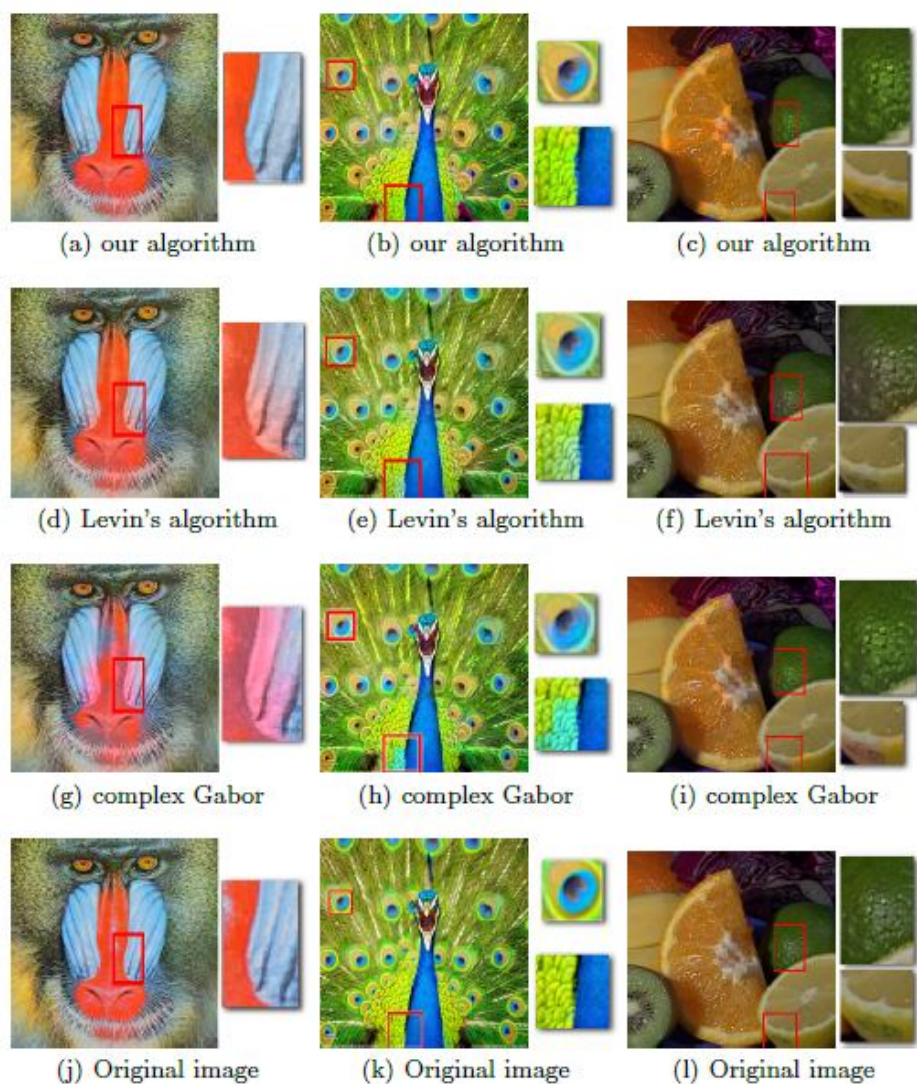


图 4-2 彩色化效果与其他方法对比

表4-1 彩色化效果PSNR对比

图像	本文算法	Levin 算法	Complex-Gabor 算法
Baboon	23.81dB	23.58dB	22.78dB
Peacock	20.88dB	21.59dB	19.85dB
Fruits	21.73dB	21.59dB	20.80dB

由上表可以看出 Levin 的结果和本文结果的 PSNR 非常相近，而使用复数 Gabor 相位的方法的 PSNR 会低大约 1dB，导致不令人满意的视觉效果。与 Levin 的算法的对比 PSNR，本文算法在 Baboon 图像中有更好的表现但在 Peacock 和 Fruits 中没有优势。但是本文算法能更好地恢复复杂彩色纹理区域，如狒狒的脸颊，孔雀的羽毛和身体。本文算法在边缘位置能实现几乎零色彩泄露。在相同的 PSNR 下，本文算法有更好的视觉效果，人类视觉系统更能适应于提取场景的结构信息，这就是为什么本文的结构保持彩色化结果具有更好的视觉效果。

果。

4.2 交互式选色界面

本文的彩色化算法同时在 matlab 平台和 C++平台上实现。Matlab 程序用于算法改进时能够快速地修改试验新算法。当算法基本成熟时将其移植到 C++平台上来。虽然失去了许多 matlab 内置工具包的支持使得代码重写非常费时费力，但是 C++增强了它的可移植性，同时用户友好的交互式选色界面也使用 MFC 实现。

用户界面的设计采用单文档窗口，最上方为传统菜单栏，用于文件、编辑、视图和帮助操作。菜单栏下方为工具栏，将一些常用按钮放在工具栏内。窗口左侧设置了导航栏，可以访问用户的所有资源，方便查找文件和文件夹。窗口中部为主面板，用于显示输入的灰度图像和自动产生的 scribble，在该面板上的鼠标动作能够实现 scribble 的选取和调用调色板为 scribble 选色的功能。最后彩色化结果也在主面板中显示。

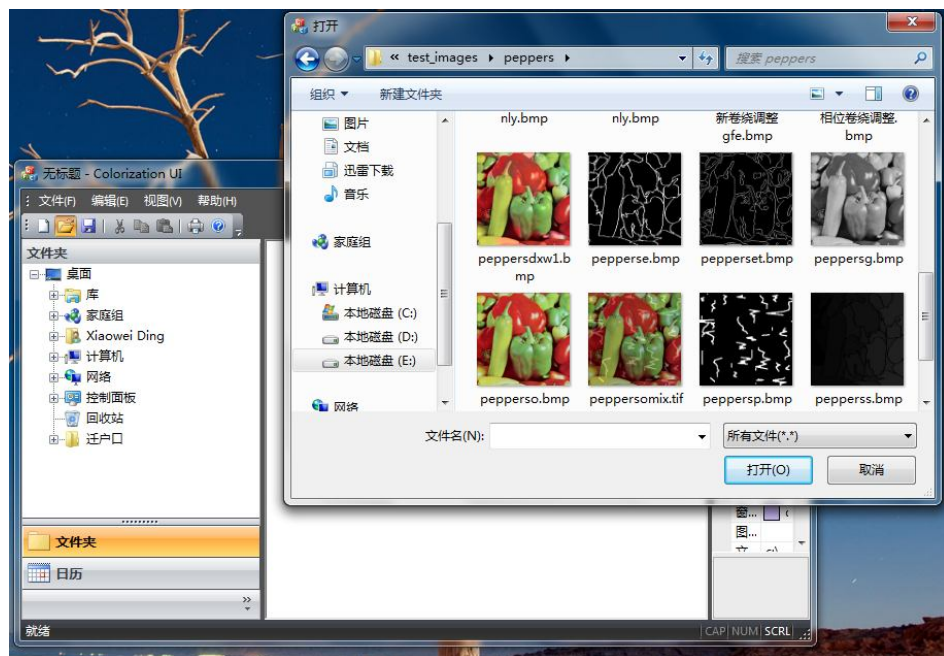


图 4-3 使用 MFC UI 载入灰度图像

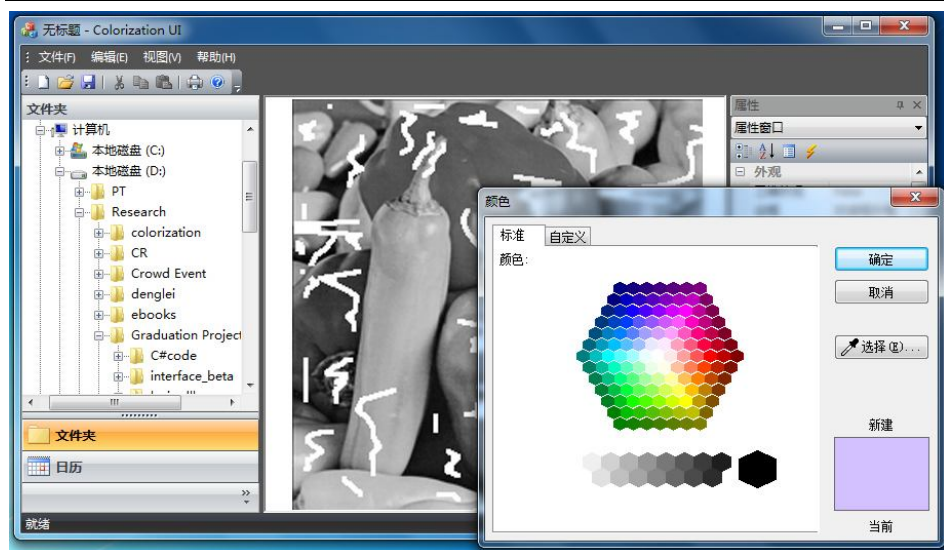


图 4-4 MFC UI 交互式选色界面

使用 UI 的流程非常简单，首先用户可以通过点击“打开按钮”载入硬盘上需要彩色化的灰度图像，如图 4-3，该图像将显示在主面板上。自动产生的 scribble 也将叠加显示在灰度图像上方为用户为 scribble 选色。然后用户可以通过用鼠标点选某条 scribble 来为该 scribble 在调色板上选色，如图 4-4 所示，选色完成后程序会开始彩色化，将最终结果显示在主面板上。整个过程非常简便直观，不需要用户有丰富的彩色化经验即可完成彩色化任务。

第五章 结论

一个出色的黑白图像彩色化算法的评价标准是它可以在尽可能少的人工干预下产生令人满意的彩色化效果。本文的工作通过应用四元代数理论和自动 scribble 产生技术实现了这一目标。使用本文算法时用户可以享受为自动产生的 scribble 进行选色的便利，而不需要手动绘制 scribble。同时，复杂彩色纹理区域的彩色化效果也因四元数彩色空间的引入和超解析信号分析而提高很多。尺度空间结构优先级理论和四元 Gabor 相位结合奠定了可靠的结构保持彩色化技术的基础。

本文算法分别在 matlab 和 C++ 两个平台上实现，matlab 程序平台便于代码调试和算法改进，当 Matlab 算法一旦验证成功，可将其移植到 C++ 平台上来。当然，失去许多 matlab 内置工具包的支持，本文 C++ 代码开发非常耗时，但是 C++ 增强程序代码的可移植性，同时用户友好的交互式选色界面可使用 MFC 实现。

本文提出的彩色化算法已在 2012 年被 International Conference on Multimedia Modeling (MMM2012) 国际会议录用。

参考文献

- [1] A. Levin, D. Lischinski, and Y. Weiss: Colorization using optimization. ACM Trans. Graph. 2004: 3, 689-694.
- [2] M.S. Drew and G.D. Finlayson: Realistic colorization via the structure tensor. In: ICIP, 2008: 457-460.
- [3] T.H. Kim, K.M. Lee, and S.U. Lee: Edge-preserving colorization using data-driven random walks with restart. In: ICIP, 2009: 1641-1644.
- [4] O. Lezoray, V. Ta, and A. Elmoataz: Nonlocal graph regularization for image colorization. In: ICPR, 2008: 1-4.
- [5] J. Heu, D. Hyun, C. Kim, and S. Lee: Image and video colorization based on prioritized source propagation. In: ICIP, 2009: 465-468.
- [6] T. Welsh. and M. Ashikhmin. and K. Mueller.: Transferring color to grayscale images. ACM Transaction on Graphics 21(3), 2002: 277-280.
- [7] R. Irony, D. Cohen-Or, and D. Lischinski: Colorization by example. In: Eurographics Symposium on Rendering, 2005: 201-210.
- [8] X. Liu, L. Wan, Y. Qu, T. Wong, S. Lin, C. Leung, and P. Heng: Intrinsic colorization. ACM Trans. Graph. 27(5), 2008: 152:1-152:9.
- [9] D. Sykora, J. Burinek, and J. Zra: Unsupervised colorization of black-and-white cartoons. In: 3rd Int. Symp. NPAR, 2004.
- [10] P.F. Felzenszwalb and D.P. Huttenlocher: Efficient Graph-Based Image Segmentation. In: International Journal of Computer Vision, 2004: 167-181.
- [11] Y. Xu, L. Song, X. Yang, L. Traversoni, and W. Lu: QWT: retrospective and new Applications. Geometric Algebra Computing for Engineering and Computer Science, Springer Verlag London 2010
- [12] E. Bayro-Corrochano: The Theory and Use of the Quaternion Wavelet Transform. Journal of Mathematical Imaging and Vision 2006: 24, 19-35
- [13] Bulow Th. Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images [D] . Kiel, Germany: Christian Albrechts University, 1999.
- [14] J. Canny: A Computational Approach To Edge Detection. IEEE Trans. Pattern Analysis and Machine Intelligence 1986: 8(6), 679-698.
- [15] Tan P H, Rasmussen L K. Tabu search multiuser detection in CDMA[A]. RadioVetenskap och Kommunikation[C]. Sweden: Stockholm, 2002: 744-748.
- [16] Tan P H , Rasmussen L K. A reactive tabu search heuristic for multiuser detection in CDMA[A]. ISIT 2002[C]. Switzerland: Lausanne, 2002: 472.
- [17] Niar S, Freville A. A parallel tabu search algorithm for the 0-1 multidimensional knapsack problem[A]. 11th International Parallel Processing Symposium (IPPS '97)[C], Switzerland, Geneva, IEEE, 1997: 512-516.
- [18] Watson J P, Whitley L D, Howe A E. A dynamic model of tabu search for the Job-Shop scheduling problem[A]. The 1st Multidisciplinary International Conference on Scheduling : Theory and Applications (MISTA 2003)[C]. Netherlands: Kluwer Academic Publishers, 320-336.
- [19] Kirkpatrick S, Gelatt C D. Optimization by simulated annealing[J]. Science, 1983, 220

- (4598):671-680.
- [20] Nourani Y , Andresen B. A comparison of simulated annealing cooling strategies[J]. Journal of Physics A: Mathematical and General, 1998, 41:8 373-8 385.
- [21] Wah B W, Chang Y-J. Trace-based methods for solving nonlinear global optimization and satisfiability problems[J]. Journal of Global Optimization, 1997,10(2):107-141.
- [22] Varanelli J M, Cohoon J P. A two-stage simulated annealing methodology [A]. 5th Great Lakes Symposium on VLSI[C]. USA: Buffalo, 1995. 50-53.
- [23] Varanelli J M. On the Acceleration of simulated annealing[D]. USA:University of Virginia, 1996.
- [24] Winston P H. Artificial intelligence[M]. New York: Addison-Wesley, 1984.
- [25] Collins G E, Loos R. Real zeros of polynomials[J]. Computing Supplementum, 1982, (4): 83-94.
- [26] Collins G E, Johnson J R, Krandick W. Interval arithmetic in cylindrical algebraic decomposition[J]. Journal of Symbolic Computation, 2002, 34:145-157.
- [27] Yang L. Recent advances in automated theorem proving on inequalities[J]. Comput Sci & Technol, 1999, 14(5):435-446.
- [28] Yang L, Xia S H. An inequality-proving program applied to global optimization[A]. Proceeding of the Asian Technology Conference in Mathematics 2000[C]. Blacksburg: ATCM Inc , 2000.40-51.
- [29] 杨路.全局优化的符号算法与有限核原理[A].林东岱.数学与数学机械化[C].山东:山东教育出版社,2001.210-220.
- [30] Hanzon B, Jibetean D. Global minimization of a multivariate polynomial using matrix methods[J]. Journal of Global Optimization, 2003, (27):1-31.
- [31] Granvilliers L. A symbolic-numerical branch and prune algorithm for solving non-linear polynomial systems[J]. Journal of Universal Computer Science, 1998, 4(2):125-146.
- [32] Csendes T. Numerical experiences with a new generalized subinterval selection criterion for interval global optimization[J]. Reliable Computing, 2003,(9):109-125.
- [33] Shary S P. A surprising approach in interval global optimization[J]. Reliable Computing, 2001,(7):497-505.

谢辞

这项工作是由高等学校博士学科点专项科研基金(No. 200802481006), NSFC-60902073, NSFC-60932006 资助的。

感谢上海交大所有给我指导的老师。本文是在徐奕老师的悉心指导下完成的, 尽管徐老师业务繁忙, 但是还是花费大量时间与我交流, 指导我理解理论、编写程序和撰写论文。在徐奕老师和杨小康老师的支持下这项工作的论文才能得以发表, 得到学术界的认可。特别感谢两位老师两年多来对我的培养和关心。

最后感谢评阅本文的老师的宝贵时间和精力。

IMAGE COLORIZATION TECHNIQUES

Colorization is an image processing technique which allows us to add colors to monochrome images with the help of manual work. The basic idea of colorization is to set up a mapping from scalar-valued intensity image to vector-valued color image using a few color scribbles provided by users. Therefore, color scribbles should be selected carefully to provide sufficient color reference for monochrome regions. Correct and rapid color propagation should also be triggered from these color scribbles using a color optimization process. In addition, scribbling can be tedious for images with complex details, and requires some skill to obtain natural-looking results.

Scribble plays a significant role in colorization. It contains all the color information that can be used in colorization process. However, images with complex structure need a large amount of careful scribbles by an experienced user, and also, hand-made scribbles will not necessarily trigger the most effective color propagation in consideration of extensive image contents. In this paper, we proposed an automatic scribble generation algorithm based on spatial distribution entropy, placing scribbles within the regions of high information density. As a result, the requested color information of each homogeneous segment is dominantly contained in the neighborhood of these scribbles.

To avoid one single scribble striding across different homogeneous segments, the image is first over-segmented into closed areas using Graph-Based Image Segmentation algorithm. Then, we perform a search in each segment for the points, around which high information density is accessible. Spatial distribution entropy is adopted to compute information density. In the end, these points are interpolated to form smooth scribbles. To ensure that there is enough color information for colorization process in a limited number of scribbles, we locate the scribbles at the places possessing high spatial distribution entropy (SDE).

Our search strategy allows backtracking of arbitrary number of steps to avoid a bad path brought by any unwise step in this procedure. In addition, it is possible that points are densely selected in a small area due to high entropy. We treat this situation as re-entry and avoid it by measuring the concentration of selected points in current scope. To direct the scribble drawing in a homogenous area, we concern about those candidate scopes whose correlation coefficient to current scope is higher than a threshold. Correlations between two scopes are measured by their correlation coefficient of grayscale histograms. Once all the possible scopes in 8 directions fail to satisfy this constraint, we need to backtrack to the previous step to redirect the searching path. Iterate this process until backtracking path is exhausted.

In this thesis, we introduce a structural priority framework based on scale-space contour 'strength' in order to measure the structural importance of global contours and local contours. In our algorithm, contours are extracted to characterize the structures of objects in an image and act as a compensation of quaternion Gabor phase pattern. Furthermore, global contours and local ones would be differently treated since contours in larger scale scope stand for more critical structures.

We define the importance of a contour as strength and propose an algorithm to calculate strength value. Our algorithm is based on multi-scale canny edge representation. Further, a

批注 [x1]: 这一段前面已经出现

contour should be assigned a large strength value if it has a long lifetime in scale space since these contours always represent the most significant structures of an image. This means that it should have higher priority in colorization process. Since multi-scale contours might not exactly match at the same location throughout the scale space, we use morphology algorithm dilating to achieve more reasonable strength estimation.

Colors often differ in different structural parts in image. Luminance distribution does carry some of the image structure information; however, it cannot act as a useful tool for image analysis under irregular illumination variations. Meanwhile, complex Gabor phases have been extensively exploited to represent image pattern due to the robustness to illumination changes and scale disturbance. In recent work, quaternion Gabor phase is found to provide a better representation of image pattern since it can realize the analysis of intrinsically 2D features (corner-like). In contrast, complex Gabor phase can only provide a powerful tool for intrinsically 1D features (edge-like) analysis. Considering that (quaternion) Gabor is a windowed (quaternion) Fourier transformation. We use quaternion Gabor phases to measure structural similarity between two neighborhoods.

The final step in colorization process is color diffusion based on quaternion phase pattern and contour strength designed above. We implement the color diffusion process by minimizing a cost function in which variables are quaternion color phases.

Vector operation in color space allows different color channels to be treated as a unity rather than channel being independently manipulated. Illumination variation has little impact on quaternion phases of color image. Hence, it is more reasonable to perform colorization by reconstructing quaternion phase vector in the minimization process. Inspired by Levin's work, we build a cost function to be minimized. As a matter of fact, the optimization problem is treated as three independent minimization problems about three phase components.

In Chapter 5, we show our colorization result, where scribbles are generated automatically. Performance comparison with other colorization algorithms is also provided in this section. It should be noted that our algorithm only needs a rough over-segmentation of the input image in automatic scribble generation process.

In the simulation, scribbles are automatically located in those regions which are supposed to contain critical color structures. No scribble strides across different colors. The original images are listed to provide visual evaluation of colorization performance. We can see that our algorithm reconstructs the missing color information very well.

we compare our algorithm with those methods depending on intensity similarity or complex Gabor phases, proving that quaternion Gabor phases and scale-space contour strength model can represent fundamental image structures. The quality of the colorization results are measured by PSNR. Levin's results and our results have very small differences in PSNR. Results by algorithm using complex Gabor phases are about 1dB lower than ours and Levin's, which leads to unsatisfying visual effects. Comparing with Levin's algorithm in PSNR, our algorithm performs better on baboon image, while a little bit worse on peacock image and fruits image. However, our colorization results promisingly recover the complex color texture in baboon's villus, peacock plumage and citrus peel. We also achieve almost zero color leakage around object contours, which cannot be achieved by other two algorithms. From the close-up image, we can see that our results have better visual effects than algorithm based on intensity similarity (Levin's algorithm) under comparable PSNR scores. Human visual perception is highly adapted for extracting structural information from a scene, which explains why our structure-preserving colorization results are

more visually satisfying.

It takes about 11s in average to generate high quality scribbles automatically for each of these input images (给出图像分辨率大小) on PC. Manual scribbles usually take at least several minutes to sketch with professional guidance. Automating this time-consuming process in colorization is user friendly and greatly shortens the overall time of colorization.

We developed colorization software on both Matlab platform and C++ platform with UI implemented by MFC.

The evaluation criterion for an outstanding colorization algorithm is whether it can achieve prospective effect requiring as little manual input as possible. Our work achieves this goal by proposing a colorization algorithm based on quaternion algebra with automatic scribble generation. Users are expected to enjoy the convenience of choosing colors for scribbles that have already been created. At the same time, colorization effects are improved in complex color texture regions under quaternion color representation and hyperanalytic signal analysis. Concept of structure priority based on scale-space contour strength together with quaternion phases forms a reliable structure-preserving colorization foundation.