

基于四元数代数的视频彩色化算法研究

摘要

视频/图像彩色化是计算机视觉领域一个引人关注并具有实用意义的问题，有着广阔的应用前景。灰度图像能传递丰富的信息，但从人眼视觉系统特性来分析，彩色图像还能增添更多为人眼所辨识的细节信息，给人带来更全面的视觉体验。从应用角度看，彩色化处理可用于对旧照片和旧的黑白电影着色翻新，不仅可以使一些图像处理软件增加趣味性，甚至可为影视制片产业带来相当可观的商业利润。

本文就基于四元数代数理论的视频彩色化方法展开研究，概括性地回顾了近年来图像/视频彩色化领域的发展背景，重点概述了几个具有代表性的经典算法，并给出了性能分析与比较；此外，文中回顾了我们小组成员前期的工作基础，并就其不完善的方面给出了算法分析和改进方案。

Anet Levin 等人提出的基于人工 scribble 的局部彩色信息蔓延法是近年比较主流的图像彩色化算法之一，这种基于灰度相关测度的优化算法，在光照稳定的情况下有着良好的视觉效果且运算快速。但是，人工交互给出 scribble 的方式存在以下不足：工作繁杂枯燥，需要一定的彩色化经验。另外，基于灰度的图像表示方式，只能利用亮度通道的相关性做颜色信息蔓延，忽略了结构的重要性，在局部光照存在变化的情形下易有误彩色化操作或欠彩色化操作。

我们小组的前期工作对 Levin 的算法进行了优化，其工作包括基于四元数代数的彩色像素表示，基于相位信息的结构提取；以及引入边界信息以改善 Levin 算法产生的不同对象间的颜色渗透问题；scribble 的自动定位等，缓解了光照不稳定情况下的颜色错误蔓延问题。选择在四元数空间进行颜色信息处理有这么几个原因：四元数空间相对 RGB 空间而言，其三个相位信息间包含的色彩信息相互间的相关性小，而且相比于其他通道相关性独立的颜色空间（如 YUV），直接从相位信息我们就可以分析图像的结构特征，因此单独提取相位信息做颜色处理是很好的选择。小组的前期工作流程可以概括为：

- (1) 用 canny 算子对灰度图进行边缘检测。
- (2) 由边缘信息进行图像重构，得到的重构图像与原灰度图像比较，求得错误图样，找出最大的误差点取为 scribble 的采样点，拟合生成 scribble。
- (3) 将图像从 RGB 彩色空间转换到 NTSC 空间处理，对亮度通道的灰度图进行 Gabor 滤波，提取各方向和频率的输出相位作为表征纹理的特征向量。
- (4) 将融入颜色信息的携带 scribble 的输入图像变换到四元数空间，提取各相位通道的图层。
- (5) 计算各像素的灰度权值 W_i ，边界权值 W_e 和纹理特征同态性权值 W_g ，共同构成最小代价方程组矩阵，根据 scribble 处的色彩信息将其在相位通道进行蔓延，最后还原到 RGB 空间，即完成最终的彩色化过程。

该工作得到了比较好的彩色化效果，在此基础上本文对其不足之处给出了进一步分析和改进，具体体现为：

- (1) 我们给出了一种新的 scribble 自动生成算法。这种想法来自于，我们期望 scribble

自动定位的产生准则是尽量分布在亮度分布不均匀的地方，而空间分布熵这个概念，既表征每个像素点携带信息量的大小又很好地在图像信息熵中融入了空间信息，使得我们可以采用 scope 的搜寻将 scribble 的信息来源点向较小空间分布熵处延伸；另外对边缘二值图进行高斯滤波后可以生成依赖于边缘的权重图，将 scribble 进一步引向近边缘区域，减小彩色化误差。具体实现过程为：根据已经完成的图像分割，在不同区域内部基于空间分布熵的计算进行目标跟踪探测，然后根据探测路线确定较小熵的离散信息点，插值形成 scribble。其中，该实现过程引入了环形中心空间结构的信息熵，其意义在于，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。事实上，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，因此我们可以将这种模型引入到 scribble 定位步骤中。

(2) 在基于相位信息分析图像的结构特征时，通过抑制相位奇点在一定程度上提高了彩色化精度。Gabor 变换最早由 Dennis Gabor 提出，该小波变换与人眼滤波特性相似，因此在图像处理领域得到了广泛的应用。我们可以调节 Gabor 滤波器的尺度参数、方向参数，以获得一组多分辨率、多方向的特征图像，基于该组图像建立的纹理描述充分反映了图像的本征结构。因此，我们构造一组倍频程的 Gabor 滤波器，在提取相位特征时通过 PCA 降维方法抑制相位奇点，降低了因相位不稳定而造成的着色误差。

为验证该彩色化工作框架的合理性，我们在公开测试图片库上测试了提出算法的性能，并与现有经典算法进行了比较分析。实验表明，我们提出的自动定位 scribble 的功能模块在大多数图片上获得了满意的彩色化效果；在结构不连续的图像边缘处，较好地抑制了不同表面之间的颜色蔓延现象。

关键词：视频/图像，彩色化，Gabor，小波，边缘检测，图像分割，信息熵

VIDEO COLORIZATION BASED ON QUATERNION ALGEBRA

ABSTRACT

Video / Image colorization is an significant problem in the field of computer vision. It has very broad application prospects. Despite of the fact that gray images can express sufficient information, however, from the human visual system characteristic, colorized pictures can add to the visual details, and greatly enrich the appreciation of its effect, which can bring more comprehensive visual experience. From the application point of view, colorization can be used to the old black and white photos and old movies coloring renovation. Colorization not only makes some image processing software more interesting, but can actually bring about considerable commercial profits in movie and TV broadcast industry as well.

One of the mainstream colorization algorithms based on manual scribble that has attracted wide attention recently is proposed by Anet Levin et. al. This algorithm is based on color based on the propagation of local color information supposing that colors of neighboring pixels are closely related to intensity. This optimized colorization method has achieved a quite good visual effect to a certain extent in cases of stable illumination, and is a pioneering fast algorithm as well. However, the way of interactive manual scribble has its shortages. For example, the scribble work is dull and doing the job requires sufficient experience of colorization. Besides, for images denoted by gray intensity, color propagation can only be done in the illuminance channel, while the significance of structure is neglected. When the illuminance is changed partly, images could be wrongly or improperly colorized.

In the previous work, our group has optimized Levin's algorithm. The work mainly includes the introduction of quaternion space, the use of phase information for structure extraction, extracting the boundary information to preserve the structures of the original image and prevent the color diffusion between different objects, and the scribble guidance module, etc. These works altogether has help relieved the color diffusion problem occurred when the illuminance is unstable.

Quaternion, as a special Clifford algebra in control theory, is widely used in signal processing, Attitude control, physics and orbital mechanics. It can precisely and easily represents rotation and orientation. In Computer Graphics (and related image analysis field), it is often used to process the problems on rotation and orientation of three-dimensional object. Besides, Quaternion Fourier Transform theory and the corresponding signal analysis problem has formed a system already.

We choose to process color information in quaternion space because of the several advantages of quaternion phase to represent color. Compared with RGB space, there is far less interrelationship between the chromatic channels. However, there is little correlation information between quaternion phases. And for YUV color space, whose channels are also independent, uneven illumination would induce side effects to UV channels. While for quaternion space, it has little impact on quaternion phases of color image, because the structure characteristics can be analyzed from the phases. Hence, it is more reasonable to conduct colorization based on quaternion phase reconstruction.

The entire flow of the colorization can be generalized as follows.

- (1) Edge detection on grayscale image using canny edge operator.
- (2) Reconstruct the image base on the edge information. Obtain the error map by comparing the reconstructed image with the original gray image, obtain error patterns. Select the maximum error point as the sampling points scribble, and plot to generate scribble.
- (3) Transform the image from RGB to NSTC space. Process the grayscale luminance channel through Gabor filter, extract the output phase in different direction and frequency as texture feature vector for each pixel.
- (4) Map the color information to scribble. Transform the input image with carried scribble and color information to the quaternion space. Extract each phase channel layer for color propagation.
- (5) Calculate the weight of each pixel including grayscale intensity weight ' W^i ', edge weight ' W^e ', and Gabor texture characteristics ' W^g '. They altogether constitute the minimum cost equations matrix. Propagate the colors in three separate phase channels by settle the matrix equation. Finally restore RGB image from quaternion space, so to finish the process of colorization.

This work has received fairly good results, but still there are some problems. In this paper we analyze its shortcomings and try to make improvements, which can be described as follows.

- (1) We designed a new scribble guidance algorithm. This idea is put forward as we wish to set some kind of criteria for the scribble guidance, that is, to explore reasonable criteria to measure the amount of information carried by each pixel. And the spatial distribution entropy, can well introduce the spatial information into the image information entropy, which is a good measure of distribution of uneven intensity, and is convenient for us to use the searching scope to extend to the points in smaller entropy area. Besides, the binary edge image that have been processed through Gaussian filter will help to guide scribbles towards the regions near edges, and thus giving more restricts near edges to further reduce color diffusion.

The implementation process in detail can be described as follows. Based on the segmented image, scope tracking is done in each area based on calculating the spatial distribution entropy. The scope always extends to a smaller entropy direction. Scribbles are formed by interpolation. During the process, a space entropy based on a ring structure is introduced, the meaning of which lies in the fact that, the visual stimulation of a given color depends largely on how centralized the spatial distribution of the color is. The more concentrated the color is distributed, the greater visual

stimulate it brings, and vice versa. The larger the spatial distribution entropy is, the more uniform and disperse the color distribution is. And the smaller spatial distribution entropy represents a concentrated distribution of the color. According to this property, this model can be used for guiding scribbles.

(2) In the process of analyzing image structural features based on phase information, we improved the colorization accuracy to a certain extent by inhibiting phase singular points. Gabor transform, which is quite similar to human visual filter function, was first proposed by Dennis Gabor. In recent years, it has been extensively brought under research and put into application in different fields, like wavelet transform, image processing, image compression, image enhancement, edge detection, feature extraction, texture analysis and so on. Gabor filters is an adjustable frequency band-pass filter, and its direction can be adjusted in several directions. By Gabor filter processing, we can extract a variety of high frequency components of two-dimensional image in different directions. By adjusting the scale parameters and orientation parameters, we can get a group of multi-resolution and multi-orientation feature images as the texture description to express the intrinsic structure of the original image. Therefore, we construct a group of Gabor filters with octave, and inhibit the phase singular points by reducing feature vector dimension using PCA. In this way, we properly reduced colorization errors brought by unstable phases.

In order to verify the rationality of this colorization frame, we've tested the algorithm on the public test gallery, and made compare and analysis with the classic algorithm. According to the experiment, our scribble guidance module has achieved satisfying colorization effect on most images. And color diffusion on boundaries of different structures is well inhibited.

Key words: Video colorization, Gabor filter, wavelet, edge detect, segmentation, entropy

目 录

第一章 绪论.....	1
1.1 概述.....	1
1.2 彩色化的研究现状.....	1
1.3 本文主要内容和安排.....	2
第二章 视频/图像彩色化综述.....	3
2.1 经典主流算法.....	3
2.2 课题组前期工作概要.....	3
2.3 四元数空间的引入.....	5
2.3.1 四元数的概念.....	5
2.3.2 四元数的应用意义.....	6
第三章 scribble 定位.....	7
3.1 空间分布熵的概念.....	7
3.1.1 图像信息熵的概念:	7
3.1.2 空间分布熵的概念.....	7
3.2 Scribble 定位算法.....	9
3.2.1 算法构思.....	9
3.2.2 算法详述.....	9
3.3 本章小节.....	17
第四章 边缘及纹理的识别.....	19
4.1 概述.....	19
4.2 Canny 算子简介.....	19
4.2.1 基本原理.....	19
4.2.2 Canny 在彩色化过程中的应用.....	20
4.2.3 基于 canny 边缘的结构分割.....	21
4.3 Gabor 滤波器.....	22
4.3.1 Gabor Filter 简介.....	22
4.3.2 Gabor 滤波器的应用.....	24
4.4 本章小结.....	32
第五章 结论.....	33
参考文献.....	34
谢辞.....	36

第一章 绪论

1.1 概述

视频/图像彩色化是计算机视觉领域一个引人关注并具有意义的问题。彩色化 (Colorization) 的概念最初由 Wilson Markle 于 1970 年提出, 用于处理阿波罗登月计划获取的月球影像。

彩色化处理指对黑白图像/视频添加色彩信息, 使之成为视觉效果更为丰满的彩色图像/视频, 广义上包括伪彩色处理和假彩色处理两种概念。^[1]而前者处理通常用于医学、遥感等领域, 对不同的对象 (可能对象原本并非彩色的) 赋予不同色彩以表达其不同特性, 目的是达到视觉增强的特殊效果; 而假彩色处理是给灰度图赋予接近真实的色彩。这里我们关注后者。图像彩色化有着相当广阔的应用前景。尽管灰度图像能传递充分的信息, 但从视觉上彩图却增添了细节信息, 并大大丰富了其欣赏效果, 给人带来更全面的视觉体验。从应用角度看, 彩色化处理可用于对旧照片和旧的黑白电影的着色翻新, 不仅可以使一些图像处理软件增加趣味性, 甚至在影视制片产业中带来可以一定的商业利润。

传统的旧电影上色需要专业的画师对黑白电影进行标色, 这种过程非常枯燥而繁杂冗长, 并且对画师的有很高的技术要求, 如果能通过计算机实现自动化的上色, 将大大降低对旧电影翻新的成本。

此外, 我们也可以设想, 如果可以仅通过传输单通道的灰度图像和少量的色彩信息来传输视频, 而将彩色处理过程在终端进行, 也可以大大降低传输的信息量。

类似于压缩过程, 视频彩色化相对于图像彩色化而言, 利用帧间的信息冗余也是很有探讨价值的, 例如将运动预测、模板等概念引入视频彩色化, 可以在一定程度上降低单帧的彩色化中不必要的重复工作。

当然, 根据不同的应用需要, 彩色化算法需要着眼于不同的代价需求, 例如强调艺术欣赏性的照片彩色化过程更注重彩色化的精细程度, 对于时间代价可以降低要求, 而实时性较高的彩色化处理过程, 例如视频传输, 则必须有相应的快速算法, 而一些次要细节由于帧间的快速切换则可以适度容忍。

目前我们仍着眼于单帧图像的彩色化处理。相信今后随着该领域研究的逐渐成熟, 人们定能将优化的彩色化算法广泛运用于各类图像和视频处理中。

1.2 彩色化的研究现状

图像彩色化时常被归结为一个“病态问题” (ill-posed problem)[5], 因为彩色图像比灰度图像增加的色彩信息是已丢失的, 需要一定的先验知识给定约束, 因而彩色化是没有定解的。

较传统的彩色化方法中, 人们通常致力于对图像进行分割 (Segmentation), 分区域给出色彩信息, 从而进行着色, 这样的方法通常对于复杂和模糊边缘 (如头发边缘) 的处理效果较差。^[2]因此视频/图像彩色化中的一个重要问题是, 如何通过尽可能少的人机交互提高计算机着色的自动化性能。近几年众多研究者致力于彩色化技术的研究, 提出了许多新的半自

动彩色化算法，这些算法总体上可以分为两类：基于颜色转移的彩色化和局部彩色化向全局扩展的方法。^[1]

基于颜色转移的彩色化算法是由 Welsh 等人于 2002 年提出的，可以说是该领域研究发展一个重大的一个突破，他首次提出了一种基于参考彩图的半自动上色法，对彩色化之后的发展奠定了一定的基础，有相当大的启发性。^[3]

另一种主流的方法是 Anet Levin 等人提出的基于标色的局部彩色信息蔓延法。^[4]他提出了一种基于灰度相关的优化算法，在一定程度上有着良好的视觉效果，也是一种开创性的算法，后人对该方法也不断进行了各种改进的研究。

我们小组的前期工作已经对 Levin 的算法进行了优化，其工作包括对四元数空间的引入，利用相位信息进行结构提取；边界信息的引入，改善 Levin 算法产生的不同对象间的颜色渗透问题；标色的自动定位等。其工作得到了比较好的效果，但也存在一些问题，我们将对其算法进行进一步的改进和优化。

1.3 本文主要内容和安排

第二章中我们将回顾近年来彩色化研究领域的发展情况，比较几种有代表性和开创性的彩色化算法，并比较其各自优缺点；回顾课题组的前期工作，给出概述，并对四元数的概念及其在彩色化处理中的应用意义做了论述。

第三章给出了一种新的基于空间分布熵的 scribble 定位算法，并对空间分布熵的概念作出阐释后给出具体的算法阐述。

第四章中我们针对在我们的算法中占重要工作地位的图像分割的方法做了研究和阐释，并分析了如何对 Gabor 滤波器的有效利用，Gabor 参数的设置以及相位卷绕问题对我们实验结果的影响。

结论部分对整个工作做了总结，并提出对后续研究的建议。

第二章 视频/图像彩色化综述

2.1 经典主流算法

2002 年, Welsh 等人首先提出了一种利用参考彩图提供色彩信息^[3], 基于像素点匹配及传递给目标灰度图着色的方法, 这种半自动的着色方法虽然在自动性能上有一定优势, 但前提必须要建立一定规模的图像数据库以保证对新的灰度图有相似的参考彩图提供彩色信息, 并且在像素点描述和匹配上仍有不足之处。

Welsh 的方法属于上述第一类的颜色转移法, 而对于第二类颜色扩展法, 又有宏观分析法和微观分析法之分。宏观分析法致力于将所有像素与邻近像素的差异总和最小化, 代表有 Levin, Horiuchi 等人。2004 年, Levin 等人提出了基于色彩线索蔓延的优化算法^[4]。其算法基于一个简单的假定: 具有相似灰度值的临近像素具有相似的色度值。利用人机交互对灰度图进行标色, 以此作为约束, 根据灰度值的相似程度确定其他像素的色度, 从而将色彩蔓延到整幅图, 完成整体着色。其中标定这些色彩“线索”的像素称为 scribble。Levin 的这种算法在有了一定量的 scribble 基础上能够很快捷地得到了比较自然的彩色化效果, 避免了繁琐的图像分割。其不足之处在于仅仅考虑了像素灰度值对颜色的影响, 而没有考虑不同对象的纹理结构差别, 因此抗噪声能力较弱, 并且容易受到光照变化的影响, 在灰度值变化平缓的不同对象间会产生颜色的渗透。这对用户标色的位置提出了一定的要求。T. Horiuchi 等人于 2004 年提出了基于概率松弛 (probabilistic relaxation) 的着色法^{[5]-[7]}。

Qu 在对漫画的彩色化算法中, 基于 Levin 的算法, 将图案 (纹理) 和灰度的一致性一并考虑在颜色蔓延的限制条件内, 优化了 Levin 的算法, 然而不足在于没有考虑结构因素, 对局部和全局的纹理进行了同等处理。^[8]

微观分析法是基于对像素颜色变化和灰度值的梯度变化存在相关性这一假设的。Yatziv 引入了“短程线距离”这一概念作为颜色蔓延的依据, 提出了将彩色化过程归结为一种求最小代价函数的方法^[9], 但这种算法需要对每个待着色的像素点寻找最短线距离, 算法代价较大, 同时对着色点的位置也比较敏感。

我们小组的工作主要还是基于 Levin 的方法做算法改进的。如前文所述, Levin 的颜色扩展仅仅考虑了灰度的影响, 一方面我们希望通过纹理结构来进一步约束色彩在不同对象间的渗透, 另一方面我们也致力于最优 scribble 的自动定位, 简化标色的过程。

Z. Yu 优化了 Levin 的算法。^[9]他首先用 canny 边缘检测对图像进行适当的分割, 基于边缘重构后的图像与原灰度图像比较产生错误图样, 将 scribble 定位在重构差错最大的区域。其彩色化过程是基于四元数空间完成的, 利用相位中携带的图像结构信息, 提取相位进行 Gabor 滤波, 提取纹理信息, 结合边界信息和灰度信息, 对色彩进行像素间蔓延, 将问题归结为建立并解最小代价方程, 最终进行四元数的相位重建得到彩图。他的算法在一定程度上限定了颜色在不同对象间的渗透, 但不足之处在于由于重建过程有依赖于边缘信息, 最初采用 canny 算子分割图像时由于一些纹理产生的伪边缘, 在重建时阻挡颜色的蔓延; 或由于灰度相近而遗漏的边缘, 由于重建时边缘、灰度、纹理三者权重共同决定颜色蔓延, 一旦遗漏边缘, 此时纹理信息对颜色渗透便会失效。

2.2 课题组前期工作概要

如前所述, Z. Yu 在基于 Levin 的优化算法的基础上, 提出了一种 scribble 自动定位的方法, 并利用 canny 边缘检测结果以及 gabor 滤波的相位结构信息对颜色蔓延在边界渗透上做了一定的限定。



(a) 原彩图



(b) scribble 定位

图 2-1 scribble 自动定位结果^[9]



(a) Z. Yu 的彩色化结果



(b) Levin 的彩色化结果

图 2-2 利用 Z.Yu 的 scribble 信息两者的彩色化效果

对比两者的上色效果,在相同条件下 Levin 的上色效果在相近的灰度区域总是有着自然的过度效果,这种情况下对象间的渗透难以避免。Z. Yu 的方法已经对一些易辨别的边界进行了良好的防蔓延的效果,但是对毛发边缘等模糊边缘仍有上色不均匀的现象。

就 scribble 定位而言,Z. Yu 通过计算机自动生成了适量的 scribble 位置提供,基本覆盖了许多易出差错的位置,达到对颜色信息较完整的覆盖。如果进一步从用户的取色需求来看,我们更希望,不同色块的 scribble 是清晰分离的,并尽可能避免大量过分细微的碎点。

总体来说,Z. Yu 处理彩色化的整个过程可以概括为:

- (1) 用 canny 算子对灰度图进行边缘检测。
- (2) 由边缘信息进行图像重构,得到的重构图像与原灰度图像比较,求得错误图样,找出最大的误差点取为 scribble 的采样点,拟合生成 scribble。
- (3) 将图像从 rgb 空间转换到 ntsc 空间处理,对亮度通道的灰度图进行 gabor 滤波,提取各方向和频率的输出相位作为表征纹理的特征向量。
- (4) 将融入颜色信息的携带 scribble 的输入图像变换到四元数空间,提取各相位通道的图层。
- (5) 计算各像素的灰度权值 W_g ,边界权值 W_e 和纹理特征同态性权值 W_f ,共同构成最小代价方程组矩阵,根据 scribble 处的色彩信息将其在相位通道进行蔓延,最后还原到 rgb 空间,即完成最终的彩色化过程。

需要指出的是,目前的色彩信息的提供尚且停留在实验阶段,即从原彩图映射 scribble 的色度值,在算法成熟后,我们小组将对 UI 界面进行进一步优化。

我们看到 Z. Yu 的前期工作得到了和原图较为相仿的结果,只有局部区域出现失真或蔓延的现象。在后面的第三、第四章中我们将会对这些现象作出详细的分析,提出改进的方案。

2. 3 四元数空间的引入

由于我们小组的彩色化算法是定义在四元数空间进行,有必要对四元数空间的概念做一下较全面的阐释。本章我们就着重讨论四元数的概念及其在图像处理中的应用价值。

2. 3. 1 四元数的概念

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。^[24]四元数的乘法与复数乘法的一个很大的区别是它不符合交换律,也就是说四元数是复数的不可交换延伸。如果把四元数的集合映射到多维实数空间的话,四元数就代表着一个四维空间,与之相对的复数为二维空间。

四元数和二维复数有些相似,但有三个相位。其定义如下:

$$IH = \{ q = s + x \mathbf{i} + y \mathbf{j} + z \mathbf{k} \mid s, x, y, z \in \mathbb{R} \} \quad (2-1)$$

式(2-1)中正交虚数 \mathbf{i} , \mathbf{j} 和 \mathbf{k} 满足式(2-2)中乘法规则。

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \quad \mathbf{k} = \mathbf{ij} = -\mathbf{ji} \quad (2-2)$$

每个四元数都是 1、 \mathbf{i} 、 \mathbf{j} 和 \mathbf{k} 的线性组合,即是四元数一般可表示为 $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ 。

两个四元数相加就像复数一样只需将相同类的系数加起来就可以。

四元数的乘法则可跟随表 2-1:

表 2-1 四元数乘法表

*	1	i	j	k
1	1	i	j	k

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

共轭的四元数由式 (2-3) 给出

$$q = s - x i - y j - zk \quad (2-3)$$

我们可以将四元数表示成模和相位的形式:

$$q = q e^{j\Phi} e^{k\psi} e^{j\theta} \quad (2-4)$$

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。

$$\psi = -\arcsin(2(q_x q_y - q_0 q_z))/2 \quad (2-5)$$

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:

$$\begin{aligned} T_i(q) &= -i q i = q_0 + q_x i - q_y j - q_z k; \\ T_j(q) &= -j q j = q_0 - q_x j + q_y i - q_z k; \\ T_k(q) &= -k q k = q_0 - q_x i - q_y j + q_z k \end{aligned} \quad (2-6)$$

判断规则为:

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

当 $\psi \in \pm\pi/4$, 则 $\phi = 0$, $\theta = \arg(q T_k(q))/2$, 或 $\theta = 0$, $\phi = \arg(q T_k(q))/2$

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

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

2.3.2 四元数的应用意义

下面就我们使用的四元数空间的变换作一下阐述。

一般许多研究人员在做图像彩色化处理(包括其他着重对图像的颜色进行处理的情况)时, YUV 空间是比较受欢迎的选择。因为单从 RGB 空间来看, 三个通道间色值其实存在很大的相关性, 有不少的冗余信息, 而变换到 YUV 空间则较好地将亮度和色度隔离开来, 并且存在较少的冗余。但 YUV 空间对光照变化比较敏感, 在这一层面上, 四元数空间又在承受光照变化时更稳定, 因此就颜色信息而言也不易受影响, 当然四元数各个通道间也携带较少的冗余信息。综上, 我们在处理彩色化问题中, 在四元数空间是最合理的。

在做空间变换时, 我们采用的是将 $q = a + bi + cj + dk$ 中的 a 置零, 将 r, g, b 三个通道分别赋给 b, c, d 。然后将其转换为如式 (2-4) 所示的模—相位表达形式。则模值记录了亮度信息, 而三个相位各自记录了图像的彩色信息, 并且三个相位通道间存在较小的相关性, 这对我们接下来的彩色化处理提供了很好的稳定性。

在我们的彩色化算法中, 四元数空间主要用于色彩的蔓延——即最小代价方程的构建

$$E(\Phi) = \sum_p \left(\Phi(p) - \frac{\sum_{q \in N(p)} W_{pq} \Phi(q)}{\sum_{q \in N(p)} W_{pq}} \right)^2, \quad \Phi = \{\phi, \psi, \theta\} \quad (2-7)$$

式 (2-7) 是最小代价方程的表达式, 其中的 Φ 是 RGB 图像变换到四元数空间后三个相位通道的集合。 W_{pq} 由 edge 边缘、Gabor 结构、intensity 灰度三个权重(详细见第四章)共同决定, 限定了颜色在像素间的蔓延深度。虽然 W_{pq} 是根据对亮度通道的灰度图像进行各种处理后得到的, 但是它仅仅提供了像素间颜色相关性的衡量尺度, 从 (2-7) 式看到, 最终相位 Φ 才是色度值的记录。四元数空间的相位的良好特性给我们提供了优质的刻录颜色信息

的平台，使得颜色信息在提取、对比、蔓延、重构各环节都几乎不受亮度和其他冗余信号的影响，对同样的算法达到更好的效果。

第三章 scribble 定位

总体来说,我们采用的新的 scribble 定位算法是根据已经完成的图像分割,在不同区域内部基于空间分布熵的计算进行的目标跟踪探测,最后根据探测路线确定离散信息点的,插值形成 scribble 的。

3.1 空间分布熵的概念

3.1.1 图像信息熵的概念:

信息熵的概念源于信息论,最初由香农提出,它由随机变量的概率计算得到,如式(3-1),设离散信源的消息符号集 $v = \{x_1, x_2, \dots, x_n\}$, 则 v 的信息熵定义为:

$$E(v) = - \sum_{i=1}^n p_i \log_2(p_i) \quad (3-1)$$

对于单个消息符号而言,其出现的概率越小,则其包含的信息量越大,信息熵是表征信源本身统计特性的一个物理量,它是信源平均不确定性的度量,是从整体的统计特性上对信源的一个客观描述。^[10]这是一维信息熵的概念。

同样,我们也可以将图像看做一个离散信源,信息熵在图像处理的应用也引起了不少人的关注和研究。图像的二维信息熵可以作如下的阐述^[11]:将图像的各颜色通道(这里的颜色通道可以是灰度通道、RGB 通道,取决于图像处理的空间)的直方图看做概率密度函数,用 h_i 表示量化后的颜色 i 的像素个数占整幅图总像素个数的百分比,设该通道内整幅图有 n 个色值,则我们可以将图像的信息熵表示为(3-2)式

$$E(H) = - \sum_{i=1}^n h_i \log_2(h_i) \quad (3-2)$$

信息熵的概念已被一些研究者被用于图像分割、图像重建和图像压缩中。例如基于 MMI 最小互信息量法^[12]和最大熵值法的图像重建^[13]。

但是单纯的从概率出发计算图像的信息熵虽然将直方图特征从 n 维降到了一维,并表达了图像的灰度/颜色分布情况,然而严重丢失了不同值的像素分布的空间信息。我们希望采取一种能体现空间分布的信息熵,来测度不同像素区域的信息量。

孙君顶等人提出了一种新的结合图像信息熵与空间分布熵的彩色图像检索方法^[14],受其启发,我们将这种融入空间信息的分布熵算法用来作为信息量的测度,以探索 scribble 的合理位置。

3.1.2 空间分布熵的概念

记 $(p_{xy})_{M \times N}$ 为 $M \times N$ 的一幅图像,其中 p_{xy} 表示像素 (x, y) 的颜色,设 $U = \{(x, y),$

$1 \leq x \leq M; 1 \leq y \leq N$ }. 假设彩色图像某一颜色通道对应的颜色数为 N , 令 B_1, B_2, \dots, B_N 为对应的颜色值, S_q 为颜色值为 B_q 的像素的集合, 并定义为: $S_q = \{ (x, y) | (x, y) \in U, p_{xy} = B_q, 1 \leq q \leq N \}$. [14]

记 $C_q = (x_q, y_q)$ 为 S_q 的质心, 这里 x_q 和 y_q 定义为

$$x^q = \frac{1}{|S_q|} \sum_{(x,y) \in S_q} x \quad ; y^q = \frac{1}{|S_q|} \sum_{(x,y) \in S_q} y \quad (3-3)$$

其中, $|S_q|$ 表示 S 集合中的元素总数, 记 r^q 为 S_q 的最大半径。

对于空间的划分, 文献[15]中采用了一种比较符合人类视觉感官特征的中心等距划分的模式。对于一个以质心 C_q 为圆心的圆形区域, 取一个正整数 M , 将半径等分为 M 个等级, 然后以 C_q 为圆心, 以 kr^q / M 为半径 ($1 \leq k \leq M$) 画圆, 得到一个分为 M 个环的圆形区域分割。由此每个环和 S_q 的交点把 S_q 分为 $R_{q1}, R_{q2}, \dots, R_{qM}$. 我们称 $R_{q1}, R_{q2}, \dots, R_{qM}$ 为 S_q 的一个分割, 在后文的scribble 定位算法中, 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3-1所示。

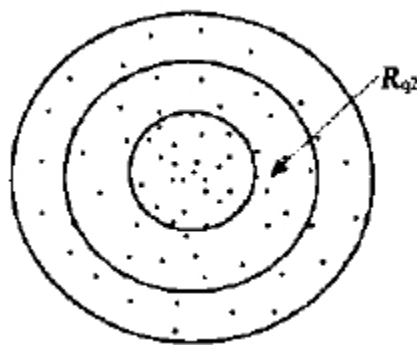


图3-1 scope 模型

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 $d = r_i$ 时为传统的直方图). 对此直方图我们便可以引出与 (3-2) 同等的信息熵计算式。

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。

颜色 i 的空间分布熵可表示为

$$E_i = - \sum_{j=1}^M p_{ij} \log_2(p_{ij}) \quad (3-4)$$

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 (3-4) 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。

3.2 Scribble 定位算法

3.2.1 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寻找能提供最多信息量的像素点集。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。

当然这种度量并不是唯一的。Z. Yu 的构思来自于图像的重建。根据重建后的图像与原图对比形成的最大误差点构成 scribble 的基本点。

类似地，根据 Levin 和 Z. Yu 的算法，基于灰度相关的彩色蔓延在灰度值变化比较敏感，在非边缘的灰度值变化较大的区域对是否同态区域容易产生误判，导致不自然的上色效果，诸如衣服的褶皱，虽然周围都是同态区域，但深色的褶皱区域像素很有可能阻挡颜色的蔓延，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。

如图 3-2 所示，我们以一定环数的圆形区域作为测算空间分布熵的基本单位，类似于生物显微镜的视野——我们称其为一个 scope，通过重复测算当前 scope 的空间分布熵，并将 scope 的中心向周围 8 个方向等距离移动后，计算各自的空间分布熵，取最小者作为移动方向，并取定为新的 scope 中心。不断重复上述过程直到找不到合适的 scope 为止。

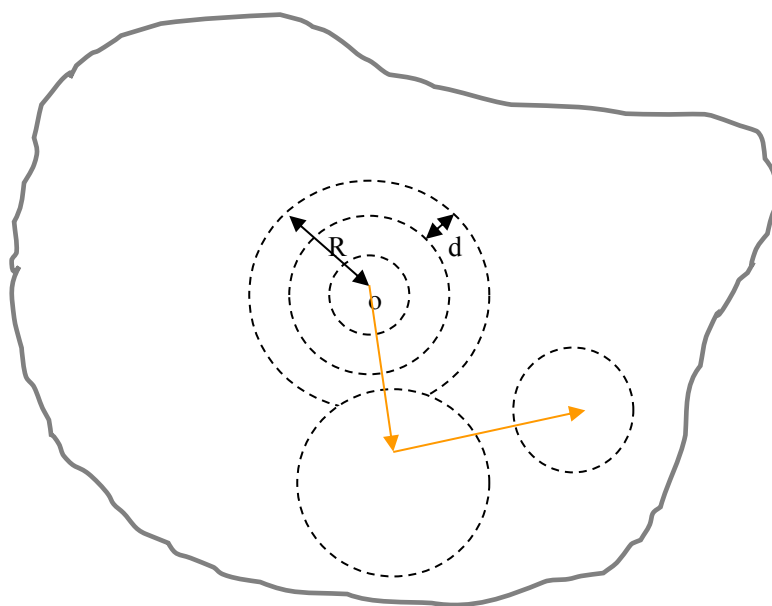


图 3-2 scope 移动示意图

3.2.2 算法详述

整个算法的框架和流程如图三所示，采用可回溯的 scope 跟踪法。（这里的流程是针对已进行分割后图像的某一区域内部进行的。）源程序见 insSearch.m。

3.2.2.1 起始点的确定

分割好的区域以 segI 作为标记图，不同区域以 0.05 的步长为灰度值标记，边缘区域记为 0。segI 如图 3-4 所示。因此在选定区域进行搜寻时，只需从 0.05 开始，循环每次依次以 0.05 的步长将对应灰度值的所有像素点归入待搜寻区域，并形成区域的 logical 二值图。

起始点的确定可以有多种选择，这里为了有足够大的空间预留给 scope 作做搜寻，因此

选择中心作为起始点。由于是连通区域即对所有的 x 和 y 分别取中间值，如 (3-5) 式。

$$x_s = \text{median}(x_i); \quad y_s = \text{median}(y_i); \quad (i = 1, 2, \dots, N) \quad (3-5)$$

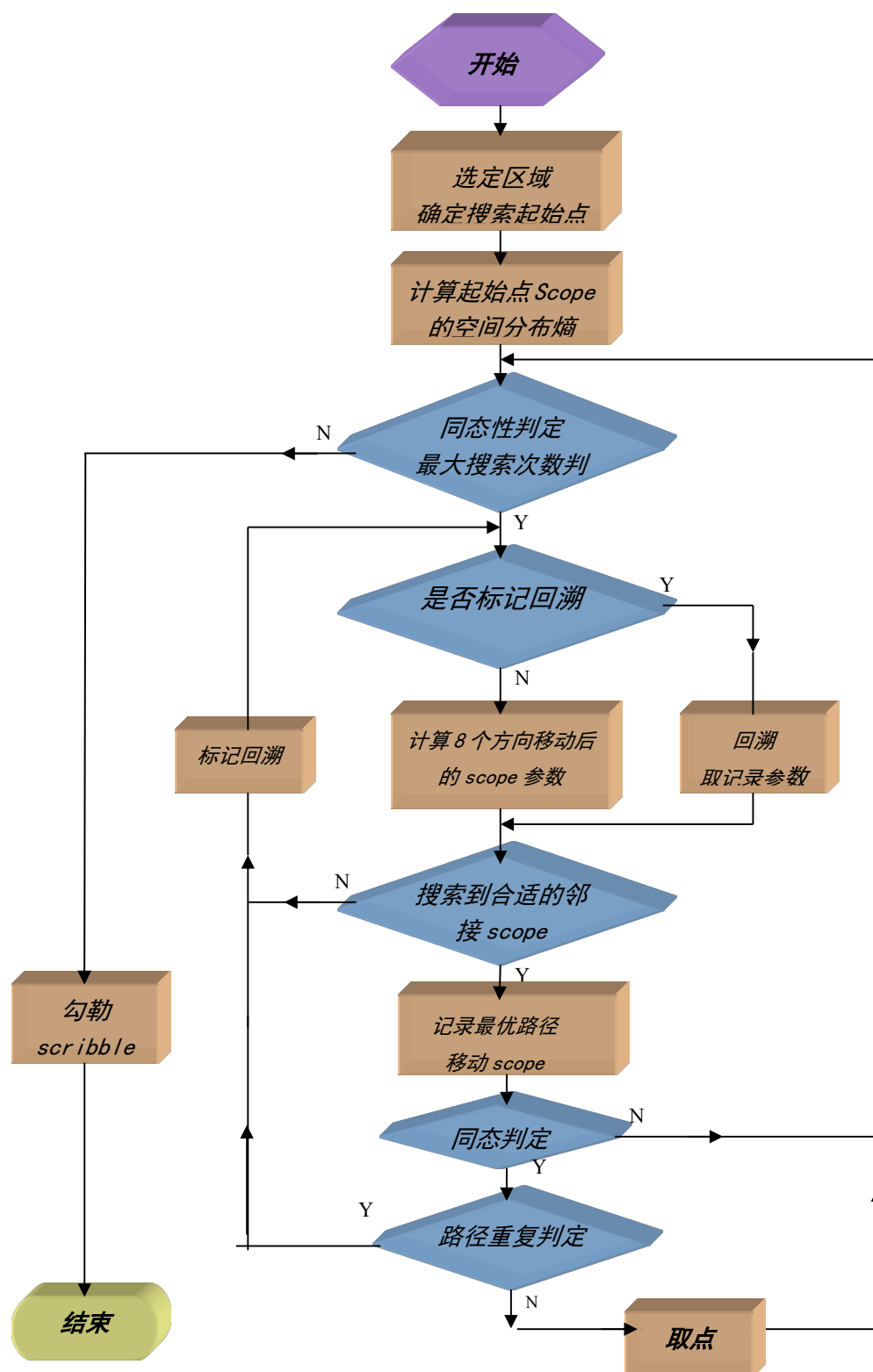


图 3-3 scribble 定位流程图流程图



图 3-4 segI 分割标记图

3.2.2.2 Scope 的参数设定

◆ 环数：

由于我们需要对不同方向的 scope 进行空间分布熵的比较，因此有必要对 scope 内的环数进行一致处理，环数过小则对于熵值的分辨率过低，环数过多则计算量大，并且导致相应的 scope 的半径必须大于一定的值才能进行计算，这样对较小的区域搜索会带来不便。综合考虑我们规定所有的 scope 环数为 4。

◆ 半径 R 初始化：

对于给定的圆心(centreY,centreX)，以及限定的圆环数，半径 R 的最低要求是必须不大于从圆心到区域边界的最小距离（最大内切）。

Maxlen.m 给出了计算最大内切半径的函数，方法是利用 θ 的旋转，求得每个角度 R 延长的最大长度，仍然以区域标记图为依据，选择最小值作为最大内切长度。

有了上述边界限定，我们最终根据内切半径的大小分情况决定 R 的大小，即确定不同情况下 $R = \text{Maxlen} / M$ 的限定中，M 的取值。对于大区域，对 M 取较大的值使 scope 的范围不至于过大，提高精准度；而对于小区域（ $\text{Maxlen} \leq 12$ ）则应取较小的 M，使得 Scope 中每一环有足够的像素供计算分布熵，例如这里我们希望至少每一环有 2~3 层的像素。具体的设定由 DefR.m 确定。

◆ Scope 间相关性系数 rlt_cal：

由于光照或其他原因会使同态对象的不同区域的亮度分布不均匀，这里采用一种简单的类似于均衡化处理的方式，对 scope 中像素分布的大致情况作线性比对。求相关系数的过程见 relate.m。

算法过程描述如下：

设取 N 个等分，分别将两个 scope 内所有像素的灰度值在最大和最小值间等分，

取各自的步长为:

$$\begin{aligned} step1 &= \frac{\max\{p_i \mid p_i \in scope1\} - \min\{p_i \mid p_i \in scope1\}}{N}; \\ step2 &= \frac{\max\{p_i \mid p_i \in scope2\} - \min\{p_i \mid p_i \in scope2\}}{N}; \end{aligned} \quad (3-6)$$

按照步长对像素归入 N 个小组统计, 得到 $bin1 = \{U_1, U_2, \dots, U_N\}$, $bin2 = \{V_1, V_2, \dots, V_N\}$ 。 U_i, V_i 分别为对应第 i 个区间内的像素个数。最后对 $bin1$ 和 $bin2$ 进行矩阵的相关系数计算, 得到的系数表征了灰度分布的线性相关程度。

这种计算是有必要的, 因为后面我们会看到, 这里对空间分布熵的计算只是对占大部分比例的灰度值进行的, 没有涵盖所有的灰度值。

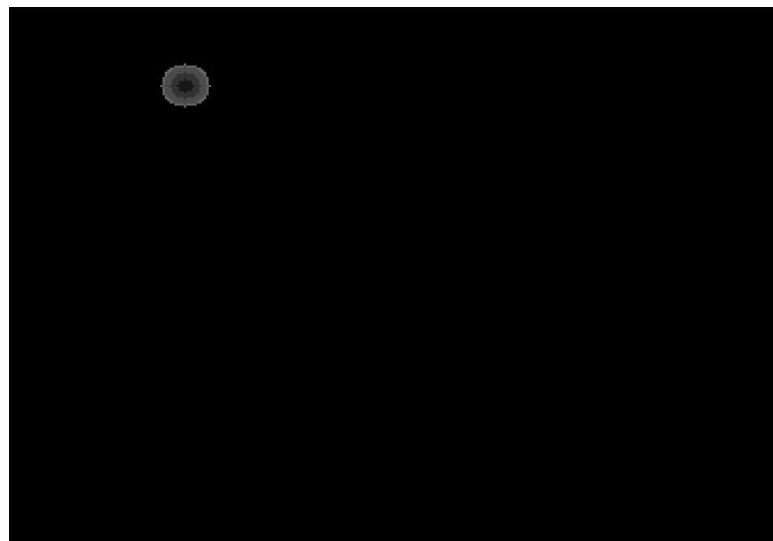
◆ 移动步长:

移动步长指预测周围八个方向 $scope$ 时所取的圆心与当前 $scope$ 圆心的直线距离。综合考虑空间有限, 相关性判定等因素, 我们令两个 $scope$ 有一定的重合, 步长取当前 $scope$ 半径的 1.8 倍。

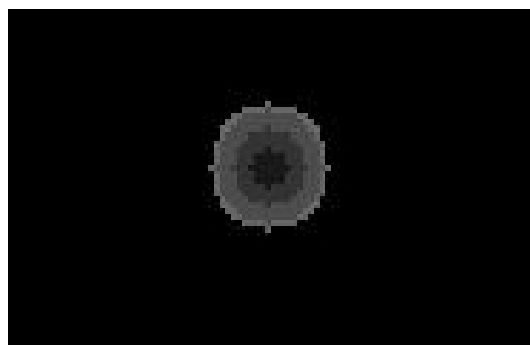
3.2.2.3 空间分布熵的计算

`CalcEntr.m` 中给出了指定圆心的 $scope$ 的空间分布熵计算, 需要不断探测是否每个环都含有分布熵的灰度值通道的像素, 若否, 则 R 以 1 个步长递减, 若 R 小于环数, 则转而将 R 复原, 对于外环没有散点分布的情况, 直接对空间分布熵作减半处理, 表明该区域灰度值变化显著, 熵值较小。如此反复试探, 对每个 R 的分布熵计算调用 `Entr.m`, 像素分布符合要求则返回熵值, 若否, 则返回 -1 标志给 `CalcEntr` 表示需要调整 R 。

`CalcEntr` 每调整好 R , 重新调用 `Entr` 之前, 都要先得到一张对应的 $Scope$ 的指示图, 即以 0.1 的步长标定每一个圆环的灰度值——`map_scope`, 图 3-5 是一张原尺寸的 `map_scope`。调用 `Scope.m` 函数。



(a) 原比例 $scope$



(b) 局部放大图

图 3-5 map_scope 标示图

Entr.m 内部对规定的 R, N, map_scope 参数采取这样的分布熵计算方法：由于这里只针对灰度图作处理，没有多个颜色通道，因此我们将一定灰度范围内的像素都作为同一个色道进行分布熵计算。这里我们将 scope 内所有的灰度值抽取按降序排列，并选择前 60% 的灰度值，再将所有属于这些 60% 灰度范围内的像素提取出来，并映射到 map_scope 的不同环区。按照(3-4)式进行计算。

最终由 Calc_Etr.m 计算并返回熵值，调整后的半径 R 以及 map_scope 的变量信息供 ins_search 决定 scope 的移动方式。

3.2.2.4 scope 移动策略

在有了空间分布熵的计算模块的基础上，整个 scope 移动策略参考图 3 所示，采取可回溯的方法，尽量避免由于单步 scope 移动不当而终止整个搜寻路线。事实上，如果采用递归迭代的方式移动，使整个 scope 路径呈树状扩散，可以获得更多的信息像素点，但是时空代价都比较大，而且过度分散的点集也会给 scribble 的勾勒带来不便。

由于当前 scope 的转移涉及许多参数的传递，我们以一个 cur(rent) 结构记录当前 scope 的所有参数，并设定 rec(ord) 结构数组来记录已经移动的每一步 scope 的参数，供当前 scope 任意步长的回溯。

Cur 结构 rec 数组单元的结构属性如下：

.radius:	scope 的最终半径
.CX, .CY:	centreX, centreY, scope 中心坐标
.entr :	scope 内的空间分布熵(entropy)
.Mentr :	Moved entropy 指周围 8 个方向的 scope 的分布熵值向量。
.Mscope :	周围 8 个方向的 scope 区域标示图
.map_scope :	当前 scope 的区域标示图
.MXY :	周围 8 个方向各 scope 的中心坐标矩阵 [x1, y1; x2, y2; ...; xN, yN]

设置 exist 逻辑变量，当置 1 表示当前没有合适的可移动的 scope，请求回溯；置 0 表示找到下一个移动方向。当检测到 exist 为 1 时将 rec 数组中的最后一个单元结构赋给 cur 结构，在 rec 数组中删除最后一个单元，并将重新赋值后的 cur.Mentr 中的最小一项置 100（也就是拒绝了原来最小熵的路径），选择次小分布熵的 scope 移动。如果不断试探回溯到 cur.Mentr

中 8 个值都为 100，说明没有可移动的方向了，则结束移动。

采用这样的回溯方法牺牲了一些内存空间代价，但是 scope 的移动范围的自由度却大大提升了。如此一来会出现的问题是，因为 scope 每移动一步都有 8 个新的自由度，很有可能迂回地回到走过的路径。为了尽可能避免这种情况，对每一步进行方向限定，但只能控制相一步内的方向不折返。因此采用了另一种判断 scope 区域重合率的方法，防止沿途重返。

S1ctPnts函数是一个实现采点功能的函数，原本用作抽取scope中的点作为scribble，现在用来将每一步的scope中密度前60%的点添入map_scribP (map of scribble points)，每选择下一步scope时需要将map_scribP和待选择的scope做映射，如果包含在该scope中的map_scribP中的点占该scope所有像素点总数的比值超过一定大小(如程序中设定6%)，则判定为路径重复，此时也应当将exist标识置1。

此外，为了防止一些意想不到的长循环，我们将scope移动步数设定了30~40的上限（包括回溯的步数）。

3.2.2.5 同态性判定

关于同态性判定的一些说明。我们在3.2.2.2中已经对相关性系数的算法做了阐述，原本做这样的判定是为了给scope的搜寻设定一个比较合适的终止点。我们对每一个新的相邻的scope相关性系数与之前所有相关性系数的均值做比较，当小于某一比例时则判定为不相关。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(3-7)式判定是否满足同态限定。

$$r_{M+1} \geq \frac{V}{M} \sum_{i=1}^M r_i \quad (3-7)$$

但是当 V 取得较大时会过度抑制 scribble 的伸展，我们以局部的 scribble 为例，对比图 6 中不同参数的效果。（源代码见 SpcEntrMax.m）

从图中可以看到，在较大的阈值下 scribble 会被限制在起始点所在的局部区域里，可能造成颜色信息覆盖面过小不能达到很好的蔓延效果，当然，蔓延效果也是和其他边缘的界定有关的。如果这里希望 scribble 的范围尽可能广，可以适当减小 V。



(a) V=0.06 时的结果



(b) $V=0.6$ 时的结果

图 3-6 不同的相关系数阈值的影响

3.2.2.6 位置权重偏移

由于基于灰度相关的彩色化过程的差错位置总是在边缘附近出现渗透蔓延，为了使这些梯度较大的位置得到更准确的彩色信息，我们有目的地将 scribble 的位置尽可能向边缘位置靠拢。因此我们根据边缘信息，对边缘的黑白二值图进行二维高斯滤波，进行模糊处理，滤波器的表达式如式 (3-8)，

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-(u^2+v^2)/(2\sigma^2)} \quad (3-8)$$



图 3-7 边缘模糊 betaI , $\sigma = 10$

滤波输出 BetaI , 其中的每个像素都可以作为边缘距离的权重因子, 如图 3-7 所示。有了 betaI , 则在比较 8 个方向的 scope 熵值时, 将熵值和对应 scope 的中心位置的 betaI 像素值相乘后再做比较。

但对于不同的边缘需要对 σ 进行适当的调整, 因为在区域分割普遍较小的情况下, σ 过大会引起边缘模糊后在区域中心的混叠, 结果适得其反; 在区域普遍较大的情况下, 应当适当增大 σ 的尺度, 否则对离边缘较远的区域几乎很少受影响。

3.2.2.7 scribble 采点

scribble 的采点来自路径 scope 中的点集, 如 3.2.2.4 节中所述的采点是一种方法, 但是无规则散点的曲线拟合比较复杂, 因此仅仅取每个 scope 的中心点作为插值点。图 3-8 显示了用 SpcEntrMax 对几个主要区域选择区域采点的结果。



图 3-8 局部采点举例 (相关性阈值约 0.16)

3.2.2.8 曲线拟合

这里 scribble 曲线的精确拟合是个比较棘手的问题, 在二维空间随机分布的散点很难一



(a) $z. \text{yu}$ 的 scribble

(b) 基于分布熵定位的 scribble
(相关性阈值取 0.05)

图 3-9 scribble 定位结果对比

一映射到x或y方向。而如果采用分段处理,由于点数节点数较少,在接口的平滑又难以处理。

这里我们以y方向为函数拟合的自变量,并对y值重合的点进行微调后做三次插值拟合,函数见ppfit.m, ppint.m, ppdiff.m等函数(借鉴jonas.lundgren等编写的函数),拟合后的scribble如图8所示。

这种方式描点的结果可能会产生一定的越界情况,我们在边缘处作了适当的截断,多余的scribble暂时保留,提供补足的色彩信息。但信息点的拟合方式还是有待考究的。

3.3 本章小节

本章详细阐述了设计的一种新的scribble自动定位算法。

总体而言,通过对其进行彩色化实验的结果表明,这种算法在合理的图像分割的基础上可以达到比较全面的关键信息点定位结果。观察图3-9容易发现,新的基于空间分布熵搜寻策略的scribble定位算法得到的scribble与原来的scribble定位结果相比更有局部整体性,除了极小的细节区域外,有较少的凌乱部分,更方便用户的整体颜色决策。并且通过把scribble引向靠近边缘的区域,而不是把scribble放在边缘线上,既大大增加了限制边缘处颜色渗透的辅助效果,同时也从用户的角度考虑,不容易发生人眼观察的歧义。

当然,它也有自身的不足之处,最显而易见的就是依赖于图像分割的效果。观察图3-4的分割结果,例如在婴儿衣服中央的小花就没有足够细化地被分割,定位时容易遗漏细节。另外,由于算法的效果本身涉及参数阈值的设定,阈值的设定没有完全做到自动计算。图3-10显示了两种相关性阈值下的scribble定位效果。



(a) 相关性阈值0.16

(b) 相关性阈值0.05

图3-10 不同参数下的scribble定位和彩色化结果

对于我们的新的scribble，后续工作还有改进的空间，这里提出尚未进行实验的建议和想法。

一方面，可以从其所依赖的图像分割着手，既然后续的彩色化算法已经运用到了Gabor小波变换的相位特征提取结构信息，联想到可以也可以顺带利用其幅度信息丰富边缘检测的依据，或者利用Gabor变换的特征值（多方向多尺度滤波后的均值和方差）所表征的纹理信息对同一幅图像中的相同结构/纹理进行聚类，进一步简化多余的scribble。

另一方面，也可以从图3-4所示的分割图出发考虑。由于我们目前只是针对非边缘区域进行scribble定位，并且scope需要一定大小的区域做空间分布熵的计算，因此距离canny边缘更近的区域可以说是盲区，建议可以从图3-4所示的边缘区的边界着手增加额外的scribble，这样可以进一步改善我们的scribble的不足。

第四章 边缘及纹理的识别

4.1 概述

由于我们的整个算法需要以边缘限定颜色的渗透，并且新的 scribble 算法也是基于图像分割基础上进行的，因此需要有相当质量的图像分割做铺垫。

常用的分割算法大致可以分为基于边缘的分割和基于区域的分割两类。常用的边缘检测算子包括 sobel 算子、Robert 算子、Prewitt 算子、Kirsch 算子、Canny 算子等，基于区域的分割包括基于直方图的阈值分割、区域生长法、形态学分水岭算法、特征空间聚类（典型的例如 K-均值聚类法、ISODATA 法、模糊 C-均值法等等）等方法。当然也有将两者结合的方法。近年来也有越来越多的关于小波变换（wavelet）在图像分割中的应用，我们所用到的 gabor 滤波就是其中之一。另外还有基于神经网络、支持向量机、基于免疫算法等等诸多人工智能领域的分割方法。更多的介绍可以参考文献[18]

而在最开始的边缘检测我们仍沿用 canny 算子进行检测，这是出于其相对精确的检测效果和成熟而广泛的应用的考虑。

本章中，我们先对采用的 canny 算子做一简单的介绍，然后将在第三节中探讨 gabor 滤波器的用来提取边缘信息和纹理信息的相关问题，第四节讨论我们对 Z. Yu 的算法的一些改进处理。

4.2 Canny 算子简介

Canny 边缘检测算子由 Canny 于 1986 年提出，由于具有良好的信噪比和检测精度，近年来在许多图像处理领域中得到广泛的应用。^[16]

4.2.1 基本原理

Canny 把边缘检测问题转化为检测单位函数极大值的问题。一个边缘在高斯噪声中即一个节约的强度变化，Canny 基于该模型给出了评价边缘检测性能优劣的三个指标：信噪比、定位精度和单边缘响应，并用数学表达式表示了上述判据，^[17]采用最优化数值方法，得到对应给定边缘模型的最佳边缘检测模板。对于二维图像，需要使用若干方向的模板分别对图像进行卷积处理，再取最可能的边缘。

其针对一维边缘中的阶跃型边缘，推导出的最优边缘检测器的形状与高斯函数的一阶导数类似。利用二维高斯函数的对称性和可分解性，则易算的高斯函数在任意方向上的方向导数与图像的卷积。

由卷积运算的线性性，实际运用中选取高斯函数的一阶导数作为阶跃边缘的次最优检测算子。

由二维高斯函数表达式：

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4-1)$$

在某方向 n 上 $G(x, y)$ 的一阶方向导数为

$$G_n = \frac{\partial G}{\partial n} = \mathbf{n}^T \nabla G \quad (4-2)$$

其中 $\mathbf{n} = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$, $\nabla \mathbf{G} = \begin{bmatrix} \partial G / \partial x \\ \partial G / \partial y \end{bmatrix}$, \mathbf{n} 是方向矢量, $\nabla \mathbf{G}$ 是梯度矢量。 \mathbf{n} 的矢量方向改变, G_n 与 $f(x, y)$ 的卷积最大时, \mathbf{n} 就是正交与检测边缘方向。

变, G_n 与 $f(x, y)$ 的卷积最大时, \mathbf{n} 就是正交与检测边缘方向。

4.2.2 Canny 在彩色化过程中的应用

本节主要阐述 Canny 边缘检测在我们的彩色化算法中的应用, 并对算法思想进行一定的描述。

由于我们小组的彩色化算法目的在于改进 Levin 彩色化过程中的颜色蔓延问题, 不同对象间的边缘结构是重要的信息。我们的目的在于通过某种限定关系将边缘信息、结构信息对色彩的控制放在和灰度信息同等重要的位置。Z. Yu 的彩色化方法最后给出的用于彩色信息在相位空间蔓延的最小代价限定方程如式 4-3 所示,

$$E(\Phi) = \sum_p \left(\Phi(p) - \frac{\sum_{q \in N(p)} W_{pq} \Phi(q)}{\sum_{q \in N(p)} W_{pq}} \right)^2, \quad \Phi = \{\phi, \psi, \theta\} \quad (4-3)$$

其中 p 表示每一个 pixel, $q \in N(p)$ 为像素 p 的 3×3 邻域内的相邻像素, W_{pq} 由三部分组成, 如式 4-4 所示

$$W_{pq} = W_{pq}^e W_{pq}^g W_{pq}^i \quad (4-4)$$

其中,

$$W_{pq}^e = \exp(-|e(p) - e(q)|^2) \quad (4-5)$$

$$W_{pq}^g = \exp(-H(p, q)) \quad (4-6)$$

$$W_{pq}^i = \exp(-(I(p) - I(q))^2 / 2\sigma_p^2) \quad (4-7)$$

分别为 edge 边缘、gabor 结构、intensity 灰度三个权重。

W^e 信息从 canny 边缘检测的结果获得, 控制在边缘处的颜色蔓延, W^i 也就是 Levin 所依据的灰度相似性程度的度量, 使得颜色能在相近的灰度像素间蔓延; W^g 的信息构成较为复杂, 它是根据将灰度图经过不同方向和不同尺度参数的 gabor 滤波器滤波后, 输出的每个各个尺度下相位组成各像素特征向量, 从而表征各像素携带的结构信息得来的。

Canny 边缘检测放在整个流程的初始进行, 为了后续算法都能根据需要利用该边缘检测结果, 仿真过程中直接将计算得到的边缘信息图以 bmp 格式保存。

在 Matlab 的库函数中存在 edge 函数, 可以通过直接调用该函数, 设置参数用 Canny, Sobel, Prewitt, Roberts 等各种算子进行边缘检测, 但是不同的参数下以不同的尺度运算得到的边缘信息层次不齐, 难以辨别哪些是主要的结构边缘, 哪些是次要的纹理边缘, 因此不能仅仅以一次检测得到的边缘信息。下面介绍整个处理过程。

- (1) 将灰度图通过高斯滤波器做平滑降噪, 这里取模板大小取 3, sigma 取 0.8。记滤波后输出图像 fg 。

- (2) 对 fg 计算每个像素的梯度，包括 x 方向、 y 方向和模值。
- (3) 设定 $canny$ 检测的深度 $deepness$ （即 σ 尺度变化范围），这里取深度为 10。也就是 10 次循环，对初值为 0.8 的 σ 每次以一定的步长递增。如 4-1 中所示，当增大 σ ，尺度变大，检测出的边缘也越粗。由 $scale_space_canny$ 函数计算并返回 10 个尺度下边缘信息，记为 ss_canny 。
- (4) $scale_map_canny$ 函数对(3)中的多尺度边缘进行整合，计算各个边缘的生命周期。从小尺度开始，依次对于每个尺度下的边缘做一定范围的膨胀处理，找到其和上一个尺度的重合部分，增加其生命周期记录。最终的结果记为 $life_time_map$ 。
- (5) $scale_space_link$ 函数根据（2）中得到的梯度信息对 ss_canny 和 $life_time_map$ 中的边缘进行膨胀，连接不同尺度下的边缘。

通过以上步骤，我们将计算得到的所有生命周期的边缘集合和最长生命周期的边缘信息保存下来。如图 4-1 所示。

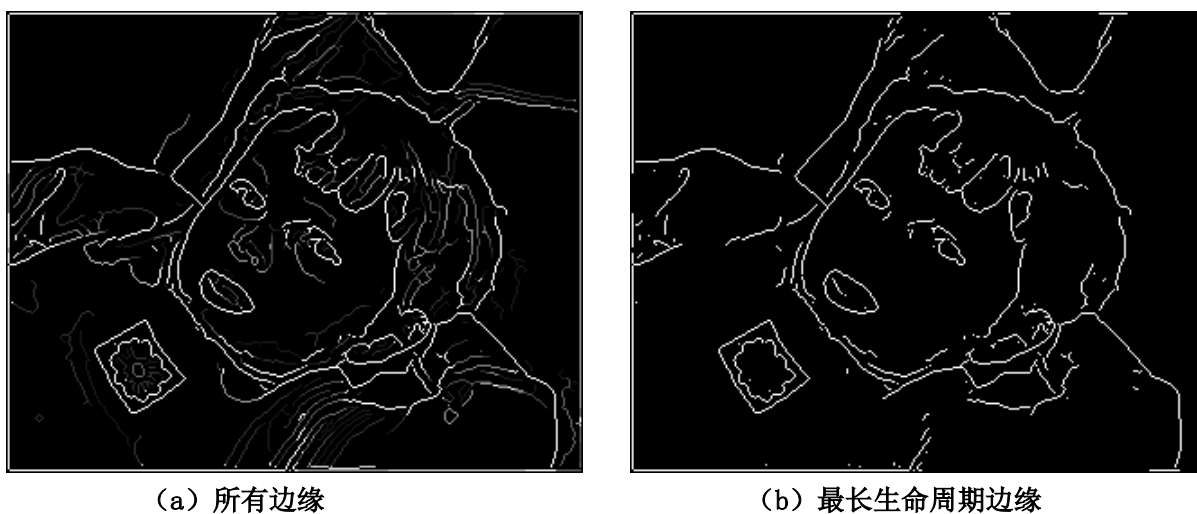


图 4-1 边缘信息

4.2.3 基于 $canny$ 边缘的结构分割

上一节得到的边缘信息一方面被用作颜色重构时的边缘权重计算，另一方面也被采用来进行进一步区域分割，为第三章所述的区域内部的基于空间分布熵的 $scribble$ 定位做准备。

这里采用一种比较巧妙的方法做分割。我们对如图 4-1 所示的边缘图进行连接处理后，对其分别用宽度为 5 像素和 3 像素的矩形结构进行膨胀处理，得到的两个二值图像再进行相减，于是得到如图 4-2 所示的副边缘二值图。这样就把整张图隔离成有限个封闭的区域，通过灰度标记的方式区分各区域——以循环的方式选择没有标定的区域，通过和边缘图的映射关系判定是否边缘区，将原来包含边缘的区域置零，而对其他区域按 0.05 的步长标记灰度值，知道再也没有剩余区域大小大于某一阈值为止，结束标定。所得的结果在图 3-4 已经给出。

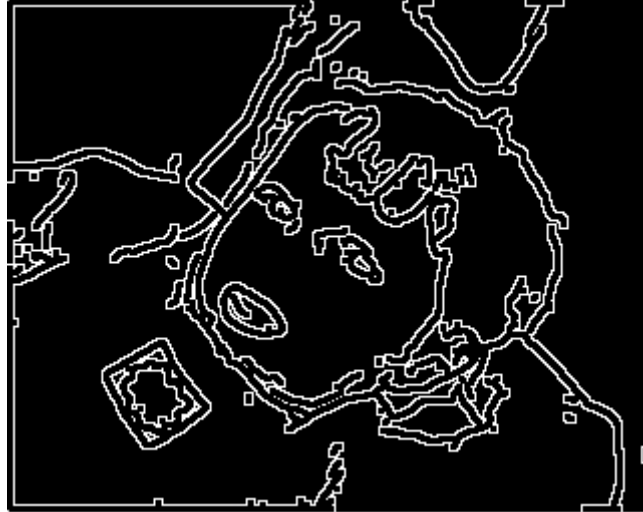


图 4-2 副边缘二值图

4.3 Gabor 滤波器

本节主要就是围绕该 W^s 能否很好地表征结构信息展开讨论的。首先先对 gabor 滤波器做一下介绍。

4.3.1 Gabor Filter 简介

Gabor 变换最早由 Dennis Gabor 提出，D. Gabor 本人还被公认为小波变换的创始人之一，并于 1946 年提出了短时傅里叶变换(Short Time Fourier Transform)。

近年来小波变换在图像处理领域，尤其是图像压缩、图像增强、边缘检测、特征提取、纹理分析等方面都得到了广泛的研究探讨和应用

小波相对于波而言具有有限宽度（窗口），其尺度对应于波中的所称的频率，我们对小波变换做如下的定义：

$$T^{wav} f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \quad (4-8)$$

其中，小波基函数为：

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (4-9)$$

小波变换的基本思想是用一系列函数去逼近一信号或函数,这一系列函数即小波基,它是由小波经过伸缩和平移得来的。而基本小波 $\psi(x)$ 具有如下性质：^[16]

波动性：有正有负，积分为 0；

衰减性：当 $|x| \rightarrow \infty$ 时， $\psi(x)$ 的衰减足够快，即平方可积；

带通性： $\psi(x)$ 在零和趋于无穷时都趋向于 0。

4.3.1.1 1-D Garbor 变换

从频域上看，一维的 Garbor 滤波器本质上是高斯函数被正弦载波调制到高频后形成的带通滤波器，时域（空域）上看就是高斯核和复数正弦函数的乘积，如 4-10 所示：^[19]

$$g(t) = k e^{j\theta} w(at) s(t) \quad (4-10)$$

其中,

$$w(t) = e^{-\pi t^2}$$

$$s(t) = e^{j(2\pi f_o t)}$$

其中的 k, θ, f 都是滤波器的参数。如果将复数滤波器分成实部和虚部两部分, 可以写作 4-11。

$$g_r(t) = w(t) \sin(2\pi f_o t + \theta) \quad (4-11)$$

$$g_i(t) = w(t) \cos(2\pi f_o t + \theta)$$

容易知道, 滤波器冲击响应的峰值出现在 f_o 的地方, 若取幅度衰减一半时为带宽, 即

$$\hat{w}\left(\frac{f - f_o}{a}\right) = e^{-\pi \frac{f - f_o}{a^2}} = 0.5 \quad (4-12)$$

则有

$$f - f_o \pm \sqrt{a^2 \log 2\pi} = 0.4697 a \approx 0.5 a \quad (4-13)$$

由此可知, a 可以用来表示带宽。

4.3.1.2 2-D Gabor 变换

Gabor 核函数的定义为:

$$g(x, y) = s(x, y) w_r(x, y) \quad (4-14)$$

其中同样的, $S(x, y)$ 是一个复数正弦函数, $W_r(x, y)$ 是一个二维的高斯函数的包络 (envelope)。其中的载波定义如 4-15 式所示,

$$s(x, y) = \exp(j(2\pi(u_0 x + v_0 y) + P)) \quad (4-15)$$

U_0, v_0 是空间上的频率, P 为相位。同样地, 我们也可以将该复数正弦函数分为实部和虚部两部分处理。并将空间频率 F_0 和方向 ω_0 定义为:

$$F_0 = \sqrt{u_0^2 + v_0^2}$$

$$\omega_0 = \tan^{-1}\left(\frac{v_0}{u_0}\right) \quad (4-16)$$

因此 (4-15) 又可以写作:

$$s(x, y) = \exp(j(2\pi F_0 (x \cos \omega_0 + y \sin \omega_0) + P)) \quad (4-17)$$

二维的高斯包络可以表示为:

$$w(x, y) = K \exp\left(-\pi\left(\frac{(x-x_0)_r^2}{a^2} + \frac{(y-y_0)_r^2}{b^2}\right)\right) \quad (4-18)$$

下标 r 表示 rotation 旋转因子，可以调节 gabor 滤波器的方向，即

$$\begin{aligned}x - x_0 &= (x - x_0) \cos\theta + (y - y_0) \sin\theta \\y - y_0 &= (y - y_0) \cos\theta - (x - x_0) \sin\theta\end{aligned}\quad (4-19)$$

a, b 作为 x,y 方向不同的高斯尺度因子，刻画了二维高斯椭圆包络的长短轴信息。K 作为高斯函数的系数，应该等于 1/ab。图 4-1[19]显示了一种参数选择下的高斯包络图形。

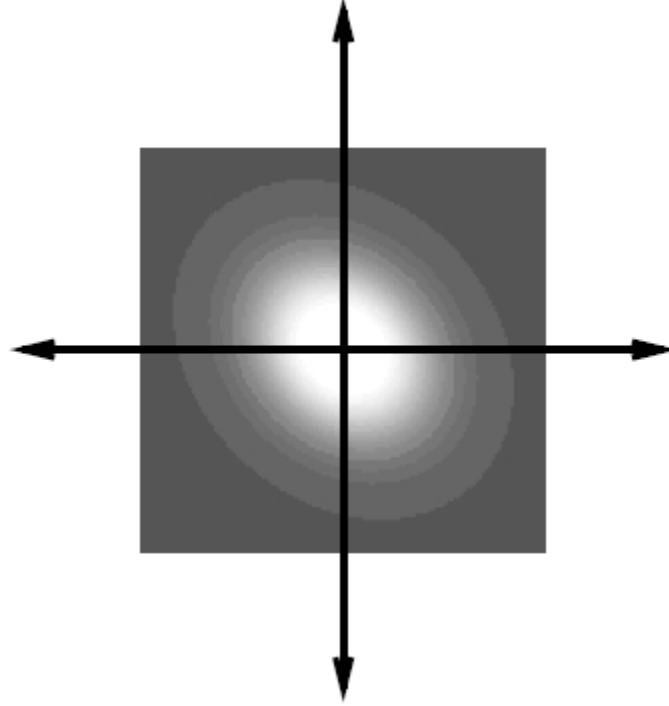


图 4-1 128*128 像素高斯包络^[19]
a=50, b=40, theta=-pi/4, x0=y0=0

但实际应用中，人们通常做如下的假定：

- (1) $a=b=1/f$;
 - (2) $x_0=y_0=0$;
 - (3) 震荡函数改写为 $s(x,y)=\exp(i*2*\pi*f_0*(x\cos\theta+y\cos\theta))-\exp(-\pi^2/2)$;
- 因此整个 Gabor 滤波器记为：

$$g(x, y, f, \theta) = \frac{1}{ab} \exp\left[-\pi\left(\frac{x_r^2}{a^2} + \frac{y_r^2}{b^2}\right)\right] [\exp(i2\pi f x_r) - \exp(-\frac{\pi^2}{2})] \quad (4-20)$$

其中最后一项 $\exp(-\pi^2/2)$ 是对 DC 分量的补偿因子，使得整个函数的积分为 0。

4.3.2 Gabor 滤波器的应用

Gabor 滤波器的诸多应用我们难以一一涵盖，这里只针对我们该课题的研究需要讨论 Gabor 滤波器的边缘检测以及纹理结构的分割两个问题进行探讨。

4.3.2.1 边缘检测

有了 gabor filter 的基本概念之后，我们希望将其为我们所用。由上述介绍我们可以看，作为具有可调节频率的带通滤波器，并且可以多方向任意调节方向，通过 Gabor 的滤波处理，我们可以找到二维空间中各种方向的高频分量——也就是我们希望提取的各种方向的边

缘信息。因此通过调节尺度和方向，我们是能够利用 Gabor 滤波器进行较高质量的边缘检测的。但是 Gabor 滤波器参数诸多，可调节范围大，既是它可被充分利用的优点，但同时也是难以调节最优参数实现最优效果的难点。

我们曾经就如何确定最优参数进行图像 Gabor 滤波这一问题展开研究。马瑜等人在其基于四元数-Gabor 的纹理分割研究中[21]，曾经使用一种多通道 Gabor 滤波的模型，其中 Gabor 滤波器的中心频率的设定是基于对输入图像的能量谱分析并用 CSA 算法[13]自动找到能量谱中几个出现最大峰值的频率点，作为多个 Gabor 滤波器的中心频率，再分别调节多个方向来采集输出图像的特征值，随后进行通道合并，得到每个像素点的一组特征向量后，用 PCA 主分量法进行特征空间降维，随后再根据先验知识确定图像分块的个数，用 K-均值法进行聚类分析。

其整个过程如图 4-2 所示。

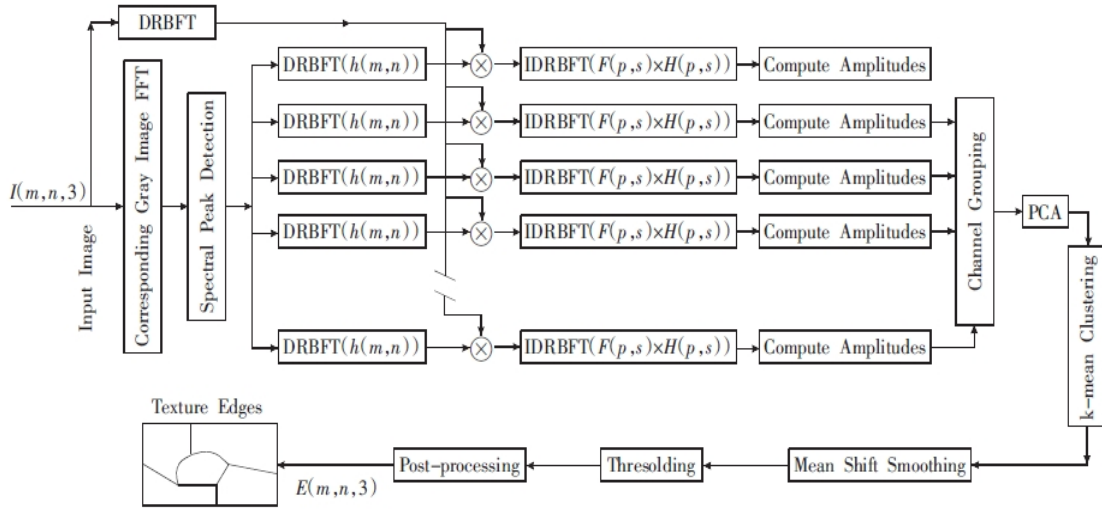


图 4-2 基于四元数-Gabor 的纹理自动分割流程图^[21]

可以说这是一种智能性相当高的参数自动选择的方法，但从整个流程来看，从分析能量谱直至多通道合并需要多步细化流程，并且对于纹理特征不是那么鲜明的图像是不适合这种需要先验知识的纹理分割型区域分割方法的。

虽然最后没有采用这种方法，但其间也对 CSA 峰值自动搜寻做了代码编写工作，可以参考附录，实验结果表明这种自动搜寻峰值的算法对少量的白噪声有一定的抑制作用，但对于集中于某一区域的大量噪声的存在可能会使结果不尽如人意。

通过和导师间的交流，我们还是放弃了这些繁琐的流程，最终采用一种修整后的 Gabor 滤波器，以变化 scale 以及通道方向的方式，取输出图像的最大响应模的一些像素点作为边缘。

修整后的 Gabor 滤波器空域表达式如 (4-21) 所示。

$$g_{uv}(x,y) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2(x^2+y^2)}{2\sigma^2}\right) \bullet [\exp(ik \bullet \begin{pmatrix} x \\ y \end{pmatrix}) - \exp(-\frac{\sigma^2}{2})] \quad (4-21)$$

这里我们以 σ/k 替换原来的 σ_0 ，而保持 σ 不变。因此在变换高斯核的尺度时，只需调整 k 的值。

其中：

$$k = \begin{pmatrix} k_x \\ k_y \end{pmatrix} = \begin{pmatrix} k_v \cos \varphi_u \\ k_v \sin \varphi_u \end{pmatrix}$$

$$k_v = 2^{\frac{v+2}{2}} \pi$$

$$\varphi_u = u \frac{\pi}{K}$$

(4-22)

其中 v 的取值决定了 Gabor 滤波的波长($\lambda=1/f$), u 的取值确定 Gabor 核函数的方向, K 表示总的方向数, 通常取 $6 \sim 8$ 。参数 σ/k 代替原来的 σ_0 决定了高斯窗口的大小, 这里取 $\sigma=1.414*\pi$ 。其中 $\exp(-\sigma^2/2)$ 仍然是 DC 部分的补偿因子

我们的程序中取 4 个频率 ($v=0, 1, \dots, 3$) 和 8 个方向 ($K=8, u=0, 1, \dots, 7$), 32 个 Gabor 核函数。图 4-3 显示了一个典型的 Gabor 滤波器 (实部) 的立体图。

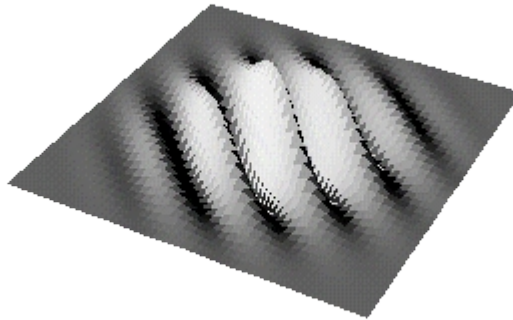


图 4-3 Gabor 实部三维图

在构造 gabor 滤波器时需要注意一些问题。

首先要注意取合适的 Gabor 滤波器的模板大小: $sz = \text{ceil}(\text{num_periods}/\text{freq})$;

x, y 分别取 $-sz:sz$, 但要值得注意的是, gabor 模板的下标还是从 1 开始的。

一次给定参数的 gabor 构造如下所示:

```
for x=-sz:sz
    for y=-sz:sz
        xh = x * cos(theta) + y * sin(theta);
        yh = y * cos(theta) - x * sin(theta);
        Greal(sz+x+1,sz+y+1)=coef*exp(-coef*xh*xh/2-coef*yh*yh/2)*...
            (cos(kh*xh)-exp(-sigma*sigma/2));
        Gimg(sz+x+1,sz+y+1)=coef*exp(-coef*xh*xh/2-coef*yh*yh/2)*...
            (sin(kh*xh)-exp(-sigma*sigma/2));
    end
end
```

这里采取的是将实部虚部分开构造的方式, 那么滤波时也应分开进行, 最后取模值, 即:

```
realgab = conv2(gI,double(Greal),'same');
```

```
imggab = conv2(gI,double(Gimg),'same');
```

当然也可以直接构造复数 Gabor 滤波器 G , 滤波后再分别提取实部和虚部。

例如:

```
gabout = conv2(gI,double(G),'same');
```

```

absI = abs(gabout);
phaseI = angle(gabout);
realI = real(gabout);
imgI = imag(gabout);
...

```

由于取不同的 θ 值和 V 的值，将所有尺度的输出图像做叠加后的图像，归一化后的总响应最大的像素点就构成了边缘。

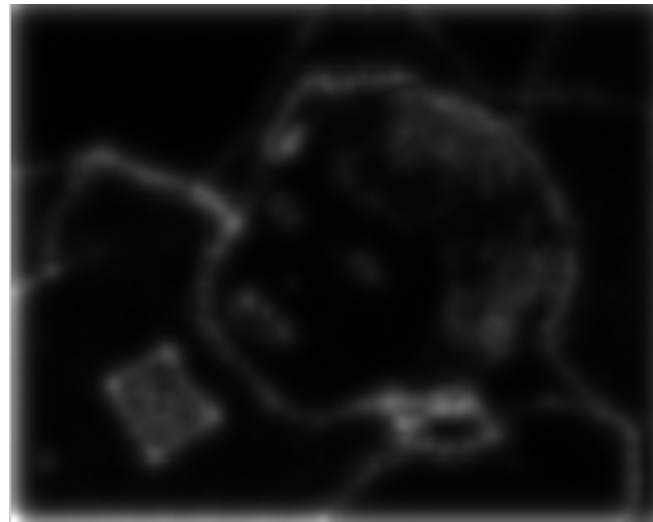


图 4-4 边缘检测结果

4.3.2.2. 结构信息提取和相位卷绕问题

(1) Gabor 变换在结构信息提取和纹理分析的应用：

Z. Yu 在式 4-3 中所提到 W^g ，事实上是利用 gabor 滤波器对变换到四元数空间后的灰度图像进行 8 个方向、4 个尺度的 gabor 滤波，分别对输出图像取相位信息，计入特征值，因而每个像素有 32 维向量作为特征值向量。

Z. Yu 认为，由于相位信息记录了结构和位置特征，相同结构的相邻像素经 gabor 滤波应当具有相似的相位，因此直接利用其各尺度下均值作为该像素的结构特征值，计算 W^g 后对结构间蔓延进行限定。

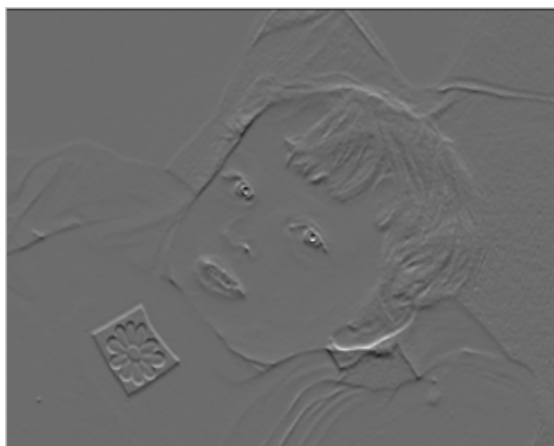
理论上这是成立的，但是为什么出现了边缘处的颜色微小失真，我们对程序代码进行深入检查和测试。

Z. Yu 所构造的 gabor 滤波器中的 k 的取值区别于 4-22，为

$$k_v = 2^{\frac{v+2}{2}} \pi \quad (4-23)$$

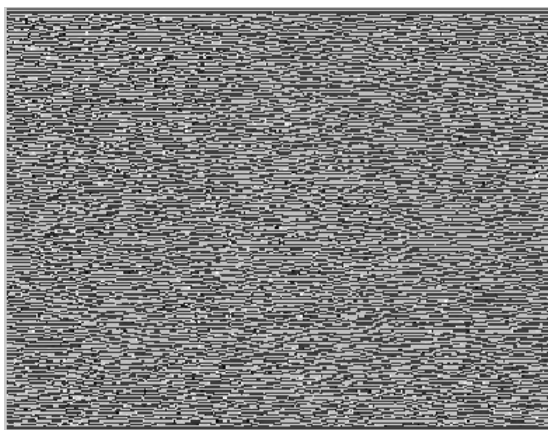
差别在于指数的符号，于是 k 的取值从 2π 开始， $f=k/2\pi < 1$ ；输出模值图基本等同于原图，而相位如图 4-5 所示，与边缘图相近。

通过实验我们也可以看出，采用这种极小尺度相位的边界限定与将式 4-3 中的 W^g 取消后的结果并无太大区别，如此相位特征值并没有发挥其应有的指示结构的作用。



**Gabor output Phase Image, $K=17.77$, $V=-7$,
 $f=2.8284$, $\theta=0$**

图 4-5 $k > 2\pi$ 时的相位结果



$V=-2, f=0.5, k=3.14$

图 4-6 $k < \pi/2$ 时的相位情况

(2) 相位卷绕问题:

当把 k 的计算式改正后，我们发现相位输出实际上有很严重的卷绕。如图 4-6 所示。这样导致了不合理的特征值聚类。在 4-3 式中引入的 Gabor 权重因此发挥了不当的蔓延效果，如图 4-7 所示。

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。

由图 4-4 可知，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 4-7 受相位卷绕影响的彩色化效果

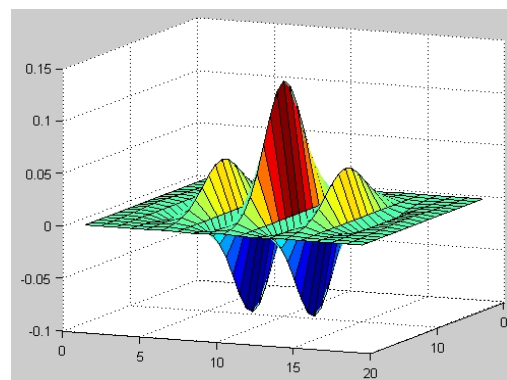
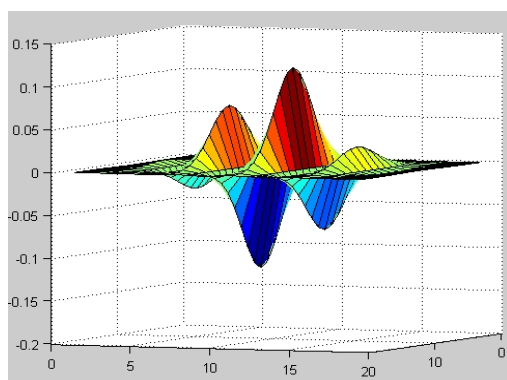


图 4-8 Gabor 滤波器的三维示意图

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。



图 4-9 奇点抑制效果图

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 4-9

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。

4.3.2.3 PCA 主分量分析

上述的多方向多尺度 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 4-10 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:

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1N} \\ y_{21} & y_{22} & \dots & y_{2N} \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1N} \\ x_{21} & x_{22} & \dots & x_{2N} \end{bmatrix} = UX \quad (4-24)$$

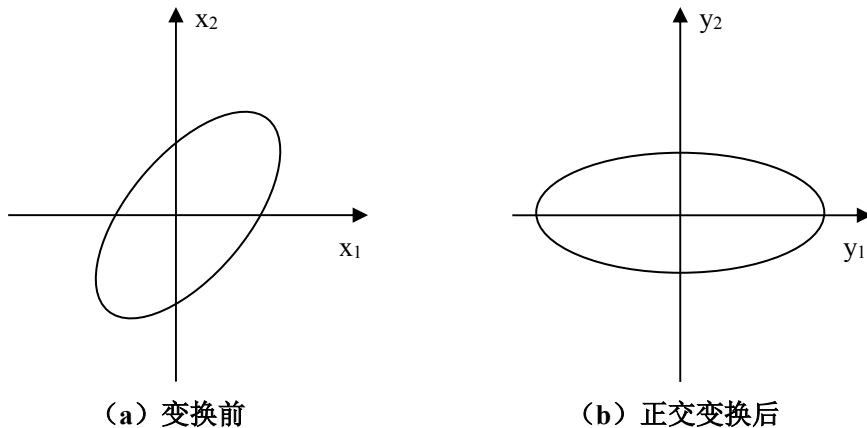


图 4-10 二维特征主成分分析示意图^[24]

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 4-10 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。

我们通过求方程组的特征根的方法容易求得 y 主分量的各个方差，以及变换矩阵和主分量关于具体的推导在此不做过多解释，可以参见相关文献。程序见 `pca.m`。

用 PCA 对 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。如图 4-11 所示，分别为取 1 个主分量、3 个主分量和 6 个主分量的效果，显然通过比较容易发现，取 3 个主分量的比较合理，只有 1 个主分量的情况下涵盖的特征信息量不够全面，而过多的 6 个分量则导致了较小方差情况下的颜色渗透。



(a) 保留 1 维主分量 (b) 保留 3 维主分量 (c) 保留 6 维主分量

图 4-11 保留不同主分量的彩色化结果



图 4-12 其他测试效果
(左原图，右彩色化效果)

4.4 本章小结

本章以 canny 边缘检测和 Gabor 纹理/结构提取为线索，阐述了我们小组彩色化工作的整体算法，分析了彩色化结果出现问题的原因所在，分析原理并探讨改进的方法，给出一些实验结果。

概括来说，多尺度的 Canny 边缘检测一方面给基于区域分割的 scribble 算法提供了图像分割的依据，另一方面提供了边缘线索权重 W^e 的计算依据。而 Gabor 小波变换具有多尺度多方向的丰富特性，其幅值输出可以作为边缘检测的另一种手段，其输出相位中记录了像素的位置结构信息，我们对传统的 Gabor 滤波器进行形式上的修正，使其尺度和方向得到数量级的量化，更适合于算法和程序设计的需要。多方向多尺度的 Gabor 滤波结果通过特征提取可以用来表征不同的纹理，但根据需求的不同，我们目前对 Gabor 滤波器的应用还是着重放在对其相位信息的利用。小组前期在理论上提出了将 Gabor 输出相位作为像素的结构特征的表示，利用这些相位信息计算结构权重 W^s ，我们的工作修正了其 Gabor 滤波器的设置错误，并发现了相位卷绕问题的存在。目前的工作进展到利用相位奇点的判定一定程度抑制边缘附近的色彩蔓延，相位卷绕问题有待进一步有针对性地解决，另外也采用 PCA 主成分分析法对相位特征向量进行降维处理，如前所述有一定的作用。

以上是目前的彩色化改进成果。由于研究时间限制，进一步的改进措施有待提出和落实，包括上述的解相位卷绕问题，以及如何更全面地发挥 Gabor 小波变换的作用，使之更好地与 Canny 边缘检测和图像分割融合在一起，等等。

第五章 结论

整个毕设过程针对 scribble 的自动定位生成以及在四元数空间中利用 gabor 滤波器分辨结构的问题进行了研究和讨论。

基于空间分布熵的 scribble 跟踪定位的方法与 Z. Yu 的 scribble 相比，有更好的集中性，除了细微区域的碎点之外，有较少的零碎点，因为当应用于 UI 界面给用户取色，零碎点通常会造成困扰；同时一方面使 scribble 尽可能靠近边缘区域，又不使得 scribble 恰好处于边界使得用户的颜色选择决策产生模糊。

这种基于空间分布熵的 scribble 定位具有信息集中定位的优势，但其缺点在于依赖图像分割的前期工作，漏分割会减少 scribble 的覆盖广度，令一方面由于计算空间分布熵所需的图像空间代价造成了边缘邻近区的部分盲区，希望通过从边界出发适当增加额外的 scribble 信息。另外我们希望在后期的工作中还能利用 gabor 的特征提取给图像区域分割聚类做支持，达到更好的效果。

就彩色化算法而言，目前基于 Z. Yu 的最小代价方程的算法最大的障碍就在于如何解决第四章中提到的相位卷绕问题。我们在挖掘了改问题后只从处理奇点的角度着手做了改进。今后将就该问题做进一步有针对性的探索。相信通过进一步研究会取得更优的彩色化结果。

参考文献

- [1] 李志永, 滕胜华, 杜坤, 邹谋炎等. 基于不平度颜色混合的图像彩色化方法[J]. 电子与信息学报, Mar.2008, Vol.30 No.3
- [2] 胡伟, 秦开怀. 高分辨率灰度图像的快速着色[J]. 计算机学报. 2009.5。
- [3] Welsh, T, Ashikhimin, M, and Mueller, K. Transferring color to grayscale images,”[C] in Proceedings of the 29th ACM International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2002, 21(3), pp. 277-280.
- [4] Levin, A, Lischinski, D and Weiss Y. Colorization using optimization.[C] In proceedings of ACM SIGGRAPH 2004. pp. 689 – 694, 2004.
- [5] T. Horiuchi, “Colorization algorithm using probabilistic relaxation,”[J] *Image and Vision Computing*, 2004, 22(3), pp. 197-202.
- [6] Takahiko Horiuchi. Estimation of color for gray-level image by probabilistic relaxation. .[C] Proc. IEEE 16th International Conference on Pattern Recognition, Quebec, Canada. 2002, (3): 867-870.
- [7] Takahiko Horiuchi and Sayaka Hirano. Colorization algorithm for grayscale image by propagating seed pixels.[C] Proc. International Conference on Image Processing, Barcelona, Spain. 2003: 457-460.
- [8] Qu, Y.G., Wong, T.T., Heng, P.A., “Manga Colorization,”[C], *ACM Transactions on Graphics* 25(3), 1214-1220.
- [9] Z. Yu, Y. Xu, X.K. Yang, L. Song, “Structure-Preserving Colorization Based on Quaternionic Phase Reconstruction,”[C], *PCM 2009*, December 15-18, 2009, Bangkok, Thailand.
- [10] 吴伟陵. 信息处理与编码[M] , 人民邮电出版社。
- [11] John Z M., “An information theoretic approach to content based image retrieval,” [D] Louisiana State University and Agricultural and Mechanical College, Phd. Thesis, 2000, 45-62.
- [12] J. Sun, L. Wu, Q. Gui, “An Improved Multiscale Maximum Entropy Image Restoration Algorithm”[C], *IEEE*, 2006
- [13] J. Sun, L.H. Wu, Q. Gui, “An Improved Multiscale Maximum Entropy Image Restoration Algorithm”[C] , 2006 *IEEE*.
- [14] 孙君顶, 丁振国, 周利华等. 基于图像信息熵与空间分布熵的彩色图像检索方法.[J] *红外与毫米波学报* 2005 年 4 月, 第 24 卷第 2 期。
- [15] Rao A B, Srihari R K, Zhang Z F, “Spatial color histogram for content-based retrieval,” [C] *Tools with artificial intelligence. Proceedings of 11th IEEE International Conference.* Washington, DC: United States, 1999, pp.183—186
- [16] 余松煜, 周源华, 张瑞, 等. 数字图像处理[M]. 上海: 上海交通大学出版社。
- [17] J.F. Canny, “A computational approach to edge detection” [J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1986, 8: 679~698.
- [18] 许新征、丁世飞、史忠植、贾伟宽等. 图像分割的新理论和新方法[J]. *电子学报* 2010

年 2 月 Vol.38, No.2A。

- [19] Javier R. Movellan. Tutorial on Gabor Filters[J/OL]
- [20] 马瑜, 王晓辉, 等. 基于 Quaternion-Gabor 滤波器的彩色纹理分割算法[J]. 计算机工程与应用, 2007, 43(28), pp. 51
- [21] Tan T N., “Texture edge detection by modeling visual cortical channels,”[J] Pattern Recognition, 1995, 28(9): 1283-1298
- [22] Zhou, J., Xu, Y., Yang, X.K., “Quaternion wavelet phase based stereo matching for uncalibrated images,”[J] Pattern Recognition Letters 28(12), 1509-1522 (2007).
- [23] 杨淑莹. 模式识别与智能计算——Matlab 技术实现.[M] 北京: 电子工业出版社
- [24] 维基百科 <http://zh.wikipedia.org>

谢辞

在论文结束之际，谨此向一直以来向我的导师徐奕老师致以衷心的感谢和崇高的敬意！感谢她让我能融入到图像信号处理这一领域的研究工作中，并一直以来对我循循诱导，使我从毫无经验到重现算法再到开创自己的想法，指导并鼓励我，使我充分体验其中的意义和乐趣所在，也让我感到四年的学习确实有很大的收获。

感谢于舟学长在彩色化课题的前期工作投入的大量精力，以及在任务转交时给我的指导和启示。没有他的认真工作也没有课题现在的进展。

同时也要感谢宋利老师对我的帮助，他让我有机会在暑假实习参与图像所的研究，为毕业设计打下良好的基础。感谢他和实验室其他学长们在平时对我的关照。

感谢大学期间给我们教授课业的所有老师，使我们在大学学习中受益匪浅。

感谢家人和朋友们的支持与理解，和一直以来对我的关心、鼓励。

感谢学术领域的研究者们不辞辛劳地以自己的智慧一点一点构筑我们学科的堡垒和人类智慧的桥梁。

感谢为评阅本文而付出时间和精力专家、教授们和教务工作者们！