

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

学士学位论文

THESIS OF BACHELOR



论文题目：基于四元相位分析的黑白图像彩色化方法

学生姓名：蔡仲锴

学生学号：5097149007

专 业：信息工程

指导教师：解 蓉

学院(系)：电子信息与电气工程学院

基于四元相位分析的黑白图像彩色化方法

摘要

黑白图像彩色化是一个具有广阔应用前景的课题。本文总结了目前主流的彩色化算法，重点讨论了课题组前期实现的基于空间分布熵的自动 scribble 生成算法和基于四元 Gabor 相位的图像彩色化方法。为了在彩色化过程中考虑到图像的边缘信息，我们引入了相位一致性原理来衡量图像的边缘强度，介绍了相位一致性的原理和实现方式，还通过实验表明，引入相位一致性后彩色化效果得到了显著的改善。此外，本文中完成了 windows 平台下算法的完整 C++ 实现，并对后续工作提出了改进方向。

关键词：彩色化，四元 Gabor 变换，相位一致性

IMAGE COLORIZATION METHOD BASED ON QUATERNION PHASE ANALYSIS

Colorization of grayscale image is a topic with broad application prospects. The paper summarizes the current mainstream algorithms of colorization, focusing on the spatial distribution of entropy-based automatic scribble generation algorithm and image colorization based on quaternion Gabor implemented previously by our research group. In order to take into account the edge information in the image colorization process, we introduced phase congruency to measure the significance of edges in the image. In this paper, we not only present the theory and implementation of phase congruency, but also show the significant improvement in colorization due to the introduction of phase congruency. In addition, we managed to implement the whole algorithm via C++ on Windows platform, and proposed follow-up directions in the end.

Key words: Colorization, Quaternion Garbor Transform, Phase Congruency

目 录

第一章 绪论	1
1.1 黑白图像彩色化问题产生的背景	1
1.2 黑白图像彩色化问题的研究现状	1
1.3 本文的主要内容和章节安排	3
第二章 基于四元相位的彩色化算法	4
2.1 Levin 的算法	4
2.2 scribble 的自动生成算法	5
2.2.1 空间分布熵理论	5
2.2.2 基于图论的图像分割算法	6
2.2.3 scribble 的生成	7
2.3 最优化算法	8
2.3.1 利用四元 Gabor 相位分析图像结构	8
2.3.2 Canny 边缘检测	8
2.3.3 最优色彩蔓延	9
2.4 本章小结	11
第三章 基于相位一致性的边缘检测	12
3.1 相位一致性简介	12
3.2 局部能量和相位一致性	12
3.3 利用小波计算相位一致性	14
3.4 噪声的影响	14
3.5 扩展到二维	17
3.5.1 二维滤波器设计	17
3.5.2 滤波器方向	18
3.5.3 二维噪声补偿	18
3.5.4 结合不同方向上的结果	18
3.6 频率扩展的重要性	19
3.7 通过高通滤波来确定尺度	21
3.8 相位一致性的实验结果与对比	23
3.9 本章小结	24
第四章 程序实现和实验结果	25
4.1 软件模块设计	25
4.2 基于相位一致性边缘检测的实现	26
4.3 实验结果和相关讨论	26
第五章 总结与展望	33
参考文献	34
谢辞	36

第一章 绪论

1.1 黑白图像彩色化问题产生的背景

图像是人类获取信息,表达信息和传递信息的重要媒介。研究表明,人类从外界获取的信息四分之三以上都来源于视觉系统,而图像、图形又是所有视觉信息的载体。

在彩色摄像技术发明之前,人们拍摄了大量的黑白图像,后来,随着彩色摄像技术的产生和普及,为这些黑白图像加上颜色,就自然成为了人们研究的课题。传统的黑白图像上色方法是由艺术家对图像进行手工上色,这种方法极为耗时费力,而且对操作者的艺术水平有较高的要求,因此,这种方法不能够用于大量黑白图片的染色,应用有限。

电子计算机问世之后,人们开始寻求借助于计算机的自动染色方法。Wilson Markle 在 1970 年第一次提出彩色化(colorization)这一术语来描述他发明的计算机辅助黑白图像上色技术,这种方法被用于处理阿波罗登月计划获取的月球影像。

黑白图像彩色化是一个具有实际意义和应用前景的课题。彩色图像相对于黑白图像来说,有着更加丰富的细节信息和更为全面的视觉体验。利用彩色化方法对旧照片和旧电影进行着色,可以降低旧照片和旧电影的翻新成本;彩色化方法还可以对医学、遥感领域的非彩色图像进行伪彩色处理,用不同的颜色表示不同特性,达到视觉增强的特殊效果;此外,如彩色化还可以作为一种压缩方法,信道中只要传输单通道的黑白图像和少量颜色信息,然后在终端进行彩色化,以此大大降低传输的信息量。

1.2 黑白图像彩色化问题的研究现状

将彩色图像转换为黑白图像,主要应用的是亮度公式 1-1 或照度公式 1-2,由于这两个公式均非可逆,因此彩色化实际上是一个病态问题(ill-posed problem),对于同一个灰度值,在采用单字节(8 比特)存储方式下,有可能存在 256×256 种不同的色彩组合。

$$Gray = (R + G + B)/3 \quad (1-1)$$

$$Gray = 0.299R + 0.587G + 0.114B \quad (1-2)$$

按照在彩色化过程中人为干预的程度,可以将彩色化方法分为三大类 [1]: 手工上色,半自动上色和自动上色。

手工上色是借助于 Photoshop、BlackMagic 等软件人为地对图像涂抹色彩。这种方法必须人为地对图像进行分割,需要较多的时间和劳动力。

半自动上色是通过一种预设的映射关系,将照度直接映射到色度,通常是借助于一个查找表(LUT)来实现。这种方法的一个典型例子就是伪彩色图像,在医学图像处理上有着广泛的应用 [2]。

自动上色方法是被研究得最为广泛的一类上色方法,涵盖了非常广泛的范畴。目前对于自动上色方法的研究还没有完善的结果,之前采用的方法或者需要大量的人工干预,或者准确性不高,或者两者兼有。我们可以把这一方法划分为以下三种类型。

- 1) 变换型: 对各个灰度值 $I_g(x, y)$ 应用某个变换方程 T_k , 得到各个色度 $I_{ck}(x, y)$ 的值。见公式 1-3。

$$I_{ck}(x, y) = T_k[I_g(x, y)] \quad (1-3)$$

- 其中, $I_{ck}(x, y)$ 表示第 k 个通道的色度值。对于每个通道 k , 都对应一个单独的变换函数 T_k , 有时候变换函数的构建还会用到更多信息, 例如像素点的位置等。
- 2) 图像匹配型: 这一方法常常采用一幅已经上好色的图像作为参考。首先, 在原灰度图像的各个像素 I_g 和参考图像的各个像素 I_s 间寻求匹配, 使得两像素间 (对于某个相似度衡量) 其距离 E 最小。然后建立一个 I_s 和 I_g 间的映射关系, 将每个灰度值映射成彩色。显然, 这一做法的关键便在于选择一幅合理的参考图像。选择参考图像的主要方法有三种: 手工选择, 自动挑选和基于聚类的方法。
 - 3) 用户选择型: 这一方法首先需要用户对灰度图像进行标记。这一标记一般是将图像的某些区域指定为彩色。用户做的标记常常称为 scribble 或 seeds。然后, 从这些 scribble 开始进行蔓延, 将相似的区域用与之匹配的 scribble 的颜色进行涂抹, 并最终达到给整幅图像进行上色的目的。这一方法也是本课题研究的类型。

在基于 scribble 的图像彩色化技术中, 大多数研究人员都从 Levin [3] 等人的工作中受到了启发, Levin 假定亮度相近的点具有相似的颜色, 基于这一设想, 可以依据灰度差别构造出一个代价函数。同时, 将 scribble 覆盖的像素点作为已知条件, 并通过对距离以及亮度相近的像素赋予较大的权重, 从而优化这个代价函数, 即可获得彩色化的效果。Horiuchi [4] 提出了基于概率松弛 (probabilistic relaxation) 的彩色化算法, 在预设图片满足局部 Markov 特性的前提下, 将问题转化为利用概率松弛解决组合优化问题。Drew [5] 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。Kim [6] 等人使同一边缘结构内的色彩一致, 产生了边缘保持更加稳定的彩色化结果。Lezoray [7] 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。Heu [8] 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。

与此相类似, Sapiro [9] 考虑灰度图像的梯度, 并将 scribble 作为边界条件求解; 最终把彩色化问题归结为 Poisson 方程的求解。

Yatzi [10] 基于最小路径的快速求解算法提出了一种快速的彩色化方法。该方法首先计算各点到 scribble 上各点的最短路径, 然后将这些路径利用特定的函数进行调和, 从而得到需要的彩色值。这一方法在不产生严重性能损失的前提下具有较快的运算速度, 而且对 scribble 的位置依赖性较低, 即对于不同的 scribble 能输出较为接近的结果。

本课题组的工作主要是基于 Levin 的方法做算法改进。在先前的工作中, Z.Yu [11] 优化了 Levin 的算法。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, 建立纹理结构描述, 同时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, 最终通过恢复的四元相位信息重建得到彩图。该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, 在重建时会阻挡颜色在同一结构内的蔓延, 或因为灰度相近而遗漏边缘。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, 有可能导致颜色在不同结构之间渗透。X.Ding [12] 提出了采用四元小波变换来衡量图像的结构差异性, 同时针对原算法中需要人工选择 scribble 的不足, 提出了一种基于空间分布熵的自动产生 scribble 的算法, 这两者加上一些细节改进, 构成了一套较为完备的彩色化方法, 实验结果显示其处理效果相较于原算法有一定进步, 避免了某些情

况下的颜色泄露问题。李粤龙 [13]在 Ding 的基础上,将原本的 MATLAB 代码移植到了 C++平台上,解决了一些细节上的问题,改进了生成 scribble 的插值方法,提高了代码效率。

1.3 本文的主要内容和章节安排

第二章主要介绍课题组前期的工作成果,从 Levin 的算法出发,重点介绍课题组已经完成的完整的 scribble 生成算法,以及采用四元小波变换和 Canny 边缘检测产生代价函数进行最小代价优化得到彩色化结果的算法体系。

第三章提出了前期工作中存在的问题,针对彩色化过程中边缘信息未有效利用的问题,引入了利用相位一致性计算边缘权重的算法,重点讲述了相位一致性的原理和特点。

第四章我们针对彩色化算法在 Windows 平台下的 C++实现,将整个算法细分为几个模块,介绍了各个模块的功能及其实现方式,给出了对不同图像的彩色化实验结果对比,并分别做了进一步讨论。

第五章是结论部分,总结了现有的工作成果,对现有算法仍然存在的局限性进行了总结和概括,并对后续工作提出了一些建议。

第二章 基于四元相位的彩色化算法

2.1 Levin 的算法

Anat Levin [3]等人提出的基于人工 scribble 的局部彩色信息蔓延法是近年来影响力较大的主流彩色化算法之一，课题组的前期工作就是对 Levin 的算法进行优化，所以本章首先介绍 Levin 的算法。

假定图像在 YUV 颜色空间，Y 是亮度通道，U 和 V 是色度通道，我们研究的彩色化问题可以重新表述为：

给定集合 $\Omega = [0, R] \times [0, C]$, $T \subset R^+$ ；给定三个通道 $Y, U, V: \Omega \rightarrow T$ 。已知灰度通道 $Y(x, y)$ 以及 scribble 区域 $\Omega_c \subset \Omega$, $|\Omega_c| \ll |\Omega|$ ；已知 $\forall U(x, y): (x, y) \in \Omega_c$ 及 $\forall V(x, y): (x, y) \in \Omega_c$ ，欲求各 $U(x, y): (x, y) \in \Omega/\Omega_c$ 及 $V(x, y): (x, y) \in \Omega/\Omega_c$

Levin 的算法不仅适用于黑白图像，也适用于黑白视频，方便起见，用粗体字母（如 \mathbf{r} , \mathbf{s} ）来表述像素点的坐标 (x, y, t) 。

Levin 的方法是基于一个前提：如果两个相邻像素 \mathbf{r} , \mathbf{s} 亮度相近，那它们的颜色也应该相近。由此出发，我们希望最小化 \mathbf{r} 处的色度 $U(\mathbf{r})$ 与它领域内像素色度的加权平均之差，即最小化公式 2-1 中的 $J(U)$ 。

$$J(U) = \sum_{\mathbf{r}} (U(\mathbf{r}) - \sum_{\mathbf{s} \in N(\mathbf{r})} w_{\mathbf{rs}} U(\mathbf{s}))^2 \quad (2-1)$$

式中 $N(\mathbf{r})$ 为 \mathbf{r} 的 8-领域； $w_{\mathbf{rs}}$ 是一个代价函数，这一类代价函数在图像分割算法中有广泛应用，它满足以下特性：

- 1) 对于 scribble 当中的点，保留其原有的颜色值。即有

$$\forall \mathbf{r} \in \Omega_c, w_{\mathbf{rs}} = \begin{cases} 1, & \mathbf{r} = \mathbf{s} \\ 0, & \mathbf{r} \neq \mathbf{s} \end{cases} \quad (2-2)$$

- 2) 若某个点的像素值接近中心点，则它将具有较大的权值；反之，权值较小。用数学公式来表达就是：若对于 \mathbf{s} 邻域内两个像素点 $\mathbf{r}_1, \mathbf{r}_2 \in N(\mathbf{s})$, $\mathbf{r}_1 \neq \mathbf{r}_2$ ，若 $\|Y(\mathbf{r}_1) - Y(\mathbf{s})\|^2 < \|Y(\mathbf{r}_2) - Y(\mathbf{s})\|^2$ ，则 $w_{\mathbf{s}\mathbf{r}_1} > w_{\mathbf{s}\mathbf{r}_2}$ 。

Levin 在文中给出了两个代价函数，见公式 2-3 和公式 2-4。

$$w_{\mathbf{rs}} \propto e^{-(Y(\mathbf{r}) - Y(\mathbf{s}))^2 / 2\sigma_r^2} \quad (2-3)$$

$$w_{\mathbf{rs}} \propto 1 + \frac{1}{\sigma_r^2} (Y(\mathbf{r}) - \mu_r)(Y(\mathbf{s}) - \mu_r) \quad (2-4)$$

其中 μ_r 和 σ_r 分别是 \mathbf{r} 附近各点像素值的平均值和方差。

Levin 的算法需要用户手动涂抹 scribble，即给出一系列点 \mathbf{r}_i 处的颜色信息： $u(\mathbf{r}_i) = u_i$, $v(\mathbf{r}_i) = v_i$ ，在这些约束条件下最小化式 2-1 中的 $J(U)$, $J(V)$ 亦同理。因为上面给出的代价函数 2-3 和 2-4 均为二次，而约束条件都是线性的，这个优化问题会产生一个大规模的，稀疏的，线性方程组，可以通过一些标准方法来求解方程组，分别得到 U, V 两个通道的值。最后，综合 YUV 通道，得到彩色化结果。

2.2 scribble 的自动生成算法

课题组原来的算法是建立在 Levin 算法的基础上。在 Levin 的算法中，首先要求用户输入一些 scribble，即在图像上涂抹一些区域，并使得这些区域的颜色和原图接近。然而，通过人工交互给出 scribble 的方式存在以下不足：工作繁琐枯燥，需要用户有丰富的彩色化经验提供细致的 scribble 才能得到令人满意的彩色化结果。为了克服这些不足，课题组前期提出了一种基于空间分布熵的自动 scribble 生成算法，自动生成的 scribble 定位在图片中信息密度较高的区域，大大减少了用户的工作量，同时优化了彩色化的效果。

2.2.1 空间分布熵理论

为了找出 scribble 合适位置，我们将融入空间信息的分布熵算法用来作为信息量的测度，这一想法是受到孙君顶等人提出的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] 的启发。

记 $(p_{xy})_{C \times R}$ 为 $C \times R$ 的一幅图像，其中 p_{xy} 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_q 为 S_q 的质心。在对图像区间的划分上，我们采用半径方向平均划分的方式，即取一个正整数 M ，对半径 M 等分，然后以质心 C_q 为圆心画 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_q 为圆心，以 d 为半径画圆 ($d = 2, 3, \dots$)，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。

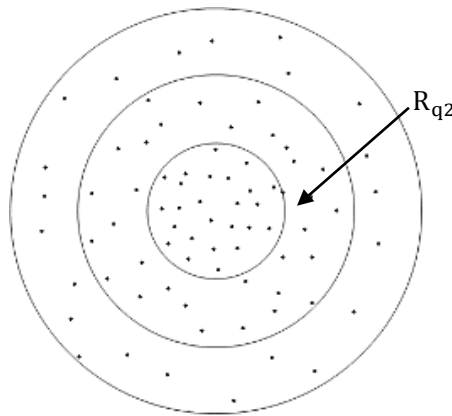


图 2-1 scope 模型

设圆 C_q 共被分割为 M 个圆环，以 $|R_{ij}|$ 表示 S_i 中颜色为 B_i 的像素落在第 j 个圆环中的总个数，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）。

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 B_i 的空间分布熵可表示为

$$E_i = -\sum_{j=1}^M p_{ij} \log_2(p_{ij}) \quad (2-5)$$

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,由公式 2-4 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。

2.2.2 基于图论的图像分割算法

为了避免一个 scribble 横跨不同的同态区域,在生成 scribble 之前,首先要应用图像分割算法对灰度图像进行预处理,将图像分割成若干个闭合区域。图像分割算法(Image Segmentation)也是数字图像处理中的一个重要课题,研究者们提出了许多基于图像相似性的分割方法,如阈值法,区域增长法,边缘检测法,模糊方法和聚类法等。

图论(Graph Theory)是数学和计算机科学的一个重要研究领域。它主要研究数学意义上的图(Graph)。图是描述物体间成对关系的数学模型,它是由节点(nodes)和边(edges)组成的,用节点表示事物,用边表示两个事物间具有这种关系。基于图论的图像分割算法有着独特的优势,它对数据类的性质的不敏感,同时聚类效果有良好的通用性。在近些年得到了广泛的关注和应用。本课题中采用图像处理中最常见的基于图论的方法——最小生成树算法,来完成图像分割。操作步骤简述如下。

1) 初始分割

将相似的像素置于同一区域,然后具体地对每一个待分割的区域找到统一的像素作为种子,将种子像素邻域内与该种子像素具有相似性质的像素加入种子像素所在的区域,再将这个新加入的像素作为种子像素,重复上述过程,直到没有满足条件的像素为止。

种子的选取是根据图像的饱和度来确定一组初始的种子,用图论中八邻域联通方案来搜索邻接像素,给定临界值 Th ,若邻域像素与种子的相似度小于这个 Th ,那么该像素所包括了该种子像素所在的区域,如果相似度大于 Th ,那么终止生长。

2) 最小生成树区域分割

在区域生长完成之后,图像中的所有的特征信息都已分布在不同的区域中,第二步是采用最小生成树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无向图 G 的一个节点,方便对像素点直接操作,节约了算法的运行时间。权函数的定义是两个节点之间的相似度,这里考虑到图像的不同性质,把性质相同的归为一个区域,设为同一类。

3) 最小生成树图像分割

无向图 $G = [X, E]$ 是具有 N 个节点的网络,本课题中采用 Prim 算法来求解最小生成树。

4) 目标函数的建立

确定图像的数据集之后,由最小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最大的类间类内的距离比,定义一个目标函数来评价不同的分类状态,如式 2-6。

$$H_t = \frac{\sum_{i=1}^m P_i (m_i - m)^2}{\sum_{i=1}^m \sum_{j=1}^t (X_i - m_i)^2 h_i} \quad (2-6)$$

实验表明,上述的方法得到的图像分割结果边缘清晰,速度快,实时处理能力强。通过这种分割算法可以把图像分割成若干个闭合区域,我们将在每个区域中搜索 scribble。

2.2.3 scribble 的生成

图 2-2 是 scribble 搜索算法的流程图。我们采用的检索思想是借助于空间分布熵来定位一些重要的支撑点，然后以这些点作为骨架，通过插值、腐蚀与膨胀等方法来描绘出连续的 scribble。算法的关键部分就在于 scribble 支撑点的定位。

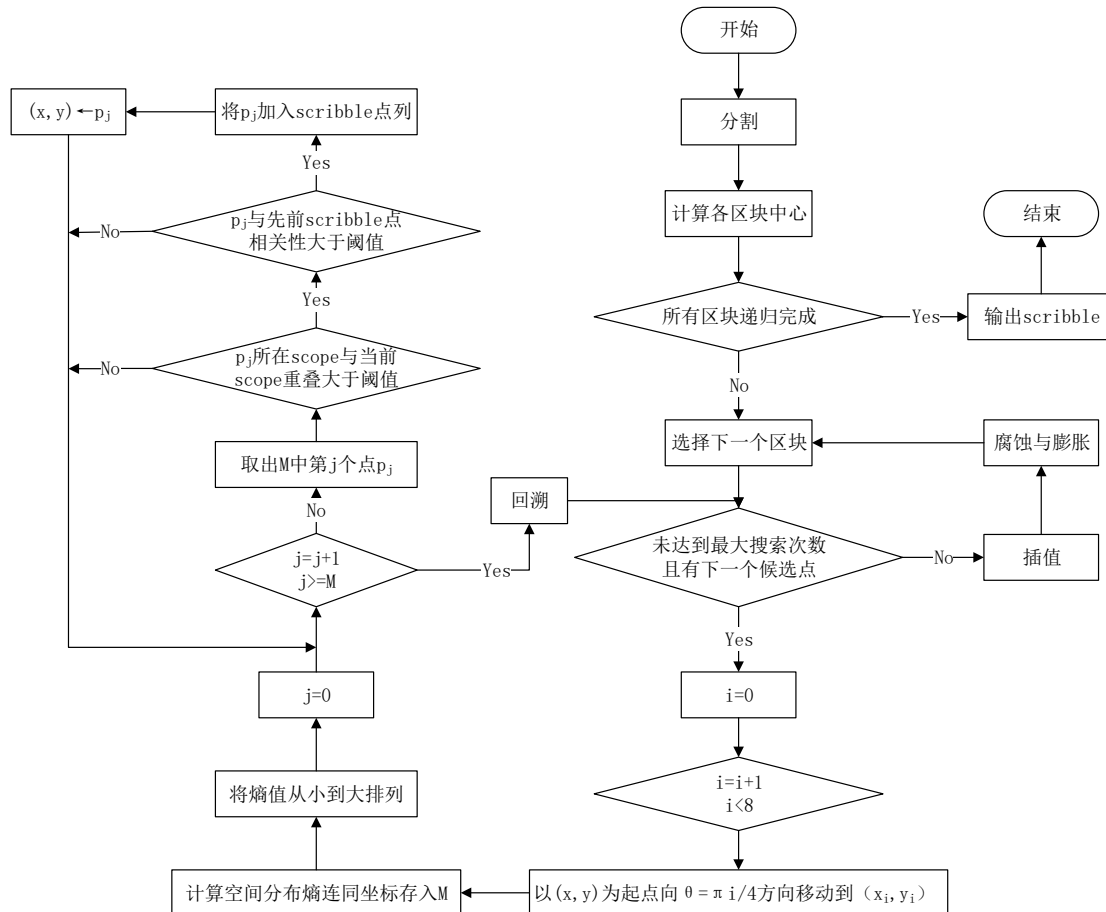


图 2-2 scribble 搜索算法流程图

我们的想法是：根据实验的结果，算法产生误判的地方常常是一些非边缘而灰度值变化又比较明显的区域。这样的一些区域，其伪边缘往往阻止颜色向其蔓延，然而这些区域又恰恰可能是同态区域。这样便极有可能导致上色的不充分。因此，能够产生较好上色结果的 scribble，应该是那些尽可能多的经过灰度值变化明显区域。用空间分布熵来表述，就是经过的区域空间分布熵尽可能小。

我们首先按照 2.2.2 中所述的分割方法对整幅图像进行分割，然后对分割所得的各个区块分别进行搜索。在这样的算法框架下，一个很自然的问题就是各个区块搜索起始点的选取。为了使得搜索得到的 scribble 尽可能的遍布各个区块，起始点应当尽量靠近各个区块的几何中心。

另外一个问题则是搜索方法问题。即：采用什么样的顺序进行搜索，和当前搜索的点，需要满足什么样的条件才能作为 scribble 的支撑点。对于任何一个点，对应有一个圆形的邻域，称作 scope。空间分布熵的计算，就在这个 scope 内进行。根据前文所述，为使 scribble

经过的范围有尽量小的空间分布熵，对给定的点 p ，我们以一定的半径（与点 p 到区域边界的距离正相关）向该点八个方向搜寻相邻的点，并采用熵值从小到大的顺序。对于当前点 p 的某个邻接点 p_i ，仅当其满足以下两个条件时，将其作为 scribble 的支撑点：

- 1) 同态性判定：计算 p 和 p_i 在各自 scope 的直方图，然后比较两个直方图得相关系数 r 。若 r 高于先前所得各支撑点的相关系数的平均值乘上某个系数，则接受 p_i 。
- 2) 最大重叠判定：若 p_i 的 scope 与先前所有点的 scope 重叠区域面积大于 p_i 的 scope 面积的某个百分比，则舍弃 p_i 。这样可以防止在某个区域内的 scope 支撑点过于密集。

2.3 最优化算法

2.3.1 利用四元 Gabor 相位分析图像结构

彩色图像中的颜色与结构信息有关，而灰度分布实际上带有一部分的结构信息，如何从灰度中提取结构信息作为彩色化的一个条件，就是我们需要考虑的问题。不幸的是，在光照变化不规则的图像中，灰度分布无法直接作为表示图像结构的有效工具。同时，因为复数 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。我们发现四元数 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 [15]，而复数 Gabor 相位只能分析一维特征 [16]。

我们引入四元 Gabor 相位，用来计算表征图像中两个相邻区域结构相似性的权重。首先，建立倍频程四元数 Gabors。

$$\mathbf{G}_{\sigma\alpha}^q(\mathbf{x}, \mathbf{u}) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} e^{-i2\pi u x'} e^{-j2\pi v y'} \quad (2-7)$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \quad (2-8)$$

式中 G 表示四元数 Gabor 卷积核，利用四元数的代数法则可以计算三个四元数 Gabor 相位。

$$\Phi_{\sigma\alpha}\{\varphi, \psi, \theta\} = \arg(\mathbf{I} * \mathbf{G}_{\sigma\alpha}^q) \quad (2-9)$$

式中 Φ 为四元数相位向量， \mathbf{I} 表示灰度图像。

现在可以定义一个估计相邻像素结构同态性的量。

$$H_{\sigma\alpha}(\mathbf{p}_a, \mathbf{p}_b) = \frac{1}{4} |\rho(\mathbf{p}_a) \varphi_{\sigma\alpha}^q(\mathbf{p}_a) - \rho(\mathbf{p}_b) \varphi_{\sigma\alpha}^q(\mathbf{p}_b)|_{2\pi} + \frac{1}{2} |\rho(\mathbf{p}_a) \theta_{\sigma\alpha}^q(\mathbf{p}_a) - \rho(\mathbf{p}_b) \theta_{\sigma\alpha}^q(\mathbf{p}_b)|_{\pi} + |\rho(\mathbf{p}_a) \psi_{\sigma\alpha}^q(\mathbf{p}_a) - \rho(\mathbf{p}_b) \psi_{\sigma\alpha}^q(\mathbf{p}_b)|_{\pi/2} \quad (2-10)$$

2.3.2 Canny 边缘检测

在本课题先前的算法中，我们采用 Canny 边缘检测提取边缘信息作为对四元数 Gabor 相位的补充。同时由于大尺度上的边缘更能表示图像的关键结构，全局边缘与局部边缘需要分别考虑。我们采用的 Canny 算法步骤简述如下。

- 1) 去噪声

第一步是对原始图像与高斯平滑模板做卷积，这样就得到了一个略微模糊的图像。由此可以减弱单个像素噪声对图像的影响。

2) 寻找图像中的亮度梯度

因为图像中边缘的方向不确定，在 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

$$L_x^2 L_{xx} + 2L_x L_y L_{xy} + L_y^2 L_{yy} = 0 \quad (2-11)$$

它在梯度方向的三阶方向导数满足符号条件

$$L_x^3 L_{xxx} + 3L_x^2 L_y L_{xxy} + 3L_x L_y^2 L_{xyy} + L_y^3 L_{yyy} < 0 \quad (2-12)$$

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 边缘检测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。

- 1) 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。
- 2) 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 边缘检测适用于不同的场合。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。原算法中 Canny 边缘检测过于耗时的问题也成为了我们引入相位一致性进行边缘检测作为替代的原因之一，在下一章会做详细讨论。

2.3.3 最优色彩蔓延

我们在构建优化模型时，依旧借鉴了 Levin [3] 的思路。首先是对代价函数的表示问题。我们采用式 2-3 所示的形式，其中的权值表示为：

$$W_{pq} = W_{pq}^{Strength} \times W_{pq}^{QGabor} \quad (2-13)$$

$$W_{pq}^{Strength} = e^{-\|E(p)-E(q)\|^2} \quad (2-14)$$

$$W_{pq}^{QGabor} = e^{-H(p,q)} \quad (2-15)$$

其中 $W_{pq}^{Strength}$, W_{pq}^{QGabor} 分别是用 Canny 边缘检测计算出的边缘权重和四元数 Gabor 变换产生的结构权重; $E(p)$ 表示点 p 的 strength; $H(p, q)$ 是各个尺度分量和角度分量的 Gabor 滤波所得结果的叠加。其算法流程如图 2-3 所示。

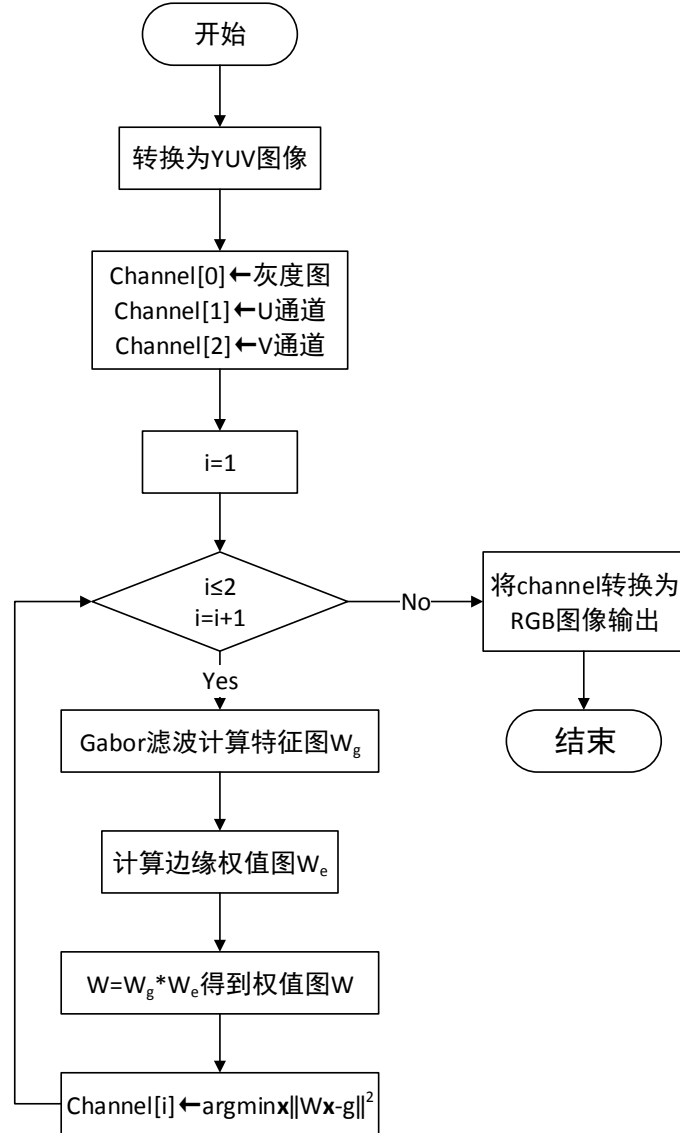


图 2-3 最优化算法流程图

最优化过程是通过求解稀疏系统来完成的。设原灰度图像 G 的尺寸为 $m \times n$; 将原图按行进行拼接, 则得到一个长为 mn 的向量 b 。在原图中坐标为 (i, j) 的点 (i, j 从 0 开始标, 即 $i = 0, 1, 2, \dots, m-1$; $j = 0, 1, 2, \dots, n-1$), 在 b 中将位于下标 $i \times n + j$ 处。现定义一个的矩阵 M , 令该矩阵的第 i 行存储下标为 i 的像素对原图中各个像素 $j (j = 0, 1, 2, \dots, n-1)$ 的权重; 即: 若设点 p, q 在向量 b 中的下标分别为 x, y , 则有 $W_{pq} = M(x, y)$ 。

不难看出, $M(x, y)$ 为一个稀疏矩阵; 且其任意一行最多有 9 个非零元素。回想我们需要求解的问题 (式 2-1), 利用这一结构即可将其改写成公式 2-16 的形式。其中 b 是由 U

或 V 通道逐行拼接得到的长度为 mn 的向量；计算得到 \mathbf{x} 之后，将其还原成 $m \times n$ 的矩阵，即得到 U 通道的彩色化结果。

$$\mathbf{x} = \arg \min_{\mathbf{x}} \|\mathbf{b} - M\mathbf{x}\|^2 \quad (2-16)$$

这一稀疏矩阵的构造方法，在 Levin 的代码中也有所体现；然而对于式 2-16 的求解，在 Levin 的代码中是直接运用 MATLAB 来求逆。而我们的目标是将算法用 C++ 实现，所以必须自己寻找这一稀疏线性系统的 solver。

我们发现，对于每个像素，只有其 8 -邻域内的点权值有可能不为零；也就是说，对于下标为 i 的像素，仅有下标为 $i-n-1, i-n, i-n+1, i-1, i+1$ ，以及 $i+n-1, i+n, i+n+1$ 这些点（如果它们存在的话）可能对其权值构成贡献，或者说， M 的第 i 行只有下标为上述数字的列可能非零。据此，我们可以断言 M 是一个带状矩阵，即在对角线的某条带状区域外所有元素均为零的一类稀疏矩阵。

需要说明的是， M 并非对称矩阵。这是因为，在需要填充彩色的 U, V 通道中，有少部分的像素位于 scribble 上。依据公式 2-4，为保证 scribble 上的点保留其原有色彩，若下标为 i 的点 \mathbf{p} 位于 scribble 上，则 M 的第 i 行除了 (i, i) 外均为零。而另一方面，若下标为 j 的某个点 $\mathbf{q} \in N(\mathbf{p})$ ，且 \mathbf{q} 不在 scribble 上，那么很可能有 $W_{\mathbf{qp}} \neq 0$ 。这样便有 $M(j, i) \neq 0$ 。据前所述 $M(i, j) = 0$ ，从而 $M(i, j) \neq M(j, i)$ 。

因此，我们需要求解的是一个非对称而带宽受限的稀疏矩阵。对于这一问题，已经有大量的研究工作和成熟的快速计算方法。这使得我们算法需要的优化能够得到较为高效的解决，也保证了我们的算法整体上的效率。

2.4 本章小结

在本章中，我们详细总结了课题组前期的工作成果，从 Levin 的算法出发，首先讲述了 scribble 自动生成算法的具体过程，并介绍了所用到的空间分布熵理论和基于图论的图像分割算法。scribble 自动生成是课题组对 Levin 算法的一个重大改进。

然后我们讨论了最优化算法，重点介绍了权值图的计算方式，即 Garbor 变换产生的结构权重与 Canny 边缘检测得到的边缘权重相乘。最终，通过求解稀疏系统的方式得到最后化的结果。

第三章 基于相位一致性的边缘检测

3.1 相位一致性简介

对于描述图像边缘显著性的参数，我们希望找到一个无量纲的量，最好是不随着图像光照和放大倍率变化的量。基于梯度的边缘检测方法，如 Sobel 算子，Canny 算子等，对光照变化都比较敏感。图像中的光强梯度的单位是 lumens/radian ，它依赖于光照、模糊、放大倍率等因素。例如，图像的尺寸加倍之后，虽然各像素的亮度不变，但所有的亮度梯度都减半了。任何基于梯度的边缘检测过程都需要一个合理修正过的阈值。但是在大部分情况下，我们不能提前知道图像的总光照水平和放大倍率，与边缘显著性相对应的图像梯度值常常是通过实际经验决定。

Morrone 和 Owens 在 [17] 中提出了一种新的特征感知模型：局部能量模型 (Local Energy Model)，这一模型不是基于局部亮度梯度，而是假定图像特征总是集中在傅立叶分量的相位极大值点，许多不同的图像特征就会叠加产生一些较高的相位一致性点，这些特征包括了台阶边缘，线形边缘和屋脊形边缘。该模型成功解释了一系列人类对图像特征感知的心理效应。这样就有可能建立一个无量纲的量来衡量一幅图像中任意点的相位一致性，相位一致性的值从 0 到 1，0 表示没有显著特征，而 1 表示非常显著的特征。由此可以指定一个阈值来在图像被看见之前就筛选出图像特征。Kovesi 在 [18] 中提出了一种使用 Gabor 小波来计算相位一致性的新方法，我们在本课题中即采用这种算法来实现相位一致性提取边缘。

3.2 局部能量和相位一致性

特征检测的局部能量模型假定人类感知的图像特征总是集中在傅立叶分量的相位极大值点处。例如，将一个方波信号展开成傅立叶级数，所有在阶跃点处的傅立叶分量都是完全同相的正弦波，角度为 0 度或 180 度，这取决于上是跳沿还是下降沿，而在方波信号的其他点处相位一致性都很低。相似地，在三角波的峰值点也能观察到相位一致性。事实上，应用相位一致性来标记图像特征让人感兴趣的一点就是，我们并不需要知道波形的具体形状，只需要找到图像中在傅立叶域中值很高的点。

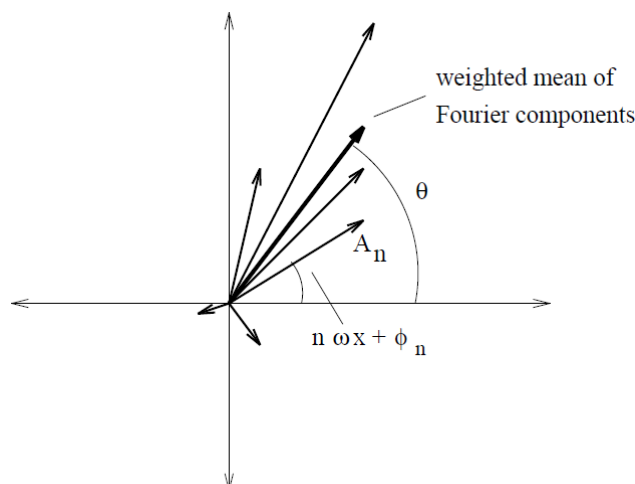


图 3-1 信号中某点的傅立叶级数分量的极坐标图

我们首先考虑一维的情况。相位一致性函数是从信号的傅立叶级数展开（式 3-1）衍生而来

$$F(x) = \sum_n A_n \cos(n\omega x + \phi_n) \quad (3-1)$$

其中 ω 是一个常数（一般是 2π ），级数中的正弦项是通过相位偏移量 ϕ_n 来表示。

我们定义相位一致性函数如式 3-2 所示。

$$PC(x) = \max_{\theta \in [0, 2\pi]} \frac{\sum_n A_n \cos(n\omega x + \phi_n - \theta)}{\sum_n A_n} \quad (3-2)$$

使右边分式取到最大的 θ 的值是点 x 处所有傅立叶级数项的相位角加权平均。对一个频率分量的实际相位角与 θ 的差值取余弦，得到的量约等于一减这个差值平方的一半（Taylor 展开，对较小的 x 有 $\cos x \approx 1 - x^2/2$ ），由此找到的相位一致性最大值约等于相位角的加权标准偏差的最小值，如图 3-1 所示。

如上所述的相位一致性是一个很难计算的量，作为替代，Venkatesh 和 Owens [19]提出相位一致性的最大值点也可以通过寻找局部能量函数的峰值来计算。一维局部能量函数的定义见式 3-3。

$$E(x) = \sqrt{F^2(x) + H^2(x)} \quad (3-3)$$

式中 $F(x)$ 是信号 $I(x)$ 去除直流分量后的函数，而 $H(x)$ 是 $F(x)$ 的 Hilbert 变换。

Venkatesh 和 Owens 指出局部能量等于相位一致性与傅立叶幅值之和的乘积。

$$E(x) = PC(x) \sum_n A_n \quad (3-4)$$

这样，局部能量直接与相位一致性成正比，所以，局部能量的峰值对应着相位一致性的峰值。

除了通过对原图像的亮度分布进行 Hilbert 变换来计算局部能量之外，还可以通过将信号与一对正交滤波器做卷积计算得到局部能量。第一个滤波器是为了除去信号中的直流分量，将这个结果保存，另外再将原信号与第二个滤波器做卷积，第二个滤波器与第一个滤波器正交（即为第一个滤波器的 Hilbert 变换），这样，就得到了两个信号，每个都是原信号的带通结果，两者之间的相位差是 90 度，这两个卷积结果然后分别平方再求和。

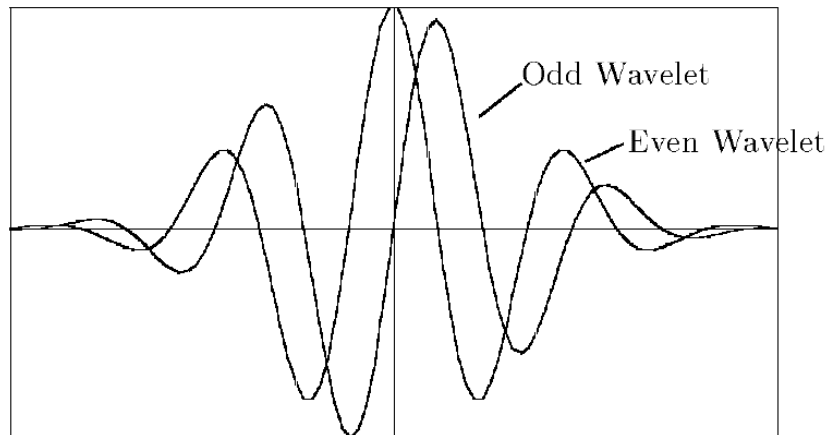


图 3-2 Gabor 小波

尽管使用局部能量函数来找到相位一致性的峰值在计算上比较方便，但是它不能提供一个表示特征显著性的无量纲的量，因为在式 3-4 中可以容易看出，局部能量带有傅立叶级数幅度和单位，即 lumens。然后，如果我们将局部能量除以傅立叶级数的幅度和，就可以得到无量纲的相位一致性函数，它的值分布在 0 和 1 之间，并且可以不随着光照和尺度放大而

变化，稳定地衡量图像的特征显著性。

原始的计算相位一致性的方法需要经历去除直流分量，做 Hilbert 变换，对 Hilbert 变换和信号交流分量求平方和，最后除以傅立叶级数幅度之和来归一化的过程。但是，上述这种途径有几个问题：首先，傅立叶变换并不适合衡量局部频率信息，需要通过加窗来控制相位一致性的尺度；其次，相位一致性的定义没有考虑到在一点处一致的频率伸展，比如，一个信号只包含一个正弦分量，这样将会在所有地方都与它自己有完全的一致性，即相位一致性的值为 1，于是下面引入小波来计算相位一致性。

3.3 利用小波计算相位一致性

我们使用小波来保证相位信息的对称性和反对称性，在本文中，采用基于复数值 Gabor 函数的小波：一个正弦波和一个余弦波，均经过高斯调制，如图 3-2 所示。用两个正交的滤波器，使得其中一个可以在给定的空间位置对一个特定的频率计算信号的幅值和相位。

对信号的分析是通过将信号分别与每个小波做卷积来完成的。 I 为信号，和分别表示在比例 n 下的偶波和奇波。在给定的小波尺度下，小波变换的幅度和相位分别由式 3-5 和 3-6 给出。

$$A_n(x) = \sqrt{(I(x) * M_n^e)^2 + (I(x) * M_n^o)^2} \quad (3-5)$$

$$\Phi_n(x) = \text{atan2}(I(x) * M_n^e, I(x) * M_n^o) \quad (3-6)$$

为了计算相位一致性，我们要计算如下的量：

- 1) 去除直流并保留尽可能多的其他频率分量的信号，记为 $F(x)$ ，需要注意的是，相位一致性只在信号频率范围比较大时才有意义，因而我们希望在信号中保留较大的频率范围；
- 2) $F(x)$ 的 Hilbert 变换，记为 $H(x)$ ；
- 3) $F(x)$ 的频率分量的幅度和，记为 $\sum A_n(x)$ 。

将原信号与偶波卷积之后累加得到信号的带通，作为 $F(x)$ ，而 $H(x)$ 可以通过累加信号与奇波的卷积而得，与 $F(x)$ 有 90 度的相位差。

$$F(x) = \sum_n I(x) * M_n^e \quad (3-7)$$

$$H(x) = \sum_n I(x) * M_n^o \quad (3-8)$$

$F(x)$ 的频率分量幅度和由式 3-9 给出

$$\sum_n A_n(x) = \sum_n \sqrt{(I(x) * M_n^e)^2 + (I(x) * M_n^o)^2} \quad (3-9)$$

由此得到上面三个量之后就可以计算相位一致性。修正后的表达式为 3-10，此处我们通过分母加上一个常数 ε 来防止 $\sum_n A_n(x)$ 非常小的情况。

$$PC(x) = \frac{E(x)}{\varepsilon + \sum_n A_n(x)} \quad (3-10)$$

式中 $E(x) = \sqrt{F^2(x) + H^2(x)}$ 。 ε 的合理值是由卷积和其他运算的精确度决定，而不是信号本身决定的。比较大的 ε 值可以抑制噪声的影响。

3.4 噪声的影响

噪声的对相位一致性的影响可见图 3-3。Kovesi 还提出了更好的补偿噪声的方法。为了确定图像中的噪声水平，需要用到如下两个观察结果：

- 1) 噪声在图像各处都存在，并且一般是一个常数；
- 2) 图像的特征产生在图像中的孤立区域。

由此可以推断出,在我们的小波滤波器组中的最小尺度滤波器组的响应几乎完全是由噪声造成的,只有特征点处的响应与背景噪声的响应不同,并且因为小波的空间延伸很小,响应特征的区域也很小。因此,从整幅图中的最小尺度滤波器对得到的响应幅度分布将会是不对称的。表示对噪声的响应的分布谱峰值就会向左边倾斜并且有个长“尾巴”延伸到分布谱的上端,作为对特征的响应。对分布谱直接求均值并不能很好地估计平均噪声响应的,因为对特征的响应会造成结果的偏差。解决的方法是用响应幅度取 \log 后的均值的指数式来估计平均噪声响应,对响应幅度取对数将会有效地减少分布谱上端“尾巴”的影响,由此,我们得到信号通过最小尺度滤波器组的平均噪声响应的估计值如式 3-11 所示。

$$\bar{A}'_0 = \exp(\overline{\log A_0(x)}) \quad (3-11)$$

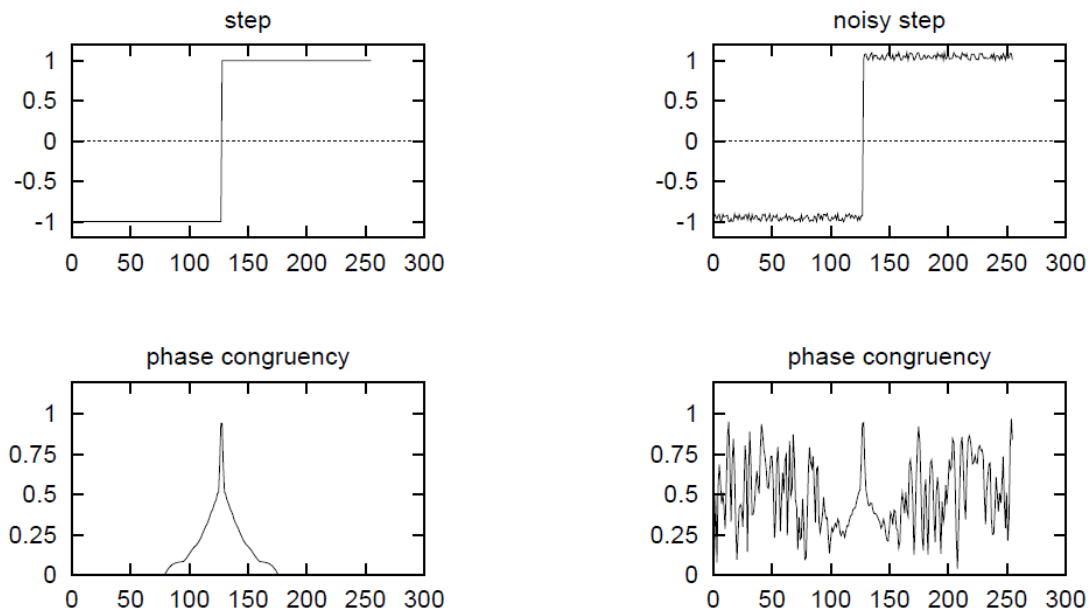


图 3-3 一个阶跃函数在有噪声和无噪声情况下的相位一致性

如果噪声分布是均匀的,典型的噪声最大响应幅值将是估计的平均噪声响应的两倍,可能的最小响应水平是 0。如果噪声分布是高斯分布,典型的最大响应幅值大约是估计的平均噪声响应的三倍,因此,我们仍然需要估计图像中的噪声在滤波器组其他尺度中的小波响应,然后推测出相位一致性计算中噪声的影响。

我们假设噪声的频率响应是平坦的,我们可以估计其他小波响应相对于最小尺度小波对的响应的值。根据 Parseval 定理,空间域中一个滤波器对的响应幅度的平方与信号的功率谱相关。如果我们假设噪声的功率是不随频率变化的常数,那么滤波器的响应幅度平方与其带宽成正比。也就是说,滤波器噪声响应的幅度值与它带宽的平方根成正比,如图 3-4 所示。

如果整幅图像的最小尺度小波对平均噪声响应用 \bar{A}'_0 ,那么所有小波尺度的估计噪声响应之和可以由滤波器带宽平方根之和乘上 \bar{A}'_0 之后得到。因此,总的噪声影响可以用式 3-12 来估计

$$T = k \bar{A}'_0 \sum_{n=0}^{N-1} \frac{1}{\sqrt{m^n}} \quad (3-12)$$

式中 N 是用到的小波尺度数量, m 是连续滤波器之间的比例因数, k 是用来估计对模态幅值

的最大噪声响应幅值的比例因数，一般 k 取 2.5 左右。

所有小波尺度的噪声响应估计之和 T 给我们提供了噪声对小波响应幅度之和 $\sum_n A_n(x)$ 的影响的上限，它是一个上限是因为一般的噪声响应不会都同相位，一些会相互抵消，又因为 $\sum_n A_n(x)$ 总是大于或等于局部能量 $E(x)$ ，所以 T 也是 $E(x)$ 的上限。如果我们从局部能量中减去估计噪声影响，再除以小波响应幅度和来进行归一化，我们就可以去除虚假的噪声响应，由此，我们将相位一致性的修正为式 3-13。

$$PC(x) = \frac{(E(x) - T)^+}{\varepsilon + \sum_n A_n(x)} \quad (3-13)$$

此处 Q^+ 表示这个函数不允许为负。

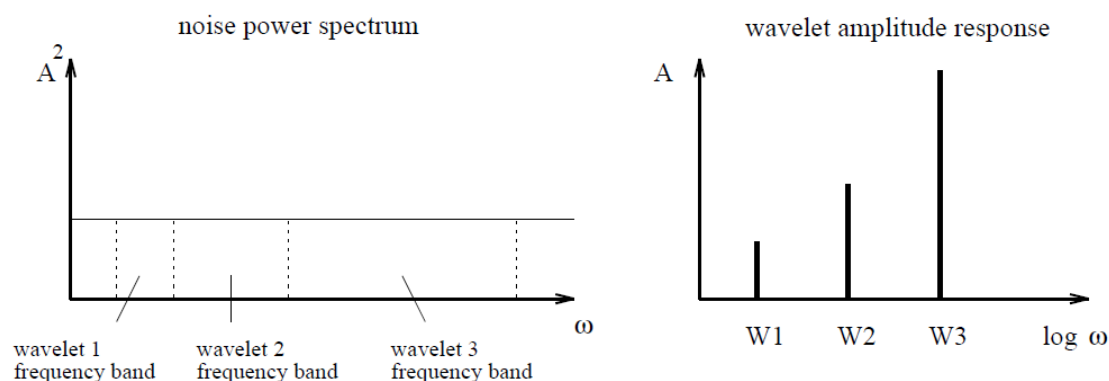


图 3-4 一维小波噪声响应与带宽平方根成正比

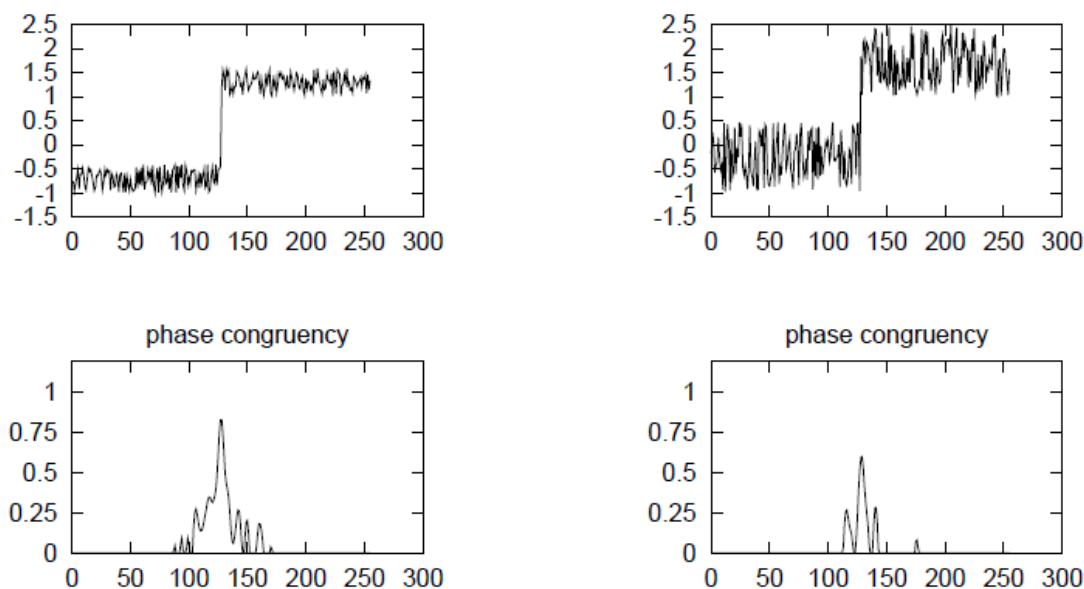


图 3-5 两个阶跃函数噪声补偿后的相位一致性，SNR 分别为 13.3 和 5.3。

图像中合理特征的相位一致性会因噪声的局部能量相对于特征的大小而减弱，上述的方法被实验证明十分有效，图 3-5 就展示了 $k = 2$ 时两个不同噪声水平的阶跃函数对应的相位一致性，左边的 SNR 为 13.3，右边的 SNR 为 5.3。

3.5 扩展到二维

上面的所有讨论都是限定在一维信号，相位一致性的计算需要对原信号进行 90 度的相移，这是通过奇对称的滤波器来实现的，为了分析二维信号，我们要讲先前的一维信号分析方法应用在不同的方向上并且将结果结合起来。我们需要讨论四个问题：二维滤波器的形状，需要分析的方向数量，每个方向上的结果最后组合的方式，二维噪声补偿。

3.5.1 二维滤波器设计

前文所述的一维滤波器可以通过直接应用一些垂直于它的方向的扩展函数来扩展到二维。最明显的可使用的扩展函数就是高斯函数。

考虑一个不与滤波器方向对齐的被台阶边缘特征卷积的二维小波，如图 3-6 所示。如果滤波器是分离的，卷积就可以通过在垂直方向上与扩展函数做一维卷积来完成。因为我们只对相位信息感兴趣，重要的就是保证与扩展函数的卷积没有损坏图像中的相位信息。

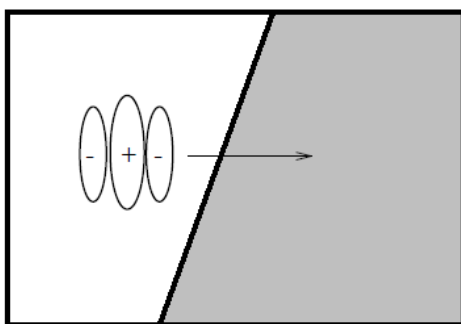


图 3-6 二维小波与角度边缘的卷积

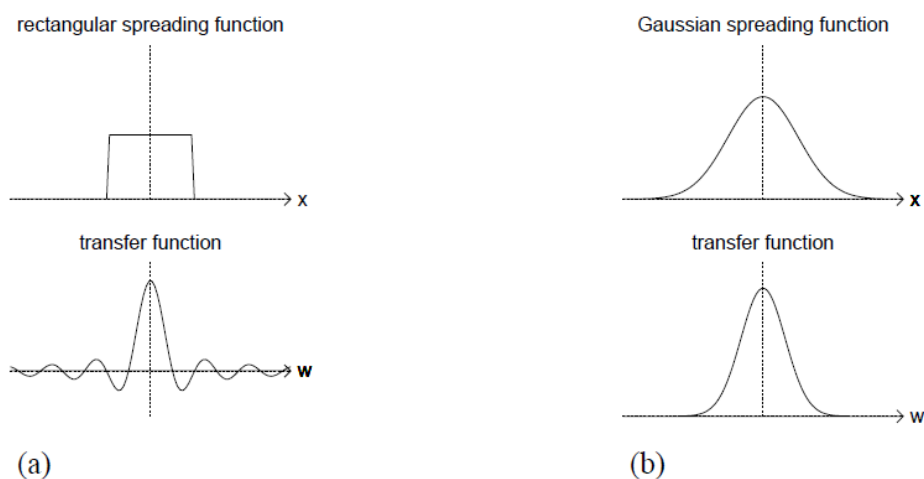


图 3-7 (a)矩形扩展函数及其传递函数；(b)高斯扩展函数及其传递函数

在此例中，图像垂直与高斯扩展函数卷积的结果将会模糊边缘，以至于在随后与小波函数的卷积中我们遇到了一个高斯平滑台阶。从频域来看，任何被高斯函数平滑的函数的幅度都会被调制，而相位却不受影响，因为高斯函数的传递函数也是高斯函数，另一方面，如果我们使用矩形的扩展函数，信号中的一些相位角就会被反转，因为传递函数有负值。

3.5.2 滤波器方向

为了检测所有方向上的特征，我们的滤波器组必须设计成在频率平面上可以均匀展开。在频率平面上，根据滤波器的空间对称性，滤波器呈现出围绕原点的二维对称或反对称高斯分布，二维小波的长宽比决定了它们的方向选择性，这个比值随着为了达到均匀覆盖二维频谱所使用的滤波器数量而改变，我们发现长宽比为 1，滤波器方向间距 30 度比较合适，最终的结果可见图 3-8。Simoncelli 等人在 [20]中介绍了一种对称滤波器设计技术，可以达到频率分布的均匀分布，这种方法在这里适用。

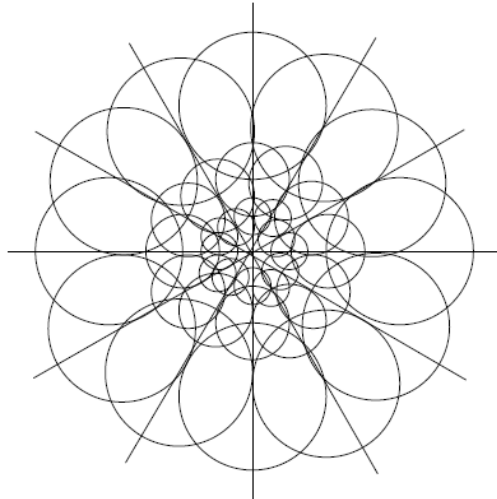


图 3-8 不同尺度方向滤波器在频率平面展开

3.5.3 二维噪声补偿

对二维信号补偿噪声，不仅要考虑滤波器的带宽，也要考虑它们的空间宽度。滤波器的空间宽度影响了它的噪声响应。在滤波器丛的设计中，为了保持方向敏感度恒定，滤波器的空间宽度随着中心频率改变。从频率平面看，滤波器的角度宽度与它们的带宽成正比。因此，空间域响应幅度的平方正比于带宽平方与噪声水平的乘积，也就是说，滤波器对噪声响应的相对幅度与它们的带宽直接成正比，由此可得噪声补偿项 T 就变成了式 3-14。

$$T = k\bar{A}_0' \sum_{n=0}^{N-1} \frac{1}{m^n} = k\bar{A}_0' \frac{1 - \left(\frac{1}{m}\right)^N}{1 - \frac{1}{m}} \quad (3-14)$$

其中 $\bar{A}_0' = \exp(\overline{\log A_0(x, y)})$ 是最小尺度二维滤波器对平均噪声响应的估计值。

3.5.4 结合不同方向上的结果

这里的问题重点是保证所有可能方向上的特征都被同等对待，并且所有可能特征的二维连接，比如角点和 T 形区域，都被一致地处理

对体现相位一致性的能量进行归一化处理必须在所有方向上能量加和之后。我们希望最后的结果是一个加权的归一化值，加权求和的各项来自各个方向，权重与各自的能量成正比，如果每个方向上的能量在加和之前就被归一化，那么每个方向在最后结果中的贡献就与它们的能量无关，这显然不是我们希望的情况。

我们采用的方式是在图像的各处计算每个方向上的能量 $E(x)$ ，然后补偿噪声的影响，再对所有方向加和，这个能量和通过除以该处所有方向所有尺度的小波响应幅度和，这样就得到了式 3-15 来计算二维相位一致性。

$$PC(x) = \frac{\sum_o (E_o(x) - T_o)^+}{\varepsilon + \sum_o \sum_n A_{no}(x)} \quad (3-15)$$

式中 o 表示方向索引。

在直接的边缘分割中，只有与边缘垂直的方向上的滤波器对结果有显著影响，其它方向上的响应将被忽略，并且对较早计算的相位一致性，上面的等式会降为一维表达式。对于复杂的特征类型，如角点和 T 形区域，一些方向会有很高的能量，但是，单个小波响应幅度之和也会增加，以此来保证仍能归一。

注意到上式 3-15 中，噪声补偿是在每个方向上独立运算的，实验证明这样做有更好的效果。从最小尺度小波对的平均响应中推测出的噪声成分可能随着方向变化而显著变化，这是因为沿着扫描线上的噪声关联经常被观察到，图像数字化中可能出现的其它过程也会导致这种情况。

3.6 频率扩展的重要性

很显然，一个点的相位一致性只有当它的频率分布在较宽的范围时才比较高，一个极端的例子就是单纯的正弦波信号，它的频率扩展很窄，其 Hilbert 变换是一个余弦函数，而它们的平方和始终是 1，因此相位一致性也始终是 1，这样的情况更普遍例子就是在经历过高斯平滑的特征处，对于这些做过平滑处理的函数，相位一致性可能在比较宽的信号范围内都很高。高斯平滑减少了信号频率成分中的高频成分，相应地，也减少了频率扩展。

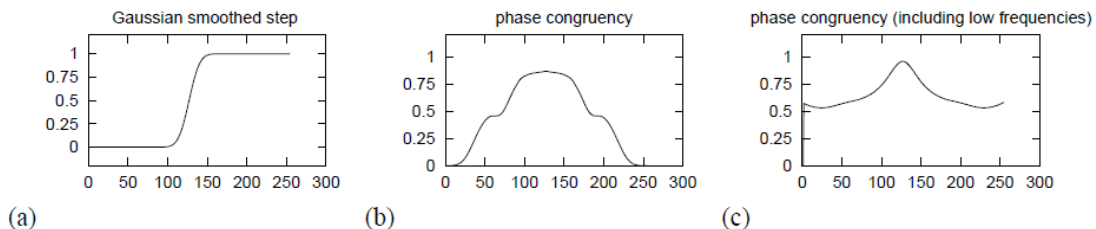


图 3-9 高斯平滑后的阶跃函数的相位一致性 (a) 平滑后的阶跃函数； (b) 考虑 65 个单位的波长后的相位一致性； (c) 考虑 150 个单位的波长后的相位一致性；

从图 3-9 中我们看出要得到收窄的相位一致性，在计算平滑函数的相位一致性的时候需要包含低频成分，这些低频成分收到平滑处理的影响最小。

我们关心两种情况：一个是在有非常窄的频率范围的地方，能量值和 $\sum_n A_n$ 相等，即相位一致性为 1；另一个是只有很窄的频率范围被考虑的情况，如 3-9 中(b)所示，这种情况下，能量和 $\sum_n A_n$ 都非常小，使得相位一致性难以计算，我们通过将相位一致性的表达式的分母中加入 ε 来处理这种情况。

因此，相位一致性作为衡量特征显著性的手段，应该把存在的频率扩展作为权重考虑进去。一个方波信号的功率谱是 $1/\omega^2$ 的形式，我们分析信号采用的每个小波都从功率谱中的上升频带获得功率，在这种形状的频谱中，每个小波获得的功率值都与它的带宽成反比。与我们分析滤波器对噪声响应幅度的方法类似，我们可以推断出滤波器对方波波形响应幅度的

期望分布,但是,不同于噪声的是,对台阶特征的响应在空间上比较集中,响应的空间宽度由滤波器的空间宽度表征,并与它的带宽成反比。所以,对每个滤波器来说,当它获得的功率随着带宽下降时,表示这些功率的响应幅度平方的空间宽度也随着带宽下降,于是,为了保持每个滤波器获得的功率,响应幅度平方的值要保持恒定,与滤波器的中心频率无关。这样,对台阶特征响应幅度的期望分布始终为 1,如图 3-10 所示,Field [21]指出在许多例子中,自然景观图像的频谱分布在下降区域与频率成反比,因为这个原因,他推荐使用几何尺度的滤波器组。在这种情况下,所有尺度上的滤波器的响应都有相同的量级,由此可以最大化利用滤波器输出所做的计算的精确度。

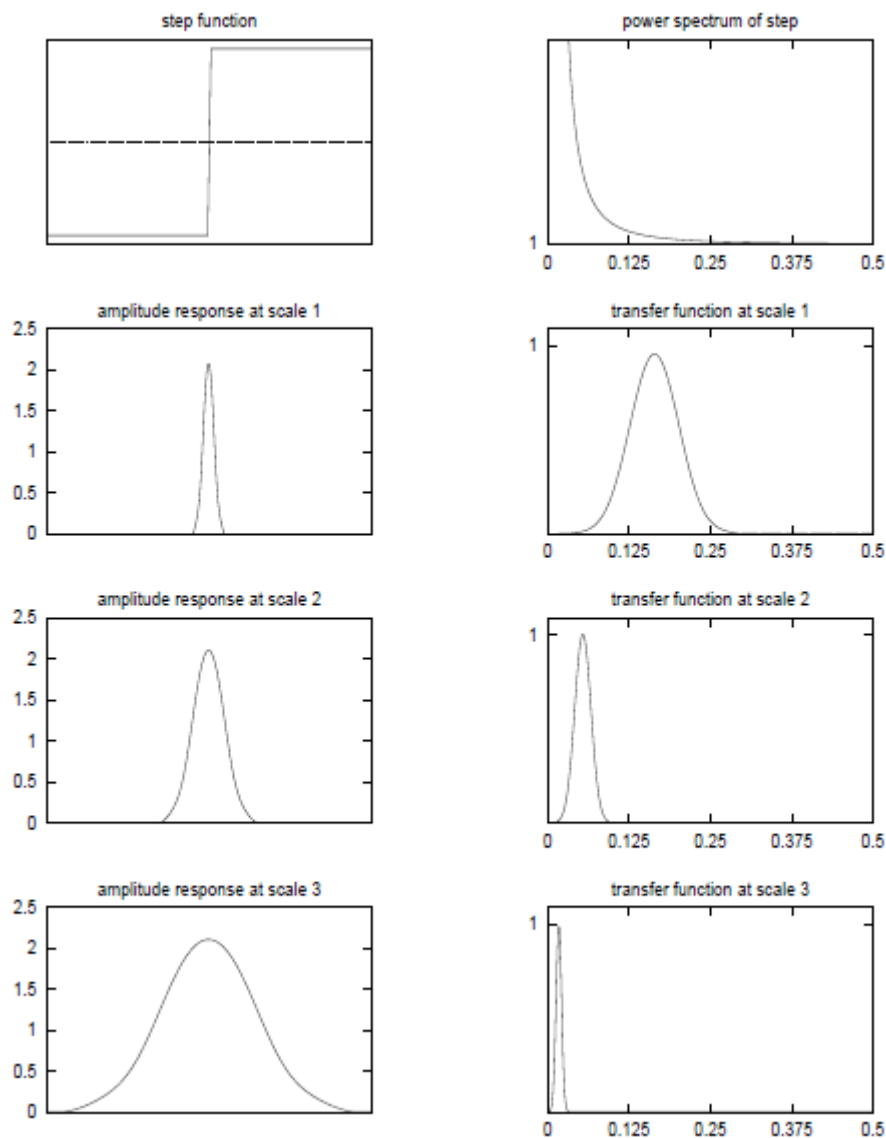


图 3-10 对阶跃特性,不同尺度的小波对的响应都有相同的幅值

需要注意的是,相对于未平滑处理过的特征,在有高斯扩展函数的二维滤波器遇到某个非正交角度处的特征时产生的特征高斯平滑并不能显著改变频率分布。其原因是高斯光滑的度数直接与每个滤波器的尺度成正比。考虑在频域的卷积过程:在某个给定中心频率的滤波

器与某个延展度的高斯调制后的特征谱做卷积，而两倍于前者频率的滤波器将会和两倍延展度的高斯调制后的特征谱做卷积。从每个滤波器的带宽来看，特征谱的相对调制是恒定的。因此，滤波器响应的总体分布也是几乎不变的。

我们需要考虑的另一类重要的特征类型是脉冲函数（对应着线性特征）和屋脊形边缘。脉冲函数的功率谱是恒定的。类似于对方波的分析，对于一个脉冲特征来说，小波滤波器的响应幅度与带宽成正比，也和中心频率成正比，这样的响应分布将偏向高频端，另一方面，对一个三角波来说，所有的特征都是屋脊形边缘，功率谱在 $1/\omega^4$ 处下降，对频率取对数，得到的响应分布会偏向低频端。

因此，我们在这里遇到的困难是没有任何理想的滤波器响应分布，考虑到脉冲函数和屋脊形边缘的预期分布是极其不同的，我们甚至不能说理想的分布是一致相位角的函数，我们只能说滤波器的响应分布不应该太窄，而因为图像中的阶跃不连续点非常普遍，均匀分布就变得十分重要。

综上所述，我们可以构建一个降低相位一致性在滤波器响应扩展很窄的位置的权重的权值函数。通过求响应幅度之和，并除以最高单个响应来衡量滤波器的响应扩展宽度，然后再除以尺度数量进行归一化，我们得到在 0 到 1 之间变化的一个分式来衡量扩展宽度，如式 3-16 所示。

$$s(x) = \frac{1}{N} \left(\frac{\sum_n A_n(x)}{\varepsilon + A_{\max}(x)} \right) \quad (3-16)$$

式中 N 是所有尺度的数量， $A_{\max}(x)$ 在 x 处有最大响应的滤波器对的幅度， ε 是用来避免分母为 0 的情况，并且在 $\sum_n A_n(x)$ 和 $A_{\max}(x)$ 都很小的情况下使结果不至于过大。

在滤波器响应扩展值中应用一个 s 形函数，可以构造出相位一致性的权值函数，如式 3-17 所示。

$$W(x) = \frac{1}{1 + e^{g(c-s(x))}} \quad (3-17)$$

此处 c 是滤波器响应扩展的下截止值， g 是控制截止锐利程度的增益因数，典型的 c 和 g 的值分别是 0.4 和 10。在这里选择 s 形函数仅仅是因为它形式简单，容易处理。

在二维图像中，通过每个点处的频率扩展对相位一致性加权需要在每个方向上独立完成。因此，权值系数在所有方向加和归一前就要在每个方向上分别乘以能量，我们修正式 3-15 得到式 3-18。

$$PC(x) = \frac{\sum_o (W_o(x)(E_o(x) - T_o)^+)}{\varepsilon + \sum_o \sum_n A_{no}(x)} \quad (3-18)$$

3.7 通过高通滤波来确定尺度

在不同尺度上分析图像的传统方法是考虑图像经过多种低通或带通之后的结果。只有低频信号的图像被认为有大尺度特征。该方法是受到大脑视觉皮质的接受区与带通滤波器的性质相似的启发，虽然这种想法是非常自然的，但是，假设大脑是利用图片通过带通滤波后的结果来进行多尺度分析是有问题的。就好像问一个人在通过低通滤波的图片中看到了什么特征，当然他会看到大尺度特征，因为图像经过低通滤波后只剩下这些特征了。

使用低通或带通滤波来进行多尺度分析的主要问题是图像中存在的特征数量和位置随着所用的尺度而变化。特征定位依赖于分析所用的尺度是我们不希望看到的。这对需要平滑处理来抑制噪声的基于差分的特征检测方法是一个问题。这也会给使用由粗到精的策略进行

特征匹配的过程带来困难。特征位置随着尺度漂移，在大尺度上的定位变差，随尺度的特征数增长也会减少。只有当特征的相对强度随着尺度变化而特征的数量和位置始终保持稳定时，由粗到精的策略才能有效。

使用相位一致性来衡量特征显著性让我们对特征尺度有了另一种理解，信号中某点的相位一致性取决于局部频率成分是怎样构成这些特征的。依据窗的大小，距离感兴趣点一定距离的特征也可能贡献局部频率成分，特征并没有被孤立地考虑，而是与其周围的特征相匹配。

因此，就相位一致性来说，导致自然尺度参数变化的量是我们进行局部频率分析的窗的大小。由于我们是用小波来计算相位一致性的，分析的尺度由小波组中最大的滤波器的空间延伸决定。采用这种方法，我们使用高通滤波来确定分析尺度，去除波长大于窗尺寸的低频成分，留下完整的高频成分。

如果我们使用一个很小的窗，每个特征都会被独立地处理，我们只需要将每个特征与临近的少量其他特征相比较，使得每个特征在局部区域显得更为重要。在窗和图片同样大小的大尺度情况下，每个特征都被认为与所有其他特征相联系，这样会得到每个特征的全局显著性。一些在小窗情况下相位一致性很高的特征，如果使用大窗，相位一致性可能会很低。一个在从小到大的广泛空间尺度上都有很高相位一致性的特征显然比在很少的空间尺度上一致的特征更加重要。

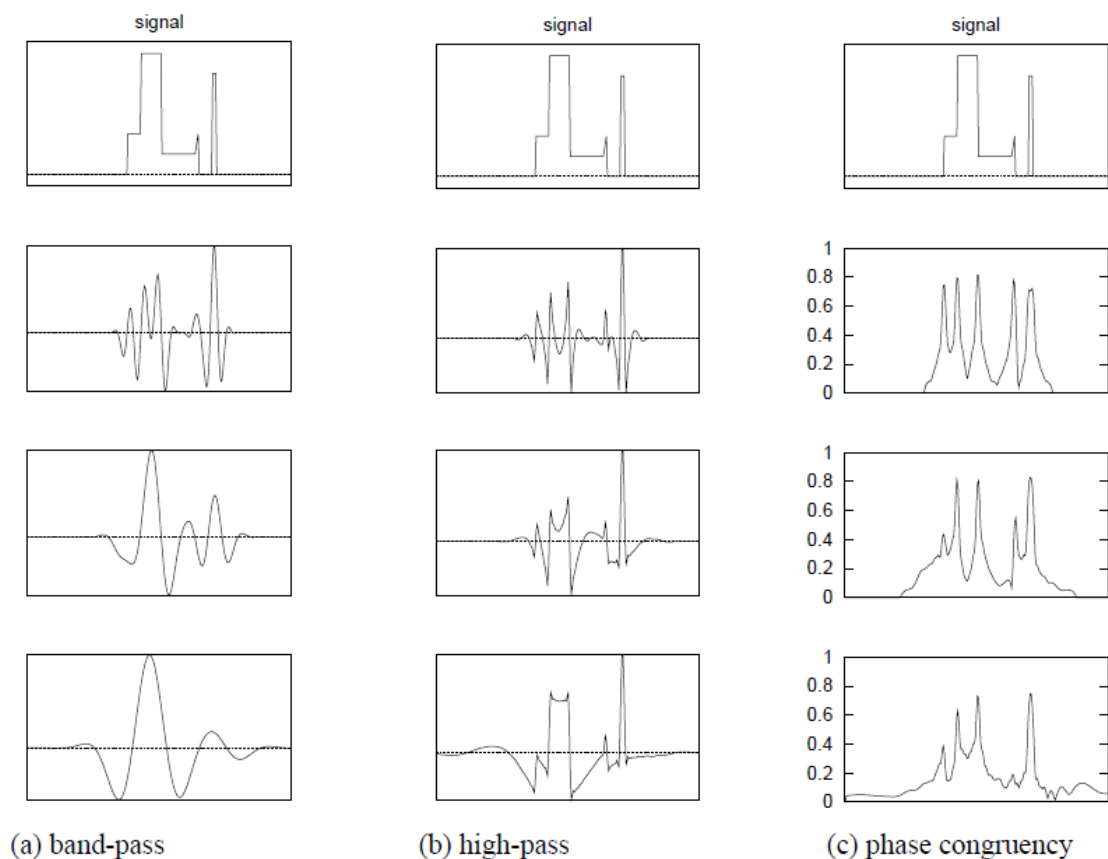


图 3-11 (a)一维信号的三个不同带通结果；(b)三个不同的高通结果；
(c)三个尺度高通滤波后的相位一致性

需要注意的是，我们最初希望找出的不依赖于图像的放大倍率的特征显著性，实际上是

无法得到的。在实际中，我们需要使用覆盖有限频谱的有限数量的空间滤波器来计算相位一致性。改变图像的放大倍率可能改变单个滤波器的相对响应，并由此改变我们所认为的相位一致性，但是相位一致性的改变远远小于对应的图像梯度的改变

总的来说，多尺度分析是通过考虑图像经过不同高通滤波器后的相位一致性来完成的。高通后的图像是通过带通后图像加和得到的，也就是从最高的频率通带加到截止频率通带。利用这种方法，不管我们考虑哪个尺度，所有的特征都被精确定位并且保持稳定，不会出现使用低通滤波时发生的漂移现象，不同尺度分析间的区别是特征的相对显著性，这就为使用由粗到细的策略提供了理想的环境。图 3-11 展示了一个一维信号在三种不同尺度的带通和高通滤波结果，以及它们在三个高通尺度下的相位一致性。

3.8 相位一致性的实验结果与对比

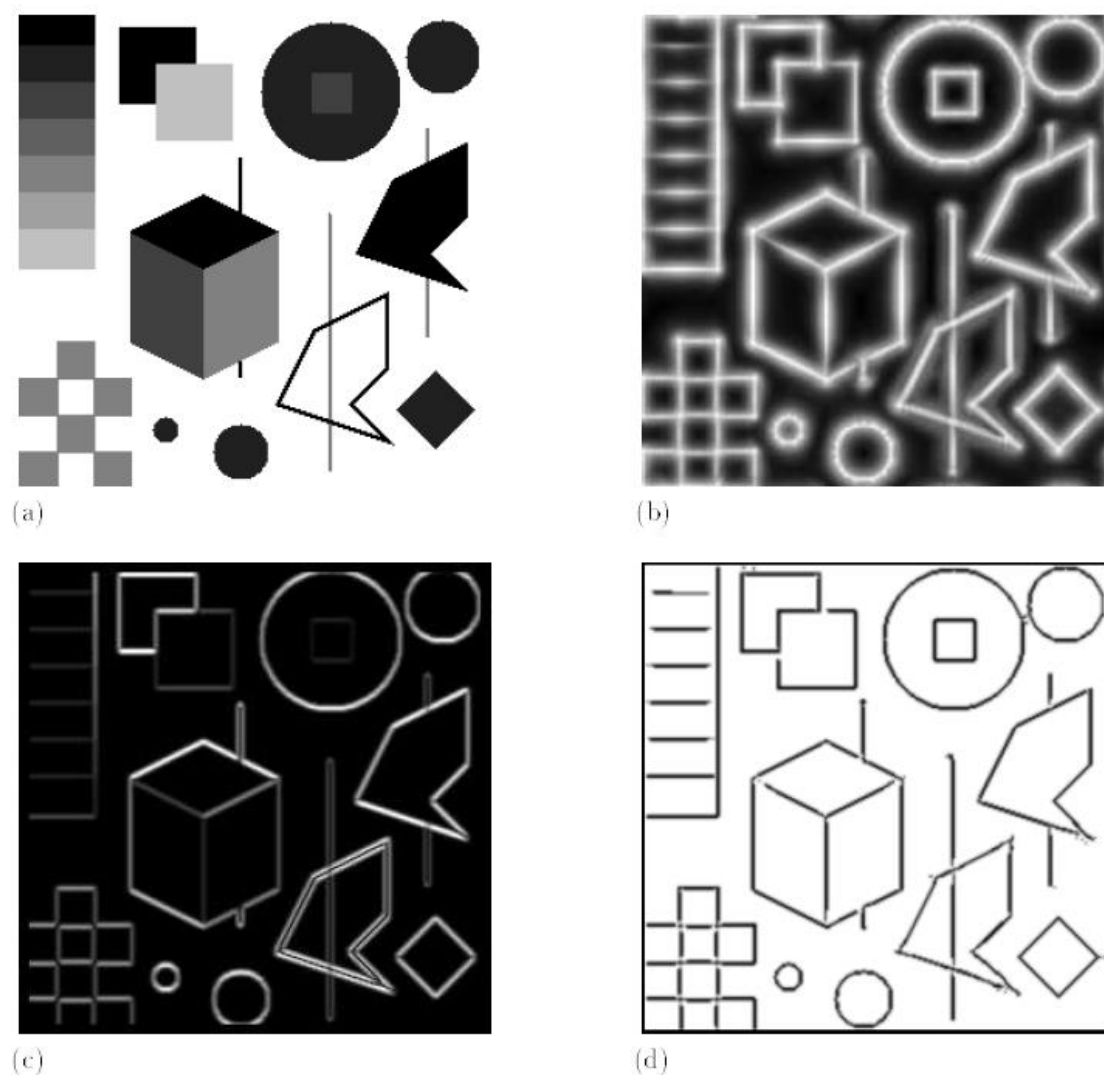


图 3-12 (a)测试图片；(b)相位一致性的原始结果；(c)使用 Canny 检测的原始结果；
(d)经过非极大抑制和设定临界值之后的相位一致性结果

相位一致性的优势是无量纲衡量特征显著性,这允许我们根据图像的类型不同设定不同的门限;相位一致性的另一个优势是可以在对图像的光照水平不做任何假设的情况下检测特征。阶跃间断,线形和屋脊形边缘都可以检测出来。

图 3-12 中展示了包含不同方向上线性特征和台阶特征的测试图片通过相位一致性进行特征检测的结果,作为对比,还给出了 Canny 检测的结果。由于两种检测方法的原理完全不同,所以这里不便采用定量分析。

这两种检测方式的主要区别在于 Canny 检测的响应值有很大的范围,这就使得临界值很难选择;另一个区别在于 Canny 检测在线性边缘的每一侧都有响应,而相位一致性检测在线性特征的中心产生响应。

3.9 本章小结

边缘检测在图像彩色化过程中有着重要应用,本章我们主要介绍了利用相位一致性进行图像特征检测的基本原理,特点和优势。

传统的边缘检测,如 Canny 检测,是基于梯度算子,而相位一致性是基于傅立叶变换理论的图像特征检测方法,其原理是图像的特征发生在相位的最大叠合处,通过计算局部能量,来度量图像各个位置的相位一致性,利用 log Gabor 滤波器组来实现。从实验结果可以看出,相位一致性可以有效地提取边缘信息。

相位一致性特征检测具有无量纲,易于设定门限值,更接近于人类视觉感知等优点,且运算速度远快于多次 Canny 检测,非常适合在本课题中用于计算边缘权重。

第四章 程序实现和实验结果

4.1 软件模块设计

为了实现第二章和第三章中的算法，形成一套完整的黑白图像彩色化软件，我们将程序分为以下几个模块：

- 1) 界面模块：与用户交互，调用其它模块计算彩色化结果并显示；
- 2) scribble 生成模块：从界面模块中获得黑白图像，自动生成 scribble 并返回；
- 3) 最优化模块：计算权值图，通过最优化得出彩色化结果，返回给界面模块显示。

在上述模块的实现过程中，还需要调用一些外部库和外部函数：

- 1) Qt 界面库：著名的跨平台图形界面库，基于 C++ 编程语言，为界面模块提供了良好的交互方式；
- 2) OpenCV：一个开源发行的跨平台计算机视觉库，scribble 生成模块和最优化模块都是使用 OpenCV 的 C++ 接口来实现基本的运算；
- 3) SISL：由挪威科技工业研究院（SINTEF）开发的插值库，scribble 生成模块调用 SISL 实现 B 样条插值；
- 4) liboctave：开源数学软件 GNU Octave 的一部分，提供了完善的 C++ 接口，在最优化模块中使用 liboctave 中的 SparseMatrix 类来求解线性稀疏系统；
- 5) 外部函数 phasecong：自编函数，基于相位一致性模型计算图像的边缘权重，在最优化模块中需要调用。
- 6) 外部函数 qgabor：自编函数，使用四元 Gabor 相位分析得到图像的结构权重，在最优化模块中需要调用。

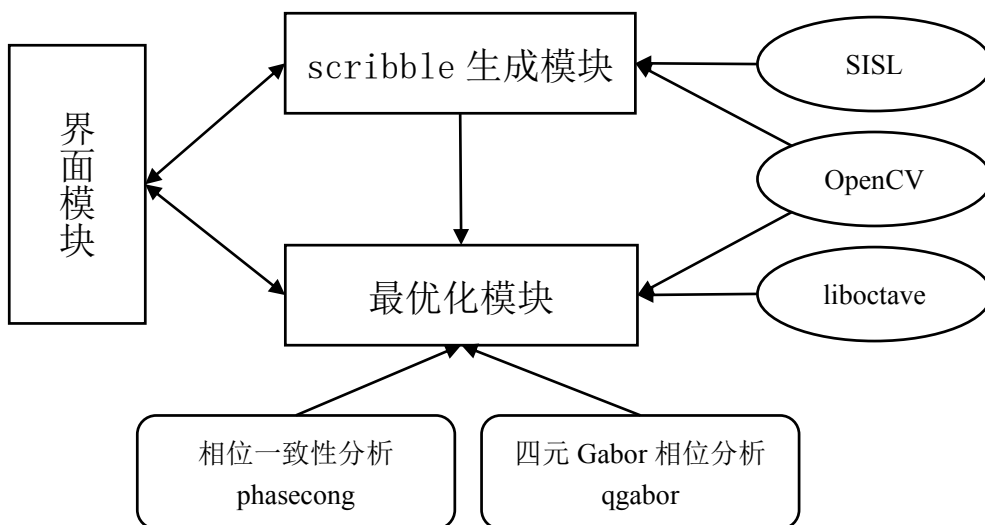


图 4-1 各模块与外部库、外部函数间的关系

图 4-1 清楚地表示出各模块相互之间以及模块与外部库、外部函数间的调用关系。

在课题组的前期工作中，已经实现了 Linux 平台下自动产生 scribble 和最优化功能，但是，在最优化模块中，未能有效实现多次 Canny 检测，边缘信息实际上未能进入权值图。

针对这一情况，我们成功地将程序移植到了更为普遍的 Windows 平台下，并用基于相位一致性的边缘检测方法实现了求得了边缘权重，并且加入到权值图中，优化了彩色化的效果和程序的运行时间，这在下面两节中会继续讨论。

4.2 基于相位一致性边缘检测的实现

在第三章中已经详细讨论了基于相位一致性的边缘检测的原理，这里总结了我们选择相位一致性的方法来替代原有的 Canny 边缘检测的原因。

- 1) 相位一致性是一个无量纲的量，它提供了特征点显著性的绝对衡量，可以很方便地应用在各种不同种类的图像中；
- 2) 相比于传统边缘检测方法，相位一致性方法的抗噪声性能更加优异；
- 3) 相位一致性的检测结果更加符合人类视觉的感知特性；
- 4) 虽然对信号的频率分解的计算过程十分复杂，但是在利用近似模型——局部能量模型之后，计算量大幅降低，实际的运算时间远低于多次 Canny 检测。

然而，现有的相位一致性代码只有 MATLAB 版本，为了在统一的平台上采用纯 C++ 语言实现算法，我们需要将 MATLAB 代码改写成 C++ 代码，并且针对 C++ 的特性做出响应的调整，以期获得更加快速而准确的运算结果。

本程序中基于相位一致性的边缘检测模块的接口函数原型如下：

`Mat phasecong(Mat_<uchar> im)`

im 为输入的灰度图像，函数返回带有边缘权重的图像，每个像素的值是 0 到 1 之间的浮点数。

在算法实现过程中，需要注意的是，我们采用对某个方向上各个点的响应幅度和求中位数的方法来估计噪声，在 MATLAB 中是直接用 median 函数来实现的，而在 C++ 中没有对应的函数，最自然的想法是使用 STL 中的 sort 函数对全图进行排序，然后取中间位置的值即为中位数，但这样做的算法平均时间复杂度为 $O(N \log N)$ ，但是，我们显然有更好的做法，通过修改快速排序算法，利用 partition 操作来找到中位数，并不需要完整的全图排序，这种方法可以将时间复杂度降低到 $O(N)$ 。

4.3 实验结果和相关讨论

本程序在 Windows 平台下已经完整实现，本节中测试的环境为 32bit Windows 8.1 系统，编译器为 Visual Studio 2013。

打开软件后，点击 File 菜单下的 Open 命令可以调用 Windows 的资源管理器来选择需要进行彩色化的图像。

如果选择的是黑白图片，则 File 菜单中的“show scribble”命令变为有效，点击“show scribble”即可看到自动生成的 scribble 直接显示在图像上，此时“Start Colorization”变为有

效，首先点击图中的 scribble 区域，软件跳出选色板，如图 4-2 所示，为每一个 scribble 选择颜色，所有 scribble 的颜色都指定后，点击“show scribble”就可以看到彩色化的结果。

如果打开的是一幅彩色图像，那么软件仍然会将其转换为黑白版本显示，用户可以将它作为黑白图像处理，对其 scribble 指定颜色后重新彩色化。软件还提供了另一种选择，即“Auto Colorization”功能，在打开彩色图像的时候，程序已经将彩色图像的另存为一个参考图像（reference image），用户点击“Auto Colorization”之后，程序会对该图像的黑白版本自动生成 scribble，然后对每个 scribble 按照参考图像上对应区域的颜色进行自动涂色，最后进行彩色化。加入这一功能是为了方便测试算法，极大地节约了测试时间。下文中的测试也是通过这一功能来完成的。

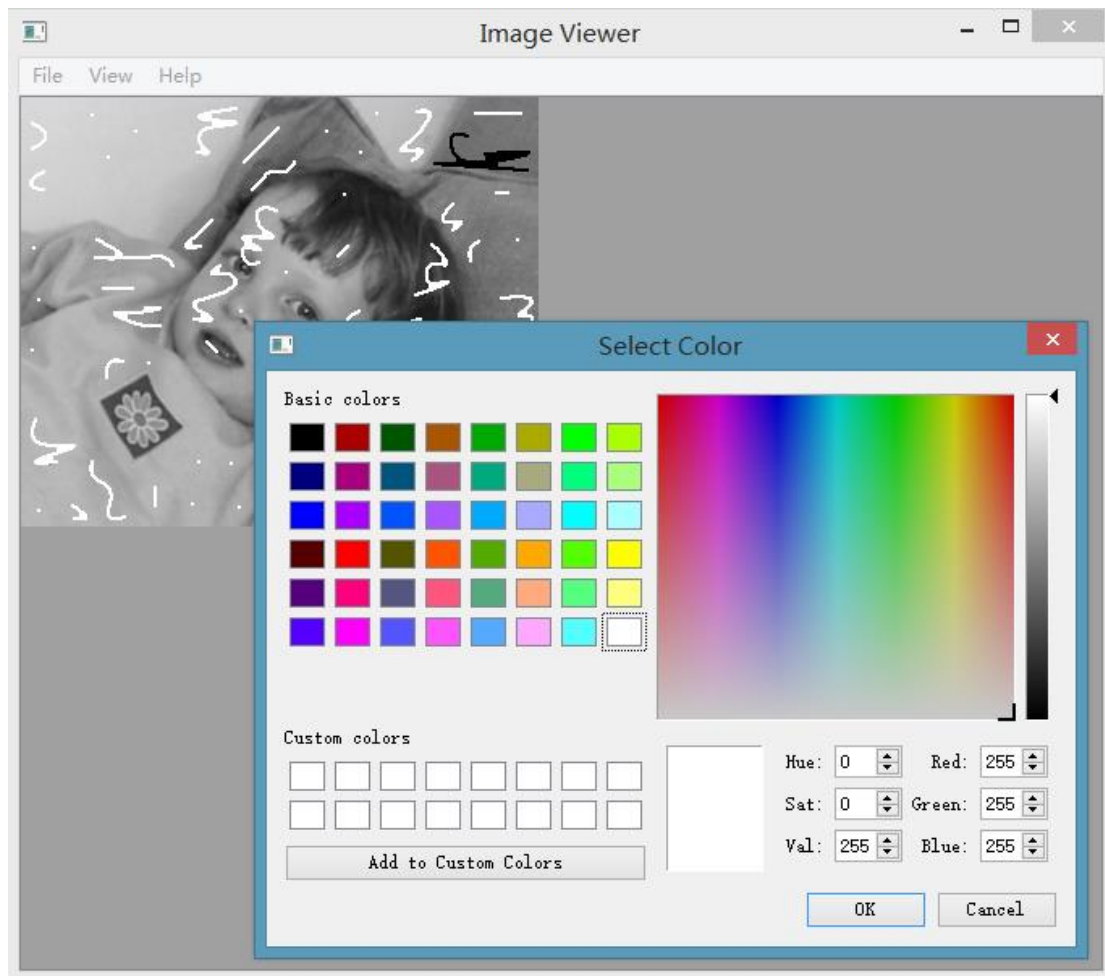


图 4-2 软件选色界面

为了便于对比，我们下面所有的测试结果都给出了彩色的原图，将原图转化为灰度图后用相位一致性检测到的边缘图，使用课题组之前的程序进行自动染色的结果以及加入相位一致性边缘权值之后的彩色化结果。

我们首先选取了一个人像图来进行测试，如图 4-3 所示，从图中可以看出，彩色化后的效果都与原图比较接近，达到了彩色化的基本目的，但仔细对比 4-3(c)和 4-3(d)，可以看出，加入相位一致性边缘之后得到的彩色化结果有着更加清晰的边缘，而在一些局部区域，比如

小孩的手肘内侧, 4-3(e)明显有袖子上的绿色过渗透的现象, 而在 4-3(f)中手肘内侧有着清晰的边界, 颜色没有渗过边界, 而我们还注意到, 虽然加上了相位一致性边缘, 手肘外侧的颜色仍然发生了过渗透, 从相位一致性的结果 4-3(b)中可以找到原因, 手肘内侧的边缘要强于手肘外侧, 甚至手肘外侧的边缘不连续, 这样就无法阻止颜色的外渗。

这给我们的后续工作带来的启发就是, 对于某些强边缘, 它们之间连结的点可能是弱边缘, 甚至是间断的, 如果通过求均值来构成一个连通的较强的边缘, 可能会有更好的阻止颜色渗透的效果。



(a) 原图



(b) 相位一致性边缘检测的结果



(c) 原程序的染色结果



(d) 加入相位一致性边缘之后的染色结果

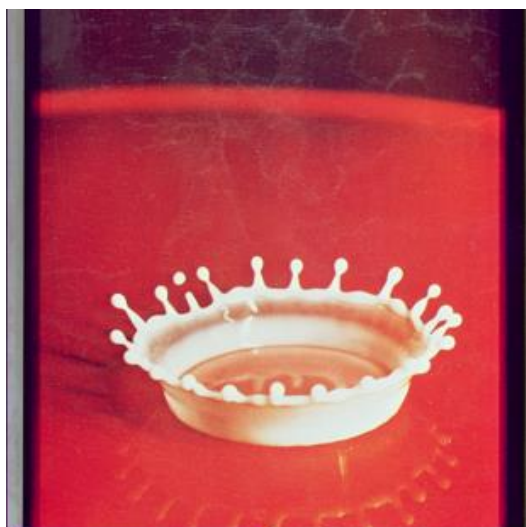


(e) 原程序结果中手肘部分的颜色

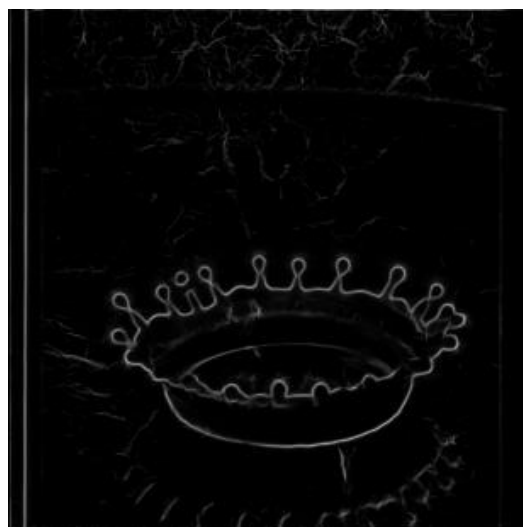


(f) 加入边缘后手肘部分的颜色

图 4-3 测试: 小孩



(a) 原图



(b) 相位一致性边缘检测的结果



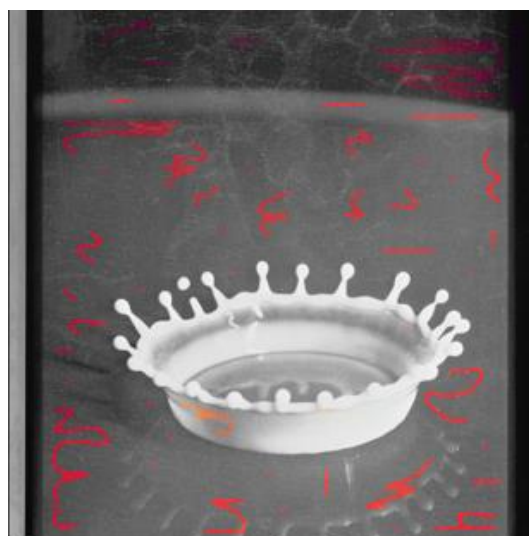
(c) 原程序的染色结果



(d) 加入相位一致性边缘之后的染色结果



(e) scribble 图



(f) 自动 scribble 上色

图 4-4 测试：飞溅

图 4-4 是一个液体飞溅的图片，它的纹理信息比较简单，边缘信息也非常明显，利用相位一致性可以很好地检测出红色和白色的边缘。对比 4-4(c)和 4-4(d)的结果，发现在红白交界的高频区域，由于纹理简单而且相似，单纯的四元 Gabor 相位分析结果不能阻止颜色扩散，4-4(c)中的白色区域明显受到红色的污染，而在应用了相位一致性边缘检测之后，红白分界处的颜色扩散受到抑制，彩色化效果明显改善。

我们还注意到的一点是，图像中间一块椭圆区域并没有被染到颜色，从 4-4(e)和 4-4(f)中可以看出，该处并没有 scribble，这与分割算法有关，后续工作可以考虑修改分割算法，使得像图 4-4 中心椭圆区域这样的边缘较明显的区域可以被独立分割成一块，在它的中心产生一个 scribble。

图 4-5 中左边长的青椒头上被辣椒染成了红色，加入了相位一致性边缘之后，红色的区域变小了，但仍然没有完全解决红色过渗透的问题，参考 4-5(b)中显示的边缘，与图 4-3 中手肘外侧类似，同样是不连续的边缘未能阻止颜色扩散。

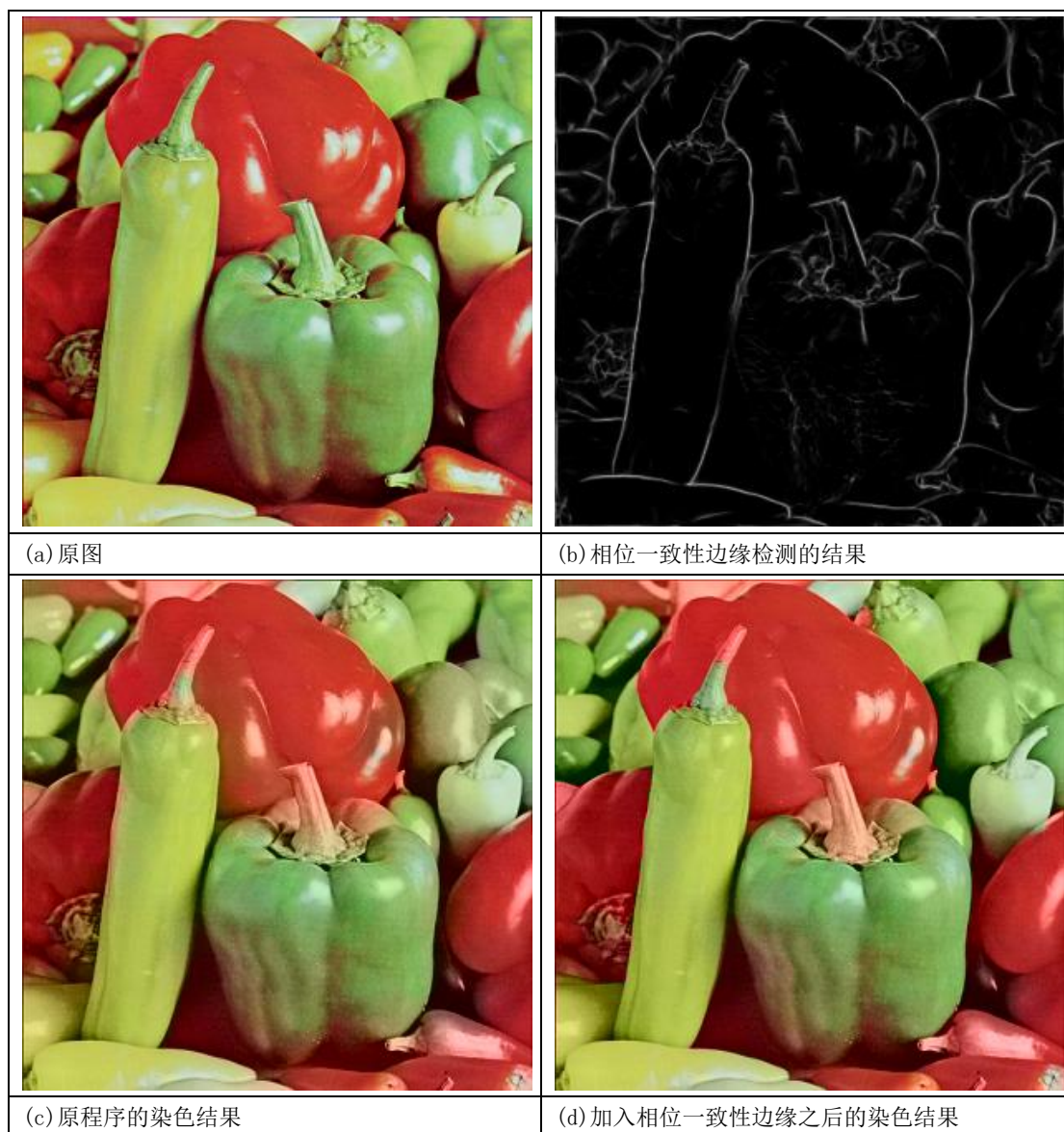


图 4-5 测试：青椒

图 4-6 选取了山魈的脸部作为测试对象，在毛发区域，有一定的过渗透现象，但是因为毛发区域高频成分过多，几乎不可能避免这种渗透，而在鼻子的区域，却有着明显的线形边缘，在加入相位一致性边缘之后，红色的鼻子向两侧蓝色区域的渗透红色的现象几乎消失了，彩色化结果更加接近原图。

图 4-7 是典型的风景图，在有天空区域的风景图中，我们对原程序的改进非常明显，在图 4-7 中可以发现，原程序的结果把城堡上端靠近天空的区域都染成了蓝色，而在我们加入相位一致性边缘检测之后，大部分被渗透的区域都被染成了应有的颜色，虽然仍有少部分区域被染成蓝色，但从观赏者的角度来看，加入相位一致性边缘之后的彩色化结果获得了更好的视觉效果，这也正是我们彩色化的目的。

通过以上这些实验，不难发现，加入了相位一致性边缘检测后的算法对各种图片都能获得一定的效果提升，边缘较明显边界的颜色过渗透现象得到了很好地解决。

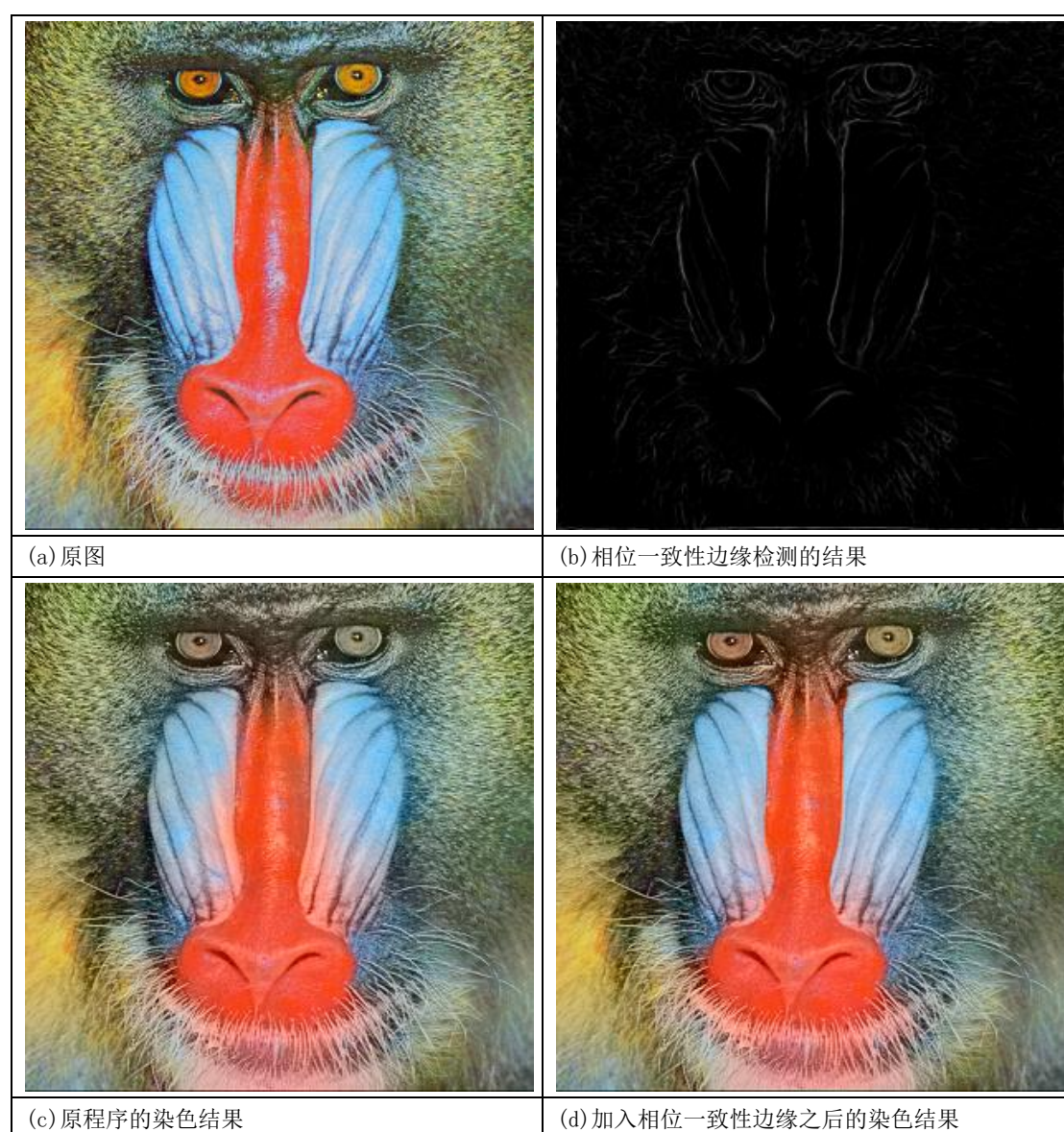


图 4-6 测试：山魈



图 4-7 测试：城堡

第五章 总结与展望

黑白图像彩色化问题是一个有趣而又实用的课题。本文总结了目前主流的黑白图像彩色化方法,介绍了课题组前期在 Levin 算法的基础上开发出的一整套彩色化算法,包括了自动 scribble 生成算法和基于四元 Gabor 相位分析的最优化算法。

针对该算法仍然存在的问题,我们引入基于相位一致性的边缘检测作为对四元 Gabor 相位分析的补充,我们详细介绍了相位一致性方法的原理和特点,通过实验结果展示了相位一致性的引入对原算法的改善,很好地解决了一些边缘上颜色过渗透的问题。

我们还继续完成了前期工作未能完成的部分,在统一的 Windows 平台下使用 C++ 语言完整实现了整套算法。

然而,现有的实验结果仍然有改进的空间,其中几个可以考虑的改进方向是:

- 1) 相位一致性边缘强度计入权值图的方式:现有的方式是直接与四元 Gabor 相位的结果相乘,后续工作可以考虑是否设定门限,是否对不连续的强边缘进行连通处理。
- 2) 估价函数的设计:现有的估价函数仍然采用的是 Levin 给出的较为简单的形式,后续工作可以测试新的估价函数;
- 3) 分割算法的进一步讨论:目前的基于图论的分割方法对参数比较敏感,在某些图片中有过分割的现象,后续工作可以探讨参数的给定方式,也可以尝试采用其它分割算法;
- 4) 考虑光流,实现黑白视频的彩色化:本课题提出的算法理论上也适用于黑白视频,后续可以在现有程序的基础,探讨黑白视频的彩色化方法。

此外,现有程序尚未经过大量的充分的测试,可能存在一些遗留的 bug,继续提升程序的鲁棒性,也是可以在后续工作中加强的方向。

参考文献

- [1] N. Smary, Image coloring Techniques and Applications, GRIN Verlag, 2012.
- [2] C.-W. Ho, C.-Y. Ke, T.-L. Wu and C.-H. Cheng, "Color adjustment circuit, digital color adjustment device and multimedia apparatus using the same". US Patent 20130044122, 21 2 2013.
- [3] A. Levin, D. Lischinski and Y. Weiss, "Colorization using optimization," *ACM Transactions on Graphics (TOG)*, Volume 23 Issue 3, pp. 689-694, 8 2004.
- [4] T. Horiuchi, "Colorization algorithm using probabilistic relaxation," *Image and Vision Computing*, pp. Volume 22, Issue 3, 197-202, 1 3 2004.
- [5] M. S. Drew and G. D. Finlayson, "Realistic colorization via the structure tensor," in *Image Processing, 2008. ICIIP 2008. 15th IEEE International Conference on*, 2008.
- [6] T. H. Kim, K. M. Lee and S. U. Lee, "Edge-preserving colorization using data-driven Random Walks with Restart," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009.
- [7] O. Lezoray, V. T. Ta and A. Elmoataz, "Nonlocal graph regularization for image colorization," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, 2008.
- [8] J.-H. Heu, D.-Y. Hyun, C.-S. Kim and S.-U. Lee, "Image and video colorization based on prioritized source propagation," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009.
- [9] G. Sapiro, "Inpainting the colors," in *Image Processing, 2005. ICIIP 2005. IEEE International Conference on (Volume:2)*, 2005.
- [10] L. Yatziv and G. Sapiro, "Fast Image and Video Colorization Using Chrominance Blending," *Image Processing, IEEE Transactions on (Volume:15, Issue: 5)*, pp. 1120-1129, 5 2006.
- [11] Z. Yu, Y. Xu, X. Yang and L. Song, "Structure-Preserving Colorization Based on Quaternionic Phase Reconstruction," *Advances in Multimedia Information Processing - PCM 2009*, pp. 847-857, 2009.
- [12] X. Ding, Y. Xu, L. Deng and X. Yang, "Colorization using quaternion algebra with automatic scribble generation," *Advances in Multimedia Modeling*, pp. 103-114, 2012.
- [13] 李粤龙, "基于四元小波变换的图像彩色化方法," 上海交通大学, 2013.
- [14] 孙君顶, 丁振国 和 周利华, "基于图像信息熵与空间分布熵的彩色图像检索方法," *红外与毫米波学报*, pp. 135-139, 第 24 卷第 2 期 4 2005.
- [15] Y. Xu, X. Yang, L. Song, L. Traversoni and W. Lu, "QWT: Retrospective and New Applications," *Geometric Algebra Computing in Engineering and Computer Science*, pp. 249-273, 2010.
- [16] E. Bayro-Corrochano, "The Theory and Use of the Quaternion Wavelet Transform," *Journal of Mathematical Imaging and Vision*, pp. 19-35, Volume 24, Issue 1 1 2006.
- [17] M. Morronea and R. Owens, "Feature detection from local energy," *Pattern Recognition*

- Letters*, p. 303–313, Volume 6, Issue 5 12 1987.
- [18] P. Kovesi, “Image features from phase congruency,” *Videre: Journal of computer vision research* 1(3), pp. 1-26, 1999.
- [19] S. Venkatesh and R. Owens, "An energy feature detection scheme," in *ICIP'89: IEEE International Conference on Image Processing: conference proceedings*, Singapore, 5-8 September 1989.
- [20] E. P. Simoncelli, W. T. Freeman, E. H. Adelson and D. J. Heeger, "Shiftable multiscale transforms," *Information Theory, IEEE Transactions on*, pp. 587-607, Volume:38 , Issue: 2 3 1992.
- [21] “Relations between the statistics of natural images and the response properties of cortical cells,” *The Journal of the Optical Society of America A*, pp. 2379-2394 , Vol. 4, Issue 12 12 1987.
- [22] P. Lagodzinski and B. Smolka, "Fast Digital Image Colorization Technique," in *Signal Processing and Information Technology, 2007 IEEE International Symposium on*, 2007.
- [23] P. Kovesi, "Phase congruency detects corners and edges," in *The Australian pattern recognition society conference: DICTA*, 2003.
- [24] Itseez, "OpenCV docs," 2014. [Online]. Available: <http://docs.opencv.org/>. [Accessed 5 2014].
- [25] SINTEF ICT, Applied Mathematics, "SISL-The SINTEF Spline Library Reference Manual," 2005.
- [26] J. W. Eaton, "Octave C++ Classes," 1993.
- [27] Digia plc, "Qt Project Documentation," 2014. [Online]. Available: <http://qt-project.org/doc/>. [Accessed 5 2014].
- [28] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient Graph-Based Image Segmentation," *International Journal of Computer Vision*, pp. 167-181, Volume 59, Issue 2 9 2004.

谢辞

在毕业论文完成之际，首先要向指导老师徐奕老师致以最衷心的感谢。徐老师对科研工作严谨认真，对学生十分关心，从大三 PRP 开始，她的这种态度就深深地感染了我，从开题到中期再到论文，徐老师都不辞辛劳地帮我们认真修改，毕设过程中一点一滴的进步都离不开徐老师的指导和帮助，能够成为在大学期间遇到徐老师这样的好老师，是我的幸运！

感谢课题组的学长学姐在前期所做的大量工作，特别要感谢李粤龙，人在美国还能不厌其烦地为我解答程序中的问题。

感谢大学期间我的所有授课老师，是他们给我打下了专业基础，让我在大学生活中受益匪浅。

感谢我的父母，一直以来都无条件地理解我支持我包容我，处处为我考虑，他们的期待与鼓励是我前进的动力。

IMAGE COLORIZATION METHOD BASED ON QUATERNION PHASE ANALYSIS

Colorization is a term generically used to describe process of adding colors to monochrome still images and movies. It is a practical topic with broad application prospects. Chromatic image has more details and provides comprehensive visual experience, compared to monochrome image. Colorization is a good way to bring those old photos or movies back to life without a high renovation cost. It is also a very useful tool for applying pseudo colors to gray images such as medical images and remote sensing images, making different colors representing different characteristics. However, since we may take numerous colors for a given intensity value, the colorization problem is regarded as “ill-posed” and thus requires certain amount of human interaction or external information.

Colorization requires considerable user intervention. A typical form of user intervention is manually annotating the image with a few color scribbles. Then, based on the assumption that neighboring pixels with similar intensities should have similar colors, a quadratic cost function is optimized, which leads to color propagation throughout the whole image by solving an optimization problem. Levin is the first person coming up with this kind of methods. Many researchers follow up Levin’s idea to seek for improvements, such as Yatziv, Sapiro, etc.

Although Levin’s method can reduce part of the time consuming manual work, people still need to perform the tedious task of drawing scribbles. In addition, as the colorization results depend largely on information provided by scribbles, the colorization results might be unstable. Finally, this method focuses on intensity variations only, thus making the ignorant of structural information. Further improvements may be obtained if structural information is considered.

The previous work of our group is also based on Levin’s algorithm. To overcome the shortcomings above, our group proposed an automatic scribble generation algorithm based on spatial distribution entropy, placing scribbles within the regions of high information density. To avoid a single scribble crossing different homogeneous segments in the image, we first apply Graph-Based Image Segmentation algorithm to the original image and then conduct a search in each segment separately. Spatial distribution entropy is adopted to measure information density. We locate the scribbles in the regions with high spatial distribution entropy.

Another problem is that Levin’s algorithm has not taken the structural and edge information of the image into consideration. To fix the problem for better colorization results, an improved weighting function was calculated in our solution, which includes the quaternion Gabor filtering coefficients and the hierarchical Canny edge strength in the weighting function. The two parts represent the structural information in various directions and the edge information separately.

In previous work, our group has implemented this algorithm via C++ on Linux platform. However, the run result of the program is far from perfect. The colored image still suffered from color leakage since the hierarchical Canny edge detection was not fully implemented in the program, making it hard to prevent color information from propagating through edges. Actually, the hierarchical Canny edge detection is not a practical solution here because it is really a time and memory consuming process.

The limitations discussed above give us the motivation to find a better edge detection instead. In this paper, we introduce the phase congruency as a measure of edge significance.

Phase congruency is a dimensionless measure of feature significance. It provides an absolute measure of the significance of feature points in an image that allows constant threshold values to be applied across wide classes of images. Thresholds can be specified in advance; they do not have to be determined empirically for individual images.

We adopt a practical implementation for the calculation of phase congruency in two-dimensional images using wavelets, which was presented by Kovess. It is shown that for a normalized measure of feature significance, such as phase congruency, it is crucial to be able to recognize the level of noise in an image and to compensate for it. A highly effective method of compensation is presented that only requires that noise be roughly constant across the image and its power spectrum be approximately constant.

We also explain the importance of weighting phase congruency by some measure of the spread of the frequencies that are present at each point in an image. This prevents false positives being marked where the frequency spread is narrow. It also improves the localization of features. While it is not possible to specify one ideal distribution of filter response amplitude with frequency, it is shown that when geometrically scaled filters are used, a uniform distribution of responses is a particularly significant one. This distribution matches typical spectral statistics of images and corresponds to the distribution that arises at step discontinuities.

A new approach to the concept of scale, also presented by Kovess, in image analysis is implemented in the program. The natural scale parameter to vary in the calculation of phase congruency is the size of the analysis window over which to calculate local frequency information. Thus, under these conditions, scale is varied using high-pass filtering rather than low-pass or band-pass filtering. The significant advantage of this approach is that feature locations remain constant over scale, and only their significance relative to each other varies.

We have chosen phase congruency feature detection to replace Canny for several reasons.

- 1) Phase congruency is a dimensionless quantity, that is invariant to changes in image brightness or contrast, hence it provides an absolute measure of the significance of feature points;
- 2) Compared to the traditional edge detection methods, the method of phase congruency has more

excellent anti-noise performance;

- 3) The test results of phase congruency is consistent with the human feature perception, making it appear more natural for human eyes;
- 4) Although the calculation of phase congruency in an image is computationally expensive, the running time is reduced significantly by applying the local energy model and other techniques in coding.

We have solved the compatibility problem and implemented the whole algorithm on Windows platform with MS Visual Studio IDE. In practice, we divide the whole software implementation into three parts: the scribble generation module, the optimization module and the graphical user interface. Several open-source libraries (OpenCV, Qt, SISL, liboctave) have been utilized in the program.

Different kinds of images have been tested in this paper. Tests show that the colorization results from the program with phase congruency calculating edge strength have much better performance than those results from the original program in recovering edges, especially for natural scenery images. The leakage of color in edge regions is prevented effectively. Consequently, the colorization results has more precise and clear edges. In addition, the running time of our algorithm is less than the original one, mainly due to the absence of time-consuming hierarchical Canny edge detection.

Despite achieving some improvements, our algorithm still has some limitations now. Color leakage may occur around weak edges. The graph-based segmentation algorithm we use is sensitive to parameters, resulting in over-segmentation for some certain images. We propose several possible enhancements to the algorithm for follower's reference in the last part of the paper.