

西安交通大学

博士学位论文

基于视觉先验的图像处理模型与算法研究

申请人：孙 剑

学科专业：应用数学

指导教师：徐宗本 教授

2009 年 4 月

**Research on Image Processing Models and Algorithms
Based on Visual Prior**

A dissertation submitted to
Xi'an Jiaotong University
In partial fulfillment of the requirements
for the degree of
Doctor of Science

By
Jian Sun
(Applied Mathematics)
Supervisor: Prof. Zongben Xu
April 2009

论文题目：基于视觉先验的图像处理模型与算法研究

学科专业：应用数学

申请人：孙剑

指导教师：徐宗本教授

摘 要

图像处理（Image processing）是模式识别和人工智能研究领域的重要分支，主要关注如何对图像进行分析、加工和处理，使其满足视觉、心理和模式识别技术的需要。图像处理被广泛地应用于人工智能技术、航空航天技术、生物医学技术、军事科学技术、公共安全技术等各个领域，具有极其重要的科学价值和应用价值。

本论文主要关注图像处理的如下基本问题：图像去噪、超分辨率、增强、填充和前景/背景分割。从数学上说，这些基本问题均为逆问题：即要求通过观测到的图像（噪音图像、模糊图像、低分辨率图像、缺失图像等）来推测原始图像或者图像前景/背景分割结果。因此从数学建模和算法设计角度，问题的最优解（即图像处理结果）往往不具有唯一性和稳定性，从而影响图像处理结果的质量。

本论文的研究总体思路是：通过发掘和建模视觉信息（即图像信息）的一般性规律（定义为视觉先验），以约束图像处理模型的解空间，使得图像处理结果满足视觉先验，因此获得更符合人的视觉感知的图像处理效果。基于这样的研究思路，本论文从图像处理每一基本问题的特异性出发，提出相应的视觉先验模型，并进而发展解决相应基本问题的优化模型及其优化算法。实验表明，所提出的图像处理模型和算法相对于已有的同类模型和算法具有更好的图像处理效果，因此具有重要的理论意义与直接而广泛的应用价值。

本论文的主要成果包括：

1. 提出了基于自然图像梯度轮廓先验的图像超分辨率重建与增强算法。通过归纳自然图像梯度场在垂直于图像边缘方向上的梯度轮廓分布，提出了一种新的基于统计的自然图像统计先验—梯度轮廓先验。该先验参数化地描述了自然图像梯度轮廓的形状和锐化度分布。基于该先验，发展了解决图像超分辨率重建和增强问题的能量函数模型，并基于变分方法设计了优化算法。理论分析与应用实践表明：该模型和算法具有良好的数学性质，而且比已有算法具有更优良的图像处理效果。

2. 提出了基于Markov随机场尺度选择模型的各向异性扩散去噪算法。通过对各向异性扩散去噪算法的尺度（即扩散停机时间）选择进行建模，提出了空间自适应的各向异性扩散滤波的尺度选择策略，并基于此发展了基于Markov随机场先验建模的尺度选择统计模型。结合该尺度选择模型，我们进而提出了新的基于各向异性扩散滤波的去噪框架。该框架可以解释为各向异性扩散尺度空间的融合算法，有效解决了已有全局同一尺度选择方法所带来去噪效果差的弊端。应用表明，所提出的基于各向异性扩散滤波的去噪算法是目前最好的图像去噪算法之一。

3. 提出了基于图像块稀疏性先验的图像填充算法。通过定义图像块结构稀疏性，进而证明结构稀疏性是刻画图像重要结构的有用指标，结构稀疏性越大，则图像块越可能处于图像边缘；基于视觉编码的稀疏性先验，在假设待填充图像块可以

由候选图像块集合的稀疏线性组合来表达的基础上，我们提出了图像块稀疏线性表达的优化模型和优化算法，并发展了基于图像块稀疏性先验的图像填充算法。相比以往的同类填充算法，所提出的新算法具有更好的填充性能。

4. 提出了基于闪光信息和MRF先验建模的前景/背景分割算法。以开/关相机闪光灯设备所提供的图像对为输入，通过数学模型建模闪光设备所带来的光照变化的先验信息，提出了基于MRF先验的图像分割能量模型，并采用Graph cut算法进行优化。与已有的分割算法相比，所提出的算法可以得到更为精确的前景/背景分割结果。该模型的重要意义在于通过相机的闪光设备，提供了一种便捷、实用的前景提取模型和算法，可期待成为新一代相机的实用技术之一。

关键词：视觉先验；图像处理；图像稀疏性；图像统计；能量模型

论文类型：应用基础

Title: Research on Image Processing Models and Algorithms Based on Visual Prior

Speciality: Applied Mathematics

Applicant: Jian Sun

Supervisor: Prof. Zongben Xu

ABSTRACT

Image processing is an important research branch of pattern recognition and artificial intelligence. It mainly concentrates on analysis and processing of the observed image to make it satisfy the needs of visual perception, psychology and pattern recognition. It is widely applied to artificial intelligence technology, aerospace technology, biomedical technology, military technology, public security technology, and so on. Therefore, it is definitely worthy of research and application.

This dissertation focuses on the following basic problems in image processing: image denoising, super-resolution, enhancement, inpainting and foreground/background segmentation. These problems are widely investigated and applied in the area of image processing and computer vision. However, these problems are all inverse problems in mathematics, i.e. to infer the unknown original image or foreground/background segmentation result from the observed image (e.g. noisy image, blur image, low-resolution image or image with missing region). It implies that the optimal solutions of models for image processing problems are not unique or stable, therefore reduce the quality of image processing results.

In this dissertation, the generic rules of visual information (i.e. natural images), which it is called visual prior, are investigated and modeled to constrain the solution space of image processing models. The well-designed visual priors force the image processing results to satisfy the generic rules of visual information, so that the results are more pleasant to visual perception. By modeling the visual priors, novel energy models or statistical models are proposed for each image processing problem. Algorithms are also designed and analyzed for optimizing these models. The experiments and comparisons show that these proposed models and algorithms achieve the state-of-the-art image processing results.

This dissertation made the following achievements.

1. Image super-resolution and enhancement algorithms using natural image gradient profile prior are proposed. Through statistically modeling the distribution of the gradient magnitudes along the direction orthogonal to image edges, a novel natural image prior, which is called gradient profile prior, is proposed. This prior parametrically models the shape and sharpness of the image gradients along edges. Using this prior, energy models are developed for the problems of image super-resolution and enhancement. It is shown that these models have excellent mathematical properties, and achieve the state-of-the-art super-resolution and enhancement results.

2. Anisotropic diffusion denoising algorithm based on the Markov Random Field scale selection model is proposed. A novel spatially variant scale (i.e. diffusion stopping time) selection scheme is designed. Using this scheme, a novel scale selection model based on Markov Random Field prior is developed and optimized using tree-reweighted message passing algorithm. Combing with this scale selection model, a novel denoising framework is designed, which can be seen as a scale space fusion algorithm. Comparisons show that it achieves the best scale selection performance and the state-of-the-art denoising performance.

3. Image inpainting algorithm using image patch sparsity prior is proposed. First, by defining the patch structure sparsity, a novel measurement is defined to measure the confidence of image structure. The higher sparsity of a patch, the more confidence in its locating at the structure. Second, patch sparse representation is designed based on the assumption that the image patch can be represented by the sparse linear combination of candidate patches. Then an optimization model for patch inpainting is proposed based on patch sparse representation, and a greedy optimization algorithm is designed to optimize this model. Compared with the previous inpainting algorithm, the proposed algorithm achieves better inpainting performance.

4. Image foreground/background segmentation algorithm using flash and Markov Random Field prior is proposed. The input images are a pair of images taken with flash on and off respectively. By modeling the flash effect on the pair of images, a novel foreground/background segmentation model is proposed based on Markov Random Field prior modeling, and optimized by graph cut algorithm. Comparison with the other algorithms show that the proposed model achieves the most accurate segmentation results. The importance of this model is that an accurate and easy-to-use foreground extraction technique is proposed, and hopefully will be applied as a practical technique of the camera in the future.

KEY WORDS: Visual prior; Image processing; Image sparsity; Image statistics; Energy model

TYPE OF DISSERTATION: Application Fundamentals

目 录

1 绪 论	1
1.1 研究背景	1
1.2 图像处理的基本问题	2
1.2.1 图像去噪	2
1.2.2 图像超分辨率重建	3
1.2.3 图像增强	3
1.2.4 图像填充	3
1.2.5 图像前景/背景分割	4
1.3 图像的先验建模	4
1.3.1 基于光滑性的先验	6
1.3.2 基于自然图像统计规律的先验	7
1.3.3 基于视觉编码稀疏性的先验	8
1.3.4 基于对象或设备的先验	9
1.4 本文的主要工作	9
2 基于自然图像梯度轮廓先验的图像超分辨率重建与增强算法	11
2.1 引言	11
2.1.1 有关图像超分辨率和增强的已有工作	12
2.2 自然图像梯度轮廓先验	12
2.2.1 梯度轮廓锐化度估计	13
2.2.2 自然图像梯度轮廓先验	14
2.2.3 高/低分辨率图像梯度轮廓的锐化度关系	16
2.3 自然图像梯度场变换	16
2.3.1 梯度场变换	17
2.3.2 基于梯度场的图像重建	18
2.4 基于梯度轮廓先验的超分辨率重建算法	19
2.4.1 超分辨率模型	20
2.4.2 更多的例子	22
2.5 基于梯度轮廓先验的增强算法	24
2.5.1 锐化度变换函数	25
2.5.2 锐化度转移	27
2.5.3 对比实验	29
2.6 小结	29
	V

3 基于Markov随机场尺度选择的各向异性扩散去噪算法	31
3.1 引言	31
3.2 各向异性扩散尺度空间	34
3.3 尺度选择模型	35
3.3.1 图像过分割	35
3.3.2 图像块尺度选择策略	36
3.3.3 Markov随机场尺度选择模型	38
3.4 基于分割块的噪声估计模型	40
3.5 实验	41
3.5.1 噪声估计比较	42
3.5.2 尺度选择比较	43
3.5.3 参数设定	47
3.5.4 去噪效果比较	49
3.6 小结	53
4 基于图像块稀疏性先验的图像填充算法	55
4.1 引言	55
4.2 算法框架	57
4.3 基于结构稀疏性先验的图像块优先度	58
4.3.1 结构稀疏性	59
4.3.2 图像块优先度	60
4.4 基于稀疏表达的图像块填充	61
4.4.1 图像块稀疏表达	62
4.4.2 优化算法	63
4.5 实验与比较	66
4.5.1 在具有虚幻边缘的生成例子上的比较实验	66
4.5.2 填充算法实验—应用于划痕和文字移除	67
4.5.3 图像块稀疏性对填充性能的提升	68
4.5.4 填充算法实验—应用于物体移除	70
4.5.5 与其它基于稀疏性的图像填充算法比较	71
4.5.6 组合系数求和约束的必要性	73
4.6 小结	73
5 基于闪光先验和MRF先验建模的前景和背景分割算法	75
5.1 引言	75
5.1.1 相关工作	76

5.2 闪光分割算法	76
5.2.1 闪光成像模型	76
5.2.2 分割能量模型	77
5.3 能量模型详细介绍	78
5.3.1 前景能量项	78
5.3.2 运动补偿的背景能量项	79
5.3.3 颜色项	81
5.3.4 模型优化	82
5.3.5 算法流程	83
5.4 实验结果	83
5.5 小结	84
6 结论与展望	87
6.1 本文主要贡献	87
6.2 进一步的研究工作	87
参考文献	89
致 谢	96
攻读博士学位期间的研究成果	97
声 明	

CONTENTS

1	Preface	1
1.1	Research background	1
1.2	Basic problems in image processing	2
1.2.1	Image denoising	2
1.2.2	Image super-resolution	3
1.2.3	Image enhancement	3
1.2.4	Image inpainting	3
1.2.5	Image foreground/background segmentation	4
1.3	Image prior modeling	4
1.3.1	Image prior based on smoothness	6
1.3.2	Image prior based on natural image statistics	7
1.3.3	Image prior based on sparsity of visual coding	8
1.3.4	Image prior related to objects or instruments	9
1.4	The outline of this dissertation	9
2	Image super-resolution and enhancement based on gradient profile prior	11
2.1	Introduction	11
2.1.1	Related work	12
2.2	Natural image gradient profile prior	12
2.2.1	Sharpness estimation of gradient profiles	13
2.2.2	Gradient profile prior	14
2.2.3	Sharpness relationship between high/low resolution images	16
2.3	Image gradient transform	16
2.3.1	Gradient transform	17
2.3.2	Image reconstruction from the transformed gradient	18
2.4	Image super-resolution based on gradient profile prior	19
2.4.1	Super-resolution model	20
2.4.2	Examples and comparisons	22
2.5	Image enhancement based on gradient profile prior	24
2.5.1	Sharpness transform model	25
2.5.2	Sharpness transfer model	27
2.5.3	Examples and comparisons	29

2.6	Brief summary	29
3	Image denoising using anisotropic diffusion with MRF scale selection model.....	31
3.1	Introduction	31
3.2	Anisotropic diffusion scale space	34
3.3	Scale selection model	35
3.3.1	Image over-segmentation	35
3.3.2	Scale selection for each segment	36
3.3.3	Markov random field scale selection model	38
3.4	Noise estimation model based on segmentation	40
3.5	Experiments and comparisons	41
3.5.1	Noise estimation comparison	42
3.5.2	Scale selection comparison	43
3.5.3	Parameter setting	47
3.5.4	Denoising comparison	49
3.6	Brief summary	53
4	Image inpainting based on image patch sparsity	55
4.1	Introduction	55
4.2	The framework of algorithm	57
4.3	Patch priority based on structure sparsity	58
4.3.1	Structure sparsity	59
4.3.2	Patch priority	60
4.4	Patch inpainting based on patch sparse representation	61
4.4.1	Patch sparse representation	62
4.4.2	Optimization algorithm	63
4.5	Experiments and comparisons	66
4.5.1	The experiments on image with illusory boundary	66
4.5.2	Comparison for application of scratch and text removal	67
4.5.3	Effect of patch sparsity for inpainting performance	68
4.5.4	Comparison for application of object removal	70
4.5.5	Comparison with the other sparsity-based inpainting algorithms	71
4.5.6	The necessity of combination coefficients constant constraint	73
4.6	Brief summary	73
5	Image foreground/background segmentation based on flash and MRF modeling.....	75

CONTENTS

5.1 Introduction	75
5.1.1 Related work	76
5.2 Flash cut model	76
5.2.1 Flash imaging model	76
5.2.2 Segmentation model	77
5.3 Items of segmentation model	78
5.3.1 Foreground flash item	78
5.3.2 Motion compensated background item.....	79
5.3.3 Color item	81
5.3.4 Optimization	82
5.3.5 Algorithm framework	83
5.4 Experiments	83
5.5 Brief summary	84
6 Conclusions and Suggestions	87
6.1 Conclusions	87
6.2 Suggestions	87
References	89
Acknowledgements	96
Achievements	97
Declaration	

主要符号表

I	图像
Ω_R	图像区域
Λ	图像函数空间
S_i	图像第 <i>i</i> 个分割块
I_n	噪声图像
I_l	低分辨率图像
I_h	高分辨率图像
I^f	开闪光灯图像
I^{nf}	关闪光灯图像
Ψ_p	以像素 <i>p</i> 为中心的图像块
N_p	像素 <i>p</i> 的邻域
$E(I)$	关于图像 <i>I</i> 的能量函数
$P(I)$	图像 <i>I</i> 的分布
∇I	图像 <i>I</i> 的梯度
$G_\sigma(\cdot)$	核宽度为 <i>σ</i> 的高斯函数

1 绪论

1.1 研究背景

图像是由视觉系统所感知到的记录客观世界中物体、场景或者现象的非结构化信息，它是视觉信息的基本表现形式。随着电子与信息科学技术的快速发展与广泛应用，图像处理技术已经成为人工智能技术、航空航天技术、生物医学技术、军事科学技术、公共安全技术等的重要组成部分。而我们所熟悉的相机、摄像头、游戏产品、医学成像系统等正在或者已经改变了我们的生活方式、工作方式和思维方式。

图像的种类有多种多样。从图像像素特征维数的角度区分，可以分为灰度图像和彩色图像。灰度图像和彩色图像可分别形式化地记为映射 $I : \Omega_R \rightarrow \Gamma^1$ 和 $I : \Omega_R \rightarrow \Gamma^3$ ，其中 Ω_R 为图像域， Γ^1 和 Γ^3 分别为灰度值空间和三维彩色通道值空间（通常取为0到255之间的整数值）；从图像域的维数角度区分，图像又可以分为二维图像（三维世界通过光学系统投影在二维平面上的图像）和三维图像。二维图像和三维图像可分别形式化地记为映射 $I : \Omega_R^2 \rightarrow \Gamma$ 和 $I : \Omega_R^3 \rightarrow \Gamma$ ，其中 Ω_R^2 和 Ω_R^3 分别为二维和三维图像域， Γ 为像素特征空间；从图像的应用背景角度来区分，图像也可以分为自然图像、卫星图像、医学图像等等。本文主要关注二维的灰度图像和彩色图像，所有可能的这类图像构成了图像函数空间 $\Lambda = \{I : \Omega_R^2 \rightarrow \Gamma^1 \text{ 或 } \Gamma^3\}$ 。

视觉系统的功能在于通过处理、分析与识别图像所记录的视觉信息，指导生物体或者人工智能系统对客观世界进行学习与响应，以实现其自身的生存与发展。如图1-1所示，计算机视觉系统通常包括三个层次 [1]：低层视觉（Low level vision）、中层视觉（Middle level vision）和高层视觉（High level vision）。低层视觉的基本任务是进行滤波、增强、恢复等图像预处理和获取图像的边缘（edge）、角点（corner）、色彩、纹理、运动等基本特征；中层视觉的基本任务是基于低层图像特征对图像进行分割或者检测，以获得图像的中层描述；高层视觉的基本任务是获取图像类别、概念等高层知识。

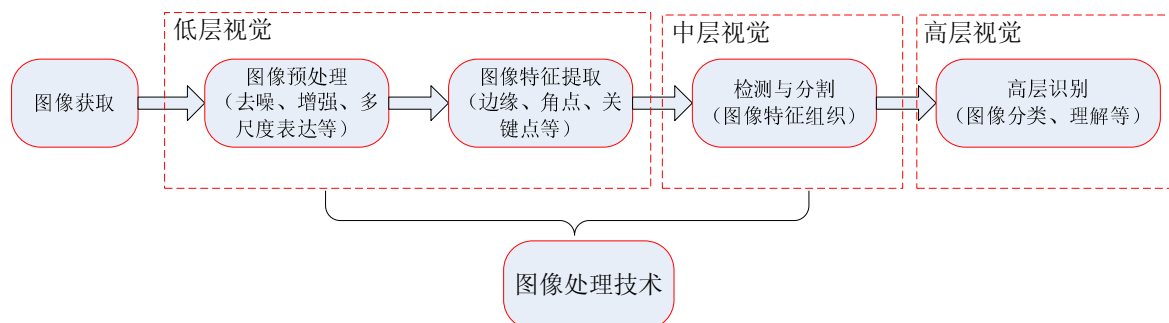


图 1-1 计算机视觉系统的三个层次：低层视觉、中层视觉和高层视觉。图像处理主要关注于解决低层和中层视觉问题。

图像处理是人工智能和模式识别技术的重要分支，它主要关注于如何对图像进行分析、加工和处理，使其满足视觉、心理和模式识别技术的需要。图像处理的主要任

务包括去噪、增强、分割、匹配、运动估计、超分辨率、填充等，相应的问题是图像处理的基本问题（如图1-2所示）。由此可知，图像处理主要关注于解决低层和中层视觉问题。图像处理技术已被广泛地应用于人工智能系统的构建、相机等图像获取设备的设计、医学图像分析、地理信息处理、数字娱乐产品的开发等。

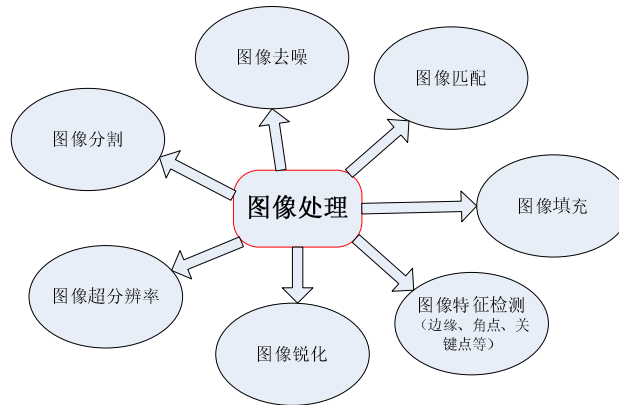


图 1-2 图像处理的基本问题

1.2 图像处理的基本问题

本论文主要关注于解决图像去噪、超分辨率、图像增强、填充和前景/背景分割五个基本问题。因此，有关这些基本问题的相关任务与背景概述如下：

1.2.1 图像去噪

图像去噪的基本任务是：输入噪声图像，去除图像中所包含的噪声信息，以利于图像的进一步分析和处理。在实际应用中，图像的噪声来源于图像采集过程中光学系统或电子系统所带来的各种误差 [2]。为处理方便，人们往往假设噪声为加性噪声，即噪声图像模型为：

$$I_n = I^* + n \tag{1-1}$$

这里， I^* 为原始的未加噪声图像， n 为噪声， I_n 为观测到的噪声图像。 n 往往被假设为均值为0的高斯分布随机变量，其均方差 σ^2 反映了噪声水平。因此图像去噪问题可概



图 1-3 图像噪声模型

述为：给定噪声图像 I_n ，如何设计模型和算法以从噪声图像 I_n 中恢复出原始图像 I^* 。

图像去噪是计算机视觉和图像处理最经典的问题，其在模式识别、高清数字电视设计、相机或摄像机设计等广泛领域有重要应用。

1.2.2 图像超分辨率重建

图像的超分辨率重建问题是指：给定低分辨率图像，如何重建出相应的高分辨率图像。如图1-4所示，给定如左图所示的图像 I_l ，我们希望获得如右图所示的高分辨率图像。很显然，如果直接放大，我们获得的是一个不清晰的图像。因而问题的关键是：如何设计模型和算法，以从观测到的低分辨率图像 I_l 获得清晰的高分辨率图像 I_h 。

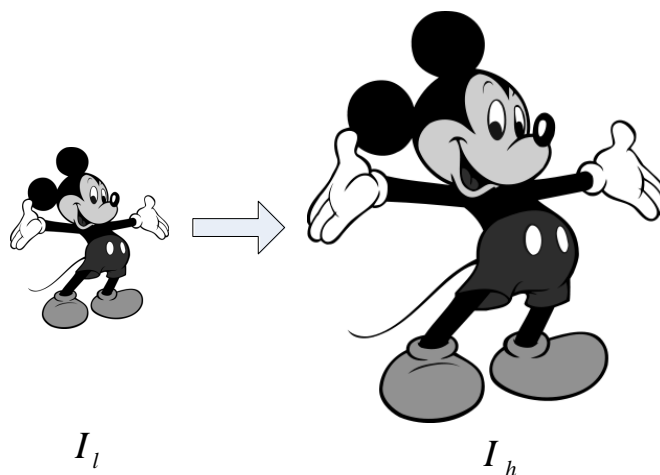


图 1-4 图像超分辨率问题。左图为低分辨率图像，右图为高分辨率图像。

图像超分辨率重建问题是计算机视觉和图像处理重要而基本的问题，其应用包括低分辨图像的放大、图像压缩、高清数字电视显示算法设计、相机数码变焦算法设计等。

1.2.3 图像增强

图像增强的主要任务是将模糊或对比度低的不清晰图像变为清晰图像。最主要的手段是图像锐化，即将模糊图像变成清晰图像。如图1-5所示：左图为由于相机运动或者对焦不准所造成的模糊图像，它丢失了图像的细节信息，图像增强的目的是获得锐化度高的清晰图像（如右图）。

图像增强也是图像处理和计算机视觉中的一个基本问题，其主要目的是通过改善图像的视觉效果，有目的地强调图像的兴趣特性，以满足进一步图像处理、分析和识别的需要。

1.2.4 图像填充

图像填充原本出现于艺术家手工修复艺术作品的污损或划痕部分。Bertalmio等人在文章 [3]中通过计算机模型和算法实现第一个自动图像填充方法，从而成为计算机视觉和图像处理的一个重要技术，其任务是通过数学模型和计算机算法，将图像中的缺

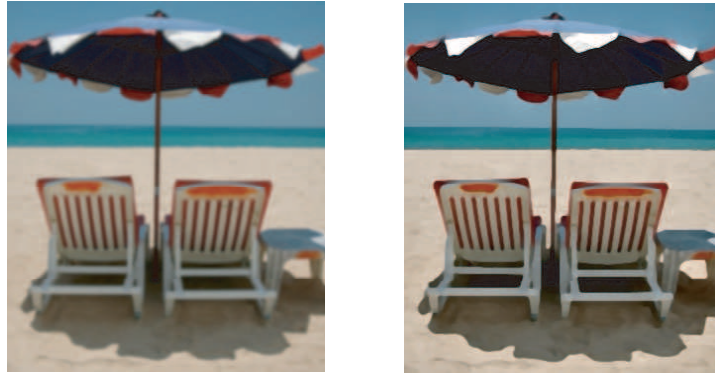


图 1-5 图像增强问题。左图为模糊图像，右图为锐化后的清晰图像。

失部分（由于污损、划痕、图像编辑、文字等造成的缺损）自动填充完整。图像填充问题如图1-6。

图像填充在如下问题中具有重要的应用：旧照片/影片的划痕去除、图像中的文字去除、图像物体移除后缺失区域的填充、图像传输中丢失块的填充等。

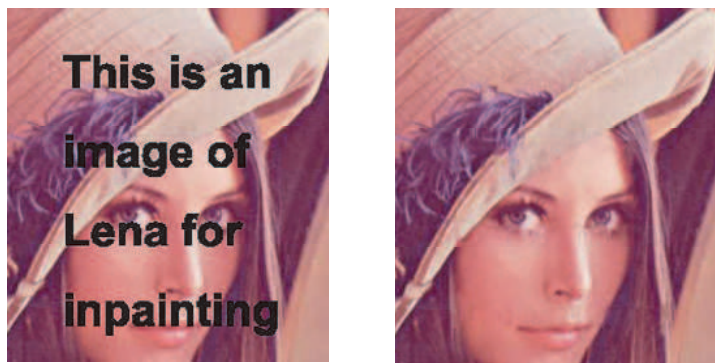


图 1-6 图像填充问题。左图为具有文字的图像，右图为将文字区域内缺失的图像信息填充完整之后的图像。

1.2.5 图像前景/背景分割

前景是指图像拍摄场景中距离相机近的物体（常常为感兴趣物体），而背景则是指处于前景物体后面的场景。前景/背景分割的任务是：将输入图像的前景物体从背景物体中分割出来。更具体地说，假设输入图像为 I ，我们需要将图像像素分为前景/背景两类，因此图像前景/背景分割问题本质上是聚类（无监督分割）或分类（半监督分割）问题。如图1-7显示了前景/背景分割的例子，只展示含有前景的图像分割结果。

图像的前景/背景分割是中层视觉的一个基本问题，其主要应用于物体的识别和编辑方面。

1.3 图像的先验建模

上述图像处理问题很显然都是逆问题，即通过已知的观测结果来推断原始图像或者图像分割结果。结合各自问题的特征，上述基本问题可进一步从逆问题的角度描述如下：



图 1-7 图像前景/背景分割问题。左图为输入图像，右图为前景/背景分割结果（只展示前景内容）。

- 图像去噪问题：通过已知的噪声图像，推测未知的原始图像；
- 图像超分辨率问题：通过已知的低分辨率图像，推测未知的高分辨率图像；
- 图像填充：通过已知的带缺失信息的图像，推测未知的完整图像；
- 图像增强：通过已知的模糊图像，推测未知的高锐化度图像；
- 图像前景/背景分割：通过已知的输入图像，推测未知的图像分割结果。

逆问题通常是病态的。根据解的适定性定义（存在性、唯一性和稳定性），图像处理中的逆问题其解往往不具有唯一性和稳定性。例如对于图像去噪问题：给定噪声图像 I_n ，满足噪声图像模型(1-1)的图像 I^* 的数目是无穷的，因此，如何约束图像处理问题的解空间以得到满足视觉感知的图像是图像处理算法研究的核心。

图像处理具有如下特点：所处理的信息为客观世界的映像，而我们所存在的客观世界是具有规律性的，因此自然界中的图像（即自然图像）也具有规律性。我们将自然图像特征（频率特征、空间组织特征或者物体类别、拍摄设备相关的特征）的经验性规律统称为**自然图像的先验信息**。自然图像先验很好地模拟了大脑系统的记忆与归纳能力，即生物体对视觉信息的处理与识别不仅依赖于视觉信息本身的特征，更依赖于大脑所归纳与记忆的有关视觉信息的一般性规律。基于这样的原理，我们以自然图像的先验信息作为指导或约束，可以很好地解决图像处理的逆问题（如图1-8）。

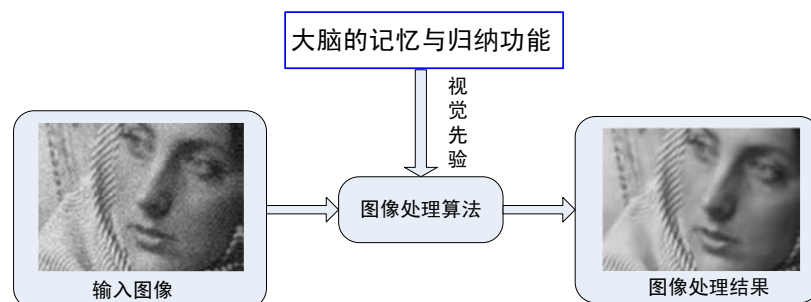


图 1-8 基于视觉先验的图像处理。通过模拟大脑的记忆与归纳功能，为图像处理算法提供视觉先验信息，以获得满足视觉感知的图像处理结果。

近年来，人们逐渐认识到自然图像先验在计算机视觉和图像处理中的重要意义，因此日益关注如何挖掘和建模自然图像的先验信息，以解决图像处理和计算机视觉中的基本问题。以下我们将简述已有的有关自然图像先验信息的研究进展情况。

1.3.1 基于光滑性的先验

光滑性先验试图描述自然图像的光滑性规律：即自然图像大部分为低频特征，而高频特征非常稀少。例如，如图1-9所示，自然图像梯度模只在图像结构（即边缘）上非零，而在大部分区域为零。对自然图像的光滑性先验进行建模可实现图像逆问题的正则化处理。

实现对自然图像光滑性建模的最基本方式是约束图像 I 的一阶或高阶导数及其非线性变换，以形成图像处理的正则化能量模型。记 $S(I)$ 为约束图像光滑性的光滑项，则基于变分方法，优化该光滑项的梯度流为：

$$\frac{\partial I}{\partial t} = -\frac{\partial S(I)}{\partial I}. \quad (1-2)$$

通过采用不同的图像导数及其非线性形式，可以导出不同形式的光滑性约束，从而构成基于变分方法的偏微分方程图像处理算法。例如，当光滑性约束取为梯度模的平方（即 $|\nabla I|^2$ ）时，其梯度流为：

$$\frac{\partial I}{\partial t} = \Delta I, \quad (1-3)$$

这导出经典的热扩散去噪算法 [4]；当光滑性约束取为梯度模（即 $|\nabla I|$ ）时，导出经典的全变差（total variational）去噪算法 [5, 6]，即

$$\frac{\partial I}{\partial t} = \operatorname{div}\left(\frac{\nabla I}{|\nabla I|}\right). \quad (1-4)$$

当光滑性约束取为梯度的非线性函数（如 $g(|\nabla I|)$ ）时，导出各向异性扩散算法 [7, 8]，即

$$\frac{\partial I}{\partial t} = \operatorname{div}\left(g'(|\nabla I|)\frac{\nabla I}{|\nabla I|}\right). \quad (1-5)$$

当光滑性约束为高阶导数模时，导出高阶各向异性扩散去噪算法 [9] 结合应用问题背景，另有其他多种光滑性约束定义，例如参见 [10]。

另一类光滑性先验建模方式是基于Markov随机场（Markov Random Field, MRF）的先验建模。Markov随机场是建模具有空间依赖关系的随机变量场概率分布的一种有效数学方法，依据Markov随机场可以用Gibbs分布来描述的原理，通常人们采用Gibbs分布来描述自然图像的光滑性，例如：

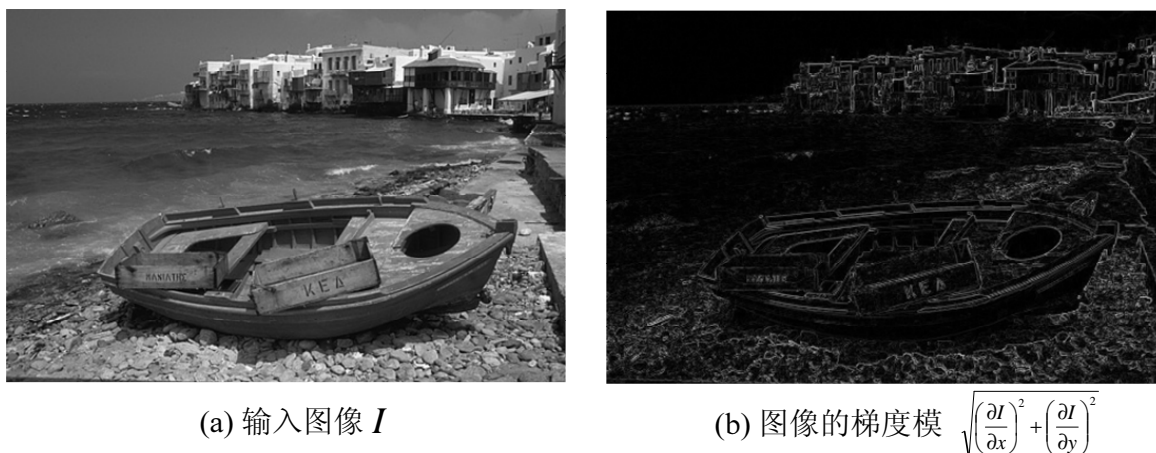


图 1-9 图像一阶导数（即梯度）的光滑性。左图为输入图像，右图为其梯度模图像，越亮代表梯度模值越大。显然，图像的梯度模值大部分都很小，只有位于图像边缘的像素具有较大的梯度模值。

$$P(I) = \frac{1}{Z} \exp(\beta_T \sum_{i,j \in N(i)} |I_i - I_j|^2) \quad (1-6)$$

建模了图像相邻像素点之间颜色的相似性，其中 $N(i)$ 为像素 i 的邻域， Z 为常数使得 $P(I)$ 的积分为1， β_T 为物体学中的温度常数。

基于光滑性先验的图像处理方法已经被广泛地应用于图像去噪、分割等基本问题，是计算机视觉和图像处理领域广泛研究与应用的重要方法。

1.3.2 基于自然图像统计规律的先验

自然图像统计先验一般是指图像滤波响应的统计分布规律。图像滤波是通过对图像进行卷积操作以提取图像的高频、低频或某一频率范围内的特征信息的过程。按其上述目的，图像滤波分别称为是高通滤波、低通滤波和带通滤波，卷积核被称为是滤波器（filter），而滤波的结果称为是滤波响应（filter response）。图1-10显示了自然图像的滤波响应及其统计分布图。

Huang等人 [11]系统研究了自然图像的导数滤波（derivative filter）响应和小波系数的统计分布，提出了用 t -分布或者指数类分布来进行统计建模，并通过自然图像集学习统计分布的参数，从而得到自然图像滤波响应的一般性统计描述。该统计描述已被应用于解决图像去模糊 [12]、去噪 [13, 14]等问题。

人们进一步提出了基于Markov随机场理论，特别是通过Gibbs分布来建模自然图像滤波响应统计规律的方法。Zhu等人 [15, 16]提出采用Gibbs分布同时建模多个滤波器的滤波响应分布，并将该模型应用于图像纹理建模。Hinton [42]和Roth等人 [17]提出了Product-of-Experts和Fields-of-Experts模型，通过Gibbs分布同时建模滤波器组（experts）和它们的滤波响应分布。这些方法被成功地应用于图像去噪和填充问题，取得了非常好的图像处理效果。

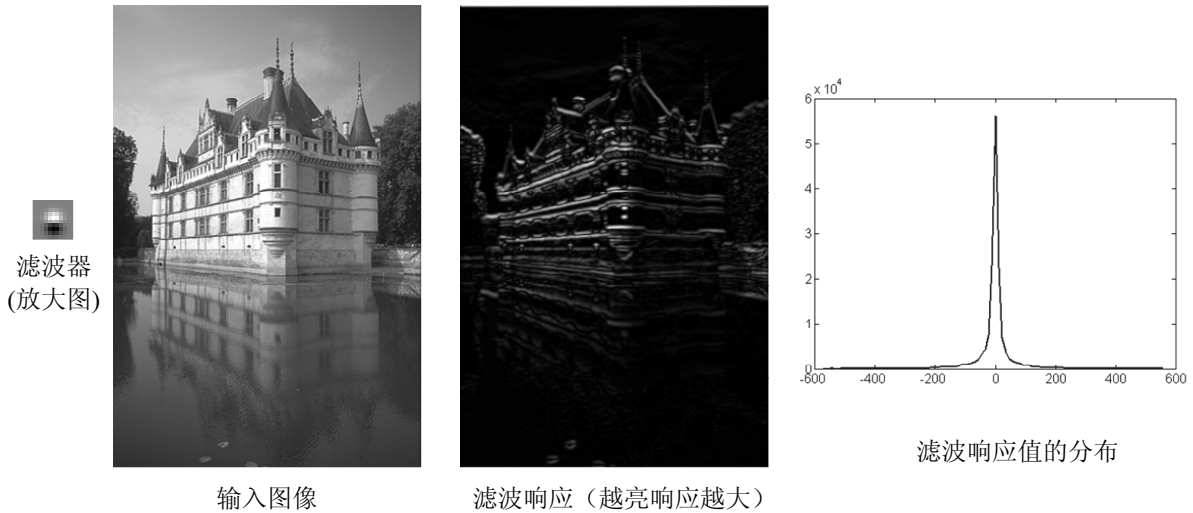


图 1-10 图像滤波与滤波响应分布。从左到右分别为滤波器，输入图像，滤波响应（越亮代表响应值越大）和滤波响应分布图。滤波响应分布图横坐标为滤波响应值，纵坐标为频数。自然图像统计即对滤波响应的分布进行统计建模（例如，可用 t -分布函数进行建模）。

不同于光滑性先验（即约束图像的局部光滑性），自然图像统计先验直接对图像滤波响应的分布进行统计建模。例如，同样建模相邻像素颜色差，不同于基于Markov随机场的光滑性先验（如公式(1-6)），自然图像统计先验关注于如何选择更好的Gibbs分布形式对像素颜色差的分布进行建模。因此，基于自然图像统计先验的Gibbs分布模型可一般地描述为：

$$P(I) = \frac{1}{Z} \exp(\beta_T \sum_{i,j \in N(i)} \phi(|I_i - I_j|)) \quad (1-7)$$

可通过自然图像训练集学习参数 β_T 和势函数 ϕ 。

1.3.3 基于视觉编码稀疏性的先验

视觉信息处理流程可概括为：视觉信息（即自然图像）通过视网膜进入初级视觉皮层进行特征提取（边缘、关键点等特征），进一步通过高级视觉皮层实现视觉信息的理解。视觉皮层的神经细胞是通过其响应模式对视觉信息进行编码的，而该响应模式具有稀疏性：即当视觉细胞群对视觉信息进行响应时，仅有极少（即稀疏）个数的神经元具有强响应，大部分细胞响应很弱或者不响应。这样的稀疏性规律被称为是视觉神经编码的稀疏性先验。

基于视觉编码的稀疏性先验，已有大量的研究工作关注于如何对视觉编码稀疏性进行建模以挖掘自然图像的本质表达。典型的工作例如包括稀疏编码（sparse coding）[18–20][21]和压缩传感（compressive sensing）[22, 23]。稀疏编码的基本思想是将视觉信息用过完全的基（basis）集合进行稀疏线性表达，这些基可以为图像的变换域[19]或图像块集合[24]等基本表达单元。此时图像的能量模型可表达为：

$$E(I) = \|I - \sum_k \beta_k b_k\|^2 + \|\vec{\beta}\|_{\ell_p} \quad (1-8)$$

这里 $B = \{b_k\}$ 为过完全基集合, $\vec{\beta} = \{\beta_k\}$ 为表达系数, 第二项为稀疏性惩罚项 ($0 \leq p \leq 1$), 即要求系数 $\vec{\beta}$ 只有少数非零。稀疏编码已经被广泛应用于图像去噪 [24]、边缘检测 [20] 等基本问题。压缩传感是近年来广受关注的研究领域, 其基本思想则是通过压缩传感矩阵将图像或者信号进行高强度的压缩, 并可以由压缩信息重建图像和信号。压缩传感已经被广泛应用于图像压缩 [22]、超分辨率 [25] 等问题。

1.3.4 基于对象或设备的先验

另外一类重要的先验信息是对象 (物体) 以及获取图像的物理设备相关的知识, 该类先验针对具体的问题背景, 利用所关注物体的规律性或者图像拍摄设备的特性来作为先验信息, 以帮助解决困难的图像处理或识别问题。

对象相关的先验信息 [26, 27] 运用所关注物体类别的规律性 (包括形状、空间位置等) 指导图像处理和分析, 例如人脸具有很强的规律性: 由眼睛、鼻子等不同部分以特定的空间关系组织在一起; 而其它类别的物体, 例如汽车、建筑物等也有其自身的规律性, 应用这些先验信息, 可以更好地解决图像处理和识别的基本问题。

设备相关的先验信息 [28, 29] 通过灵活运用相机、摄像机等设备的自身功能或特性得到具有特殊效果的图像或者多张不同拍摄方式的图像, 并应用数学方法分析和建模图像中所蕴含的先验信息, 进而帮助解决图像处理的逆问题。由于设备相关的先验建模可以非常有效地解决图像处理或计算机视觉中的逆问题, 因此相关研究备受关注, 人们特别希望通过此方面研究指导或革新相机等图像采集设备的功能设计。现有的工作包括利用闪光设备相关的先验信息 [28–32] 和相机光圈设备相关的先验信息 [12, 33] 等建模图像, 这些方法通过建模闪光灯所带来的光照变化和不同大小光圈所带来的模糊变化以实现图像分割、增强等功能。

1.4 本文的主要工作

本论文集集中于图像处理基本问题的建模与算法研究。所涉及的基本问题主要包括: 图像去噪、图像增强、超分辨率、填充和前景/背景分割等。研究目标是系统地提出解决这些基本问题的数学模型及其算法, 以发展出当前最好的图像处理方法与技术。

本论文的整体研究思路是: 针对基本问题的特征与背景, 以统计方法、变分法和正则化方法为基本手段, 发掘和建模自然图像的视觉先验信息, 并设计出解决图像处理基本问题的有效算法。所取得的主要成果包括:

(1) 提出了基于自然图像梯度轮廓先验的图像超分辨率重建与增强算法。通过挖掘自然图像的梯度场沿图像边缘垂直方向的轮廓分布, 提出了一种新的自然图像统计先验—梯度轮廓先验 (Gradient Profile Prior, GPP)。梯度场轮廓先验参数化地描述了自然图像梯度场轮廓的形状和锐化度分布。基于该先验, 通过控制梯度场轮廓的形状和锐化度参数, 提出一种新的梯度场变换, 并进而发展出了非常有效的图像超分辨率重建和增强算法。实验表明, 新算法处理效果显著优于当前的同类算法。

(2) 提出了基于MRF尺度选择的各向异性扩散去噪算法。通过研究各向异性扩散去噪算法的尺度 (即扩散停机时间) 选择问题, 提出了基于MRF光滑性先验建模的空间自适应尺度选择模型, 该模型很好地克服了传统尺度选择策略为全局选择同一个尺

度的缺陷，尺度选择性能远远优于传统尺度选择策略；结合该尺度选择模型，提出了新的基于各向异性扩散滤波的去噪算法。应用表明，其性能与当前最好的去噪算法相当甚至更好。

(3) 提出了基于图像块稀疏性先验的图像填充算法。通过引进图像块稀疏性概念（包括结构稀疏性和图像块表达稀疏性），提出了基于例子的新型图像填充算法。我们证明，图像块的结构稀疏性很好地度量了图像块是否处于图像的结构（即边缘）之中，图像块的结构稀疏性越大，它越可能处于图像边缘；图像块稀疏表达则基于视觉编码的稀疏性先验，假设待填充图像块可以由候选图像块集合的稀疏线性组合来表达，进而发展出了图像稀疏线性表达的优化模型和优化算法。图像块结构稀疏性和图像块稀疏表达很好地解决了基于例子的图像填充算法的两个基本问题：图像块选择和图像块填充。所发展的新的填充算法相比以往的填充算法具有更好的填充性能。

(4) 提出了基于闪光先验信息和MRF光滑性先验建模的前景/背景提取算法。以开/关相机闪光灯设备所提供的图像对为输入，通过数学模型建模光照变化的先验信息，提出了基于MRF光滑性先验的图像分割能量模型，很好地解决了图像前景提取问题。与已有的前景提取算法相比，所提出的算法可以得到更为精确的分割结果。该模型的重要意义在于通过相机的闪光设备，提供了一种便捷、实用的前景提取模型和算法，期待可作为新一代相机的基本功能之一。

本论文的结构安排如下：第2章介绍基于自然图像梯度轮廓先验的图像超分辨率重建与增强算法；第3章介绍基于MRF尺度选择的各向异性扩散去噪算法；第4章介绍基于图像块稀疏性先验的图像填充算法；第5章介绍基于闪光先验信息和MRF先验建模的前景/背景分割算法；第6章总结全文并提出进一步值得研究的问题。

2 基于自然图像梯度轮廓先验的图像超分辨率重建与增强算法

2.1 引言

自然图像的统计先验建模是计算机视觉中广受关注的研究方向 [8, 15, 17, 34]。图像统计先验应用统计学方法对自然图像的一般性规律进行建模，广泛应用于图像超分辨率重建 [35–38]、去噪 [13, 14, 17]、填充 [39]、去模糊 [12, 40]、透明图层分割 [41] 等等。由于图像处理问题的病态性，自然图像先验可以很好地正则化图像处理模型的解空间，以得到合乎自然图像统计规律的图像处理结果。

自然图像先验通常建模自然图像滤波响应的分布规律，被称为是一般性图像先验 (**generic image prior**)。用于建模自然图像先验的分布包括Laplacian分布 [11]、广义高斯分布 [11]、混合高斯分布 [34] 等，它们建模自然图像高频滤波响应的重尾分布；另一方面，Product-of-Experts模型 [42] 和Field-of-Experts模型 [17] 等通过Markov随机场对滤波器组和滤波响应分布同时进行建模和学习。

传统的图像统计先验仅关注图像滤波响应在全图范围内的边际分布 (**marginal distribution**)，即直方图分布，而忽略滤波响应的空间分布规律。然而，滤波响应大的梯度沿着图像边缘的空间分布形状常呈现一定的规律性。因此，本章将挖掘和建模自然图像梯度场沿图像边缘垂直方向的轮廓 (定义为梯度轮廓) 分布，并以此为基础提出一种新的一般性图像统计先验—梯度轮廓先验 (**Gradient Profile Prior**)。梯度轮廓先验将参数化地描述自然图像梯度轮廓的形状和锐化度。我们发现，该先验的形状参数非常稳定于图像的分辨率。基于梯度轮廓先验，我们将通过控制梯度轮廓的形状和锐化度，提出一种梯度场变换以将原始图像梯度场变换为目标图像梯度场。因为形状先验参数稳定于图像分辨率，我们只需根据不同的问题背景发掘原始图像和目标图像之间的锐化度关系。

将梯度轮廓先验应用于单张图像的超分辨率重建和图像增强。单张图像的超分辨率重建问题是由输入的低分辨率图像重建出高分辨率图像。该问题是个逆问题：即给定低分辨率图像，其对应的未知高分辨图像是不唯一的。因此，为提出一个好的图像超分辨率算法，必须设计好的图像先验以约束高分辨率图像的空间，使其满足自然图像的特征分布。我们从高/低分辨率图像训练集中学习高/低分辨率图像之间的锐化度关系，进而通过梯度场变换由低分辨率图像的梯度场推理出高分辨率图像的梯度场。结合梯度场约束和图像重建约束，可以很好地估计出高分辨率图像，且结果中很少引入锯齿 (**jaggy**，即图像边缘不平滑，含有锯齿) 或者环晕 (**ringing**，即图像边缘周围具有晕状效应) 等伪影效应 (**artifacts**，即由算法引起而非自然图像本身含有的结构或特征)。图像增强问题是将输入的模糊图像增强为高锐化度图像。我们定义了一种锐化度映射函数以建模模糊图像的梯度轮廓锐化度和高锐化度图像的梯度轮廓锐化度之间的关系，进而通过梯度场变换由模糊图像的梯度场推理出高锐化度图像的梯度场。高锐化度图像可以通过变分方法从估计的高锐化度图像梯度场重建出来。

相对于传统的图像先验，梯度轮廓先验具有如下的优势：不同于光滑性先验，梯度轮廓先验并不约束自然图像梯度场尽量接近于零，因此小尺度和大尺度的图像细节

信息均可以被重建；通过在梯度场进行变换和重建操作，可以避免以往超分辨率算法或增强算法带来的伪影效应。

2.1.1 有关图像超分辨率和增强的已有工作

目前主要有三种单张图像超分辨率算法：基于差值的算法、基于重建的算法和基于学习的算法。基于差值的算法通过低分辨率图像的已知点进行插值得到高分辨图像中的未知点，包括最近邻（nearest neighbor）插值算法、双线性（bilinear）或者双立方（bicubic）插值算法等。基于插值的算法 [43, 44, 46] 实现简单，但容易模糊化高频信息或者在图像边缘引入锯齿效应。基于重建的算法 [47–50] 通过极小化重建约束：

$$\|I_l - \{I_h * G_\sigma\} \downarrow\|^2 \quad (2-1)$$

来实现超分辨率重建，即要求未知的高分辨率图像（ I_h ）进行高斯卷积且进行下采样后与已知的低分辨率图像（ I_l ）尽可能相近。Back-projection [47] 算法是一种典型的基于重建的算法。如果没有先验进行约束，该优化问题的解会在图像边缘引入锯齿或者环晕效应。因此需要设计好的图像先验对模型进行正则化。基于学习的算法 [25, 35, 36, 48, 51–56] 构建高/低分辨率图像块对的训练集合，那么已知低分辨率图像块对应的高分辨率图像块的信息可以通过训练集合推理出来。论文 [36] 提出只对图像初始简图（Primal sketches，例如边缘、角点等）进行特别的关注，而其他区域通过插值方法进行超分辨率重建，这是因为人的视觉系统更加关注于图像的边缘等特征。基于学习的算法特别依赖于训练集和测试集合的相似性，而且不清楚多大规模的训练集合是足够的。

Unsharp masking (UM) 算法是最基本的图像增强算法 [57]。在该方法中，图像首先与其低通滤波后的模糊化图像进行相减得到高频信号，然后将部分（以百分比记）高频信号加到原始图像，从而实现图像锐化。该方法可以很好地增强图像的锐化度和对比度，但会在图像边缘引入晕状效应。Shock 滤波 [58, 59] 是另一类重要的图像增强算法，该方法通过求解逆扩散方程实现对图像细节信息的增强。Osher 和 Rudin [60] 提出如下方程增强模糊图像 I_0 ：

$$\frac{\partial I(\cdot, t)}{\partial t} = -\text{sign}(\Delta I) |DI|, \quad I(\cdot, 0) = I_0(\cdot). \quad (2-2)$$

其中， $\text{sign}(\Delta I)$ 可视为边缘检测器：当 $\Delta I < 0$ 时，该方程退化为 $I_t = |DI|$ ，等价于膨胀形态学操作；当 $\Delta I > 0$ 时，该方程退化为 $I_t = -|DI|$ ，等价于腐蚀形态学操作。随着方程的演化，增强图像倾向于演化为分片常数图像。论文 [59, 61, 62] 中提出了基于更多类型边缘检测器的微分方程方法。

本章将提出基于梯度轮廓先验的一类新的图像超分辨率算法和增强算法，以下第 2.2 节首先定义梯度轮廓先验并进行参数学习；第 2.3 节提出基于梯度轮廓先验的梯度场变换；第 2.4 节和 2.5 节分别将梯度轮廓先验应用于图像超分辨率问题和增强问题，提出新的算法并给出实验结果；最后在第 2.6 节小结本章工作。

2.2 自然图像梯度轮廓先验

本节将研究自然图像梯度场垂直于图像边缘的轮廓的统计规律，我们首先给出自然图像梯度轮廓及其先验的定义。

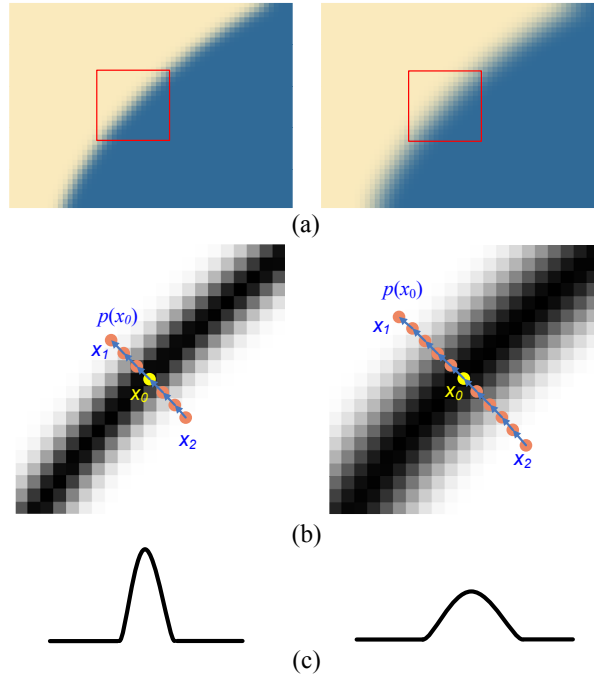


图 2-1 梯度轮廓。(a) 锐化度不同的两条边缘图像。(b) 梯度模图(亮度越小, 梯度模越大)。由边缘点 x_0 出发, 沿着梯度方向或负梯度方向逐点延长直到梯度模不再降低, 从而得到轮廓 $p(x_0)$ 。(c) 轮廓 $p(x_0)$ 上的梯度模曲线, 即定义为梯度轮廓。

记图像梯度 $\nabla I = m \cdot \vec{N}$, 其中 m 为梯度模值, 即

$$m = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2}, \quad (2-3)$$

\vec{N} 为梯度方向。梯度模的局部极大值点称为是图像边缘点 (edge pixel)。图2-1(a)显示了两条具有不同锐化程度的边缘图像。图2-1(b)显示了相应的梯度模值图。图像2-1(b)中像素 x_0 为图像边缘点, 从 x_0 出发, 沿着梯度方向和负梯度方向逐点延长直到梯度模不再降低, 可得到一条一维的路径, 我们称梯度模值在该路径上的分布为**梯度轮廓**。图2-1(b)(c)显示了经过 x_0 的梯度轮廓。

2.2.1 梯度轮廓锐化度估计

我们用如下的方法来度量梯度轮廓的锐化度:

$$\sigma(p(x_0)) = \sqrt{\sum_{x \in p(x_0)} \frac{m(x)}{M(x_0)} d^2(x, x_0)}, \quad (2-4)$$

其中,

$$M(x_0) = \sum_{s \in p(x_0)} m(s), \quad (2-5)$$

$d(x, x_0)$ 表示像素点 x 和 x_0 之间沿着梯度轮廓的曲线长度。注意到, 图像边缘锐化度越高, 标准差 σ 越小, 我们将 σ 定义为**梯度轮廓的锐化度**。

独立地为每条梯度轮廓估计锐化度显然对噪声是不鲁棒的，为了更好地估计图像边缘的锐化度，我们假设沿着同一条边缘相邻的梯度轮廓的锐化度具有一致性，即假设锐化度具有光滑性。基于该假设，我们构建一个图 G ，其中每个图节点为图像边缘点，位于同一条图像边缘且距离小于给定距离（例如可设为5）的两个边缘点（即图节点） i, j 之间具有连接边，边的权重 w_{ij} 定义为：

$$w_{i,j} = \exp(-\zeta_1 \cdot |\nabla u_i - \nabla u_j|^2 - \zeta_2 \cdot d(i, j)^2), \quad (2-6)$$

其中，指数中的第一项为梯度相似性，第二项为 i, j 之间的欧氏距离。对于每个边缘点 i ，我们使用公式(2-4)估计其锐化度，然后通过极小化如下能量函数估计结合光滑性先验的图像锐化度：

$$E(\{\sigma_i\}) = \sum_i [(\sigma_i - \hat{\sigma}_i)^2 + \gamma \sum_{j \in N(i) \cap p(i)} w_{i,j} \cdot (\sigma_i - \sigma_j)^2], \quad (2-7)$$

这里， $N(i)$ 为节点 i 的邻接点， $p(i)$ 为经过 i 点的梯度轮廓。我们通过梯度下降法优化该能量函数，由于能量函数(2-7)具有凸性，因此容易求得其全局最优解。在优化求解这一能量函数时，可取初始锐化度为每个梯度轮廓独立估计的锐化度，在实现中可设 $\gamma = 5$ ， $\zeta_1 = 0.15$ ， $\zeta_2 = 0.08$ 。

2.2.2 自然图像梯度轮廓先验

现在我们通过统计方法来建模自然图像梯度轮廓的规律性。如图2-1(c)所示，自然图像的梯度轮廓具有类高斯分布形式，因此我们可用最一般的指数族分布—广义高斯分布对自然图像梯度轮廓来进行建模：

$$g(x; \sigma, \lambda) = \frac{\lambda \alpha(\lambda)}{2\sigma \Gamma(\frac{1}{\lambda})} \exp\{-[\alpha(\lambda) |\frac{x}{\sigma}|]^\lambda\}, \quad (2-8)$$

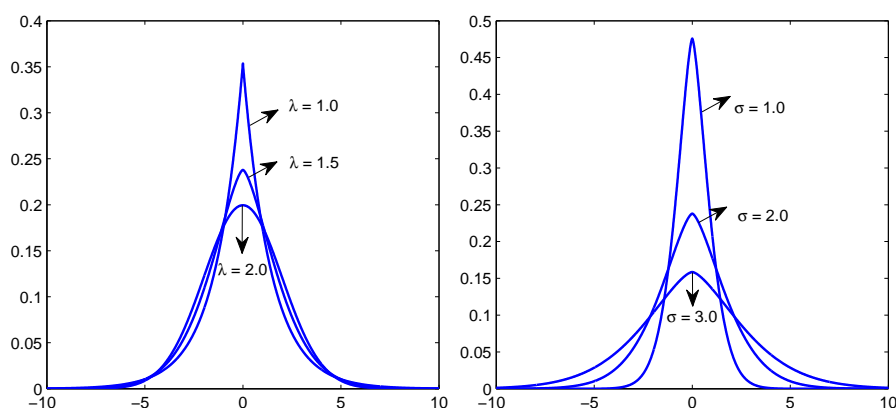


图 2-2 广义高斯分布在不同参数下的分布函数形状。(a) $\sigma = 2.0$ ， λ 取为 1.0, 1.5, 2.0。(b) $\lambda = 1.5$ ， σ 取为 1.0, 2.0 和 3.0。参数 σ 控制分布的整体形状。 $\lambda = 2$ 时，该分布为高斯分布；当 $\lambda = 1$ 时，该分布为 Laplacian 分布。参数 σ 为分布的标准差。

其中, $\Gamma(\cdot)$ 为gamma函数; $\alpha(\lambda) = \sqrt{\Gamma(\frac{3}{\lambda})/\Gamma(\frac{1}{\lambda})}$ 为尺度因子, 目的是使得广义高斯分布的二阶矩等于 σ^2 , 其中 σ 可以通过2.2.1节中所介绍的方法来估计, λ 为控制该分布整体形状的形状参数。显然, 当 $\lambda = 2$ 时, 该分布为高斯分布; 当 $\lambda = 1$ 时, 该分布为Laplacian分布。图2-2显示了广义高斯分布在不同参数下的分布函数形状。

下面我们说明如何通过学习来确定自然图像梯度轮廓先验分布的参数, 以反映自然图像梯度特征的一般性规律。我们从专业的图像论坛随机下载了1000幅自然图像, 所有图像均处于原始分辨率。我们从每一幅图像中随机收集了1000条梯度轮廓以构成集合 Ω_1 (共1百万个梯度轮廓)。我们还通过将原始图像缩小 2×2 倍、 3×3 倍和 4×4 倍, 并分别采集1百万个梯度轮廓构成三个梯度轮廓集合 Ω_2 、 Ω_3 、 Ω_4 。显然, 集合 $\{\Omega_i\}_{i=1}^4$ 分别来自于同一图像集在不同分辨率下的图像集合。下面我们将通过这4个梯度轮廓集合研究梯度轮廓先验的形状参数, 以及不同分辨率下的锐化度参数关系。

为建模自然图像梯度轮廓, 我们通过所采集的梯度轮廓集合拟合先验分布公式(2-8), 以优化其中的形状参数。这样所得到的先验分布及其参数可以很好地描述自然图像梯度轮廓的分布, 即实现了对自然图像梯度特征的建模。我们用Kullback-Leibler (KL)距离 [45]度量拟合误差, 因此最优形状参数 λ^* 满足:

$$\lambda^* = \operatorname{argmin}_{\lambda} \left\{ \frac{1}{|\Omega|} \sum_{p \in \Omega} \operatorname{KL}(p, g(\cdot; \sigma_p, \lambda)) \right\}, \quad (2-9)$$

其中, σ_p 为梯度轮廓集合中的轮廓 p 的锐化度参数 (如2.2.1节中求得)。Kullback-Leibler (KL)距离定义为:

$$\operatorname{KL}(p, g(\cdot; \sigma_p, \lambda)) = \sum_{x \in p} \frac{m(x)}{M(p)} \log \frac{m(x)/M(p)}{g(x; \sigma_p, \lambda) / \sum_{s \in p} g(s; \sigma_p, \lambda)}, \quad (2-10)$$

其中, x 和 s 为梯度轮廓 p 上的像素点。我们分别在4个梯度轮廓集合上计算了不同形状参数下的拟合误差, 如图2-3所示。我们发现, 最优形状参数为1.6, 且稳定于图像分辨率。这意味着自然图像的梯度轮廓形状参数与图像分辨率无关或者关系不大。

我们使用Pearson- χ^2 假设检验度量梯度轮廓先验对自然图像梯度轮廓的拟合程度。对梯度轮廓 $p(x_0)$ 的 χ^2 假设检验定义为

$$\chi^2(p) = \sum_{x \in p(x_0)} \frac{[m(x) - E(x)]^2}{E(x)}, \quad (2-11)$$

其中,

$$E(x) = \frac{g(d(x, x_0))}{\sum_{s \in p(x_0)} g(d(s, x_0))} \cdot \sum_{s \in p(x_0)} m(s). \quad (2-12)$$

对于显著性水平 κ 和自由度 $n - 1$ (n 为梯度轮廓 p 上的像素个数), 如果 $\chi^2(p) < \chi_{(\kappa, n-1)}^2$, 那么梯度轮廓 $p(x_0)$ 满足梯度轮廓先验的假设不能被拒绝, 即意味着梯度轮廓先验可以很好地描述梯度轮廓 $p(x_0)$ 。我们取显著性水平 $\kappa = 0.01$, 那么对于4个梯度轮廓集合, 梯度轮廓拟合的 χ^2 值与 $\chi_{(\kappa, n-1)}^2$ 的差值的平均值分别为 -2.22, -1.90, -1.50, -1.20, 这意味着整体上自然图像的梯度轮廓能很好地由梯度轮廓先验描述。

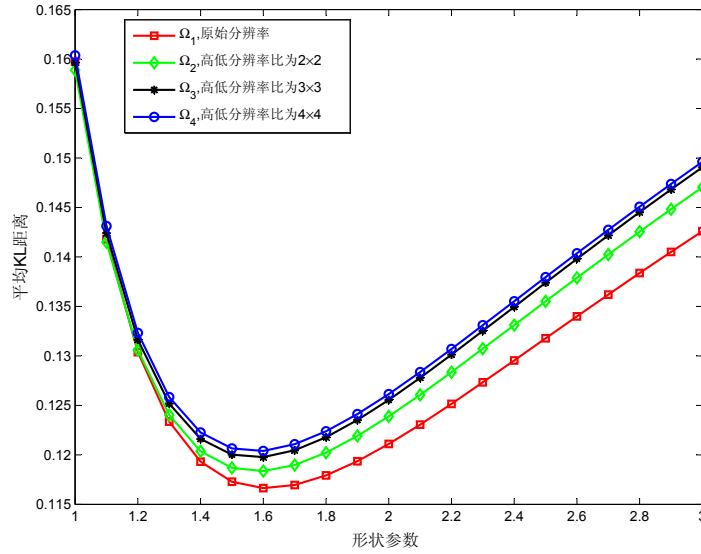


图 2-3 在不同形状参数 λ 下，用广义高斯先验分布拟合图像梯度轮廓的平均KL误差。在4个集合上，最优的 λ 均为1.6。

2.2.3 高/低分辨率图像梯度轮廓的锐化度关系

现在我们研究高分辨率图像与低分辨率图像之间的梯度轮廓锐化度关系。类似于[36, 51, 54]中的做法，我们研究低分辨率图像 I_l 的上采样图像（分辨率与高分辨率图像一致） I_l^u 与高分辨率图像 I_h 的梯度轮廓锐化度关系。这样做的目的是使得高、低分辨率图像的边缘能够更好地对齐，进而方便研究它们之前的锐化度关系。在算法实现中，我们用双立方插值对低分辨率图像进行上采样操作。

对于 I_l^u 中的每个梯度轮廓，我们获得它在高分辨率图像 I_h 中相匹配的梯度轮廓。因为两图像中的边缘像素点不一定能完全对齐，我们通过测度边缘像素点之间的距离和方向差来选择最匹配的边缘像素点。因此，给定图像 I_l^u 中的边缘点 e_l ，在图像 I_h 与其最匹配的边缘点 e_h 为：

$$e_h = \operatorname{argmin}_{e \in \mathcal{N}(e_l)} \{ \|e - e_l\| + 2\|\vec{N}(e) - \vec{N}(e_l)\|\}, \quad (2-13)$$

其中， $\mathcal{N}(e_l)$ 为高分辨率图中 e_l 的 5×5 近邻点。

为了估计高/低分辨率图像锐化度关系，我们将锐化度 σ 离散化。我们取离散化间隔为0.1，则 σ 被离散化为： $\sigma = 0.0, 0.1, 0.2, \dots$ 。我们统计 I_l^u 中所有锐化度为 σ_l ($\sigma_l \in \{0.0, 0.1, 0.2, \dots\}$)的梯度轮廓所对应的图像 I_h 中的梯度轮廓的平均锐化度。图2-4显示了在高/低分辨率倍数为 2×2 、 3×3 和 4×4 情况下，高/低分辨率图像锐化度对应关系。X-坐标轴显示上采样的低分辨率图像梯度轮廓锐化度，Y-坐标轴显示相匹配的高分辨率梯度轮廓的平均锐化度。

通过以上方法，我们概括出了自然图像梯度场沿着图像边缘梯度轮廓的一般分布规律（梯度轮廓模型(2-8)及其最优形状参数和不同尺度下的锐化度关系），这一规律称为**梯度轮廓先验**。

2.3 自然图像梯度场变换

应用梯度轮廓先验，我们在本小节提出梯度轮廓变换，目的是通过变换梯度轮廓

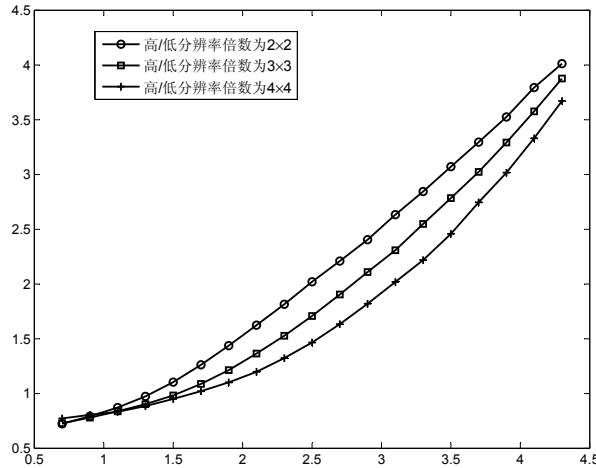


图 2-4 高/低分辨率图像梯度轮廓的锐化度关系。X-坐标轴显示上采样的低分辨率图像梯度轮廓锐化度，Y-坐标轴显示相匹配的高分辨率梯度轮廓的平均锐化度。

的形状和锐化度将原始梯度场变换为目标梯度场。我们也说明如何从梯度场恢复重建出图像。

2.3.1 梯度场变换

首先，我们考察如何将原始梯度场中的梯度轮廓 p_s （形状参数和锐化度参数为 $\{\lambda_s, \sigma_s\}$ ）变换为目标梯度场中的梯度轮廓 p_t （形状参数和锐化度参数为 $\{\lambda_t, \sigma_t\}$ ）。首先计算梯度轮廓 p_s 与 p_t 之间的变换率(Transform Ratio):

$$\begin{aligned} r(d) &= \frac{g(d; \sigma_t, \lambda_t)}{g(d; \sigma_s, \lambda_s)} \\ &= c \cdot \exp\left\{-\left(\frac{\alpha(\lambda_t) \cdot |d|}{\sigma_t}\right)^{\lambda_t} + \left(\frac{\alpha(\lambda_s) \cdot |d|}{\sigma_s}\right)^{\lambda_s}\right\}, \end{aligned} \quad (2-14)$$

其中，

$$c = \frac{\lambda_t \cdot \alpha(\lambda_t) \cdot \sigma_s \cdot \Gamma(1/\lambda_s)}{\lambda_s \cdot \alpha(\lambda_s) \cdot \sigma_t \cdot \Gamma(1/\lambda_t)}, \quad (2-15)$$

d 为沿着梯度轮廓与图像边缘点的距离。因此，目标梯度轮廓 p_t 可以由原始梯度轮廓与变换率 $r(d)$ 相乘得到。给定原始梯度轮廓参数 $\{\lambda_s, \sigma_s\}$ ，目标梯度轮廓参数 $\{\lambda_t, \sigma_t\}$ 可依据不同应用的问题背景通过学习或者定义方式得到。

进一步，我们应用变换率(2-14)进行梯度场变换，即将原始梯度场 ∇I_s 变换为目标梯度场 ∇I_t^T :

$$\nabla I_t^T(x) = r(d(x, x_0)) \cdot \nabla I_s(x), \quad (2-16)$$

其中，下标 t 代表目标梯度场，上标 T 代表通过梯度变换所估计的梯度场。 x_0 为通过 x 的梯度轮廓上的边缘点，而 $d(x, x_0)$ 为 x 与 x_0 沿着梯度轮廓的距离。在算法实现中，我们从每个非零梯度的象素 x 出发，沿着梯度模增大的方向（梯度方向或者负梯度方向）直到达到边缘点 x_0 （梯度模的局部极大值点），然后用公式(2-16)调整其梯度值。

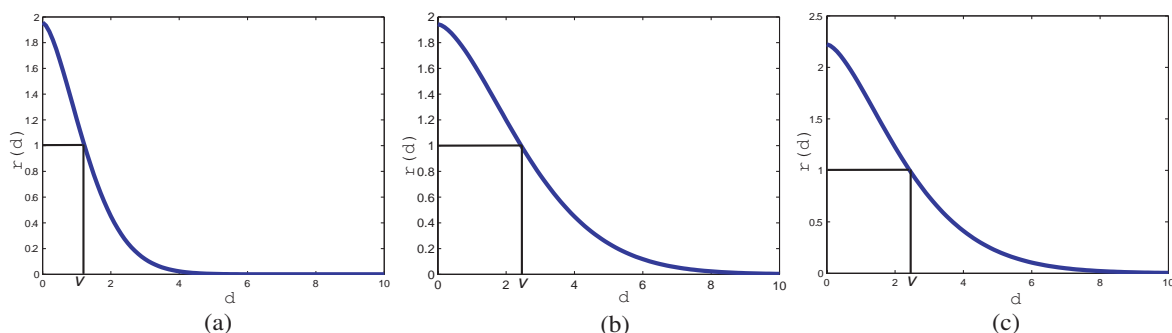


图 2-5 变换率图，如公式(2-14)。(a) $\lambda_t = 1.6, \sigma_t = 1.0, \lambda_s = 1.6, \sigma_s = 2.0$ 。(b) $\lambda_t = 1.6, \sigma_t = 2.0, \lambda_s = 1.6, \sigma_s = 4.0$ 。(c) $\lambda_t = 1.6, \sigma_t = 2.0, \lambda_s = 2.0, \sigma_s = 4.0$ 。当像素距离边缘点的距离大于 v 时，变换率大于1，否则变换率小于1。用变换率与梯度轮廓相乘时，与梯度轮廓边缘点距离小于 v 时，梯度模被增大，否则被减小。

通过对原始梯度场的每个梯度轮廓进行如上的梯度变换，我们可以将原始梯度场变换为目标梯度场。

图2-5显示了具有不同参数的变换率的形状。通过与梯度轮廓相乘，与边缘点距离小于 v 的像素点梯度模被增大，否则被减小。

2.3.2 基于梯度场的图像重建

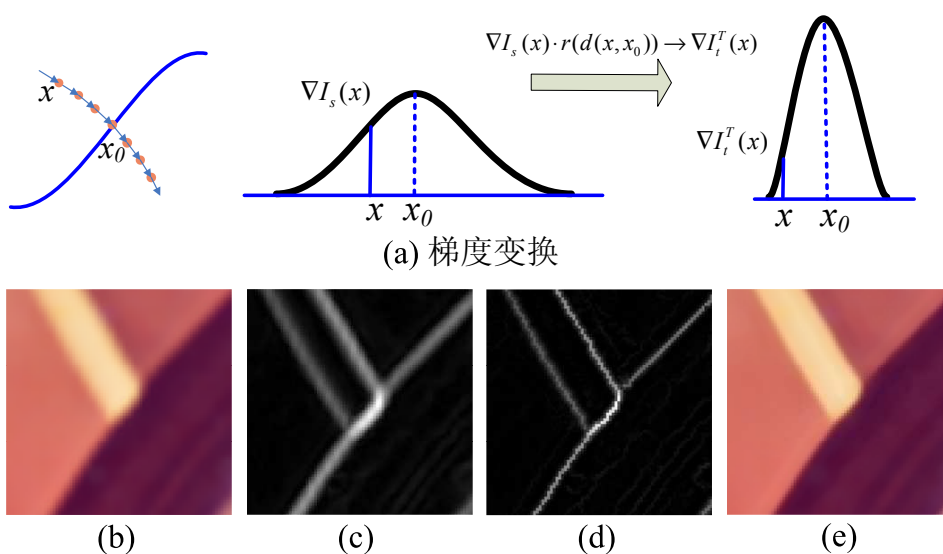


图 2-6 梯度场变换，图像为Lena图像的一部分。(a) 左图和中图显示通过点 x 和 x_0 的梯度轮廓。像素点 x 进行梯度变换后的梯度为原始梯度与变换率 $r(d(x, x_0))$ 相乘得到；(b)和(c)显示原始的图像及其梯度场；(d)和(e)为梯度场变换后的梯度场和重建图像。

给定变换后的梯度场，进一步的问题是如何从变换后的梯度场重建图像 [63, 64]。最直接的方法是对梯度场进行积分，这要求梯度场具有零卷曲率 (curl) 以使其可积。但是我们无法保证变换后梯度场是可积的，故我们采用变分方法来实现图像重建。

在变分法的框架下，目标是使得重建图像 I 的梯度场与给定的目标梯度场 $\nabla I_t^T(x)$ 接近，即极小化能量：

$$E(I; \nabla I_t^T) = \int |\nabla I - \nabla I_t^T|^2. \quad (2-17)$$

为优化上述能量泛函，我们证明如下定理：

定理 2.1 能量泛函(2-17)的全局极小解为如下泊松方程的解：

$$\Delta I = \Delta I_t^T. \quad (2-18)$$

证明 能量泛函 $E(I; \nabla I_t^T)$ 是关于 I 的二次泛函且为凸。其局部最优解即为全局最优解。

E 的局部极小解可以通过其Euler - Lagrange方程得到。记

$$L = |\nabla I - \nabla I_t^T|^2 \quad (2-19)$$

$$= [I_x - (I_t^T)_x]^2 + [I_y - (I_t^T)_y]^2, \quad (2-20)$$

其中，下标 x （或者 y ）分别代表关于 x （或者 y ）的偏导数。 E 的Euler - Lagrange方程为：

$$\frac{\partial L}{\partial I} = \frac{\partial}{\partial x} \frac{\partial L}{\partial I_x} + \frac{\partial}{\partial y} \frac{\partial L}{\partial I_y} \quad (2-21)$$

$$= 2[I_{xx} - (I_t^T)_{xx}] + 2[I_{yy} - (I_t^T)_{yy}] \quad (2-22)$$

$$= 2\Delta I - 2\Delta I_t^T, \quad (2-23)$$

其中， Δ 为Laplace算子，取 $\frac{\partial L}{\partial I} = 0$ ，则 E 的全局最优解满足：

$$\Delta I = \Delta I_t^T. \quad (2-24)$$

此即(2-18)式。（证毕）

根据定理2.1，能量泛函(2-17)可以通过解泊松方程(2-18)来获得极小解，其解 I 即为由变换梯度场 ∇I_t^T 所重建的图像。

为在定义域离散的数字图像上求解泊松方程，我们可以通过差分方法对泊松方程进行差分离散化，差分网格取为数字图像像素所构成的网格，进而将连续形式的泊松方程转化为数字图像像素网格上的线性方程组，并用松弛Gauss-Seidel方法求解离散线性方程组，所得的解即为重建的图像。

图2-6显示了梯度场变换与图像重建的一个例子。其中，(a)显示梯度轮廓的变换，(b-e)显示了真实图像的梯度场变换和图像重建，(c)为图像(b)的梯度场，(d)为变换后的梯度场，(e)为通过解泊松方程重建的图像。从该图可以看到，重建的图像与原始图像相比锐化度更高，而且结果中没有晕状效应。

2.4 基于梯度轮廓先验的超分辨率重建算法

本节我们将梯度轮廓先验应用于图像超分辨率问题，即给定低分辨率图像，设计基于梯度轮廓先验的超分辨率模型与算法以估计输入图像的高分辨率图像形式。图像

超分辨率问题的难点在于它为典型的逆问题，我们将应用梯度轮廓先验来约束问题的解。

应用梯度轮廓先验到高分辨率图像提供了如下的约束：（1）所重建的高分辨率图像的梯度轮廓形状参数接近于1.6；（2）高/低分辨率图像的梯度轮廓的锐化度满足2.2.3节所学习到的锐化度关系。为实现这样的约束，我们提出如下模型。

2.4.1 超分辨率模型

给定低分辨率图像，首先使用公式(2-16)变换输入低分辨率图像的上采样图像 I_l^u 的梯度场，以估计高分辨率图像的梯度场。公式(2-16)中的形状参数 λ_s 与 λ_t 设置为第2.2.2节中所学习到的形状参数。梯度变换率（公式(2-16)）中的锐化度参数 σ_s 可由图像 I_l^u 直接估计，而目标梯度轮廓的锐化度参数 σ_t 可由第2.2.3节中所学习到的锐化度关系设置。

我们用由低分辨率图像所估计出的高分辨率图像梯度场 ∇I_h^T 约束高分辨率图像重建。给定低分辨率图像 I_l ，为重建高分辨率图像 I_h ，我们在图像函数空间 Λ 中极小化：

$$E(I_h|I_l, \nabla I_h^T) = E_i(I_h|I_l) + \beta E_g(\nabla I_h|\nabla I_h^T), \quad (2-25)$$

其中， $E_i(I_h|I_l)$ 为超分辨率问题的重建约束， $E_g(\nabla I_h|\nabla I_h^T)$ 为梯度场约束。

图像域重建约束度量已知低分辨率图像 I_l 与所求高分辨率图像 I_h 的低分辨率图像（高斯卷积后下采样的图像）之间的接近程度，即：

$$E_i(I_h|I_l) = |(I_h * G) \downarrow - I_l|^2. \quad (2-26)$$

其中， G 为空间滤波器， $*$ 为卷积算子， \downarrow 为下采样算子。我们用高斯滤波 G 作为空间滤波器，当放大倍数为 2×2 、 3×3 和 4×4 时，其核宽度分别设置为0.8、1.2和1.6。由于高斯滤波和下采样操作仅适用于离散函数形式，因此假定 I_h 和 I_l 均为数字图像所展开的向量形式（例如： $m \times n$ 分辨率图像可按列展开为 $mn \times 1$ 的向量）。

梯度场约束要求所估计高分辨率图像的梯度场尽可能逼近由低分辨率图像所估计的高分辨率图像梯度场 ∇I_h^T ：

$$E_g(\nabla I_h|\nabla I_h^T) = |\nabla^d I_h - \nabla^d I_h^T|^2, \quad (2-27)$$

其中， ∇I_h 为所求高分辨率图像 I_h 的梯度场。为使得能量(2-25)中的两项定义统一为离散形式，因此定义中采用离散梯度算子 ∇^d ，即为梯度的离散差分算子。通过该约束，我们要求高分辨率图像 I_h 的梯度轮廓具有与自然图像一致的梯度场统计分布。

为优化上述能量泛函，我们可以证明如下定理：

定理 2.2 能量泛函(2-25)的全局极小解可以由如下的偏微分方程求解：

$$\frac{\partial I_h}{\partial t} = -\frac{\partial E}{\partial I_h}, \quad (2-28)$$

这里，

$$\frac{\partial E(I_h)}{\partial I_h} = ((I_h * G) \downarrow - I_l) \uparrow * G - \beta \cdot (\Delta^d I_h - \Delta^d I_h^T). \quad (2-29)$$

Δ^d 为离散Laplacian算子，即Laplacian算子的离散差分形式。

证明 能量泛函 $E(I_h|I_l, \nabla I_h^T)$ 包含两项 $E_i(I_h|I_l)$ 和 $E_g(\nabla I_h|\nabla^d)$ 。由于第一项中的卷积操作和下采样算子均可以由矩阵表达，而 I_h 为向量形式，因此极小化 E 的问题是 最小二乘问题，且 E 具有凸性；第二项中 ∇^d 为离散差分算子且 I_h 和 I_l 为向量，那么 ∇I_h 和 ∇I_h^T 均可以写为矩阵与向量相乘的形式，因此第二项也为最小二乘问题，具有凸性。因此， $E(I_h|I_l, \nabla I_h^T)$ 局部最优解即为全局最优解。

E 的局部极小解可以通过其梯度下降流得到。记

$$L = |(I_h * G) \downarrow - I_l|^2 + |\nabla^d I_h - \nabla^d I_h^T|^2 \quad (2-30)$$

那么， E 关于 I_h 的梯度下降流方程为：

$$\frac{\partial E(I_h)}{\partial I_h} = ((I_h * G) \downarrow - I_l) \uparrow * G - \beta \cdot (\Delta^d I_h - \Delta^d I_h^T). \quad (2-31)$$

上式中， Δ^d 为离散Laplace算子形式。因此能量泛函(2-25)的全局极小值可以由梯度下降法求得。（证毕）

我们可采用如下公式迭代求解微分方程(2-29)：

$$I_h^{t+1} = I_h^t - \tau \frac{\partial E(I_h)}{\partial I_h}. \quad (2-32)$$

迭代步长例如可设定为 $\tau = 0.2$ ，参数 $\beta = 0.5$ ，且用低分辨率图像的上采样图（双立方插值） I_l^u 作为迭代的初始条件。这种情况下，基于梯度轮廓先验的超分辨率算法可总结为如下算法。

算法2.1：基于梯度轮廓先验的超分辨率算法

输入： 低分辨率图像 I_l 和放大倍数 $m \times m$ 。

步骤1： 估计高分辨率梯度场 ∇I_h^T ：

- 用Bicubic插值对低分辨率图像 I_l 放大 $m \times m$ 倍得到 I_l^u ；
- 计算 I_l^u 的梯度场，计算梯度场的局部极大值作为边缘点，进而计算通过边缘点与图像边缘垂直的梯度轮廓；
- 根据梯度轮廓先验，对梯度轮廓进行梯度场变换，从而估计出高分辨率梯度场 ∇I_h^T 。

步骤2： 计算高分辨率图像：以 I_l^u 为初始高分辨率图像，迭代（2-32）直到收敛，收敛结果即为高分辨率图像。

输出： 高分辨率图像 I_h 。

图2-7显示了图像超分辨率的结果。第一行为输入低分辨率图像。(a)为最近邻差值结果和上采样（双立方插值）图像的梯度场；(d)为真实的高分辨率图像和它的梯

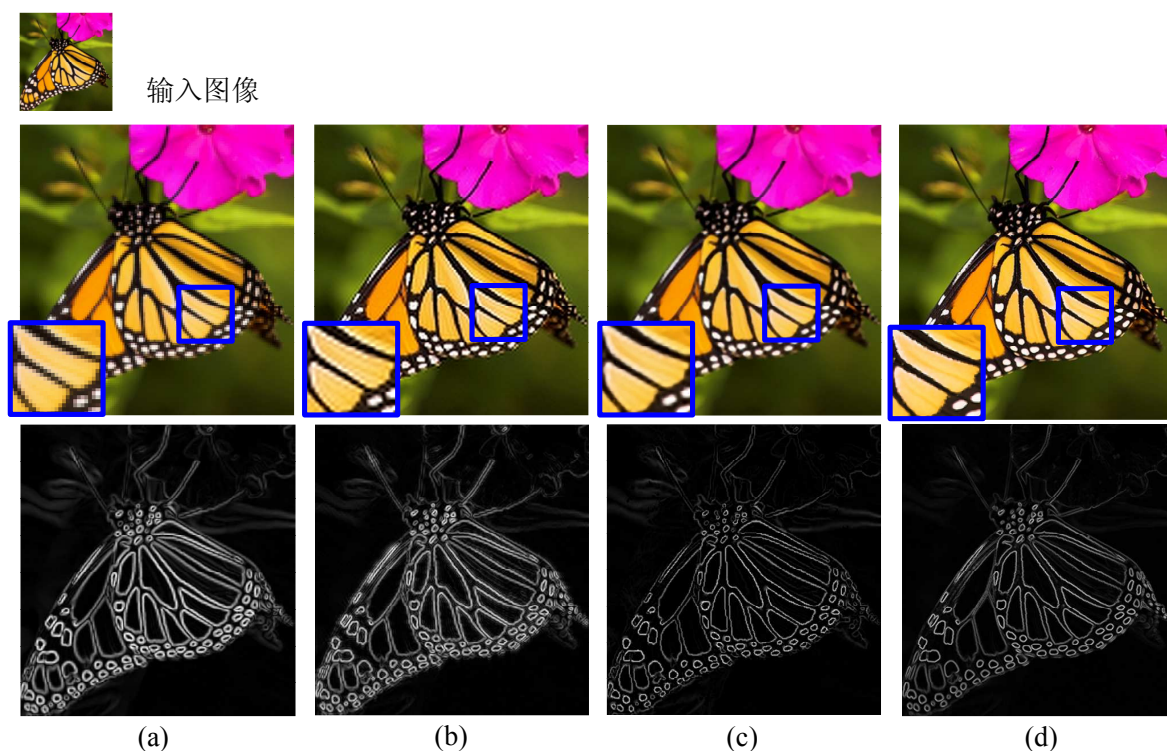


图 2-7 超分辨率模型(高/低分辨率倍数 3×3)。第一行为输入低分辨率图像。(a) 最近邻插值的图像和上采样(双三次插值)图像的梯度场。(b) Back-projection算法的超分辨率图像及其梯度场。(c) 我们的超分辨率结果及其梯度场。(d) 真实的高分辨率图像及其梯度场。与Back-projection算法结果的梯度场相比,我们算法的变换梯度场与真实高分辨率图像的梯度场更一致,而我们的重建结果很少具有锯齿状和晕状效果。

度场;(b)为back-projection [47]算法结果及其梯度场,该算法只用到了图像域的重建约束;(c)的下图为通过梯度场变换所估计的高分辨率图像梯度场,由图可见,该估计梯度场与真实梯度场非常接近;(c)中的上图为我们的超分辨率结果,相比back-projection算法,我们的算法由于引入了梯度轮廓的约束,结果很少带有晕状效果,更符合视觉感知。

图2-8还显示了在生成图像上的处理结果,我们的方法可以非常好的重建出高分辨率图像,结果具有很高的锐化度,且很少引入锯齿状或晕状效果。

2.4.2 更多的例子

我们将本章所提出的该超分辨率算法应用到更多的图像例子中。对于彩色图像,仅在亮度通道运用所提出的超分辨率操作,这是因为人眼对图像的亮度信息更为敏感,而彩色通道可以由双立方插值算法进行超分辨率操作。实验图像为公用的例子图像(例如Lena图像)或来源于专业的图像库。

在图2-9中,我们将所提出的算法与双立方插值算法、增强的双立方插值算法、back-projection [47]算法和梯度场重建算法(即从梯度变换所估计的高分辨率梯度

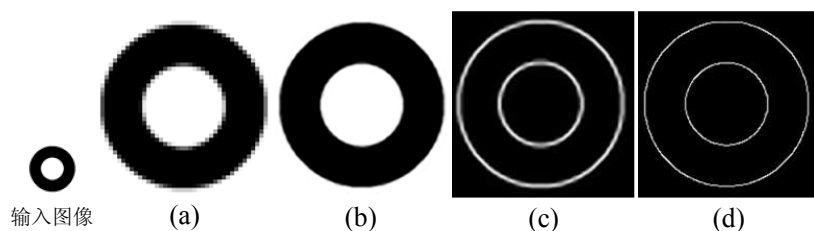


图 2-8 合成图像的超分辨率结果，放大倍数为 4×4 。最左端为输入低分辨率图像。(a) 最近邻插值结果。(b) 重建的高分辨率图像。(c) 上采样图像（双三次插值）的梯度场。(d) 梯度场变换的高分辨率梯度场。

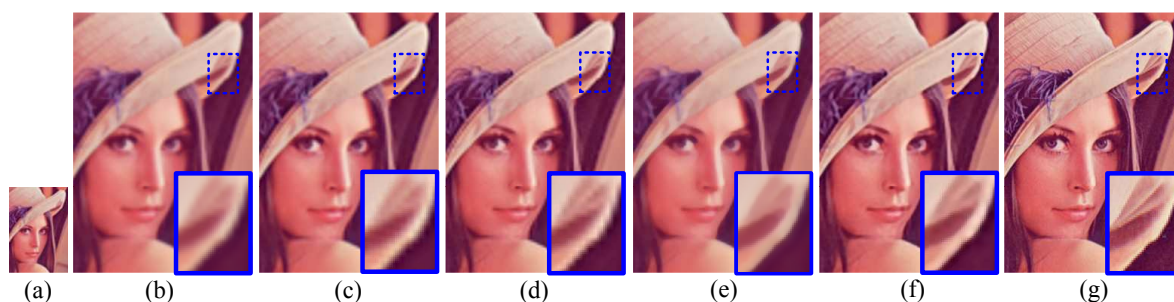


图 2-9 超分辨率算法比较（放大倍数为 3×3 ）。(a) 低分辨率图像（Lena图像的一部分）。(b) 双立方插值结果。(c) Unsharp Masking算法增强的双立方插值结果。(d) back-projection算法结果。(e) 梯度重建算法结果，即直接由变换梯度场重建的高分辨率图像。(f) 我们的算法结果。(g) 真实高分辨率图像。显然，我们的结果(f)与真实图像(g)更接近。

场重建高分辨率图像）进行比较。双立方插值算法的结果过于光滑化，例如矩形方框中所显示的图像区域。增强的双立方插值算法和back-projection算法结果在图像边缘周围引入晕状和锯齿状现象。通过估计的高分辨率梯度场重建高分辨率图像的结果具有很少的晕状和锯齿状现象，但是颜色对比度与真实的高分辨率图像有一定的差距。通过结合梯度场约束和图像域重建约束，我们的算法结果最真实。

图2-10显示了我们的方法与基于学习的算法 [51]和alpha通道超分辨率算法 [37]的

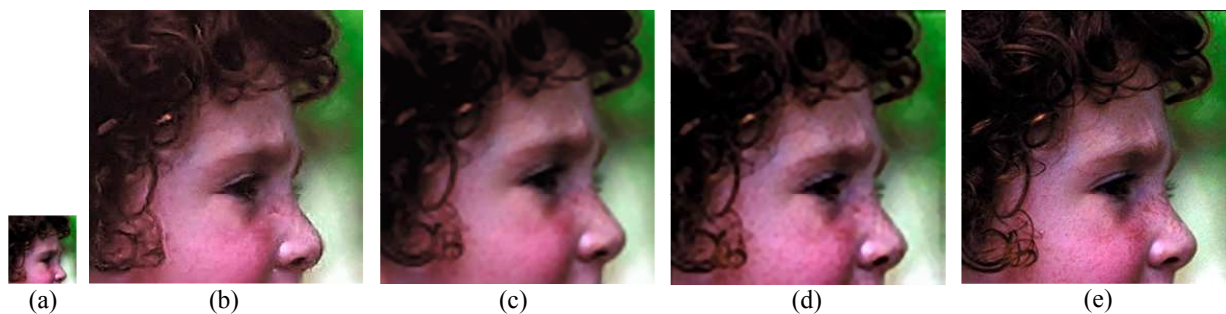


图 2-10 超分辨率算法比较（放大倍数为 4×4 ）。(a) 低分辨率图像。(b) 基于学习的算法 [51]的结果。(c) alpha通道超分辨率算法 [37]的结果。(d) 我们的算法结果。在我们的超分辨率结果中，大尺度边缘和小尺度纹理都能很好地重建出来。

结果比较。可以看出，基于学习的算法结果具有很高的锐化度，但是会引入高频的伪影效果，例如沿着鼻子边缘的晕状效果。 α 通道超分辨率算法结果的图像大尺度边缘具有很强的锐化度，但是小尺度的图像结构，比如人脸上的斑点没有很好的恢复出来。这是因为很难对图像弱边缘进行 α 通道值估计。与这些结果相比，我们的算法能很好地恢复大尺度和小尺度的图像结构。

图2-11显示了三个例子，放大倍数分别为 8×8 ， 8×8 和 16×16 。其中高分辨率图像是通过每次放大 2×2 倍，放大多次得到。由于放大倍数很大，原图相对放大结果过小，因此输入的低分辨率图像用其最近邻放大的结果进行显示，并放大矩形框中的区域以便更好地比较超分辨率结果。所有结果都显示：我们的算法可以稳健地恢复图像的大尺度和小尺度图像结构，且很少在结果中引入锯齿或晕状伪影效果。

我们还以RMS（Root Mean Square）和ERMS [36]（Edge Root Mean Square）量化测度超分辨率结果的优劣。表2-1显示了在Monarch (图像2-7), Lena (图2-9) 和Head (图2-10)三个例子上的超分辨性能量化比较。与双立方插值算法和back-projection算法相比，我们的算法结果具有更小的RMS和ERMS值。在3.0 GHz CPU的个人电脑平台上，例子Monarch(原始分辨率为 399×423)，Lena(原始分辨率为 500×500)和Head(原始分辨率为 280×280)的运算时间分别为7.4s、8.7s和3.5s。

2.5 基于梯度轮廓先验的增强算法

生物视觉系统对图像的边缘（即亮度变化较大的区域）具有更强的敏感性，边缘锐化度很高的图像在视觉感知上具有更好的视觉效果。但是，图像往往由于相机光学系统的低通滤波操作或者图像处理软件（例如Photoshop）的滤波操作而产生模糊化效果。因此，为了提高图像的视觉效果，我们将应用该梯度轮廓先验增强图像的锐化

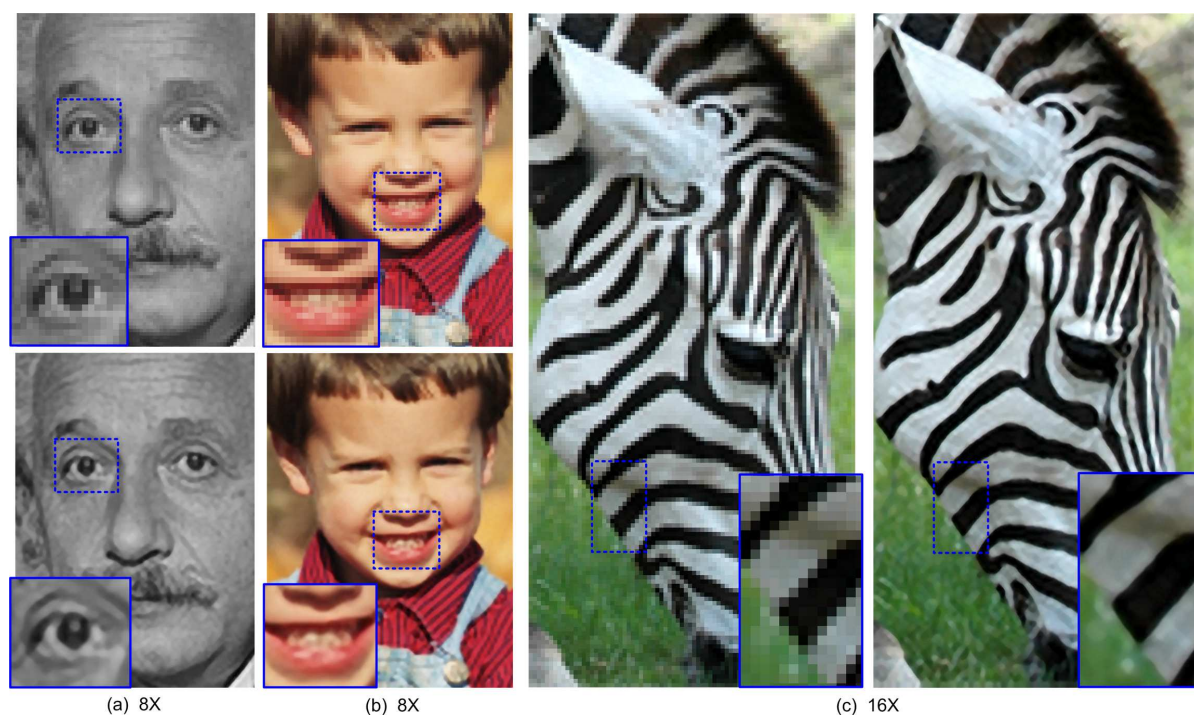


图 2-11 超分辨率结果（放大倍数为 8×8 和 16×16 ）。(a)(b)的上图和(c)的左图为最近邻插值的结果，(a)(b)的下图和(c)的右图为我们的结果。

表 2-1 超分辨率结果量化比较

测试图像	双三次插值		back-projection		我们的算法	
	RMS	ERMS	RMS	ERMS	RMS	ERMS
Monarch	16.4	26.0	13.6	21.3	13.2	20.9
Lena	8.8	11.5	8.2	10.8	7.8	10.1
Head	8.7	10.9	8.6	10.6	8.4	10.3

度。

注意到，在图像增强问题中，我们并没有关于模糊图像与高锐化度图像之间的锐化度关系的先验信息。因此我们只能借助于人-机交互或者数学模型设置模糊图像和目标图像之间的锐化度依赖关系。我们定义如下两种方式：首先，定义锐化度变换函数，用函数形式参数化地描述锐化度关系，其中的参数可以通过人-机交互由用户根据需要设定；其次，提出锐化度转移算法，从而实现将指定的高锐化度图像的锐化度分布转移到模糊图像上，以实现模糊图像的自动增强。

2.5.1 锐化度变换函数

算法2.2: 基于梯度轮廓先验的图像增强算法

输入: 模糊图像 I_b 和参数 μ 。

步骤1: 估计锐化度增强图像的梯度场 ∇I_s :

- 计算 I_b 的梯度场，计算梯度场的局部极大值作为边缘点，进而计算通过边缘点与图像边缘垂直的梯度轮廓；
- 根据梯度轮廓先验，用锐化度变换函数对梯度轮廓进行梯度场变换，从而估计出锐化的梯度场 ∇I_s 。

步骤2: 从 ∇I_s 重建出高锐化度图像 I_s ，重建方法如第2.3.2节。

输出: 高锐化度图像 I_s 。

首先定义锐化度变换函数 F ，该函数建模了锐化度增强图像的梯度轮廓锐化度与模糊图像的梯度轮廓锐化度之间的函数关系，定义为：

$$F(\sigma) = [1 - \exp(-\mu\sigma)]\sigma, \quad (2-33)$$

其中， μ 为控制该函数形状的参数。显然， $F(\sigma) < \sigma$ ，即增强之后的梯度轮廓锐化度相比模糊图像的梯度轮廓锐化度更高（ σ 越小，锐化度越高）。图2-12 显示了该函数

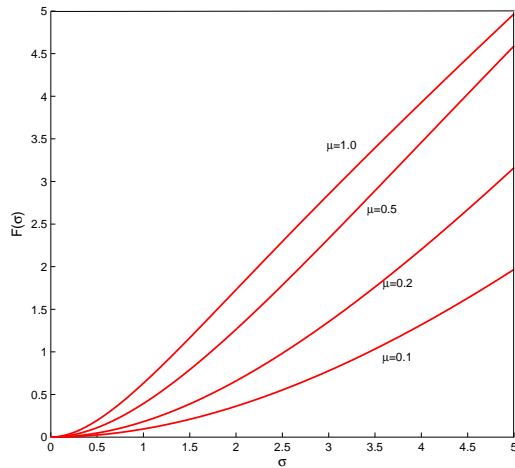


图 2-12 不同参数 μ 下的图像边缘锐化度变换函数(公式(2-33))

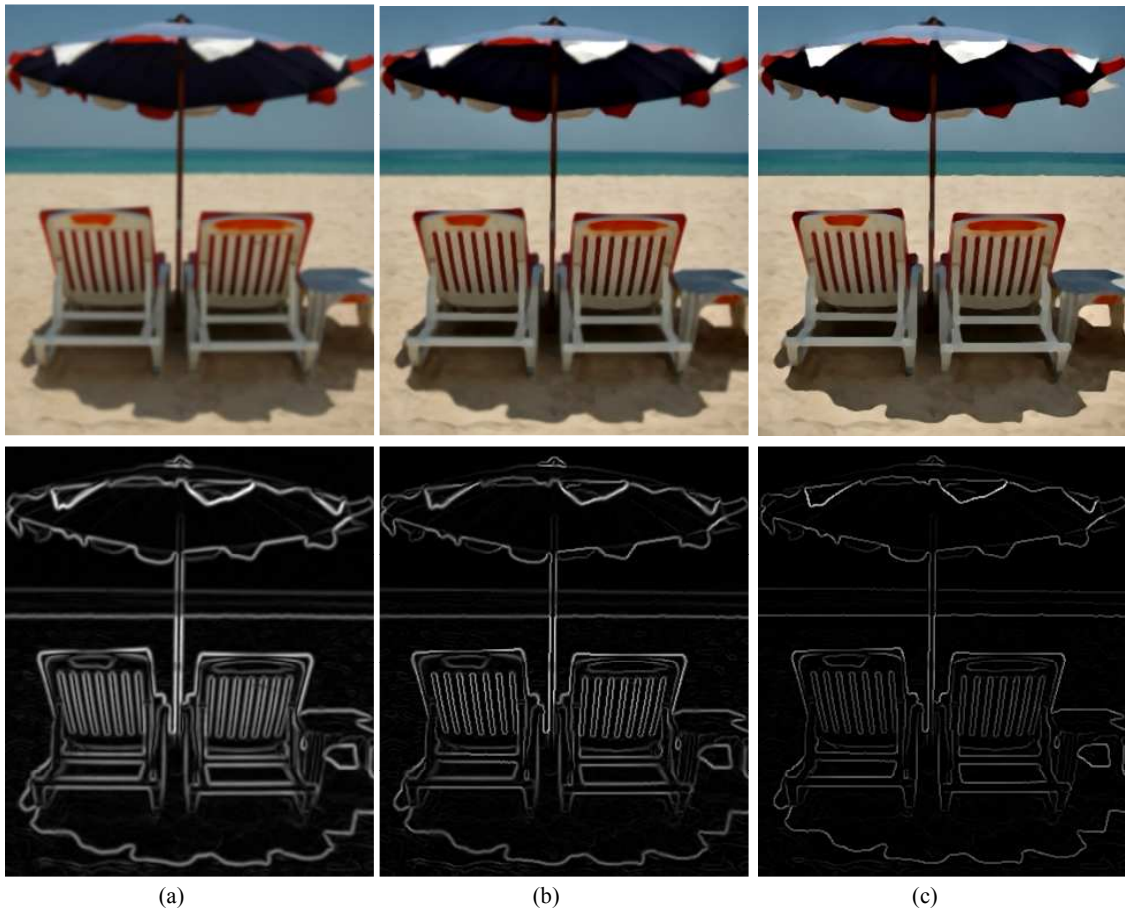


图 2-13 应用锐化度变换函数进行图像增强的结果（公式(2-33)）。(a) 模糊图像和它的梯度图。(b) 增强的梯度场(下图)和重建的增强图像（上图），其中 $\mu = 0.5$ 。(c) 增强的梯度场(下图)和重建的增强图像（上图），其中 $\mu = 0.1$ 。显然，(c)中的增强图像比(b)中的增强图像具有更高的锐化度，而(b)中的增强图像比(a)中的增强图像具有更高的锐化度。

在不同参数 μ 下的形状。 μ 值越小，边缘锐化度增强的程度越大。在该定义下，梯度变换率定义为：

$$r_s(x, x_0) = \frac{g(d(x, x_0); F(\sigma(x)), \lambda^*)}{g(d(x, x_0); \sigma(x), \lambda^*)}, \quad (2-34)$$

其中， x_0 和 $\sigma(x)$ 为边缘像素点和通过点 x 的梯度轮廓的锐化度， λ^* 取为1.6，即为从自然图像中所学习到的梯度轮廓最优参数。基于梯度轮廓先验的图像增强算法框架如算法2.2。



图 2-14 由聚焦不准引起的模糊图像的增强结果。左图为模糊图像，右图为增强图像， μ 取为0.2。

图2-13显示了用不同参数 μ 的锐化度变换函数(2-33)的图像增强结果。图2-13(a)显示了模糊图像和它的梯度图；图2-13(b)(c)是参数分别取为 $\mu = 0.5$ 和 $\mu = 0.1$ 时的图像增强结果。图2-13(b)(c)的第二行和第一行是增强的梯度图和重建的增强图像。显然，图2-13(c)的增强结果比图2-13(b)的增强结果具有更高的锐化度。参数 μ 可以由用户通过人-机交互方式设定，通常可以设置为0.1和0.5之间的值。

图2-14显示了真实的由于相机聚焦不准所引起的模糊例子。图2-14中左图显示输入模糊图像，右图显示图像增强结果，参数 μ 设置为0.2。

2.5.2 锐化度转移

图像转移(Image transfer)算法是计算机图形学与图像处理中常用的一项技术 [65, 66]。基本想法是：将原始图像特征（例如颜色、亮度等）的统计规律作用于目标图像上，使得目标图像具有与原始图像相似的特征分布。

基于图像转移的基本思想，我们提出一种自动的图像锐化度增强算法，即锐化度转移算法。基本思路是将给定的原始图像 I_s （通常为高锐化度图像）的锐化度分布转移至输入的目标图像 I_t （通常为输入的模糊图像），使得模糊图像 I_t 具有与高锐化度图像 I_s 相同的锐化度分布，进而实现图像增强。首先我们计算原始图像和目标图像的锐化度分布，然后将该原始图像的锐化度分布作用到目标图像上。假设 I_s 和 I_t 的锐化度累计直方图分别为 H_s 和 H_t ，那么对于 I_t 中的某个梯度轮廓（锐化度为 σ ），在锐化度转移之后的锐化度为：

$$\sigma' = H_s^{-1}(H_t(\sigma)), \quad (2-35)$$

其中, H_s^{-1} 为 H_s 的逆。

基于梯度轮廓先验的锐化度转移算法框架如算法2.3。图2-15显示了一个图像锐化度转移的例子。图2-15(a)为输入的高锐化度图像（原始图像），(b)为输入的模糊图像（目标图像），(d)(e)为两个图像的锐化度分布。显然，在原始图像中的梯度轮廓的锐化度大都分布在1.0左右，但是目标图像的锐化度大都分布于2.4左右。这与我们的视觉感知是一致的：即原始图像相对于目标图像在视觉感知上锐化度更高一些。图2-15(f)为锐化度分布转移的映射关系 $H_s^{-1}H_t$ ，它将原始图像的锐化度分布转移作用到目标图像上。图2-15(c)为锐化度转移之后的增强结果。显然，图像增强结果具有很高的锐化度，且沿着图像边缘具有很少的幻影效果。

算法2.3: 基于梯度轮廓先验的锐化度转移算法

输入: 高锐化度图像 I_s 和输入的模糊图像 I_t 。

步骤1: 分别求 I_s 和 I_t 的梯度轮廓锐化度累计直方图 H_s 和 H_t ，并计算 H_s 的逆 H_s^{-1} ；

步骤2: 求 I_t 的梯度轮廓及其锐化度 σ ，然后求其锐化度转移后新的锐化度 σ' ，公式见(2-35)；

步骤3: 根据梯度轮廓先验和锐化度转移前后的锐化度关系对模糊图像梯度场进行梯度场变换，变换方法如第2.3.1节；

步骤4: 从梯度场变换后的梯度场重建出高锐化度图像，重建方法如第2.3.2节。

输出: 高锐化度图像 I_s 。

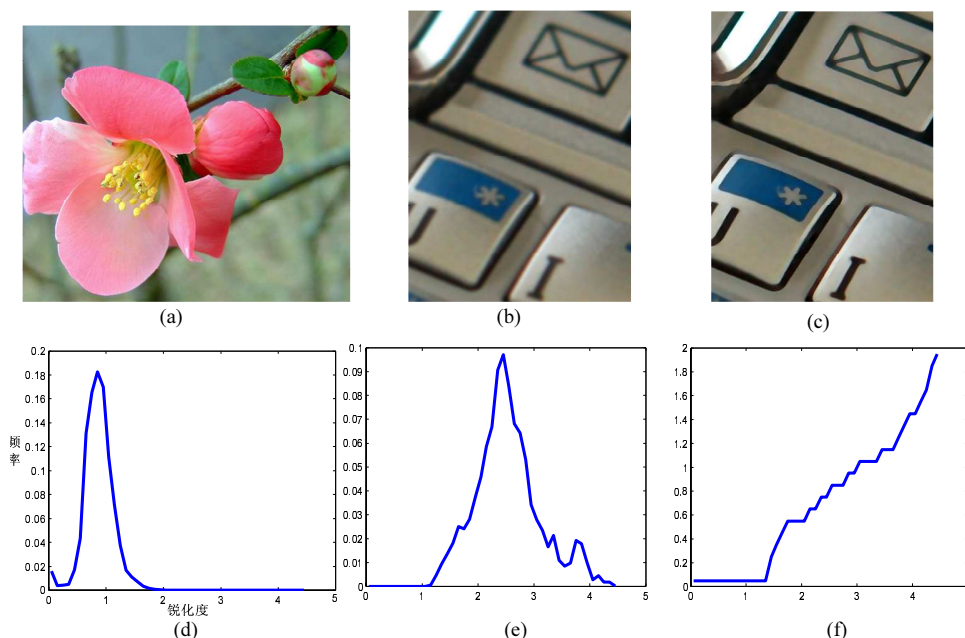


图 2-15 图像锐化度转移。(a) 原始图像。(b) 目标图像。(c) 将原始图像的锐化度分布转移到目标图像的图像增强结果。(d) 原始图像(a)的锐化度分布。(e) 目标图像(b)的锐化度分布。(f) 从(d)到(e)的锐化度映射关系。

2.5.3 对比实验

现在我们将本节所提出的增强算法与经典的图像增强算法进行性能比较，比较的算法包括：Unsharp masking (UM)算法和shock滤波算法 [59–61]，它们是最基本和最常用的图像增强算法。Unsharp masking增强算法的基本思想是：将输入的模糊图像加上高频的细节信息以增强其锐化程度，而该高频信息是由输入图像与它的高斯卷积图像的差得到。Osher-Rudin的shock滤波算法 [60]和Kramer的shock滤波算法 [59, 61, 154]是两种基本的shock滤波算法。两种算法的区别在于：前者用图像的Laplacian变换域检测图像边缘，而后者是用图像的二阶方向导数检测边缘；两种算法的相同点在于：图像边缘增强的过程均采用逆扩散方程增强所检测到的图像边缘，而该方程等价于沿着图像边缘进行形态学操作（膨胀和腐蚀操作）。

图2-16为图像增强算法的比较结果。图2-16 (a)为模糊图像，(b)(c)(d)(e)分别为所提算法的增强结果、Osher-Rudin的shock滤波结果、Kramer的shock滤波结果和Unsharp masking的结果。对于Unsharp masking算法，高斯核宽度取为2.0，且所添加的高频信息的百分比率为100%。由图可见，Unsharp masking算法可以很好地恢复图像的对比度，但是增强结果沿着大尺度边缘有明显的晕状幻影，而且边缘锐化度不够高。Osher-Rudin的shock滤波算法和Kramer的shock滤波算法都可以产生锐化度很高的增强结果。但是，这两种方向倾向于将所有检测出的边缘锐化成阶梯型的边缘，即边缘不存在颜色过渡。例如，男孩的脸上的细节和阴影也被增强为阶梯型边缘。与这些结果相比，所提出的基于梯度轮廓先验的图像增强算法能够自适应地增强大尺度和小尺度的边缘锐化度，而不会将所有边缘均增强为阶梯型边缘。我们的结果（ μ 取为0.2）具有最自然的图像增强结果，且结果中很少有晕状或过光滑的现象产生。



图 2-16 图像增强结果比较。(a) 输入模糊图像。(b) 我们的增强结果， $\mu = 0.2$ 。(c) Osher-Rudin的shock滤波增强结果。(d) Kramer的shock滤波结果。(e) Unsharp masking的增强结果其中，高斯核宽度取2.0，添加高频信息的百分比率为100%。

2.6 小结

本章提出了一种新的一般性自然图像先验：梯度轮廓先验。应用该先验，定义了一种自然图像的梯度场变换，即通过调整图像梯度轮廓的形状和锐化度参数（调整方

式由具体的应用决定)实现梯度场的变换。应用变换后的梯度场约束图像超分辨率问题中的高分辨率图像和图像增强问题中的高锐化度图像。我们在自然图像和合成图像上进行了图像超分辨率和图像增强的实验和对比。实验表明,所提出的基于梯度轮廓先验的图像超分辨率和增强算法可以很好地锐化图像的边缘和细节信息,同时很少引入晕状或锯齿效应;对比表明,所提出算法相对于已有的同类算法具有更好的图像处理效果。

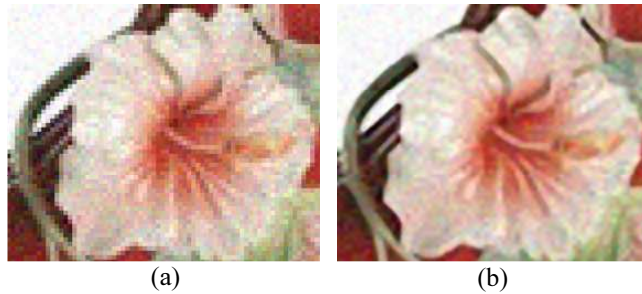


图 2-17 在噪声图像上的超分辨率实验(放大倍数为: 4×4)。(a)带噪声的低分辨率图像的最近邻放大结果;(b)所提出算法的放大结果:带噪声的低分辨率图像首先由non-local去噪算法 [67]进行去噪,然后将去噪的图像由所提出的超分辨率算法进行放大,而噪声通过双线性插值进行放大并加到去噪图像的超分辨率结果中,从而得到最终的超分辨率结果。

对于噪声图像,由于噪声的影响,梯度轮廓的估计可能会不准确、不稳健。以超分辨率问题为例,我们做如下处理:首先对噪声图像进行去噪,然后对去噪图像进行超分辨率操作,并将上采样的噪声图像加回到去噪图像的超分辨率结果中,从而得到最终结果。这样可以避免噪声对梯度轮廓估计的影响。如图 2-17为例,带噪声的低分辨率图像首先由non-local去噪算法 [67]进行去噪,然后将去噪的图像由我们的超分辨率算法进行放大,而噪声通过双线性插值进行放大并加到去噪图像的超分辨率结果中,从而得到最终的超分辨率结果。

在下一步的工作中,我们计划将该图像超分辨率和增强算法应用到视频中,并通过设计加速算法,以期该方法可以应用于高清电视、摄像机等工业产品中。我们还关注于如何将梯度轮廓先验应用于更多的图像处理问题中。

3 基于Markov随机场尺度选择的各向异性扩散去噪算法

3.1 引言

图像去噪是计算机视觉和图像处理的经典问题，基本的去噪方法包括：基于滤波的去噪算法和基于变换（尤其是小波变换）的去噪算法。

基于滤波的去噪算法通过滤波器与噪声图像的卷积实现去除图像的高频信息。高斯滤波算法是最基本的滤波去噪算法，该算法用邻域像素点距离的高斯函数值作为权重的带权平均值更新中心像素的颜色值，该算法在去除噪声的同时会模糊化图像边缘。双边滤波（Bilateral Filtering）算法 [68] 通过颜色、位置两类信息计算带权平均的权重进而进行图像滤波，它克服了高斯滤波算法不能很好保持图像边缘的缺点。非局部平均（Non-local Means）滤波算法通过计算相邻像素点所在图像块的相似性作为带权平均的权重，该算法具有比双边滤波更好的滤波性能。另一类重要的去噪算法是基于扩散的滤波去噪算法，该类算法基于偏微分方程方法和变分方法，将图像看作为函数并建立关于图像微分的能量函数，通过求解偏微分方程优化能量函数，以实现图像的滤波去噪。热扩散方程是最基本的扩散去噪算法，等价于图像的高斯滤波。为使得扩散方程具有保持图像边缘的特性，人们设计了大量的非线性扩散方程 [8, 69–71]，例如各向异性扩散方程 [7]、全变差扩散方程 [5] 等等。

基于变换的去噪方法通过图像变换（例如小波变换等）将图像分解为不同方向、不同尺度下的特征，并对这些特征分别进行去噪，然后重构出去噪图像。小波收缩算法（Wavelet Shrinkage）算法 [72] 通过小波变换将图像分解，并在变换域中进行阈值化处理以实现去噪。小波域高斯尺度混合（Gaussian Scale Mixture Model）模型通过对小波变换域的系数进行高斯尺度混合模型建模，并通过贝叶斯推理得到去噪后的小波系数，进而重建出去噪图像，该算法是目前最好的去噪算法之一。

尺度空间方法是计算机视觉中广泛研究与应用的方法。主要应用于图像滤波 [73]、边缘检测 [74–76]、特征提取 [77–79] 以及数据挖掘 [80]。我们知道，扩散尺度空间方法是一种重要的滤波去噪算法，由基于偏微分扩散方程的滤波算法导出。它以噪声图像为扩散方程的初始条件，通过迭代求解扩散方程，噪声和图像细节信息将会被逐渐去除。其中，扩散方程的迭代次数与迭代步长的乘积被称为当前的滤波尺度。随着尺度的不断增加，滤波图像构成了滤波尺度空间 [73, 81]。各向同性的扩散尺度空间 [82] 由热扩散方程导出，它对图像信息的扩散不具有方向性，即各个方向的扩散程度一致，因此该方程会模糊化图像边缘。Perona 和 Malik [7] 提出了各向异性的扩散方程，其扩散系数是空间自适应的：对于图像边缘区域，其扩散系数较小；对于图像平滑区域，其扩散系数较大。Rudin 等人 [5] 设计了全变差扩散策略，该策略极小化图像的全变差（图像所有像素点的梯度模之和）的同时约束所去除的噪声满足已知的噪声统计 [6, 83]。

本章关注各向异性扩散尺度空间，它由如下扩散方程诱导 [84]：

$$\frac{\partial I(x, t)}{\partial t} = \operatorname{div}[g(|\nabla I|)\nabla I]. \quad (3-1)$$

扩散方程应用于图像去噪的一个关键问题是：如何选择最优尺度（即停机尺度）以得到最好的去噪结果，这个问题即为扩散方程的尺度选择问题。该问题可如下形式化描述：给定由扩散方程所诱导的尺度空间 $\{I(x, t)\}_{0 \leq t < +\infty}$ ，目的是找到最优的停机尺度 t^* ，使得在该尺度下，去噪的图像 $I(x, t^*)$ 与原始未加噪声的图像最接近。当选择小的停机尺度时，噪声没有被充分去除，但是当选择大的停机尺度时，噪声被去除的同时，图像的细节信息也被去除。所以选择合适的尺度显著影响扩散滤波的去噪效果。我们需要选择合适的停机准则以平衡扩散滤波的去噪能力和保持细节的能力。

为扩散滤波选择合适的停机尺度是个很困难的问题，已经有很多工作研究如何解决这个问题。Dolcetta 和 Ferretti [85] 将尺度选择问题建模为能量极小化问题，该能量耦合了计算复杂度和停机代价两方面因素。实践中，需要仔细地平衡两方面因素以得到合适的停机尺度。Weickert [86] 提出了基于信噪比（Signal to Noise Ratio, SNR）的尺度选择准则以选择最优尺度 t^* ，在最优尺度下去噪图像 $I(x, t^*)$ 满足：

$$\frac{\text{Var}(I(x, t^*))}{\text{Var}(I(x, 0))} = \frac{1}{1 + 1/\text{SNR}}, \quad (3-2)$$

其中，SNR为已知的噪声图像信噪比。由于要求信噪比已知，所以该准则可以很好地选择最优停机尺度。Solo [87] 设计了基于重建质量的二次测度以选择尺度，可以通过SURE (Stein's Unbiased Risk Estimator) 算法进行计算。Mrazek [88] 提出了基于相关性的尺度选择策略，即选择尺度使得在最优尺度下，扩散图像与去除的噪声之间的相关性最小：

$$t^* = \underset{\{t \geq 0\}}{\text{argmin}} \frac{\text{Cov}(I(x, 0) - I(x, t), I(x, t))}{\sqrt{\text{Var}(I(x, 0) - I(x, t))\text{Var}(I(x, t))}}. \quad (3-3)$$

该策略简单且不需要假设任何的噪声分布或者噪声图像的信噪比，但该相关性准则并不能很好地反映去噪质量。Papandreou [89] 发展了基于交叉验证的尺度选择策略。该方法将图像区域分为训练集和测试集，最优尺度选择的准则为使得测试集上的 L_1 或者 L_2 误差测度最小。该策略具有很好的统计学习基础，但很难保证所选择尺度是最优的。全变差算法（Total Variation, TV）模型 [5] 使用噪声分布约束全变差扩散过程，使其收敛于如下稳定状态：即去除的噪声满足已知的噪声统计分布。该算法也可以看作是类似于信噪比的尺度选择策略。

已有工作均为整幅图像选择同一个停机尺度，这样的全局尺度选择策略有如下问题：以图3-1 为例，我们为整个图像区域、光滑区域（实线矩形区域内）和纹理区域（点线矩形区域内）分别选择最优尺度。尺度选择的准则为极小化扩散图像与原始未加噪声图像之间的均方差（Mean Square Error, MSE）。由于误差评测时的参考图像为原始的未加噪图像，因此所选尺度为真实的最优尺度。图3-1(c)显示光滑区域的最优尺度为42.2，而纹理区域的最优尺度为14.8，远小于光滑区域的尺度。这意味着：在纹理区域，扩散方程需要具有较小的停机尺度，以保持图像的细节信息；在平滑区域，扩散方程需要具有较大的停机尺度，以充分去除噪声。如果为整幅图像选择统一的尺度，那么最优尺度为25.0，处于纹理区域和平滑区域的最优尺度之间，这是由于图像是由纹理区域和平滑区域混合而成，该尺度平衡了图像不同区域的尺度选择。

上述观察启发我们设计一种基于图像分割的空间自适应尺度选择模型。我们的基本想法是：将图像分割为多个小图像块（图像块内像素具有相似的颜色和纹理），并

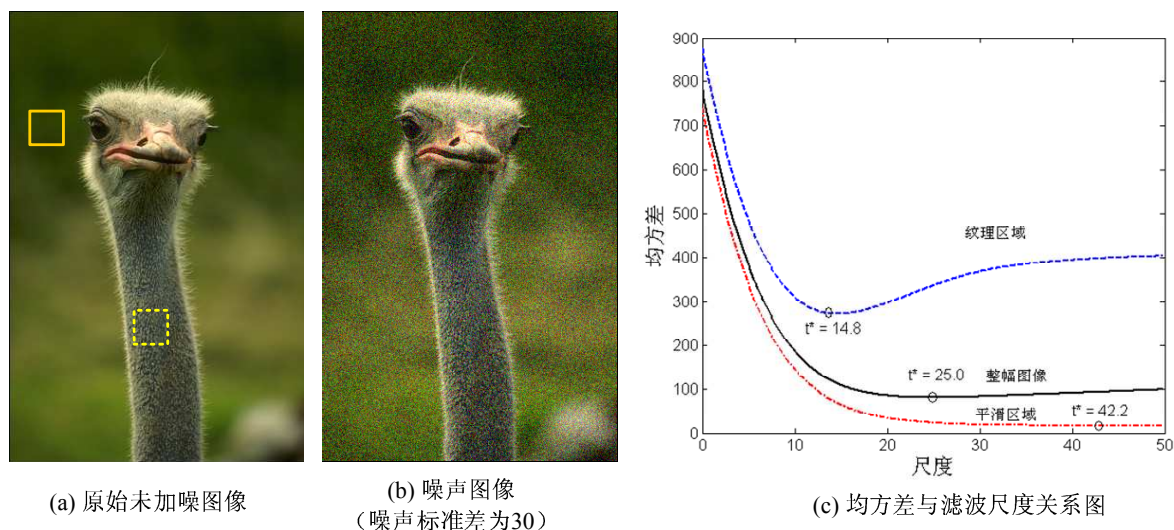


图 3-1 图像及其子区域的最优尺度。(a) 原始图像。(b) 噪声图像，噪声的标准差为30。(c) 滤波图像与原始图像之间的均方差相对于滤波尺度的变化关系。图(a)中的纹理区域（点线矩形边框内区域）、整个图像区域和平滑区域（实线矩形边框区域）的最优尺度（即均方差最小时所对应的尺度）分别为14.8、25.0和42.2。

为不同的图像块选择不同的尺度，因此最终的扩散图像为来自于不同最优尺度下的扩散图像块的组合，即整个扩散尺度空间的融合。这样的空间自适应选择策略有如下优点：纹理区域和光滑区域被分开为不同的图像块，并分别选取尺度，可克服为整幅图像选择同一尺度所引起的不同区域尺度选择的不均衡效应；其次，基于分割的图像表达可以帮助我们设计更好的噪声估计模型。

考虑相邻图像块之间的尺度一致性，我们可提出基于Markov随机场(Markov Random Field, MRF)光滑性先验的尺度选择模型。熟知，MRF模型是一种基于图的概率建模方法，可证明给定随机场为MRF等价于它满足Gibbs分布。近年来，该模型被广泛应用于图标记（graph labeling）问题以解决图像分割[30, 90–92]、数字蒙太奇（digital photomontage）[93]、立体视觉[94]、纹理合成[95]等计算机视觉问题。利用该框架，我们可以自然地将相邻尺度块之间的尺度一致性引入到模型中，即尺度选择模型倾向于对相似且相邻的图像块选择相同的尺度，同时允许为不相似且相邻的图像块选择不同的尺度。与以往的尺度选择策略相比，所提出的MRF尺度选择模型显著提高了各向异性扩散方程用于图像去噪的性能。

结合MRF尺度选择策略，我们提出新的基于各向异性扩散滤波的去噪算法，其流程如图3-2所示。首先，用各向异性扩散方程对图像进行逐步地扩散以得到扩散尺度空间，并对最大尺度 T 下的去噪图像进行过分割（即分割为小图像块集合），每个图像块具有相近的空间坐标和颜色值。然后用基于图像块的噪声估计算法估计噪声分布，进而用MRF尺度选择模型对图像进行尺度选择。最终输出的去噪结果为整个尺度空间的融合：即其中的每个图像块均来自于其最优尺度下的扩散图像中的相应图像块。应用于自然图像去噪的实验结果显示：各向异性扩散方程结合MRF尺度选择策略，可以更好地保持图像纹理，同时噪声可以被充分去除。以峰值信噪比（Peak Signal to Noise Ratio, PSNR）为测度，算法可以与目前最好的去噪算法可比甚至更好。

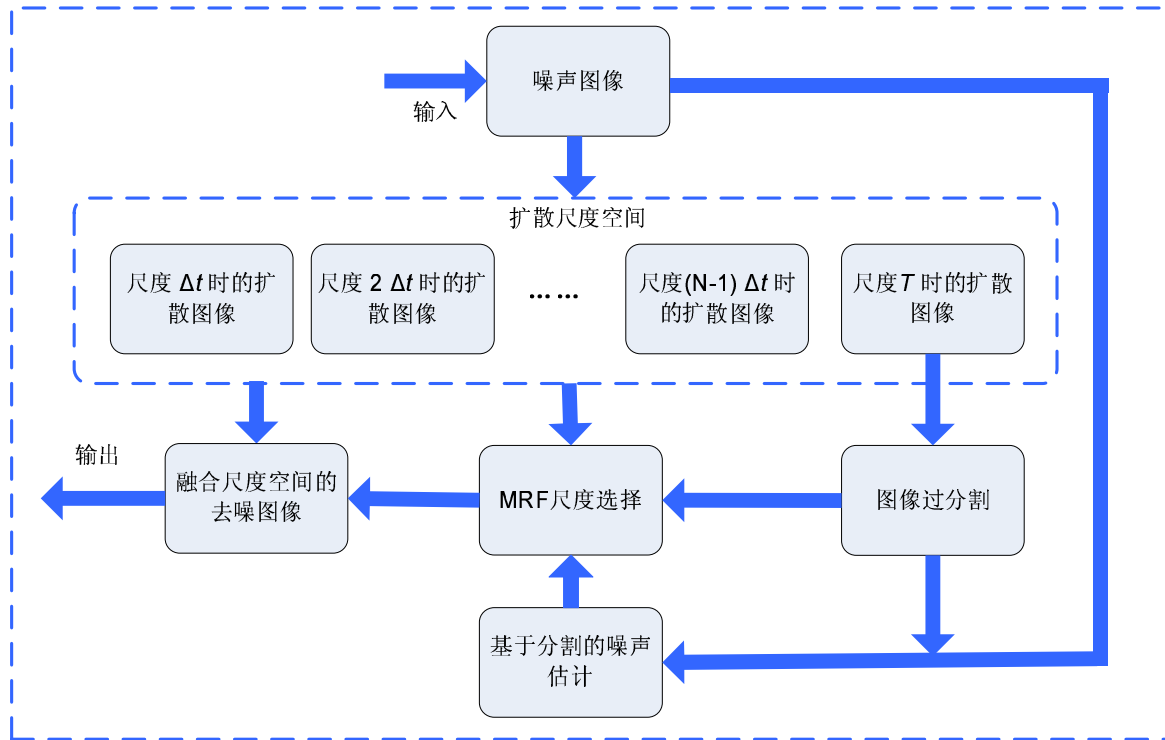


图 3-2 算法流程。首先将噪声图像进行扩散以生成扩散尺度空间，并将尺度 T 下的扩散图像进行分割，然后用基于分割的噪声估计模型进行噪声估计。应用该估计，用 **MRF** 尺度选择模型进行空间自适应的尺度选择，最终输出的去噪图像为整个尺度空间的融合，即每个图像块来自于其最优尺度下的扩散图像中的相应图像块。

我们以下将顺序介绍算法流程中的各个组成部分。第3.2节概述各向异性扩散方程的基本知识和各向异性扩散尺度空间的构造方法；第3.3节首先给出图像过分割算法，然后介绍每个分割块的尺度选择方法，最后介绍**MRF**尺度选择模型；第3.4节提出基于分割块的噪声估计算法；第3.5节提供与其他的尺度选择模型比较结果；第3.4节总结该工作。

3.2 各向异性扩散尺度空间

各向异性扩散方程可以看作为极小化如下能量泛函的梯度下降流：

$$E(I) = \rho(|\nabla I|^2). \tag{3-4}$$

其梯度下降流可以由如下的偏微分方程控制：

$$\frac{\partial I(x, t)}{\partial t} = \text{div}[\rho'(|\nabla I|^2)\nabla I]. \tag{3-5}$$

当 $g(|\nabla I|) = \rho'(|\nabla I|^2)$ 时，它与方程(3-1)完全一致，即为各向异性扩散方程。以输入的噪声图像为其初始条件，且使用反射边界条件。 $g(\cdot)$ 选择为递减函数，且满足：当 $r \rightarrow +\infty$ 时， $g(r) \rightarrow 0$ 。

随着尺度 t 的增加, 扩散方程实现了对图像的多尺度分析。相对于尺度 t 的扩散图像集合 $\{I(x, t)\}_{0 \leq t < \infty}$ 构成了一个尺度空间, 称为**各向异性扩散尺度空间**, 其关于尺度的极限扩散图像为常值图像或者分片光滑图像[96, 97]。

各向异性扩散方程的数值实现是如下的迭代过程:

$$\begin{aligned} I^{m+1}(i, j) = & I^m(i, j) + \tau[g(|\nabla I^m(i + 1/2, j)|) \times (I^m(i + 1, j) - I^m(i, j)) \\ & + g(|\nabla I^m(i - 1/2, j)|) \times (I^m(i - 1, j) - I^m(i, j)) \\ & + g(|\nabla I^m(i, j + 1/2)|) \times (I^m(i, j + 1) - I^m(i, j)) \\ & + g(|\nabla I^m(i, j - 1/2)|) \times (I^m(i, j - 1) - I^m(i, j))]. \end{aligned} \quad (3-6)$$

因此, 扩散图像 I^m 的尺度为 $t = m \times \tau$, 算法中通常取 $\tau = 0.15$ 。不断扩散输入噪声图像直到尺度 T 为止, 使得在该尺度下扩散图像过光滑(如图3-3(a)为例)。那么尺度 T 之前的扩散图像构成了一个各向异性扩散尺度空间。尺度 T 的选择标准是: 所在尺度下去除的噪声标准差为已知或估计的噪声标准差的1.1倍。在此, 我们用小波域噪声估计模型[72]进行噪声标准差估计, 而在第3.4节中将提出一种更为精确的噪声估计模型用于尺度选择。

进一步, 等间距地离散采样尺度区间 $[0, T]$ 可得到 N 个扩散图像, 其中尺度间隔为 $\Delta t = T/N$ 。进而构造一个空间和时间均为离散的扩散尺度空间:

$$S = \{I(x, t)\}_{t=\Delta t, 2\Delta t, \dots, N\Delta t}, \quad (3-7)$$

我们将以此为基础进行扩散尺度空间的尺度选择。

3.3 尺度选择模型

本小节将提出基于分割的尺度选择模型。其基本想法是: 首先, 将图像过分割, 即分割为小图像块集合, 使得图像块内像素具有相似的空间位置和颜色值; 然后为每个图像块设计尺度选择准则; 最终在图像块所构成的图结构上建立Markov随机场(MRF)尺度选择模型, 并用tree-reweighted消息传递(message passing)算法进行模型优化。

3.3.1 图像过分割

由于噪声的影响, 直接对输入的噪声图像进行分割并不稳健, 因此我们对尺度 T 下的扩散图像进行图像分割, 以消除噪声对分割的影响。图像像素点用5维特征(r 、 g 、 b 颜色通道和 x 、 y 位置坐标)进行表达, 然后用 k -means算法对图像进行聚类, 每一类即为一个分割块[98, 99]。为了初始化 k -means算法, 我们将图像初始分割为大小为 16×16 的矩形图像块集合, 然后迭代地更新聚类结果直到收敛。图3-3(b)显示了一个图像分割的例子。

记图像区域为 Ω_R , 每一个图像块为 S_i , $i = 1, \dots, K$, K 为分割块总数。显然, $\Omega_R = \cup_i S_i$ 。给定基于分割的图像表达, 我们将发展一个基于MRF的尺度选择模型。

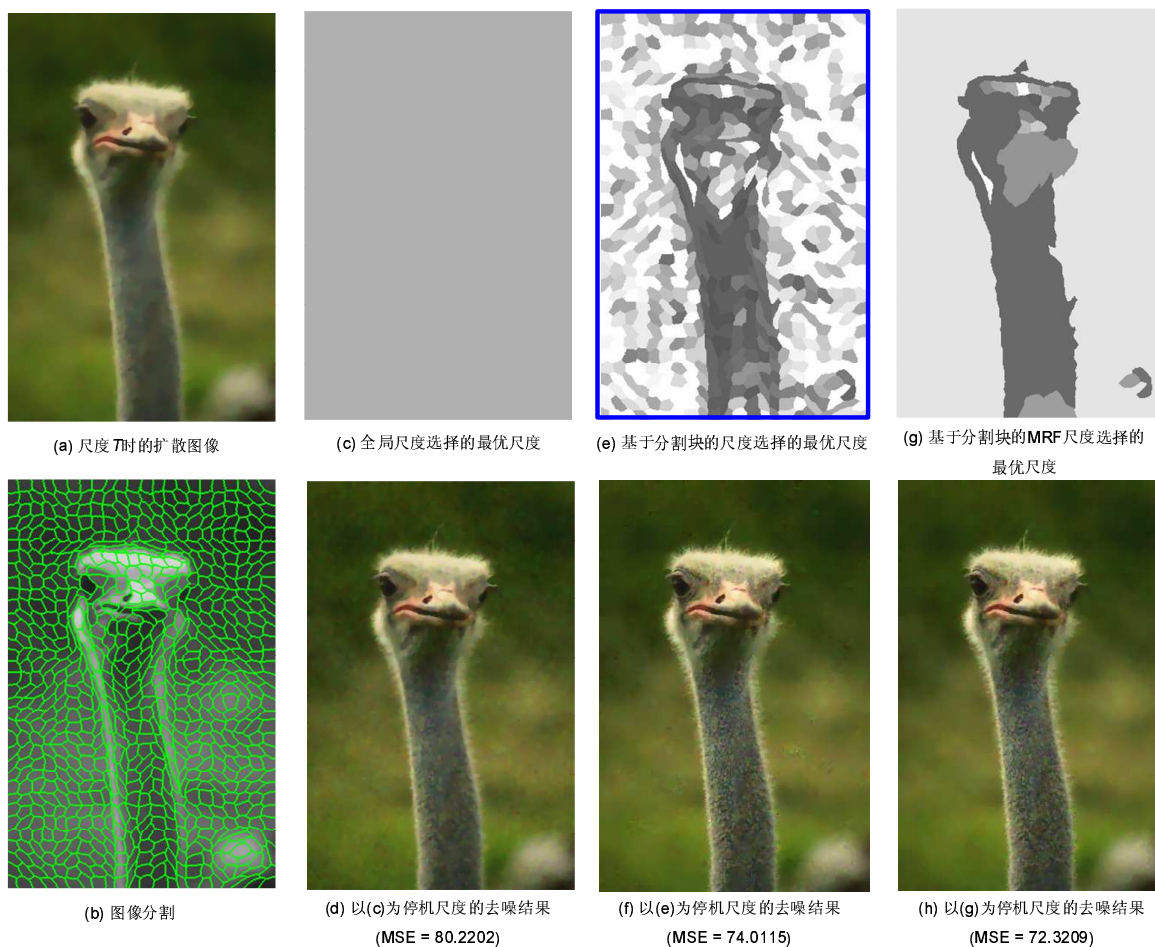


图 3-3 全局尺度选择策略和基于分割的尺度选择策略对比（所加噪声标准差为35，MSE为均方误差）。

3.3.2 图像块尺度选择策略

现在考虑如何为每个图像块选择尺度。由于图像块为关于象素颜色和位置特征的聚类块，因此每个块中的象素具有相似的颜色值，所以我们可以合理地假设每个图像块中的颜色是满足高斯分布的。

本章研究满足高斯分布的加性噪声，因此图像噪声模型为：

$$I_n = I^* + n, \tag{3-8}$$

其中， I^* 为原始的未加噪图像， I_n 为噪声图像， n 为加性的高斯噪声。在每个图像分割块 S_i 中，记 I_n 的方差和均值分别为 Σ 和 μ ，噪声 n 的方差为 Σ_n （本节中假设已知）， I^* 的方差和均值为 Σ_s 和 μ_s 。基于这样的图像噪声模型(如公式(3-8))，我们可计算出：

$$\Sigma = \Sigma_s + \Sigma_n, \tag{3-9}$$

$$\mu_s = \mu. \tag{3-10}$$

为了进行尺度选择，我们首先应该对给定尺度下的扩散图像的去噪性能进行度

量。为此我们首先分析尺度 t 时的扩散图像与原始未加噪图像之间的误差测度：

$$\text{Err}(I^*(x), I(x, t)) = \mathbf{E}\{(I^*(x) - I(x, t))^2\}, \quad (3-11)$$

其中， $E(\cdot)$ 为期望。尺度选择的目标即为极小化该误差测度。但是原始未加噪图像是未知的，所以我们分析扩散图像 $I(x, t)$ 和噪声图像 $I_n(x)$ 之间的误差测度：

$$\begin{aligned} \mathbf{E}\{(I(x, t) - I_n(x))^2\} &= \mathbf{E}\{(I(x, t) - I^*(x) + I^*(x) - I_n(x))^2\} \\ &= \mathbf{E}\{(I(x, t) - I^*(x))^2\} + \mathbf{E}\{(I^*(x) - I_n(x))^2\} \\ &\quad + 2\mathbf{E}\{(I(x, t) - I^*(x), I^*(x) - I_n(x))\}. \end{aligned} \quad (3-12)$$

在偏差/方差分析中， $I^*(x)$ 被假定为常数，因此上式的协方差项为0。与此不同，我们假设 $I^*(x)$ 在每个分割块中均为高斯分布的随机变量。因此为极小 $\mathbf{E}\{(I(x, t) - I^*(x))^2\}$ ，即希望：

$$\begin{aligned} &[\mathbf{E}\{(I^*(x) - I_n(x))^2\} - \mathbf{E}\{(I(x, t) - I_n(x))^2\}] \\ &\quad + 2\mathbf{E}\{(I(x, t) - I^*(x), I^*(x) - I_n(x))\} = 0. \end{aligned} \quad (3-13)$$

我们期望公式(3-13)中的两项均为0：

$$\mathbf{E}\{(I_n(x) - I^*(x))^2\} - \mathbf{E}\{(I_n(x) - I(x, t))^2\} = 0, \quad (3-14)$$

$$\mathbf{E}\{(I(x, t) - I^*(x), I^*(x) - I_n(x))\} = 0, \quad (3-15)$$

其中， $\mathbf{E}\{(I_n(x) - I^*(x))^2\} = \Sigma_n$ 。当去除的噪声 $I_n(x) - I(x, t)$ 分布满足已知的噪声分布（即 $N(0, \Sigma_n)$ ）时，第一项等于0。而当 $I(x, t) = I^*(x)$ 时，第二项等于0。但由于 $I^*(x)$ 未知， $I(x, t) = I^*(x)$ 是无法实现的，因此我们要求 $I(x, t)$ 与 I^* 的分布是一致的。我们将上述两种要求作为分割块的尺度选择准则，即要求在所选最优尺度下：

准则1：去除的噪声 $I_n(x) - I(x, t)$ 分布满足已知的噪声分布（即 $N(0, \Sigma_n)$ ）；

准则2：扩散图像块的颜色分布与原始未加噪声图像块颜色分布一致。

概括起来说，我们为每一个图像块选择尺度的准则为：

$$I_n(x) - I(x, t) \sim N(0, \Sigma_n), \quad (3-16)$$

$$I(x, t) \sim N(\mu_s, \Sigma_s), \quad (3-17)$$

其中， $\mu_s = \mu$ 且 $\Sigma_s = \Sigma - \Sigma_n$ 。我们用Kullback - Leibler (K-L) 距离测度估计的分布和目标分布的距离。以第二个准则为例：估计图像块 S_i 中的颜色和去除噪声的高斯分布参数为：

$$\hat{\mu}_s = \frac{1}{|S_i|} \sum_{x \in S_i} I(x, t),$$

$$\hat{\Sigma}_s = \frac{1}{|S_i|} \sum_{x \in S_i} (I(x, t) - \hat{\mu}_s)(I(x, t) - \hat{\mu}_s)^T,$$

其中， $|\cdot|$ 为像素个数。那么，在尺度 t 时，扩散图像的颜色分布与真实的颜色分布之间的K-L距离为：

$$\text{KL}_s(t_i; S_i) = \sum_{x \in S_i} [G(I(x, t); \hat{\mu}_s, \hat{\Sigma}_s) \log_2 \frac{G(I(x, t); \hat{\mu}_s, \hat{\Sigma}_s)}{G(I(x, t); \mu_s, \Sigma_s)}], \quad (3-18)$$

其中, $G(\cdot)$ 为高斯分布的概率密度函数。相似地, 我们可以为准则一一定义相应的距离测度 $\text{KL}_r(t_i; S_i)$ 。

将两个准则所对应的距离测度融合在一起, 即可得为图像块 S_i 选择尺度 t_i 的代价为:

$$D_i(t_i; S_i) = \frac{1}{2} \frac{\text{KL}_s(t_i; S_i)}{\sum_{t'} \text{KL}_s(t'; S_i)} + \frac{1}{2} \frac{\text{KL}_r(t_i; S_i)}{\sum_{t'} \text{KL}_r(t'; S_i)}. \quad (3-19)$$

因此, 图像块 S_i 的最优尺度应极小化 $D_i(t_i; S_i)$, 即

$$t_i^* = \operatorname{argmin}_{t_i} \{D_i(t_i; S_i)\}. \quad (3-20)$$

用上述公式为每个图像块选择最优尺度, 最终的去噪结果为所有来自其最优尺度下的扩散图像块的融合结果。

图3-3(e)和(f)为应用该尺度选择模型选择的尺度结果和相应的去噪结果。图3-3(c)和(d)显示了全局选择同一个最优尺度(即为整幅图像选择同一个最优尺度, 使得该尺度下扩散图像与原始未加噪图像最接近)的尺度选择结果。与全局选择同一尺度的结果相比, 我们的结果与原始图像的差别最小(以均方差测度)。

我们发现, 对每个图像块独立的选择尺度会产生尺度的不一致性, 即相邻且相似的图像块的尺度有较大差别, 如图3-3(e)所示, 在相同纹理内部, 相邻的图像块所选择的尺度差别较大, 会导致所得到的最终去噪结果在视觉感知上不自然。尺度不一致性产生的原因在于: 对图像块进行尺度选择的标准是测度去除噪声或者扩散图像颜色的分布与目标分布之间的相似性, 而一个图像块内往往只有几十或者几百个像素点, 从统计意义上说, 有限的样本无法估计出真实的分布, 因此会引起尺度选择的不准确。

3.3.3 Markov随机场尺度选择模型

为了克服独立为图像块选择尺度所带来的尺度不一致性, 我们提出如下的光滑性先验: 相邻且相似的图像块应具有相近的尺度, 但允许相邻但不相似的图像块具有不同的尺度。因此, 在图像同一纹理区域内部, 图像具有一致的尺度, 而在不同纹理区域具有不同的尺度, 从而与我们的视觉感知相一致。

我们将用Markov随机场 [100, 101] 模型对尺度选择进行光滑性约束。在该框架下, 相邻图像块之间的尺度一致性被建模并融入所建立的MRF能量模型中。MRF随机场尺度选择模型建立于图 $G = (\mathcal{V}, \mathcal{E})$ 上, 其中 $\mathcal{V} = \{v_1, v_2, \dots, v_K\}$ 为一系列的图结点, 本章假定每个图像块定义为一个图结点。因此, 每个图结点 v_i 代表一个图像块 S_i , 且图结点的总数为图像块的总数 K 。本章关注逐对Markov网络 (Pairwise Markov Network), 因此 \mathcal{E} 为一系列相邻图像对之间的邻接关系。该Markov图结构如图3-4。

将每个图像块 S_i (即图结点) 关联一个随机变量 X_i , 所有随机变量 $X = \{X_1, X_2, \dots, X_K\}$ 构成了一个随机场。每个随机变量 X_i 取尺度空间中扩散图像的尺度下标集合 $\{1, \dots, N\}$ 中的某一个值 t_i 。尺度标记集合 $t = \{t_1, t_2, \dots, t_K\}$ 为随机场 X 的一个观测, 在给定图像 I 的前提下, 其后验概率分布定义为:

$$\Pr(X = t|I) \propto \mathbf{P}(I|X = t)\mathbf{P}(t)$$

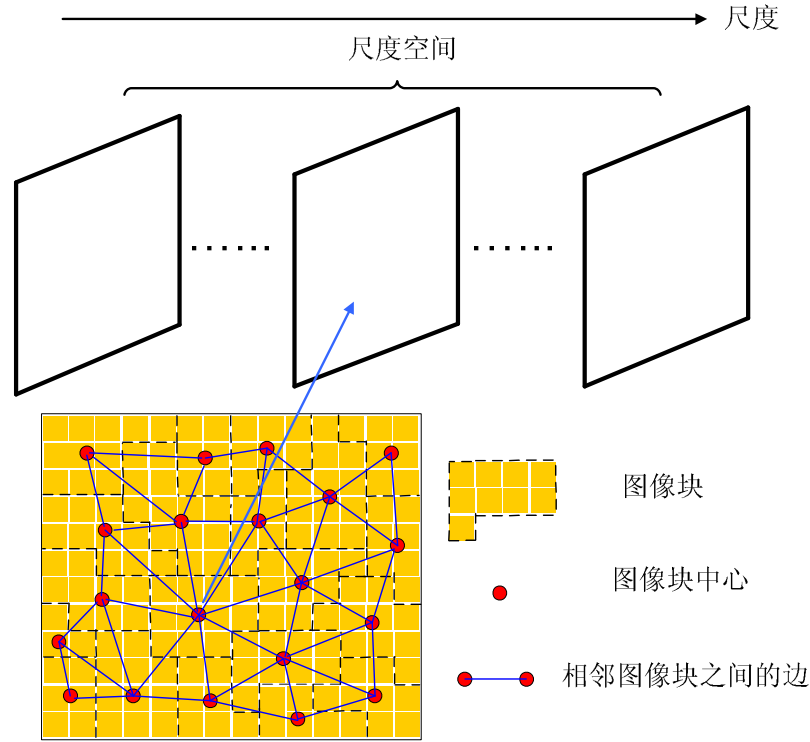


图 3-4 基于图像块的MRF图结构。图结点为图像分割块，图的边为相邻图像块之间的邻接关系。在尺度选择模型中，为每个图像块即图结点选择尺度空间中的某个尺度。

$$\propto \prod_{i \in \mathcal{V}} \exp(-D_i(t_i; S_i)) \prod_{(i,j) \in \mathcal{E}} \exp(-\lambda V_{i,j}(t_i, t_j; S_i, S_j)). \quad (3-21)$$

其中第一项为图像块尺度选择的似然项，第二项为尺度光滑性先验项。

我们将图像尺度选择问题建模为极大化该后验概率分布，这等价于极小化其负对数，即如下的能量函数：

$$E(t) = \sum_{i \in \mathcal{V}} D_i(t_i; S_i) + \lambda \sum_{(i,j) \in \mathcal{E}} V_{i,j}(t_i, t_j; S_i, S_j). \quad (3-22)$$

能量函数第一项定义为公式(3-19)，该项度量了将图像块 S_i 赋予尺度 t_i 的代价。第二项为相邻图像块之间的尺度光滑性测度，该项度量了相邻的图像块（ S_i 和 S_j ）的尺度（ t_i 和 t_j ）一致性。

对于相邻的图像块 S_i 和 S_j ，我们定义其尺度光滑性函数为：

$$V_{i,j}(t_i, t_j; S_i, S_j) = [1 - \delta(t_i, t_j)] \exp\left(-\frac{|\mu(S_i) - \mu(S_j)|^2}{2 \langle (\mu(S_i) - \mu(S_j)) \rangle^2}\right), \quad (3-23)$$

其中，

$$\delta(t_i, t_j) = \begin{cases} 1 & \text{if } t_i = t_j \\ 0 & \text{否则} \end{cases}. \quad (3-24)$$

$\mu(S_i)$ 为图像块 S_i 的平均颜色或者灰度值，且 $\langle \cdot \rangle$ 为期望。该定义的意义为：对两个相邻且相似的图像块 S_i 和 S_j ，约束它们的尺度具有强相似性；对两个相邻且相似的图

像块 S_i 和 S_j ，允许它们的尺度具有弱相似性。公式(3-23)中的指数项定义测度了图像块之间的相似性。类似的光滑性定义也被广泛应用于图像分割等问题[30, 90–92]。

这样，我们所提出的MRF尺度选择模型可最终定义为：

$$E(t) = \sum_i \left[\frac{1}{2} \frac{\text{KL}_s(t_i; S_i)}{\sum_{t'} \text{KL}_s(t'; S_i)} + \frac{1}{2} \frac{\text{KL}_r(t_i; S_i)}{\sum_{t'} \text{KL}_r(t'; S_i)} \right] \quad (3-25)$$

$$+ \lambda \sum_{j \in N(i)} \left[(1 - \delta(t_i, t_j)) \exp\left(-\frac{|\mu(S_i) - \mu(S_j)|^2}{\langle (\mu(S_i) - \mu(S_j)) \rangle^2}\right) \right].$$

通过在函数空间 $\{t = \{t_1, \dots, t_k, \dots, t_K\}; \forall k, t_k \in \{1, \dots, N\}\}$ 中极小化该能量函数，图像块 S_i 的最优尺度为 $t_i \times \Delta t$ 。那么在最终的去噪结果中，图像块 S_i 的内容来源于尺度空间中其最优尺度下的滤波图像的相应图像块。

能量函数(3-25)为MRF理论框架中典型的图标记问题（graph-labeling），该类问题被认为是NP-难的。近年来，人们提出了图切算法（graph cut）[102, 103]、有环置信传递算法（loopy belief propagation）[104, 105]和tree-reweighted信息传递（message passing）算法[106, 107]等，这些算法都可以有效、快速地求解该优化问题。实验证明，图切算法和tree-reweighted信息传递算法均可获得高精度的优化结果[108]。

本章中，我们使用tree-reweighted信息传递算法优化能量函数(3-25)。该算法的研究起源于Wainwright等人[106]对MRF能量函数的研究，他们得到了MRF能量函数的下界，并通过tree-reweighted max-product信息传递算法极大化该下界从而优化能量函数。该算法通过一系列的信息传递子过程以发现具有相同最优值的树结构分布的集合。Kolmogorov[107]改进和发展了该算法，在每步迭代中保证下界都会增加。我们使用该改进的tree-reweighted信息传递算法¹进行模型优化。

图3-3(g)和(h)显示了用MRF随机场尺度选择模型进行尺度选择的结果。由于考虑了相邻图像块的尺度一致性，我们得到了更加合理的尺度选择结果。它在图像纹理一致区域具有相似的尺度，并且在光滑区域具有较大的尺度，而在纹理强的区域具有较小的尺度。因为约束了尺度一致性，且分割块的边缘与图像边缘很好地对齐，因此最终的去噪结果很自然，很少引入幻影。最终的去噪结果相对于原始未加噪图像的均方差也进一步降低。

3.4 基于分割块的噪声估计模型

类似于基于信噪比的尺度选择方法，基于分割块的MRF尺度选择方法也假设噪声的方差是已知的。但在实际应用中，我们并不知道噪声的分布。于是，本小结中我们提出小波域中的基于图像分割块的噪声估计算法。

首先让我们回顾已有的噪声估计算法：一些商用软件（例如NeatImage）通过在图像中选取一块光滑的区域，然后将该区域的噪声方差作为整个图像的噪声方差；Donoho[72]提出了一种小波域的噪声估计方法，用小波分解系数（HH频段）的中值绝对方差（Median Absolute Deviation, MAD）估计噪声分布，该方法是图像噪声估计的一种常用方法。Stefano[109]提出了一种基于训练的小波域噪声估计模型，显然噪声估计的精度依赖于训练集合的选择；Liu等人[99]提出了一种基于分割的噪声估计

¹<http://www.adastral.ucl.ac.uk/~vladkolm/papers/TRW-S.html>

模型，但该模型侧重于处理相机CCD噪声，即与图像亮度相关的噪声。本小节，我们将发展一种简单但有效的噪声估计模型，实现在小波域中估计加性噪声的标准差。

首先介绍基于MAD的噪声估计方法。该方法用图像在小波分解域中的HH子带中的小波系数进行噪声估计，这是因为该子带中的小波系数为图像的高频信息，因此噪声大都被反映在该子带的小波系数中，而原始图像信息在该子带中反映的相对较少。那么噪声的标准差估计公式为：

$$\sigma = \text{MAD}(W)/0.6745, \quad (3-26)$$

其中， W 为HH子带中的小波系数。该方法的弱点是倾向于过高估计噪声的标准差，这是因为：自然图像的高频信息本身也会在HH子带中带来非零的小波系数。特别地，当噪声方差相对于图像高频信息的方差小得多的时候，噪声会被严重的过高估计。

MAD噪声估计算法可能带来噪声过估计的缺点可以用基于图像块的图像表达来解决。我们知道，自然图像的小波系数的边际分布（即直方图）满足非高斯的重尾分布[14, 18, 110]。该稀疏性先验意味着大量的小波系数值为0，而非0的只占小部分。因此我们可以将自然图像的高频部分去掉，而只用低频区域的小波系数进行噪声估计，从而得到更精确的噪声估计结果。

该噪声估计的想法可以通过如下的步骤实现：首先用MAD噪声估计方法对每个图像分割块进行噪声方差估计，将方差远远大于图像块的平均方差的图像块去除，然后将剩余的图像块的平均噪声方差作为图像的噪声方差。不断地迭代上述过程，直到收敛或者达到预设的最大迭代数（实现中设置为10）。基于图像块的小波域噪声估计算法可概括为算法3.1。

算法3.1：基于分割块的噪声估计算法

输入： 噪声图像 I_n 和它的分割块集合 $\{S_i\}, i = 1, 2, \dots, K$ 。

步骤1： 将图像用Daubechies小波进行小波变换，记其HH子带的小波系数为 W ；

步骤2： 用公式(3-26)为每个图像块 S_i 估计噪声方差 σ_i^2 。然后计算所有图像块噪声方差（即： $\{\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2\}$ ）的平均值 μ_σ^2 和标准差 τ_σ 。初始化 $\sigma = \mu_\sigma$ ，设 $l = 1, L = 10$ ；

步骤3： 迭代直到收敛或达到最大步数 L ：

步骤3.a： 设 $\sigma_{\text{old}} = \sigma$ ，将噪声方差 σ_i^2 满足 $|\sigma_i^2 - \mu_\sigma^2| > 2\tau_\sigma$ 的图像块去除，然后在剩余块上重新估计 μ_σ^2 和 τ_σ 。

步骤3.b： 设 $\sigma = \mu_\sigma, l = l + 1$ 。

输出： 输出估计的噪声方差 σ^2 。

3.5 实验

本小节，我们将用自然图像测试基于分割的MRF尺度选择模型。测试图像来源于

去噪的标准测试图像集合¹和Berkeley分割数据库²。Berkeley分割数据库是由人进行分割标记的自然彩色图像所组成。近年来，由于这些图像丰富的颜色和纹理，也被广泛应用于图像去噪 [17, 35, 99] 和填充[17]。图3-5列出了我们所采用的测试图像。

我们将比较MRF尺度选择模型与已有尺度选择策略的尺度选择性能，同时还将比较结合MRF尺度选择策略的各向异性扩散去噪与当前流行的去噪算法的性能。对于彩色图像，假设R、G、B通道的值是独立的，那么 Σ_s 和 Σ_n 为3-对角矩阵。

3.5.1 噪声估计比较

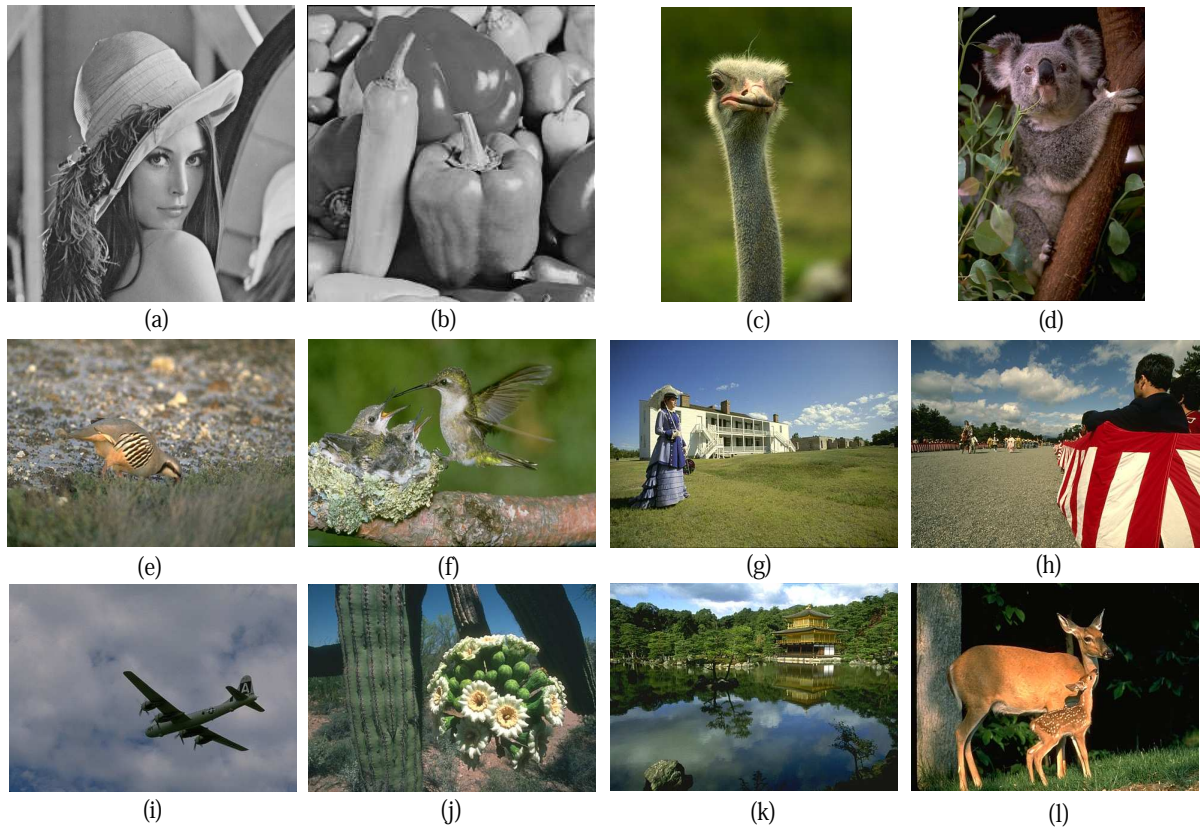


图 3-5 测试图像

首先评估基于分割块的噪声估计算法。为测度噪声估计的精确度，我们定义估计的噪声标准差与真实的噪声标准差之间的差异，差异用平均绝对方差（Mean Absolute Deviation, MAD）进行度量。当所加的真实噪声标准差为 σ_0 时，测试集上估计的噪声标准差的MAD误差测度定义为：

$$\text{MAD}(\sigma_0) = \frac{\sum_{i=1}^{12} |\sigma_i - \sigma_0|}{12}, \quad (3-27)$$

其中， σ_i 为测试图像集中第*i*个图像估计的噪声标准差。

我们比较基于分割块的噪声估计算法（记为Wav_seg）和经典的小波域MAD噪声估计算法(记为Wav)的性能。表3-1列出了在加不同噪声水平时（标准差从5到50）的噪

¹<http://decsai.ugr.es/~javier/denoise/index.html>

²<http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/>

声估计MAD误差测度值，图3-6为噪声估计误差图。由表3-1和图3-6可见，当噪声标准差增加时，基于分割块的噪声估计算法Wav_seg相对于Wav算法更加精确。因此，所提出算法显著地降低了噪声估计的误差，尤其是当加的噪声标准差比较小的时候。

表 3-1 噪声估计比较（估计误差以MAD值测度）

方法	噪声标准差								
	5	10	15	20	25	30	35	40	50
Wav	1.3536	1.0447	0.8296	0.6721	0.5504	0.4501	0.3730	0.3113	0.2303
Wav_seg	0.8210	0.6500	0.5480	0.4476	0.3595	0.2835	0.2246	0.2140	0.2198

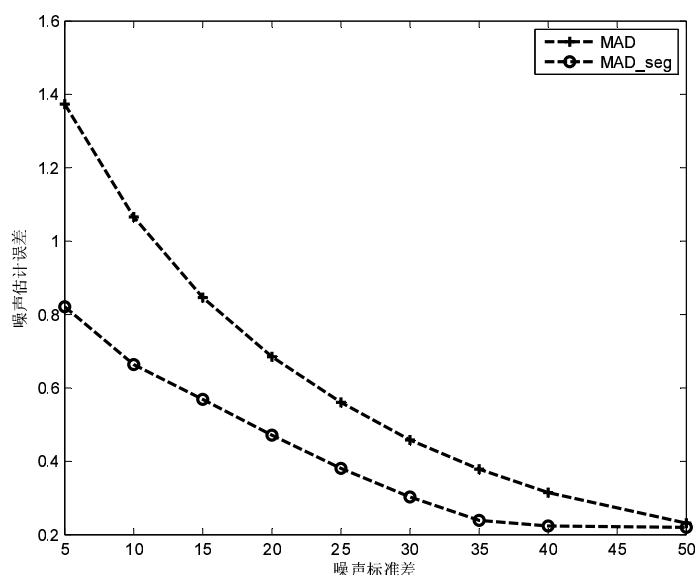


图 3-6 基于分割块的噪声估计模型(Wav_seg)与MAD噪声估计模型(Wav)的性能比较图。X-坐标为噪声标准差，Y-坐标为噪声估计误差（估计误差以MAD值测度）。

现在的问题是：用该基于分割块的噪声估计模型估计噪声分布的标准差对MRF尺度选择模型的性能影响情况如何？我们将在下小节中给出答案。

3.5.2 尺度选择比较

在本小节中，我们将比较MRF尺度选择模型与以往的全局尺度选择策略的性能。主要比较两种典型的全局尺度选择策略：基于信噪比的尺度选择算法（记为SNR）和基于相关性的尺度选择方法（Corr）。我们还将比较全局选择尺度的最理想情况（记为Opt），即选择扩散图像与原始未加噪图像最接近的尺度。显然，该情况为全局尺度选择的极限情况，因为实际中原始未加噪的图像是未知的，因此所有的全局尺度选择策略的尺度选择性能只能逼近该理想情况，而不可能优于该理想情况。为测

试所提出的尺度选择方法，我们将实验基于真实噪声方差的MRF尺度选择模型（记为MRF1）的性能和基于估计的噪声方差（噪声估计如第3.4节）的MRF尺度选择模型（记为MRF2）的性能。

实验中，采用各向异性扩散方程所广泛使用的两种典型扩散项 $g(\cdot)$ 的定义方式[7, 89],

$$\text{类型1: } g(|\nabla I|) = \frac{1}{\sqrt{|\nabla I|^2 + \varepsilon}}, \quad (3-28)$$

$$\text{类型2: } g(|\nabla I|) = \frac{1}{1 + |\nabla I|^2/Q^2}. \quad (3-29)$$

设置 $\varepsilon = 0.01$ ，使得第一类型扩散项的分母项非0。第二类型扩散项的参数 Q 设置如[7]。带有这两种扩散项的各向异性扩散方程等价于极小化能量函数(3-4)，其中

$$\rho(|\nabla I|^2) = |\nabla I|, \quad (3-30)$$

$$\rho(|\nabla I|^2) = \log(1 + \frac{|\nabla I|^2}{Q^2}). \quad (3-31)$$

以如下方式评价尺度选择策略的优劣：给定尺度选择策略下，用各向异性扩散滤波在所选择的最优尺度下的去噪图像与原始未加噪图像之间的均方差(MSE)测度尺度选择性能，其定义为：

$$\text{MSE}(I) = \frac{1}{|\Omega|} \sum_{p \in \Omega} (I(p) - I^*(p))^2, \quad (3-32)$$

其中， I 为去噪图像， I^* 为原始未加噪声图像， p 为图像域 Ω 中的一个象素点。MSE测度越低，意味着尺度选择的性能越高。我们还定义均方差比 R 来测度给定尺度选择方法与全局尺度选择的理想情况的性能比， R 定义为：

$$R = \frac{\text{MSE}_{\text{den}}}{\text{MSE}_{\text{opt}}}, \quad (3-33)$$

其中， MSE_{den} 为用给定尺度选择策略进行尺度选择的去噪均方差。 MSE_{opt} 为全局尺度选择策略理想情况下的去噪均方差。 $R < 1$ 意味着该尺度选择策略性能优于全局尺度选择策略的理想情况。显然，对于所有的全局尺度选择策略，其 $R \geq 1$ 。

表3-2和3-3显示了随着噪声水平（即噪声标准差）的增加，不同尺度选择策略的去噪图像的平均均方差。在每个噪声水平下，平均均方差定义为测试图像的去噪均方差的平均值。表3-2和3-3列出了各向异性扩散项为两种类型时的平均均方差值。在每个表中，我们比较了5种不同的尺度选择策略的性能，包括：SNR、Corr、Opt、MRF1和MRF2（方法说明见上文）。图3-7为平均均方差值图。图3-7的左图和右图分别为第一类型和第二类型扩散项的各向异性扩散滤波的尺度选择性能比较。表3-2、3-3和图3-7显示：首先，SNR和Corr尺度选择策略的平均均方差高于全局尺度选择策略的理想情况的均方差；其次，MRF尺度选择模型(MRF1和MRF2)与SNR和Corr的尺度选择策略相比具有更小的平均均方差；第三，MRF1和MRF2性能上都优于全局尺度选择的理想情况。

表3-4和3-5显示了在不同噪声水平下，分别具有第一类扩散项和第二类扩散项的各向异性扩散滤波的去噪结果的均方差比。我们实验和比较

表 3-2 对于第一类型扩散项的尺度选择平均均方差

方法	噪声标准差								
	5	10	15	20	25	30	35	40	50
SNR	16.258	44.144	73.950	103.556	131.045	157.090	182.585	206.726	250.402
Corr	55.094	64.255	79.535	124.524	139.796	154.457	168.390	181.588	220.706
Opt	12.412	32.656	53.715	73.988	93.084	110.977	127.791	143.627	172.796
MRF1	12.612	31.131	50.004	68.577	86.741	103.945	120.833	137.108	169.261
MRF2	13.396	31.719	50.504	68.741	86.369	103.627	112.421	137.394	169.930

表 3-3 对于第二类型扩散项的尺度选择平均均方差

方法	噪声标准差								
	5	10	15	20	25	30	35	40	50
SNR	13.999	37.491	70.640	106.106	137.615	172.433	208.219	244.051	317.195
Corr	16.457	42.785	71.849	117.948	141.083	166.934	187.550	201.046	293.102
Opt	12.323	34.521	59.715	87.827	111.811	137.992	164.459	191.419	245.767
MRF1	11.909	32.511	54.834	78.844	98.945	120.530	141.695	162.591	206.032
MRF2	12.138	32.653	54.842	76.836	98.403	119.639	140.629	161.689	208.264

了SNR, Corr, MRF1和MRF2尺度选择策略的尺度选择性能。每个表的最后一列列出了这些尺度选择策略在不同噪声水平下的平均均方差比。图3-8的左图和右图显示了具有类型一和类型二扩散项的各向异性扩散滤波结果的均方差比。从表3-4和3-5可观察到,对于第一类扩散项,SNR方法和Corr方法的平均均方差比为1.452和1.724;对于第二类扩散项,SNR方法和Corr方法的平均均方差比为1.256和1.191。这些值均大于1,这是因为这两种方法均为全局尺度选择策略,必定会比全局选择策略的最理想情况要差。但是,MRF1和MRF2两种尺度选择方法的平均均方差比均小于1。这意味着我们的MRF尺度选择策略(MRF1和MRF2)在平均性能上都会远远好于全局尺度选择策略的最理想情况。

我们还发现,MRF2方法(即用基于分割块的噪声估计算法估计的噪声方差作为输入的MRF尺度选择方法)的性能与MRF1方法(即用真实的噪声方差作为输入的MRF尺度选择方法)很接近。如表3-4和3-5所示,对于第一类扩散项的各向异性扩散滤波,MRF1和MRF2尺度选择的去噪结果的平均均方差比为0.954和0.958;而对于第二类扩散项的各向异性扩散滤波,MRF1和MRF2尺度选择的去噪结果的平均均方差

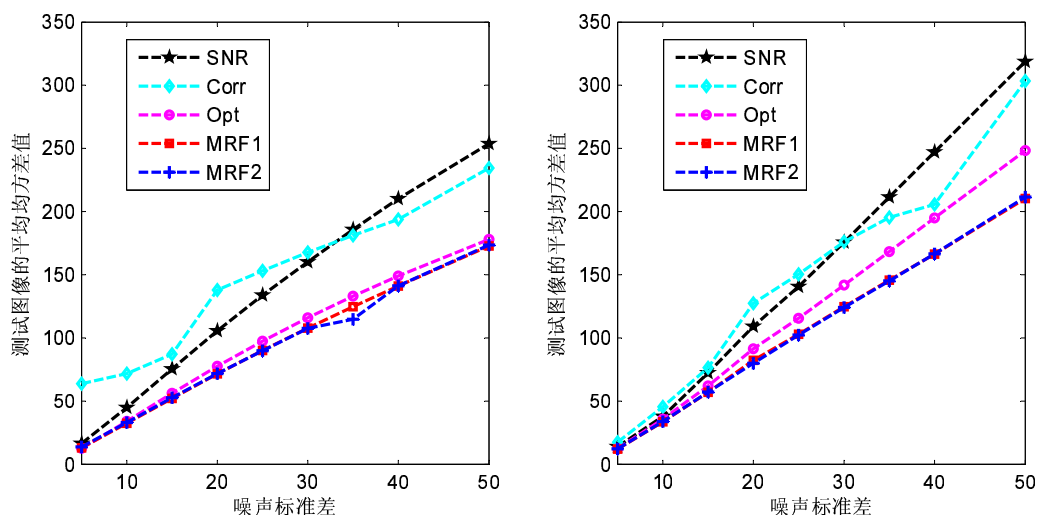


图 3-7 应用不同类型的尺度选择模型，各向异性扩散方程去噪结果的平均MSE误差。左图和右图分别为扩散项取为第一类型扩散项和第二类型扩散项的结果。

表 3-4 具有第一类扩散项的各向异性滤波的去噪结果平均均方差比

尺度选择方法	噪声水平									平均
	5	10	15	20	25	30	35	40	50	
SNR	1.366	1.421	1.441	1.466	1.466	1.469	1.476	1.482	1.483	1.452
Corr	3.979	1.793	1.391	1.669	1.486	1.375	1.300	1.246	1.276	1.724
MRF1	1.002	0.945	0.928	0.926	0.933	0.941	0.952	0.962	0.994	0.954
MRF2	1.053	0.959	0.935	0.930	0.933	0.940	0.900	0.972	1.003	0.958

表 3-5 具有第二类扩散项的各向异性滤波的去噪结果平均均方差比

尺度选择方法	噪声水平									平均
	5	10	15	20	25	30	35	40	50	
SNR	1.166	1.103	1.233	1.259	1.281	1.298	1.312	1.318	1.330	1.256
Corr	1.270	1.196	1.169	1.333	1.238	1.183	1.120	1.043	1.171	1.191
MRF1	0.967	0.945	0.919	0.898	0.882	0.871	0.860	0.846	0.837	0.892
MRF2	0.982	0.947	0.920	0.883	0.882	0.866	0.857	0.847	0.851	0.893

比为0.892和0.893。这意味着：由于噪声估计的误差所引起的尺度选择性能的下降是非常小的。用我们的噪声估计模型估计的噪声方差为输入，MRF尺度选择模型性能上

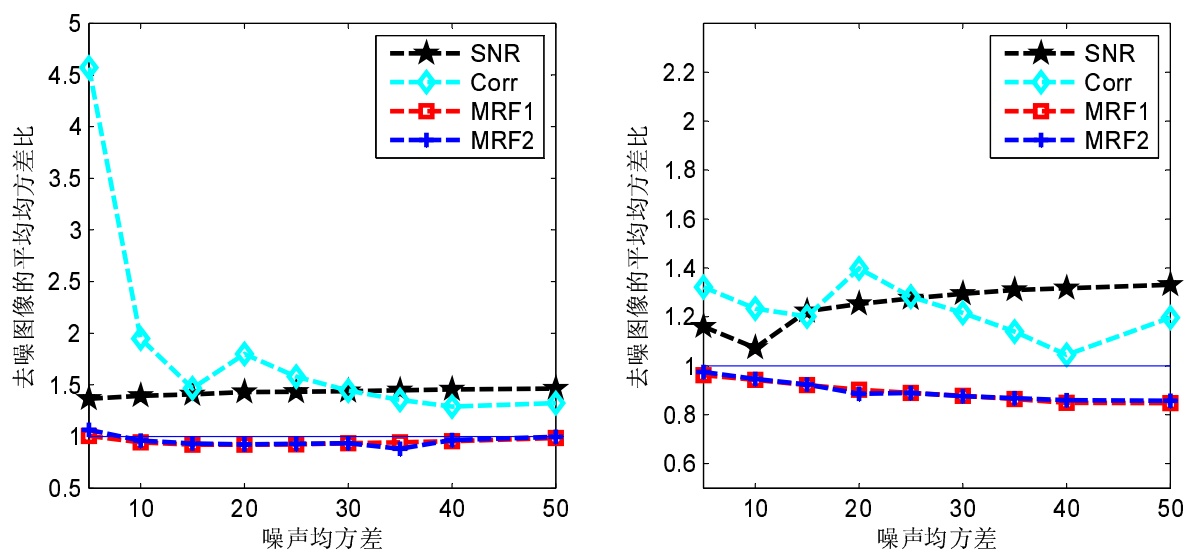


图 3-8 应用不同类型的尺度选择模型，各向异性扩散方程去噪结果的平均均方差比。左图和右图分别为扩散项取为第一类型扩散项和第二类型扩散项的结果。

还是远远优于全局尺度选择的最理想情况。

3.5.3 参数设定

本节将讨论MRF尺度选择模型的参数选择问题。第一个参数是尺度空间的滤波图像数 N ，第二个参数是正则化参数项 λ （公式(3-25)中）。

1) 尺度空间滤波图像数

首先讨论尺度空间滤波图像数 N 。设定 N 等价于设定尺度空间的连续两个滤波图像之间的尺度间隔 Δt ，它们之间的关系为 $\Delta t = T/N$ ，其中 T 的设置已经在第3.2节中讨论。

我们用测试集的噪声图像和原始图像对来设定合适的参数 N 。基本原则是：对于给定的噪声图像/原始图像对集合，尺度间隔 Δt 的选取应该能够区别噪声图像中的不同图像块的尺度差别，在此，图像块的最优尺度定义为所选尺度下去噪图像与原始未加噪图像在均方差测度下最接近。基于这个想法，我们在如下集合上建立了图像块对之间的尺度差的直方图：

$$\{|t_{gt}(S_i) - t_{gt}(S_j)|; i, j \in [1, K], i \neq j, \text{ 且 } t_{gt}(S_i) \neq t_{gt}(S_j)\}, \quad (3-34)$$

其中， $t_{gt}(S_i)$ 为图像块 S_i 的最优尺度。那么，我们设置 Δt 为直方图的bin，使得直方图在该bin值以下的积分为0.05。那么尺度空间滤波图像数 $N = \lceil T/\Delta t \rceil$ 。这样，噪声图像的图像块之间尺度差的95%均大于 Δt ，因此以此间隔构造的尺度空间可以区别绝大部分图像块之间的尺度差。

我们估计了测试集在不同噪声水平下的平均尺度间隔和平均滤波图像数，数据见表3-6。对类型一和类型二扩散项，平均尺度间隔随着噪声的增大而增大。但是平均滤波图像个数相对稳定于噪声水平，这是因为尺度空间最大尺度 T 也会随着噪声水平的

增大而增大。因此我们设定 N 为不同尺度和不同噪声水平下的平均滤波图像个数，对具有类型一扩散项的各向异性扩散滤波，平均滤波图像个数为35；对具有类型二扩散项的各向异性扩散滤波，平均滤波图像个数为41。

表 3-6 不同噪声水平下的平均尺度间隔和平均滤波图像个数

扩散项类型		噪声水平								
		5	10	15	20	25	30	35	40	50
类型一	平均 Δt	0.140	0.267	0.467	0.547	0.760	0.940	1.120	1.307	1.500
	平均 N	35	38	34	39	35	35	35	34	38
类型二	平均 Δt	0.072	0.126	0.176	0.216	0.288	0.324	0.376	0.400	0.410
	平均 N	42	47	45	42	38	38	38	38	44

我们研究了滤波图像数目 N 对尺度选择性能的影响。图3-9显示了MRF尺度选择性能（平均均方差）相对于 N 的关系。设定扩散方程具有类型一扩散项，且加噪声水平为30。当 N 从5增加到100时，去噪图像的平均均方差对 N 并不敏感，处于间隔[111.34, 112.78]之间。并且当 N 很大的时候，性能并不是一定增高，这是因为：大的 N 或者小的 Δt 会造成尺度选择结果对每个分割块尺度选择精度的高敏感性，这种现象类似于数据分类和回归中的过拟合现象。

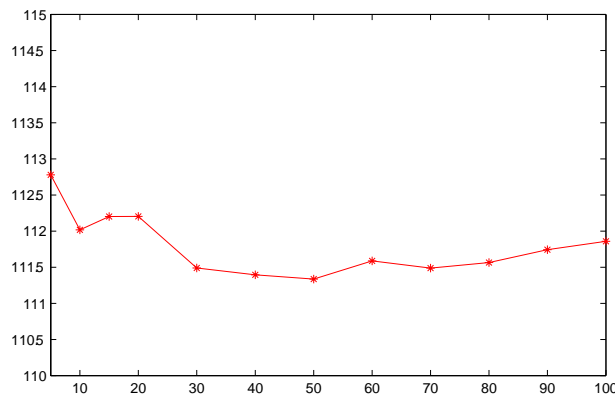


图 3-9 尺度空间滤波图像数目对去噪结果性能的影响。噪声标准差为30，扩散方程使用类型一扩散项。X-坐标为尺度空间滤波图像数目，Y-坐标为去噪图像的平均MSE。

2) 正则化参数 λ

我们进而分析如何设定公式(3-25)中的正则化参数 λ 。该参数平衡每个图像块的尺度选择和相邻图像块之间的尺度一致性。当设定小的 λ 时，会引起相邻尺度块之间的尺度不一致性，进而影响最终的去噪结果。如果设定大的 λ 将会产生更强的尺度一致

性，会使得两个尺度不同的图像分割块被赋予相同的尺度。问题是如何设置合适的正则化参数以更好地平衡两项之间的关系呢？

我们用如下公式选择不同噪声水平 σ 下的最优正则化参数 $\lambda^*(\sigma)$ ：

$$\lambda^*(\sigma) = \min_{\lambda} \left\{ \frac{1}{12} \sum_{i=1}^{12} \text{MSE}(I_i(x; \lambda, \sigma)) \right\}, \quad (3-35)$$

其中 $I_i(x; \lambda, \sigma)$ 为正则参数设为 λ 且噪声标准差为 σ 时，第 i 个测试图像的去噪结果。最优的参数 λ^* 选择为使得测试图像的去噪结果与原始图像之间的平均均方差最小的正则化参数。这种方法为经常使用的判别式参数设定方法。表3-7给出了在噪声标准差为2到70之间时的最优的正则化参数 λ^* 。显然，随着噪声标准差的增大， λ^* 值会逐渐变大。因此，对任意的噪声水平 σ_0 ，最优正则化参数 $\lambda^*(\sigma_0)$ 可以由表3-7中的数据进行内插或者外插得到。

表 3-7 不同噪声水平下的最优 $\lambda (\times 10^{-3})$

噪声水平	2	5	10	15	20	25	30	35	40	50	60	70
λ (类型一)	0.01	0.05	0.09	0.30	1.4	1.6	2.0	2.7	4.4	6.4	8.3	9.6
λ (类型二)	0.12	0.3	0.8	1.4	1.6	2.2	2.4	2.7	3.0	3.6	4.2	5.1

3.5.4 去噪效果比较

现在我们比较采用MRF尺度选择方法的各向异性扩散去噪方法与目前最流行的去噪算法的性能。

首先介绍目前几种流行的去噪算法。全变差去噪算法(Total variation denoising)[5, 6, 9]是一种经典且广泛使用的基于偏微分方程的滤波算法。它极小化图像全变差的同时约束去除噪声的统计量。Non-local denoising (NL)滤波算法[67]是最近发展起来的一种新的滤波算法。它通过邻域内图像块的带权平均值来估计中心像素点的颜色或灰度，而权重通过像素点周围的图像块的相似性来测度。实验中，我们设定用来求平均颜色（灰度）的像素区域为中心像素点周围 21×21 的窗口，而用来比较相似性的图像块大小为 7×7 ，且用于计算权重的高斯带宽参数为10倍的噪声标准差。小波域的高斯尺度混合去噪模型（Wavelet GSM）[14]使用高斯尺度混合模型来拟合小波系数的重尾分布，并通过贝叶斯最小二乘法恢复原始图像小波系数。我们使用steerable pyramidal小波分解（使用5个尺度）以得到图像的小波分解结果[14]。小波域高斯尺度混合模型去噪算法被认为是目前最好的图像去噪算法之一。

图3-10, 3-11和3-12展示了三个例子以比较不同去噪算法的性能。全变差算法(TV)的结果在纹理区域产生过模糊效应，比如图3-10中动物的身体部分、图3-11中的树林部分和图3-12中的草地部分，但是在光滑区域噪声没有被充分去除。Non-local denoising算法的结果在纹理区域和光滑区域都有幻影效果。小波域的高斯尺度混合模型可以很好的平衡平滑区域的去噪和纹理区域的保持细节的要求，但是在恢复的图像

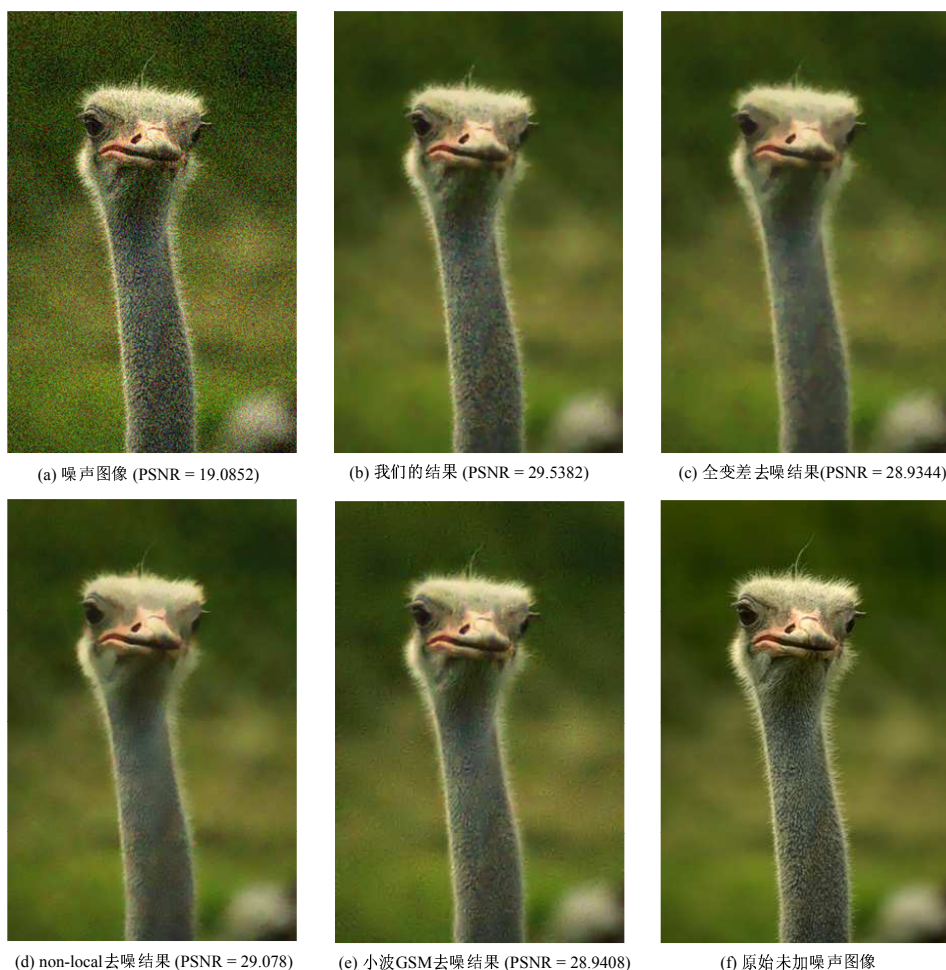


图 3-10 图像去噪结果比较（噪声图像的PSNR值为19.086）

中有明显的晕状幻影结构。对于我们的算法，即结合MRF尺度选择模型的各向异性扩散滤波，我们展示了图像过分割结果、尺度选择结果(越亮意味尺度越大)和最终的去噪结果。结果显示，所提出的方法可以在平滑区域充分去除噪声，同时在纹理区域很好地保持细节信息。

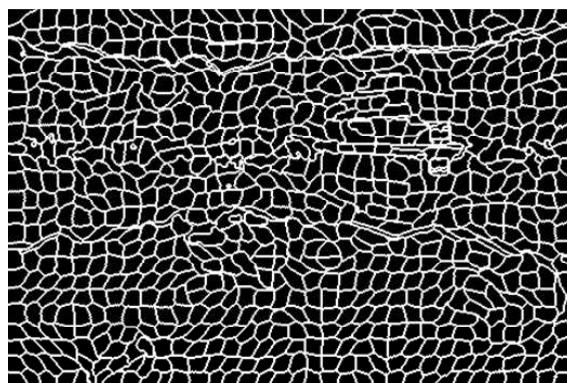
我们还用PSNR(Peak Signal-to-Noise Ratio)量化测度了在测试图像上的去噪性能。给定原始未加噪图像 I^* ，其噪声图像的去噪结果 I 的PSNR值定义为：

$$\text{PSNR}(I; I^*) = 10 \log_{10} \frac{255 \times 255}{\text{MSE}(I, I^*)}, \quad (3-36)$$

其中，MSE为均方差。表3-8给出了全变差（TV）去噪、Non-local滤波（NL）、小波域的高斯尺度混合模型去噪（wavelet GSM）和结合MRF尺度选择模型的各向异性扩散滤波（采用第一类扩散项）的去噪性能。表3-8的最后一列展示了这些去噪算法在测试图像上的平均性能。由表3-8可见：结合MRF尺度选择模型的各向异性扩散滤波和小波域的高斯尺度混合模型可以得到最好的去噪效果。当输入噪声图像的PSNR为22.11和19.08时，我们的方法在12个例子中有6个例子结果最好，且平均性能最高。全变差去噪算法可以视为一种各向异性扩散滤波方法，而尺度选择问题自然融合到了其能量约束中。与其相比，我们的尺度选择方法显著地提高了各向异性



(a) 噪声图像



(b) 分割结果



(c) 尺度图



(d) 我们的结果 (PSNR = 28.3360)



(e) 全变差去噪结果 (PSNR= 27.2463)



(f) non-local 去噪结果 (PSNR = 27.2672)



(g) 小波GSM去噪结果 (psnr = 28.2628)

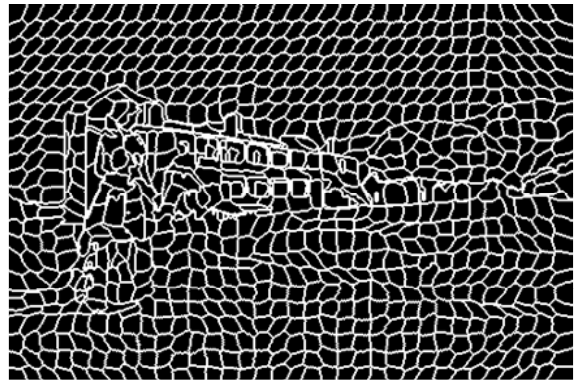


(h) 原始未加噪图像

图 3-11 图像去噪结果比较 (噪声图像的PSNR值为22.110)



(a) 噪声图像



(b) 分割结果



(c) 尺度图



(d) 我们的结果 (PSNR = 28.5537)



(e) 全变差去噪结果 (PSNR = 27.7884)



(f) non-local去噪结果 (PSNR = 28.0657)



(g) 小波GSM去噪结果 (PSNR = 28.6408)



(h) 原始未加噪图像

图 3-12 图像去噪结果比较 (噪声图像的PSNR值为22.110)

表 3-8 不同去噪算法的性能比较 (PSNR)

image	PSNR = 22.110				PSNR = 19.086			
	TV	NL	GSM	MRF	TV	NL	GSM	MRF
(a)	30.400	31.358	32.667	31.650	29.473	29.239	30.910	29.932
(b)	28.853	29.883	30.305	30.084	27.644	27.666	28.220	27.981
(c)	30.862	30.888	31.113	31.730	28.934	29.078	29.225	30.145
(d)	28.840	29.218	29.441	29.873	27.806	27.061	27.204	28.153
(e)	29.163	29.665	31.567	30.788	28.187	27.637	29.493	28.934
(f)	27.898	27.838	29.266	28.920	26.406	26.065	27.072	27.080
(g)	27.788	28.066	28.767	28.554	26.553	26.525	27.040	26.853
(h)	27.699	28.035	27.997	28.167	26.235	26.167	25.731	26.202
(i)	36.917	35.549	36.984	37.505	35.137	32.756	34.728	35.393
(j)	27.652	27.405	28.359	28.328	26.214	25.848	26.436	26.562
(k)	27.246	27.267	28.246	28.406	26.067	25.606	25.839	26.549
(l)	28.529	28.314	28.518	29.241	27.067	26.464	27.700	27.459
平均	29.360	29.456	30.269	30.271	28.052	27.509	28.300	28.436

扩散滤波的去噪性能，在噪声图像PSNR为22.110和19.08时，平均PSNR值分别提高了0.911和0.384。

3.6 小结

本章提出了一种空间自适应的各向异性滤波的尺度选择模型。在此模型下，最终的去噪结果为整个尺度空间的融合，因此该算法也可以视为尺度空间融合的一种有益探索。由于考虑了相邻图像块的尺度相似性这一视觉先验，我们设计了一种基于Markov随机场的尺度选择模型，并由tree-reweighted信息传递算法进行优化。在这个模型中，纹理区域具有较小的停机尺度，因此细节信息会比较好的保持；而在光滑区域具有较大的停机尺度，因此噪声会被充分去除。将算法应用于自然图像去噪，我们的尺度选择结果远远优于以往的全局尺度选择策略的结果，并且优于全局尺度选择结果的最理想情况。结合该尺度选择策略，各向异性扩散方程的去噪性能被显著提高，平均性能优于目前流行的算法结果，例如：non-local去噪算法、全变差算法和小波域高斯尺度混合算法。

尽管图像分割块可以很好地区别图像光滑区域和纹理区域，但一些小的结构（例如头发）还是很难用分割块表达，这使得我们的尺度选择方法不能很好地为小的结构

选择合适的尺度，进而影响最终结果，因此有必要研究更为细致的图像分割算法。另一方面，凡是具有尺度的去噪算法均可以构成不同意义上的尺度空间，可以应用类似的尺度选择模型解决其尺度选择问题，因此研究更为广泛的尺度空间尺度选择问题是我们进一步研究的方向。

4 基于图像块稀疏性先验的图像填充算法

4.1 引言

图像填充 (image inpainting) 的任务是: 将图像中丢失区域的信息填充完整, 使得填充的图像在视觉感知上具有较高的质量。图像填充是计算视觉和图像处理的一个基本任务, 主要应用包括: 老旧照片或者影片的修复、图像传输中丢失块的填充和图像中区域移除后的区域填充等。

目前应用中最基本的图像填充方法是基于扩散的填充方法 [3, 83, 111]。该类方法通过将缺失区域边缘的图像信息逐点扩散到缺失区域内部直到缺失区域被填充完整。该类方法具有较好的理论背景, 方法大多基于偏微分方程方法和变分方法。Bertalmio 等 [3] 提出沿着图像的 isophote (即等灰度/颜色值的轮廓) 从缺失区域边缘向内部扩散图像信息。进一步, 他们还将流体动力系统中的 Navier-Stokes 方程引入到图像填充问题中 [111]。Chan 和 Shen [83] 提出了一个基于全变差框架的填充算法, 图像缺失区域可以通过极小化图像全变差来进行填充。基于此框架, 他们进一步提出了曲率驱动的扩散方程, 该方程实现了视觉感知的连续性原理, 而该原理在全变差框架中并没有被考虑 [10]。Ballester [112] 提出对图像 isophote 方向和灰度值同时进行填充的办法, 也以变分方法为基础设计了能量函数。近年来, 人们还通过对自然图像进行统计建模以解决图像填充问题 [13, 17, 39]。基于扩散的填充算法在图像划痕去除、文字去除等方面已经取得了非常成功的结果。整体来说, 该类算法可以很好地应用于非纹理或者狭小缺失区域的填充, 但是在填充纹理区域或者大范围缺失区域时会引入光滑效应。

第二类方法是基于例子 (Exemplar) 的图像填充算法。从算法角度看, 该类算法也是一种基于扩散的填充算法, 不同之处是该类算法进行扩散的单位是图像块, 而不是图像像素点。该类算法的基本思想是: 以图像块为单位, 将图像已知区域的信息不断地扩散到图像的未知区域, 图像已知区域的图像块被称为例子 (exemplar)。该算法思想起源于纹理合成的工作 [113]: 纹理可以通过在已知区域中采样最相似的纹理块进行合成。该算法对纯粹的纹理合成效果很好, 但很难处理既有纹理又有结构的区域。因此, Bertalmio 等人 [114] 提出了将图像分解为纹理和结构两层, 然后对结构层用基于扩散的方法进行填充, 而对纹理层用纹理合成的方式进行填充。该算法可以克服基于扩散算法的过光滑效果, 但是很难填充大的缺失区域。Criminisi 等人 [115] 设计了一种基于例子的方法, 该算法结合了基于扩散和纹理合成的算法思想。为了使算法能够填充既有纹理又有结构的区域, 该算法定义了图像块优先度, 处于图像结构上的图像块具有高优先度, 因此更先被选择进行填充。Wu 等人 [116] 受到各向异性扩散滤波的启发, 提出一种 cross-isophotes 的基于例子的图像填充算法。Wong 等人 [117] 提出了一种基于 non-local 均值的算法, 可以更稳健、精确地填充图像。该算法用待选例子的 non-local 均值 [67] 对图像块进行填充, 而不是选择一个最匹配的图像块。人们还设计了更多类型的基于例子的填充算法 [118, 119]。与基于扩散的填充算法相比, 该类算法的优势在于可以很好地填充大范围的图像缺失区域。

最近, 图像的稀疏表达被引入到图像填充问题中 [120–123]。基于稀疏性先验的图像填充算法的基本想法是将图像用一个过完全的变换集 (over-complete transforms) 进

行稀疏线性表达，变换包括小波变换、contourlet变换、DCT变换等。然后缺失图像信息可以通过自适应更新稀疏表达来实现填充。Guleryuz [121–123]提出了自适应的图像稀疏表达，并成功应用于图像划痕、丢失块的填充。Elad 等人 [120]改进了该方法，将图像分解为纹理和结构两部分，分别用两个不相容的过完全变换集合进行稀疏表达。该算法可以比较有效地填充既有纹理又有结构的图像区域，但是与基于扩散的算法相似，会在大范围缺失区域填充时引入过光滑效果。

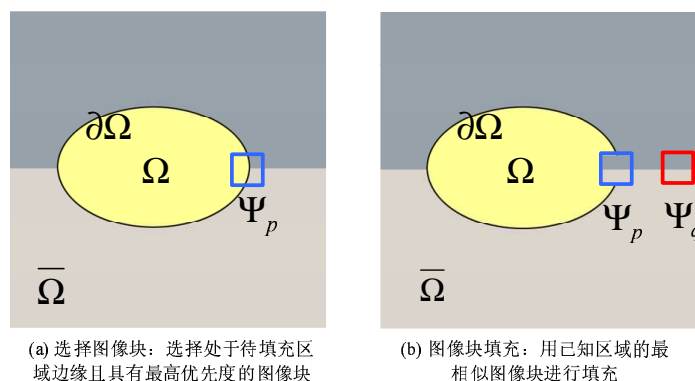


图 4-1 经典的基于例子的图像块扩散填充算法。 Ω 为待填充区域， $\bar{\Omega}$ 为已知区域，从待填充区域边缘 $\partial\Omega$ 开始，以图像块为单位逐渐向图像内部进行填充。填充过程为迭代以下两个子步骤：(a)选择图像块，即选择处于待填充区域边缘且具有最高优先度的图像块 Ψ_p 进行填充；(b)图像块填充，即用已知区域中与所选图像块最相似（已知像素的相似性）的图像块 Ψ_q 对所选图像块中的未知像素进行填充。

我们主要关注基于例子的图像填充算法研究，该类算法的基本步骤如图4-1所示。所涉及的主要问题是：

- 如何设计图像块优先度，以控制处于待填充区域边缘图像块被选择填充的顺序？
- 如何用已知图像区域中的例子图像块对待填充图像块进行填充？

经典的基于例子的图像填充算法 [115, 124]采用基于isophote的优先度定义以控制图像块的选择次序，具体地，用图像块中心点所在的待填充区域边缘法线方向与图像isophote（等颜色/灰度的轮廓）方向的内积测度图像的优先度；并采用最匹配的例子图像块 [124]或者多个例子图像块的带权平均 [117]进行图像块填充。

为了更好地解决基于例子的图像填充算法的两个基本问题，我们将定义两类图像块的稀疏性先验，并分别用于解决图像块优先度定义和图像块填充，据此设计一种新的基于例子的图像块扩散填充算法。两类图像块的稀疏性先验为：

- 结构稀疏性：通过定义图像块与其周围图像块的相似性稀疏程度来度量其处于图像结构上的置信程度。结构稀疏性被用于定义图像块的优先度，结构稀疏性强的图像块具有较高的可能性处于图像结构上，因此被优先选择。我们还将说明该定义的合理性。

- 图像块稀疏表达：通过已知区域图像块（**exemplar**）的稀疏线性组合来填充所选择图像块。我们约束该线性组合为稀疏组合，即仅有非常少的几个图像块的线性组合系数非零。该方法即提高了填充的稳健性，同时避免了填充结果中可能产生的光滑效应。

我们将结构稀疏性和图像块稀疏表达统一称为**图像块稀疏性**。与以往的基于扩散和稀疏性的填充算法相比，基于例子的填充算法可以很好地填充大范围的缺失区域，且结果具有较高锐化度。与传统的基于例子的图像块扩散算法相比，我们的算法可以更好地填充图像区域，且填充结果与其周围已知区域的图像内容更为一致。

结构稀疏性的想法来源于近年来人们对自然图像的稀疏性先验大量的相关研究工作。我们知道，自然图像对滤波（梯度滤波、小波变换等）的响应直方图具有很强的稀疏性，表现为非零的响应非常稀疏，即自然图像滤波响应的稀疏性。这类稀疏性先验可以很好地描述自然图像的一般性统计规律，被广泛应用于图像去噪 [13, 125]、超分辨率 [35, 38]、填充 [13, 17, 39]、去模糊 [12, 40]等。图像块稀疏性与此类似，它描述了另一类自然图像的稀疏性先验：即处于图像结构上的图像块与其相邻图像块的非零相似性具有强稀疏性，即非零的相似性非常稀疏。因此，结构稀疏性通过度量图像块与其相邻图像块的相似性的稀疏程度来反映其是否处于图像结构上。

图像块稀疏表达的思想来源于近年来人们对信号稀疏表达 [18]的大量相关研究工作。稀疏表达已成为图像、信号处理关注的热点，它假定图像、信号可以由一个过完全的基/变换集合的稀疏线性组合来表达，其组合系数假定满足 $\ell^p(0 \leq p \leq 1)$ 的稀疏性约束。稀疏表达具有较好的生物解释：即大脑的视觉皮层细胞集合可以看作为一个过完全的基/变换集，其对视觉信息的编码具有很强的稀疏性。稀疏表达框架被广泛应用于图像去噪 [24, 126]、边缘提取 [20]、识别 [127, 128]、超分辨率 [25]、纹理合成 [129]等，并取得了目前最好的结果。在图像块稀疏表达中，我们将稀疏表达的思想引入到基于例子的图像填充问题，即假定自然图像块可由已知区域图像块的稀疏线性组合表达，并基于问题建立了约束优化模型和优化算法。

以下在第4.2节，我们将整体介绍所提出的填充算法框架；第4.3节和第4.4节分别阐述填充算法框架中所涉及的两个具体问题：结构稀疏性和图像块稀疏表达；第4.5节提供实验说明新算法的有效性以及与已知算法的比较；第4.6节总结本章工作。

4.2 算法框架

给定图像 I ，假定其缺失区域为 Ω ，已知区域为 $\bar{\Omega}$ ， $\partial\Omega$ 为缺失区域的边缘，记 Ψ_p 为以像素点 p 为中心的图像块。

本章将提出的图像填充算法框架如图4-2所示，算法本身可描述为一迭代算法，每次填充过程可分为：待填充图像块选取和图像块填充，直到缺失区域被填充完整。

待填充图像块选取，即在缺失区域的边缘上选择某个图像块进行待填充。图像优先度设计方式应使处于图像结构上的图像块先被填充，而位于纹理内部的图像块后被填充，这样既有纹理又有结构的图像缺失区域可以被完美地填充 [124]。我们提出图像块的结构稀疏性建模以定义图像块的优先度。结构稀疏性可以鲁棒地测度图像块是否处于图像结构，具体定义为图像块与其相邻图像块相似性的稀疏程度。结构稀疏性

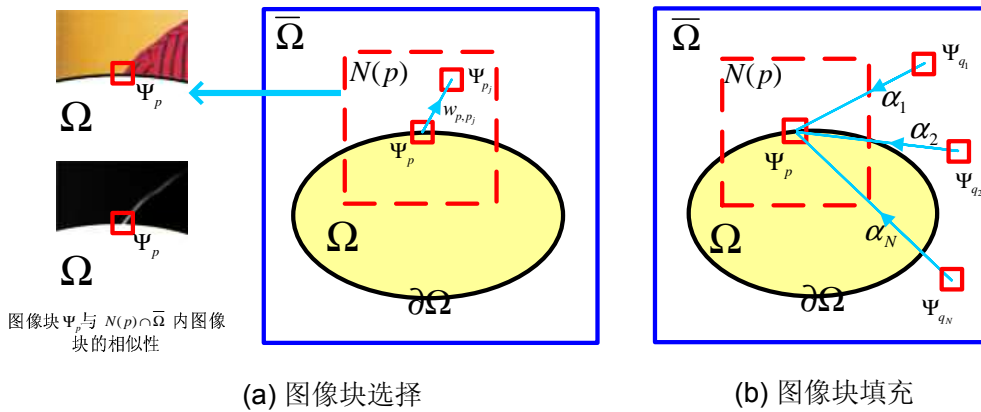


图 4-2 基于图像块稀疏性的填充算法框架。 Ω 为待填充区域， $\bar{\Omega}$ 为已知区域， $\partial\Omega$ 为缺失区域的边缘。(a)图像块选择：处于 $\partial\Omega$ 上的图像块优先度越高则越被优先选择。优先度通过图像块处于图像结构上的置信度来定义，定义为图像块与其相邻且处于 $\bar{\Omega}$ 区域的图像块的相似性的稀疏程度。例如，(a)中左图显示 Ψ_p 与图像块集合 $\{\Psi_{p_j}\}_{p_j \in N(p) \cap \bar{\Omega}}$ 中的图像块的相似性具有很强的稀疏性，这是因为它处于图像的边缘结构上。(b)图像块填充：对于所选图像块 Ψ_p ，用已知图像块 $\{\Psi_{q_1}, \Psi_{q_2}, \dots, \Psi_{q_N}\}$ 的稀疏线性组合来填充 Ψ_p 中的未知像素，该填充模型还受到其局部邻域 $N(p) \cap \bar{\Omega}$ 中的图像内容的约束。

越强，处于图像结构上的可能性越大，因此被选择的优先度越高。如图4-2(a)左上图为例：图像块 Ψ_p 位于图像的衣服纹理和背景纹理之间的边缘上，显然处于强结构之上。通过计算该图像块与其相邻图像块的相似性（如图4-2(a)的左下图所示， Ω 为已知区域，剩余区域越亮代表相似性越强），显然相似性图具有很强的稀疏性，即只有稀疏的非零元素存在。因此该图像块具有较高的优先度，应被优先选取进行填充。

图像块填充，即填充上一步所选的待填充图像块内的缺失信息。图像块填充的基本思想是用已知区域中的例子图像块对待填充图像块进行填充。我们提出图像块稀疏表达模型，即假定待填充图像由已知区域中的例子图像块集合的稀疏线性组合来表达，我们通过约束组合系数为最稀疏形式（ ℓ^0 正则化约束）且待填充图像块与相邻的已知图像块具有一致性。如图4-2(b)所示，待填充块 Ψ_p 用例子图像块集合 $\{\Psi_{q_k}\}_{k=1}^N$ 的稀疏线性组合进行表达，组合系数为 $\{\alpha_k\}_{k=1}^N$ 且仅有稀疏的非零元素存在，并约束所填充的图像块应与其周围的图像块（属于集合 $N(p) \cap \bar{\Omega}$ ）的颜色相近。

在下一节中，我们将详细介绍如何定义图像块优先度和图像块稀疏表达。

4.3 基于结构稀疏性先验的图像块优先度

图像块优先度的定义对于填充既有纹理又有结构的图像区域至关重要。优先度定义的原则是：先结构后纹理，即处于图像骨架部分（即图像结构）的图像块被优先选择和填充，而处于纹理内部的图像块具有较低的填充优先度，这样既有纹理又有结构的图像缺失区域可以被很好地填充完整 [115, 124]。

4.3.1 结构稀疏性

图像块优先度定义应能鲁棒地刻画图像块处于图像结构上的置信度。本章中，我们定义了基于结构稀疏性的优先度定义，可以很好的满足要求。由于自然图像中的结构相比纹理具有更高的稀疏度，而处于结构（或纹理）上的图像块的相邻且相似的图像块也一定处于同一结构（或纹理）上，因此可以通过测度图像块与相邻图像块的非零相似性的稀疏程度来反映该图像块处于图像结构上的置信度，即结构稀疏性的基本思想。

我们首先说明如何度量图像块相似性。以图4-2(a)为例， Ψ_p 为待填充区域边缘 $\partial\Omega$ 的图像块，假定 P 为对角矩阵，提取图像块的缺失信息。记 Ψ_p 为列向量，用来存储图像块内的灰度/颜色值，那么 P 定义为对角矩阵： $P = \text{diag}(\{a_{i,i}\}_{i=1}^M)$ ，其中 M 为图像块维数， $a_{i,i}$ 定义为：

$$a_{i,i} = \begin{cases} 1, & \text{第}i\text{个像素为待填充像素,} \\ 0, & \text{否则.} \end{cases} \quad (4-1)$$

$$(4-1')$$

类似地，定义 \bar{P} 为提取图像块的已知信息的对角矩阵，满足： $\bar{P} + P = I$ 。

定义图像块 Ψ_p 的相邻图像块 Ψ_{p_j} 为中心点 p_j 属于如下集合的图像块：

$$N_s(p) = \{p_j : p_j \in N(p) \text{ 且 } \Psi_{p_j} \subset \bar{\Omega}\}, \quad (4-2)$$

其中， $N(p)$ 为像素 p 的邻域，那么 Ψ_p 和 Ψ_{p_j} 的相似性为：

$$w_{p,p_j} = \frac{1}{Z(p)} \exp\left(-\frac{\mathbf{d}(\bar{P}\Psi_p, \bar{P}\Psi_{p_j})}{\sigma^2}\right), \quad (4-3)$$

其中， $\mathbf{d}(\cdot, \cdot)$ 表示向量之间的均方距离， $Z(p)$ 为正规化参数，使得 $W = \sum_{p_j \in N_s(p)} w_{p,p_j} = 1$ ， σ 设置为5.0。

给定图像块的相似性，进一步的问题是如何度量 Ψ_p 与 $\{\Psi_{p_j}\}_{p_j \in N_s(p)}$ 的非零相似性的稀疏程度。对于图像块 Ψ_p ，我们定义它与相邻且已知图像块相似性的稀疏程度为：

$$\rho(p) = \sqrt{\frac{[\sum_{p_j \in N_s(p)} w_{p,p_j}^2] \cdot |N_s(p)|}{|N(p)|}}, \quad (4-4)$$

这里， $|\cdot|$ 为元素个数。为方便表达，我们称该稀疏度度量为**结构稀疏性**。在该表达中，相似性平方和越大或者区域 $N_s(p)$ 越大，结构稀疏性越大。下面将证明，该定义可以很合理地度量非零相似性的稀疏程度。

定理 4.1 图像块 Ψ_p 的结构稀疏性 $\rho(p)$ 当在 $N_s(p)$ 中仅存在一个图像块与其相似性非零时取得最大值 $\sqrt{\frac{|N_s(p)|}{|N(p)|}}$ ；而当 Ψ_p 与 $N_s(p)$ 中所有图像块相似性均非零且均等于 $\sqrt{\frac{1}{|N(p)|}}$ 时，取得最小值 $\frac{1}{|N_s(p)|}$ 。

证明 可以观察到， $\rho(p)$ 相对于 $W = \sum_{p_j \in N_s(p)} w_{p,p_j}^2$ 单调递增。为找到 $\rho(p)$ 的极大、极小值，我们仅需要计算 W 的极大、极小值。

首先计算 \mathbf{W} 的极小值。问题可以转化为在约束 $\sum_{p_j \in N_s(p)} w_{p,p_j} = 1$ 的前提下极小化 W ，根据拉格朗日乘子法理论，极大化：

$$E(p) = -W + \lambda \left(\sum_{p_j \in N_s(p)} w_{p,p_j}^2 - 1 \right) \quad (4-5)$$

容易证明，对任意 $p_j \in N_s(p)$ ， $w_{p,p_j} = 1/|N_s(p)|$ 时， W 取到极小值。

然后考虑极大化 W 。由于 $0 \leq w_{p,p_j} \leq 1$ 一定成立，因此 $W \leq 1$ 。等式成立当且仅存在一个 w_{p,p_j} 非0（即等于1），而其他所有相似度都等于0。因此当仅存在一个 $p_j \in N_s(p)$ 使得 w_{p,p_j} 非零时， W 达到极大值。

将 W 的极大、极小值代入 $\rho(p)$ ，以得到 $\rho(p)$ 的极大、极小值。（证毕）

定理4.1说明：当图像块的非零相似性的分布最稀疏和最不稀疏时，结构稀疏性值分别达到最大和最小值。

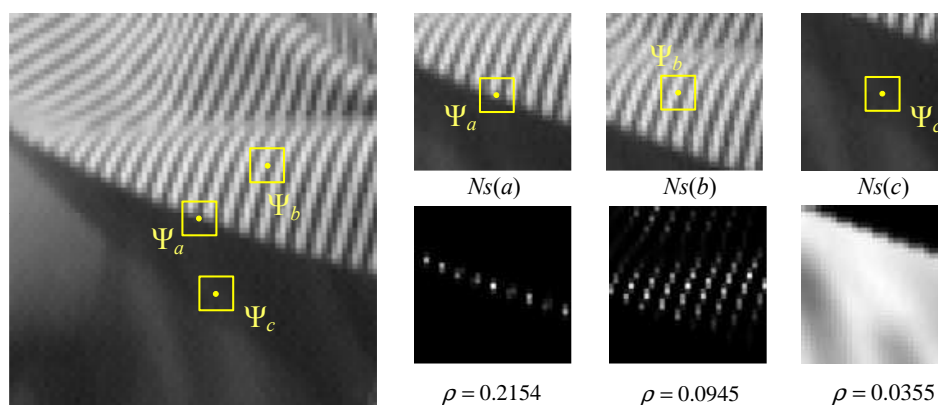


图 4-3 结构稀疏性。分别计算了图像 Ψ_a 、 Ψ_b 和 Ψ_c 的结构稀疏性。图像块大小和局部窗口的大小为 7×7 和 51×51 。 Ψ_a 、 Ψ_b 和 Ψ_c 的结构稀疏性大小为0.2154、0.0945和0.0355，位于结构上的图像块 Ψ_a 的结构稀疏性最高。

以图4-3为例，我们计算了其中图像块 Ψ_a 、 Ψ_b 和 Ψ_c 与其相邻图像块的相似性。结果见图4-3右三列的下图，其中越亮意味着相似性越强，黑点意味着相似性为零。从图中可看到，位于图像结构上的 Ψ_a 具有最稀疏的非零相似性，而位于纹理内部的 Ψ_b 与 Ψ_c 的相似性图稀疏性较低（因为更多的相邻图像块具有非零相似性）。图像块 Ψ_a 、 Ψ_b 和 Ψ_c 结构稀疏性分别为0.2154、0.0945、0.0355，处于结构上的图像块 Ψ_a 具有最大的结构稀疏性。

4.3.2 图像块优先度

最后，我们图像块的优先度为结构稀疏性项与图像块可信度项的乘积：

$$P(p) = \rho(p)C(p), \quad (4-6)$$

其中， $C(p)$ 为图像块 Ψ_p 的可信度，定义与[115]相同，即

$$C(p) = \frac{\sum_{q \in \Psi(p) \cap \bar{\Omega}} c(q)}{|\Psi(p)|}, \quad (4-7)$$

其中, $c(q)$ 测度像素 q 的颜色或灰度信息的可信度, 在图像未知区域初始化为 0, 在图像已知区域初始化为 1。在图像填充的每次迭代中, 所填充的新像素点的可信度被赋予所在图像块中心像素点的可信度。这样, 随着填充过程的进行, 新更新的填充边缘上的图像块可信度会逐渐降低。通过结构稀疏性项与图像块可信度的乘积, 处于填充边缘上具有较高结构稀疏性且较大图像块可信度的图像块被优先选取。

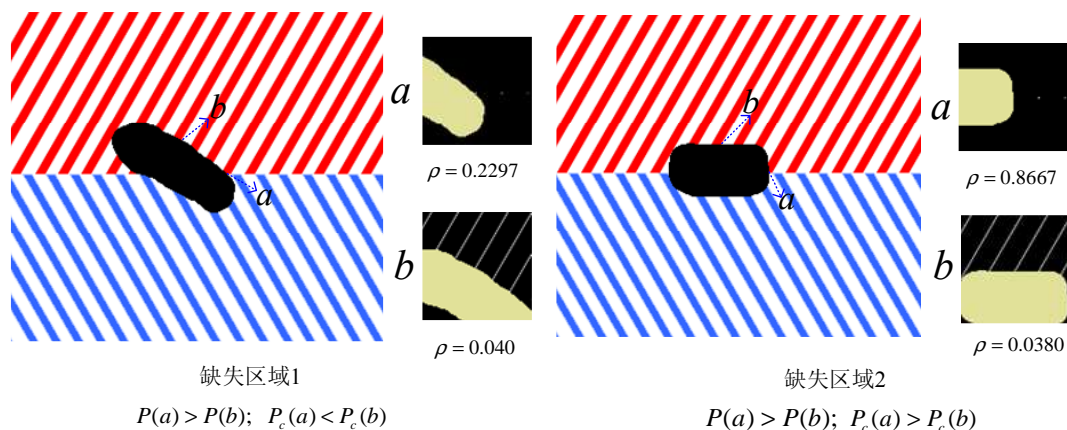


图 4-4 图像块优先度对比。 P 为所提出的图像块优先度, P_c 为 Criminisi 提出的优先度。

我们提供一组实验比较基于结构稀疏性的图像优先度 P 与 Criminisi 的优先度 P_c [115, 124] 的稳健性。注意到, 在 Criminisi 的定义中, 用等灰度方向 (即 isophote 方向) 和缺失区域边缘的法向方向之间的内积测度是图像块否处于结构上。图 4-4 展示了既有纹理又有边缘 (虚幻边缘, 即人感知到但不真实存在的边缘) 的例子。像素 a 、 b 分别处于纹理之间的虚幻边缘和纹理内部。针对 Criminisi 的优先度, 在左图中, 以 a 为中心图像块的优先度满足: $P_c(a) < P_c(b)$, 但在右图中, $P_c(a) > P_c(b)$, 这说明 Criminisi 的优先度不稳健于图像填充区域边缘的方向。针对我们的优先度, 在左、右两图中, 均有 $P(a) > P(b)$, 这说明我们定义的图像块优先度只与图像块本身是否处于结构上有关系, 而与缺失区域的形状无关, 因此我们的定义更为稳健。

4.4 基于稀疏表达的图像块填充

在上一节中, 我们说明了具有最高图像块优先度的图像块 Ψ_p 应该如何被优先选择。由于被优先选择的图像块处于缺失区域边缘, 因此图像块中既含有未知像素又含有已知像素, 这一小节, 我们将阐述如何对所选择图像块的未知像素进行填充。在传统的基于例子的图像填充算法中 [115, 124], Ψ_p 通过在已知区域中采样与其最相近的图像块进行填充 (用采样图像块像素点填充待填充块中的相应缺失像素)。与此不同, Wong [117] 提出了用已知区域中多个图像块的 Non-local 带权平均 [67] 图像块进行填充, 权重用图像块与待填充图像块的相似性测度。而在图像超分辨率问题中, 流行学习的 Locally Linear Embedding (LLE) [130] 被应用到如何从低分辨率图像块推理出高分辨率图像块。该算法通过分析已知低分辨率图像块与其近邻图像块的线性表达关系, 并将该关系应用到高分辨率以推理未知的高分辨率图像块。这两种改进的算法 (Non-local 均值、LLE) 均使用了多个已知的图像块去推理未知的图像块, 而不是用

一个最匹配的图像块，因此能更稳健地估计出未知信息从而产生更好的结果。但是，这些方法均用固定个数的图像块的组合进行图像块填充，一方面没有考虑到针对不同图像块的自适应性，另一方面，如果参与组合的图像块数目过大，会在填充结果中产生过模糊现象。

下面将提出一种新的基于信号稀疏表达的图像块填充模型，即假设待填充图像块可以由已知图像块的稀疏线性组合来推理。该模型与以往方法的不同在于：一方面，在稀疏表达框架下自适应的选择参与线性组合（即线性组合系数非零）的图像块数目，该数目相对于不同图像块是自适应且稀疏的，因此图像块填充更为灵活且很少引入过模糊效应；另一方面，在约束优化模型中，我们考虑了所填充图像块与邻域内图像块的纹理或结构一致性约束，使得所填充的内容与周围的内容更为一致。

4.4.1 图像块稀疏表达

给定待填充图像块 Ψ_p ，我们首先从已知区域中采样多个例子图像块（*exemplar*），记为： $\{\Psi_q\}_{q=1}^N$ ，其中 N 为图像块 Ψ_p 的*exemplar*数目。与以往工作相似[117, 130]，例子图像块选取为与 Ψ_p 最相似的 N 个图像块，我们设置 N 为20。记 P 为获取图像块 Ψ_p 中未知元素的矩阵，见公式(4-1)。我们假设： Ψ_p 可以由 $\{\Psi_q\}_{q=1}^N$ 的线性组合来表达：

$$\hat{\Psi}_p = \sum_{q=1}^N \alpha_q \Psi_q. \quad (4-8)$$

从而，图像块 Ψ_p 的未知象素点可以由 $\hat{\Psi}_p$ 中相应的象素点来填充：

$$P\Psi_p = P\hat{\Psi}_p. \quad (4-9)$$

其中组合系数 $\vec{\alpha} = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$ 可以在稀疏表达的框架下，通过求解约束优化问题来获得。因为假设图像块 Ψ_p 为 $\{\Psi_q\}_{q=1}^N$ 的稀疏线性表达，因此约束优化的目标为极小化组合系数 $\vec{\alpha}$ 的 ℓ^0 范数。下一步，将介绍该约束优化问题的约束项。

我们设计如下两类约束。第一类用于约束稀疏线性组合 $\hat{\Psi}_p$ 的颜色/灰度值。首先，在待填充图像块中已知颜色/灰度的象素点上，所估计的图像块 $\hat{\Psi}_p$ 的颜色/灰度应该逼近待填充图像块中相应的已知颜色/灰度：

$$\|\bar{P}\hat{\Psi}_p - \bar{P}\Psi_p\|^2 < \epsilon. \quad (4-10)$$

其中， ϵ 为控制逼近误差的容忍度参数。其次，在待填充图像块中的未知象素点（即需要填充的象素点）上，所填充的象素颜色/灰度与其周围已知的图像块颜色/灰度应具有一定的相似性：

$$\beta \|P\hat{\Psi}_p - P \sum_{p_j \in N_s(p)} w_{p,p_j} \Psi_{p_j}\|^2 < \epsilon. \quad (4-11)$$

其中， w_{p,p_j} 的定义与公式(4-3)一致。该约束度描述了所填充的图像块与其周围已填充完或已知图像块的相似程度。参数 β 用于平衡约束(4-10)和(4-11)之间的强度，在实现中设置为0.25。上述定义的约束称为局部图像块一致性约束（Local Patch Consistency Constraint）。

第二类约束为关于组合系数的约束，即约束线性组合系数 $\vec{\alpha} = \{\alpha_i\}_{i=1}^N$ 的和为1：

$$\sum_i^N \alpha_i = 1. \quad (4-12)$$

该约束称为组合系数求和约束（Coefficients Summation Constraint, CSC）。在该约束下，当向量 $\vec{\alpha}$ 只存在一个元素系数非零时，该非零系数一定为1，此时该模型与传统的基于例子的填充算法一致，即只用一个最匹配的已知图像块进行填充。

最终，线性组合系数 $\vec{\alpha}$ 可以通过优化如下的 ℓ^0 约束优化问题进行求解：

$$\min \{ \|\vec{\alpha}\|_0 \} \quad (4-13)$$

$$\text{s.t. } \|\bar{P}\hat{\Psi}_p - \bar{P}\Psi_p\|^2 < \epsilon \quad (4-14)$$

$$\beta \|P\hat{\Psi}_p - P \sum_{p_j \in N_s(p)} w_{p,p_j} \Psi_{p_j}\|^2 < \epsilon$$

$$\sum_i^N \alpha_i = 1.$$

其中， $\|\vec{\alpha}\|_0$ 为 $\vec{\alpha}$ 的 ℓ^0 泛数，即向量中的非零元素个数。

局部图像块一致性约束可以写为更简洁的形式：

$$\mathcal{E} = \|D\hat{\Psi}_p - \Psi_T\|^2 < \epsilon, \quad (4-15)$$

其中，

$$D = \begin{bmatrix} \bar{P} \\ \sqrt{\beta}P \end{bmatrix} \text{ 且 } \Psi_T = \begin{bmatrix} \bar{P}\Psi_p \\ \sum_{p_j \in N_s(p)} w_{p,p_j} \Psi_{p_j} \end{bmatrix}. \quad (4-16)$$

那么，约束优化问题(4-13)可以进一步写为：

$$\min \{ \|\vec{\alpha}\|_0 \} \quad (4-17)$$

$$\text{s.t. } \|D\hat{\Psi}_p - \Psi_T\|^2 < \epsilon$$

$$\sum_i^N \alpha_i = 1.$$

该约束优化问题也可以写成一个能量极小化问题：

$$\vec{\alpha}^* = \min_{\vec{\alpha}} \{ \gamma \|\vec{\alpha}\|_0 + \|D\hat{\Psi}_p - \Psi_T\|^2 + \eta \left\| \sum_i^N \alpha_i - 1 \right\|^2 \}. \quad (4-18)$$

当选择合适的 γ 和 η 时，该问题等价于约束优化问题(4-17)。

4.4.2 优化算法

本节讨论如何设计算法求解问题(4-17)。一般来说， ℓ^0 优化问题本质上是组合优化问题，很难进行求解。人们已先后提出Matching Pursuit (MP)算法、Orthogonal Matching Pursuit (OMP)算法 [131, 132] 和Basis Pursuit (BP) [133]等算法，以贪婪的方式

求解算法的近似最优解。另外一类优化 ℓ^0 问题的基本想法是将 ℓ^0 问题转化为凸优化问题，即通过优化一个 ℓ^1 正则化问题来求得近似最优解。在统计与机器学习领域， ℓ^1 正则化问题得到了广泛的研究，Lasso算法 [134]是一种非常流行的算法。然而，一般来说，很难确定使用哪种正则化算法进行优化（ ℓ^1 泛数或者 ℓ^0 泛数）更好 [127]。在应用中，由于OMP算法很简单且有效，被广泛用于图像的稀疏表达求解中，并应用于图像去噪 [24, 126]、编码 [135]、边缘检测 [127]、音频信号分离 [136]等等。

优化问题(4-17)和(4-18)与经典的稀疏正则化问题不同，增加了关于 $\vec{\alpha}$ 的组合系数求和约束项。因此，已往的通过优化 ℓ^1 范数或者 ℓ^0 范数正则化稀疏表达的算法均不能直接应用。我们提出一种新的类似于Matching Pursuit的算法来求解优化问题(4-17)。

该算法的基本思想是：以系数非零图像块集合为空集开始，每次从已知图像块集合 $Q = \{\Psi_q\}_{q=1}^N$ 中选择一个新的非零元素，进而重新构造满足组合系数求和约束且使得局部图像块一致性误差（如公式(4-15)）最小的线性组合方式，直到误差(4-15)达到约束要求。该算法以贪婪的方式找到满足图像块两类约束的最稀疏线性组合方式，具体算法如下。

假定在第 m 步时，我们已经选择了 m 个非零的已知图像块(记为 $S^m = \{\Psi_{q_1}, \Psi_{q_2}, \dots, \Psi_{q_m}\}$)，那么当前的稀疏表达为：

$$\hat{\Psi}_p^m = \sum_{i=1}^m \alpha_i^m \Psi_{q_i}. \quad (4-19)$$

在下一步 $m+1$ 时，我们从剩余的已知图像块集合（ $Q - S^m$ ）中选择一个新的图像块 $\Psi_{q_{m+1}}$ 。选择方法是：使得在增加图像块 $\Psi_{q_{m+1}}$ 后，新的稀疏表达具有最好的局部图像块一致性，即

$$\begin{aligned} \{\Psi_{q_{m+1}}\} &= \min_{\Psi \in Q - S^m} \{\mathcal{E}(\Psi_{q_1}, \Psi_{q_2}, \dots, \Psi_{q_m}, \Psi)\} \\ &= \min_{\Psi \in Q - S^m} \left\{ \left\| D \left(\sum_{i=1}^m \alpha_i^{m+1} \cdot \Psi_{q_i} + \alpha_{m+1}^{m+1} \cdot \Psi \right) - \Psi_T \right\|^2 \right\}. \end{aligned} \quad (4-20)$$

α_i^{m+1} 中的上标代表迭代次数，下标代表图像块标号。对于任意图像块 $\Psi \in Q - S^m$ ，可如下更新组合系数：

$$\begin{aligned} \alpha_i^{m+1} &= \min_{\vec{\alpha}} \left\{ \left\| \left(\sum_{i=1}^m \alpha_i \cdot D\Psi_{q_i} + \alpha_{m+1} \cdot D\Psi \right) - \Psi_T \right\|^2 \right\} \\ \text{s.t.} \quad &\sum_{i=1}^{m+1} \alpha_i = 1. \end{aligned} \quad (4-21)$$

记Gram矩阵为：

$$G = (\Psi_T \vec{\mathbf{1}}^T - X)^T (\Psi_T \vec{\mathbf{1}}^T - X), \quad (4-22)$$

其中 X 为矩阵，其列为 $\{D\Psi_{q_1}, D\Psi_{q_2}, \dots, D\Psi_{q_m}, D\Psi\}$ ，且 $\vec{\mathbf{1}}$ 为由1所组成的向量，则该问题具有显式解：

$$\vec{\alpha} = \frac{G^{-1} \vec{\mathbf{1}}}{\vec{\mathbf{1}}^T G^{-1} \vec{\mathbf{1}}}. \quad (4-23)$$

按照上述方式迭代直到局部图像块一致性约束(如公式(4-15))满足。

总之，利用图像块稀疏表达进行图像块填充的算法可概括为算法4.1。基于该图像块填充算法，最终的基于图像块稀疏性的图像填充算法如算法4.2。

算法4.1: 基于图像块稀疏表达的图像块填充算法

输入: Ψ_p 为待填充图像块。 $Q = \{\Psi_q\}_{q=1}^N$ 为已知图像区域 $\bar{\Omega}$ 中，与填充图像块最接近的 N (设置为20)个图像块（相似性定义为图像块已知像素上的平方误差测度）。

初始化: 设置 $m = 0$, $S^m = \{\}$, $\epsilon = 5$ 。

算法迭代: 迭代如下操作:

- 对于每个图像块 $\Psi \in Q - S^m$ ，通过公式(4-23)计算图像块集合 $S^m \cup \{\Psi\}$ 的稀疏线性组合系数向量 $\vec{\alpha}$ 。
- 从集合 $Q - S^m$ 中选择图像块 $\Psi_{q_{m+1}}$ ，使得集合 $S^m \cup \{\Psi_{q_{m+1}}\}$ 上的新的线性表达具有最小 \mathcal{E} ，如公式(4-15)。
- 如果 \mathcal{E} 下降，那么 $S^{m+1} = S^m \cup \{\Psi_{q_{m+1}}\}$, $m = m + 1$ ；当 $\mathcal{E} < \epsilon$ 或者 \mathcal{E} 增加时，跳出循环，进入下一步。

填充图像块: 图像块 Ψ_p 中的未知元素用稀疏线性表达 $\sum_{i=1}^m \alpha_i \Psi_{q_i}$ 中相应的已知元素进行填充。

算法4.2: 基于图像块稀疏性的图像填充算法

输入: 带有未知区域 Ω 的图像 I 。

步骤1: 对于未知区域 Ω 边界 $\partial\Omega$ 上的任意像素 p ，计算以 p 为中心图像块 Ψ_p 与其周围相邻已知图像块 $\{\Psi_{p_j}\}_{p_j \in N_s(p)}$ 的相似性。然后通过结构稀疏性计算图像块优先度。

步骤2: 迭代如下过程，直到未知区域被填充完毕:

- 选择 $\partial\Omega$ 上具有最高优先度的图像块 Ψ 。
- 通过图像块稀疏表达算法计算已知图像块相对于所选图像块的稀疏表达（如算法4.1），并用它填充所选择图像块 Ψ 中的未知像素。
- 更新未知区域 Ω ，对于边界 $\partial\Omega$ 上新出现的像素点，计算其上图像块的结构相似性和图像块优先度。

输出: 填充图像。

图4-5显示了图4-8中第三个例子的待填充区域内200个图像块稀疏表达中的非零元素个数。由图可见，不同的图像块的非零元素个数是动态变化的。这与以往的单个非零元素和固定个数非零元素的算法明显不同。

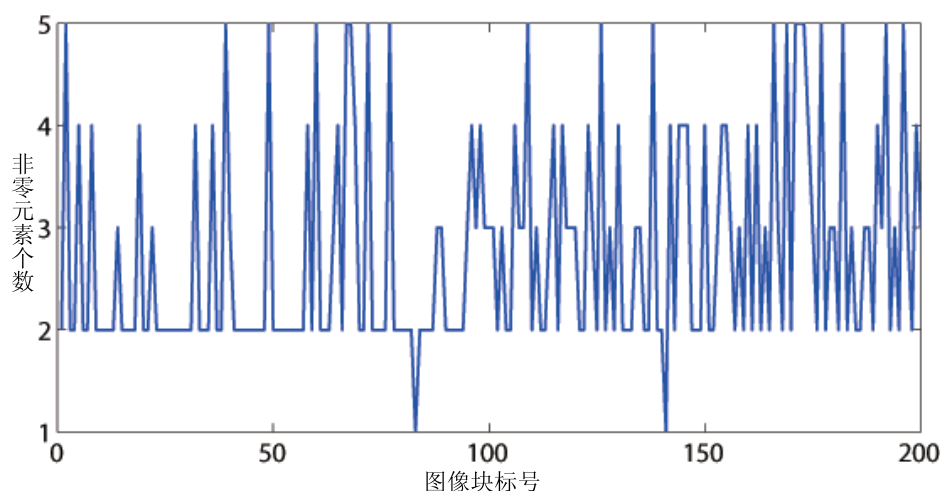


图 4-5 图4-8中第三个例子的待填充区域内200个图像块稀疏表达中的非零元素个数。X-坐标为图像块标号，Y-坐标为非零元素个数。

4.5 实验与比较

本节中，我们将在多个生成图像或自然图像上测试所提出的基于图像块稀疏性的填充算法性能。我们将算法应用于图像缺失块填充、划痕/文本移除和物体移除。在这些例子中，我们将测试所提出的图像块稀疏性对基于例子的图像填充算法的改进程度，并与基于扩散的填充算法、基于例子的同类填充算法和其他基于稀疏性的填充算法进行对比。在下面例子中，如果没有特殊说明，设定图像块大小为 7×7 ，图像块邻域窗口大小为 51×51 ，且误差容忍度参数 ϵ 设置为5.0。

4.5.1 在具有虚幻边缘的生成例子上的比较实验

图4-6展示了在生成的具有虚幻边缘的图像上的填充过程和结果。输入图像是由条状纹理和平滑纹理构成，且两类纹理之间存在虚幻的边缘（即由人所想象的边缘）。图4-6的上一行和下一行分别展示了我们的填充算法和Criminisi的填充算法[115, 124]进行图像填充的中间结果和最终结果。在Criminisi的基于例子的填充算法中，图像块的优先度定义基于isophote方向和未知区域边缘法向方向的内积。如图4-6的第二行为例，沿着条状纹理的方向为isophote方向，在缺失区域右上角，缺失边界与isophote方向垂直，因此具有最高的优先度。由第二行可见，图像缺失区域从右上角开始逐渐向左下角填充，导致最后填充结果中，条状纹理扩散到了平滑纹理区域内，如图4-6(f)下图。

我们算法的图像填充过程如图4-6的第一行所示。由于处于虚幻边缘上的图像块具有更为稀疏的邻域图像块相似性，因此优先度更大，它们更优先地被选取进行图像块填充。观察我们算法的图像填充过程：相对于处于纹理内部的图像块，沿着虚幻边缘的图像块被更快速的填充，使得图像结构被先于纹理填充完毕。由于我们还限制了

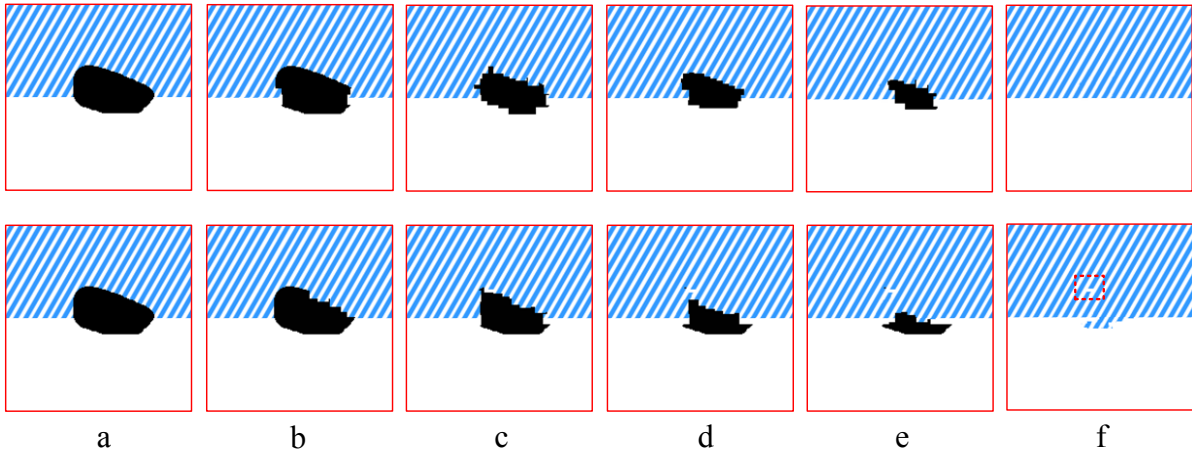


图 4-6 图像填充过程比较。图像块大小取为 11×11 。第一行和第二行分别展示了我们的算法和Criminisi的基于例子的算法 [115, 124] 的图像填充过程。(a) 输入的具有缺失区域的图像。(b)-(e) 算法填充过程的中间结果，可见随着算法进行，缺失区域逐渐被填充。(f) 最终的填充结果。

图像块的局部一致性，所填充的图像块不会出现与其周围图像块差距过大的现象，因此在Criminisi的算法结果中出现的位于矩形框内的错误并没有在我们的算法结果中出现。

4.5.2 填充算法实验一应用于划痕和文字移除

现在将填充算法应用于划痕和文字填充，并与其他典型的填充算法进行性能比较。首先与基于扩散的填充算法进行比较。基于扩散的填充算法可以很好地恢复图像划痕和文字移除区域的图像结构，但其缺点在于不能很好地填充纹理较强的区域。

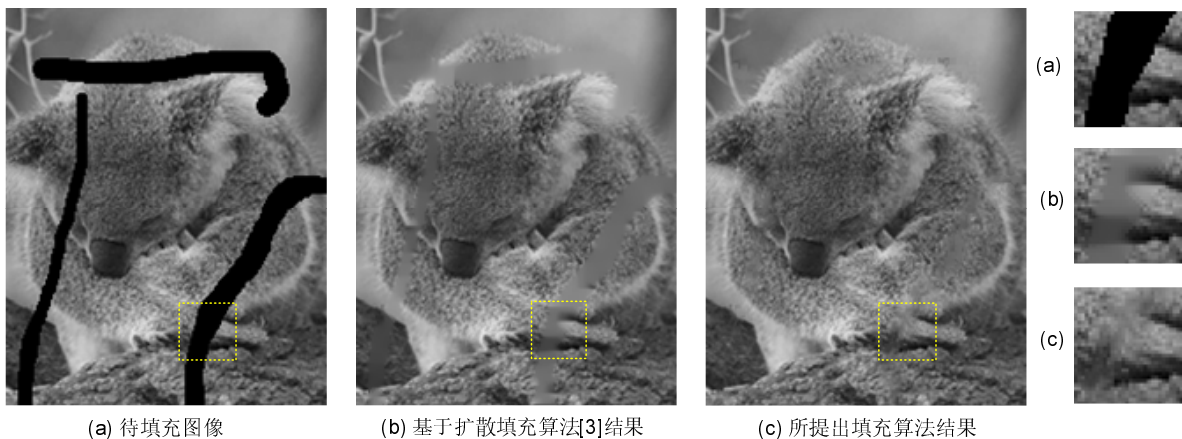


图 4-7 与基于扩散的填充算法比较。(a) 输入的待填充图像。(b) 基于扩散的填充算法[3]的结果。(c) 我们填充算法的结果。基于扩散的算法能很好地填充纹理平滑的划痕区域，但不能很好地填充纹理较强的区域。例如矩形框内的图像块（最右侧一列为其放大显示），基于扩散的填充结果过于光滑，而我们的结果可以很好地恢复皮毛的纹理。

如图4-7所示，我们比较了所提出的新算法与基于扩散的填充算法[3]。由应用中可看到，基于扩散的算法能很好地填充纹理平滑的划痕区域，但不能很好地填充纹理较强的区域。例如矩形框内的图像块（最右侧一列为其放大显示），基于扩散的填充结果过于光滑，而我们的结果可以很好地恢复皮毛的纹理。

下面将与其它基于例子的图像填充算法进行比较。我们将与如下算法进行比较：图像结构与纹理分别填充算法[114]、Criminisi的基于例子的填充算法[115][124]和Wong的基于例子的填充算法[117]。图像结构与纹理分别填充算法基本思想是将图像分解为结构层和纹理层，结构层通过基于扩散的填充算法填充[114]；而纹理层通过纹理合成的方式进行填充[113]。在实现该算法过程中，我们用Criminisi的基于例子的填充算法进行纹理填充，该算法优于原文[114]中所采用的纹理合成算法[113]。Criminisi的基于例子的填充算法[115, 124]和Wong的基于例子的填充算法[117]均通过图像块扩散实现图像填充，不同在于：前者采用最匹配的例子图像块进行填充；而后者采用多个例子图像块的Non-local均值进行图像块填充。

图4-8展示了对典型5个例子进行应用的比较结果，前三个例子应用于划痕移除，后两个例子应用于文本移除。第一行展示了原始的输入图像，在其他行中，第一列为输入的具有缺失区域的图像，第二列到第五列分别展示了结构和纹理分别填充算法的处理结果[114]、Criminisi的基于例子的填充结果[115, 124]、Wong的基于例子的填充结果[117]和我们算法的处理结果。我们用缺失区域中内填充结果与原始完整图像之间的峰值信噪比(Peak Signal to Noise Ratio, PSNR)定量测度图像填充质量。

Criminisi的填充算法（如图4-8(c)）可以产生锐化度很高的填充结果。但是由于只用了一个最匹配的例子图像块进行填充，因此结果不稳健，在结果中产生了很多不自然的冗余结构。例如，在图4-8(c)的第二行第三列的结果中，树干上出现冗余的树枝结构，而该结构在原始完整图像中不存在。Wong的填充算法（如图4-8(d)）由于使用了多个例子图像块的组合进行图像块填充，因此算法更为稳健，填充结果更好。例如，上面提到的冗余的树枝结构在Wong的填充算法结果中明显削弱了。结构与纹理分别填充算法可以很好地恢复图像的结构和纹理，且产生很高的PSNR值。但是，该算法的结果在图像结构上产生一定的模糊效应，该效应在文章[115]也被观测到。对于我们的算法，首先图像优先度定义更为稳健，而且多个例子图像块通过稀疏表达自适应组合且满足所填充图像内容与周围图像内容具有一致性，因此填充结果最好，且具有最高的PSNR值。

4.5.3 图像块稀疏性对填充性能的提升

下面我们验证图像块稀疏性对基于例子的图像填充算法的性能改进程序。本章分别提出了结构稀疏性和图像块稀疏表达分别解决基于例子的图像填充算法中的图像块选取和图像块填充两个基本问题，为验证所提出的两类图像块稀疏性对填充性能的影响，我们将对如下算法进行比较：（1）经典的基于例子的填充算法 [115, 124]，即结合基于isophote的图像块优先度进行图像块选取和最匹配图像块进行图像块填充的算法（记为Crim）；（2）改进图像块选取，即采用基于结构稀疏性的优先度进行图像块选取，且仍用最匹配图像块进行图像块填充的算法（记为Spar_Crim）；（3）改进图像块填充，即采用图像块稀疏表达进行图像块填充，且仍用基于isophote的图像块优先度进行图像块选取的填充算法（记为Crim_Spar）；（4）改进图像块选取和图像块填

充，即采用基于结构稀疏性的优先度进行图像块选取，且采用图像块稀疏表达进行图像块填充的填充算法（记为Spar）；（5）Wong提出的改进图像块填充算法，即采用基于isophote的图像块优先度进行图像块选取，且采用多个匹配图像块的Non-local均值进行图像块填充的算法（记为Wong）。

表 4-1 基于不同类型的图像块选取和图像块填充方法的图像填充性能（以PSNR值测度）。

方法	Wong	Crim	Crim_Spar	Spar_Crim	Spar
(a)	20.89	16.47	24.48	18.22	23.66
(b)	20.48	18.40	20.31	18.81	20.57
(c)	23.32	21.55	22.65	21.69	23.55
(d)	24.57	23.40	26.12	25.64	26.46
(e)	24.80	21.21	25.07	23.31	25.77
平均值	22.81	20.21	23.73	21.53	24.02

表4-1列出了应用上述5种算法在图4-8中5个例子上的填充性能PSNR值。由表可见，与经典的基于例子填充算法为基准，对于Spar_Crim算法，由于采用了基于结构稀疏性的优先度进行图像块选取，填充结果相对于经典方法具有更高的PSNR值；对于Crim_Spar算法，采用了基于图像块稀疏表达的图像块填充，填充结果的PSNR值也更优于经典方法的结果，同时也优于Spar_Crim算法的结果；对于Spar算法，采用结构稀疏性和图像块稀疏表达分别改进了图像块选取和图像块填充，填充结果平均性能最高；而对于Wong的基于例子算法，由于改进了图像块填充方法，使得填充性能相对于经典算法有了明显的改进，但平均性能低于Crim_Spar和Spar算法的结果。这说明基于图像块稀疏表达的图像块填充优于多个例子图像块Non-local均值的性能。

4.5.4 填充算法实验一应用于物体移除

本小节将针对物体移除应用比较我们的算法与已有基于例子的填充算法的性能。物体移除是一种重要的图像编辑任务，目的是将所感兴趣物体从图像中移除，并对图像上所留下的缺失区域进行填充。图像填充算法在物体移除中的应用在于对物体移除后所留下的缺失区域进行填充。由于物体往往占据图像较大面积，因此填充该类缺失区域具有很大的挑战性。如图4-9所示，第一列和第二列为原始完整图像和移除物体后所留下的图像，第三列到第五列分别为Criminisi的填充算法结果、Wong的基于例子的填充算法结果和我们的填充算法结果。在Criminisi算法的填充结果中，所填充的图像内容与其周围的图像纹理不一致，例如在第一个例子中的树纹理出现在水纹理中，而第四个例子中的草地纹理出现在岩石纹理中。这种现象产生的原因是：采用唯一最匹配例子图像块进行填充无法稳健的推理出未知的图像纹理或者结构。Wong的算法使用多个例子图像块的平均图像块进行图像填充，其结果在划痕和文字移除应用上已经展示较好的结果。在填充大块的移除区域时，该算法具有比较好的稳健性，例如

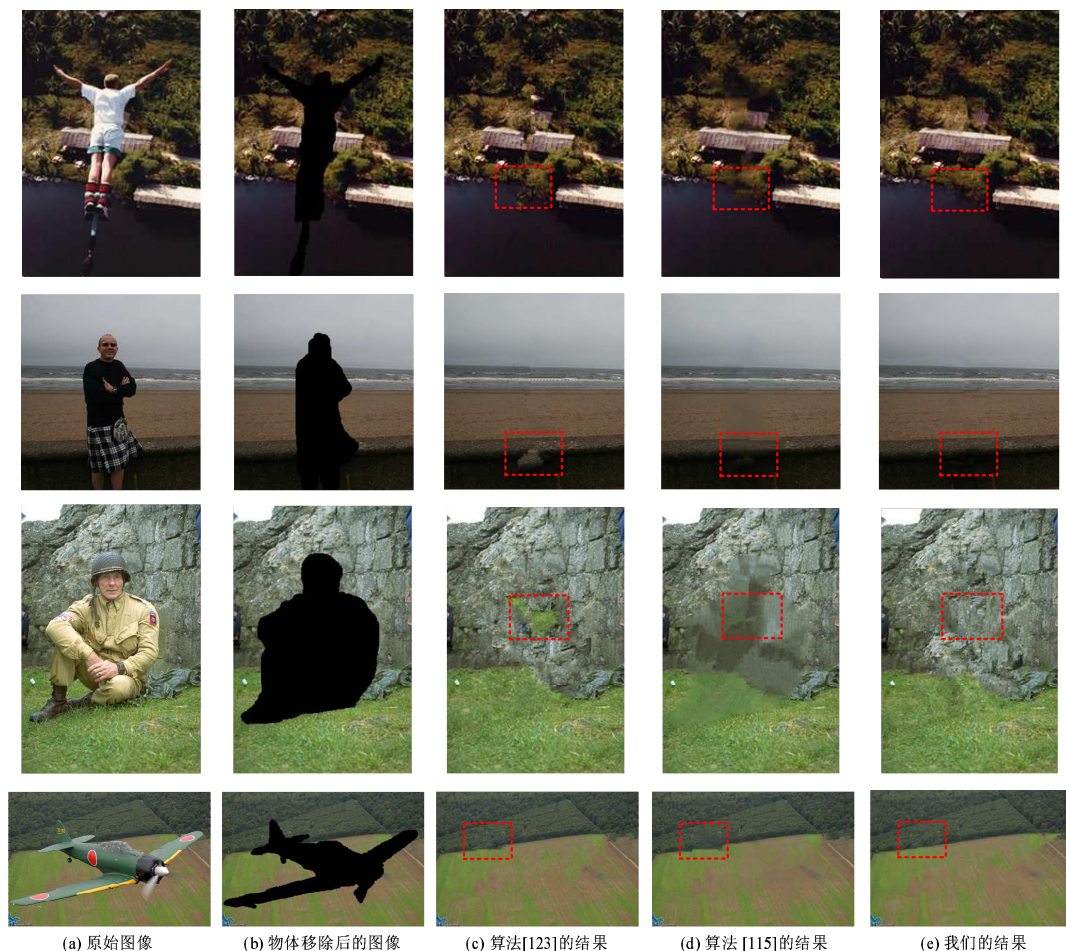


图 4-9 物体移除应用的填充算法性能比较。(a) 原始图像。(b) 物体移除后的图像。(c)-(e) Criminisi 的填充算法[123]结果、Wong 的填充算法[115]结果和我们的填充算法结果。我们算法所填充的图像内容与其周围图像内容具有更好的一致性。

在Criminisi填充结果中出现的图像内容不一致性问题在该算法结果中得到削弱。但该算法产生明显的过光滑效应。在我们的算法结果中，由于约束了图像邻域信息一致性，因此所填充图像内容与其周围内容具有更好的一致性；同时，由于稀疏性约束，只有稀疏数目个例子图像块对图像块的线性表达有贡献，因此填充结果很少产生过光滑效应。

4.5.5 与其它基于稀疏性的图像填充算法比较

我们还将所提出的算法与其它基于稀疏性的填充算法进行了比较。与我们提出的图像块稀疏性不同，以往基于稀疏性的填充算法的基本想法是对整个图像或大图像区域（而不是小的图像块）进行稀疏表达，基函数为图像过完全变换集合。Guleryuz [122, 123]提出通过对图像进行自适应的稀疏线性表达来恢复其中的未知区域，其自适应体现在不断迭代地更新图像的稀疏表达以及表达基的集合。Elad 等人 [120]设计了一种改进的算法，即将图像通过Morphological Component Analysis (MCA)方法分解为纹理和结构两部分，分别对两部分用不相容的两个基集合进行稀疏表达。

为进行性能比较，我们从作者主页上下载了Guleryuz的算法¹和Elad的算法²的源程序，并应用于丢失图像块的内容填充。对于Guleryuz的算法，我们用图像的离散余弦变换（Discrete Cosine Transform）作为基进行稀疏表达，变换块大小为 16×16 。对于Elad的算法，我们运行MCALab库代码进行纹理/结构分解 [19]和图像填充。设置纹理层的基为局部离散余弦变换（变换块大小为 32×32 ），而结构层的基采用curvelet变换（5种分辨率）。

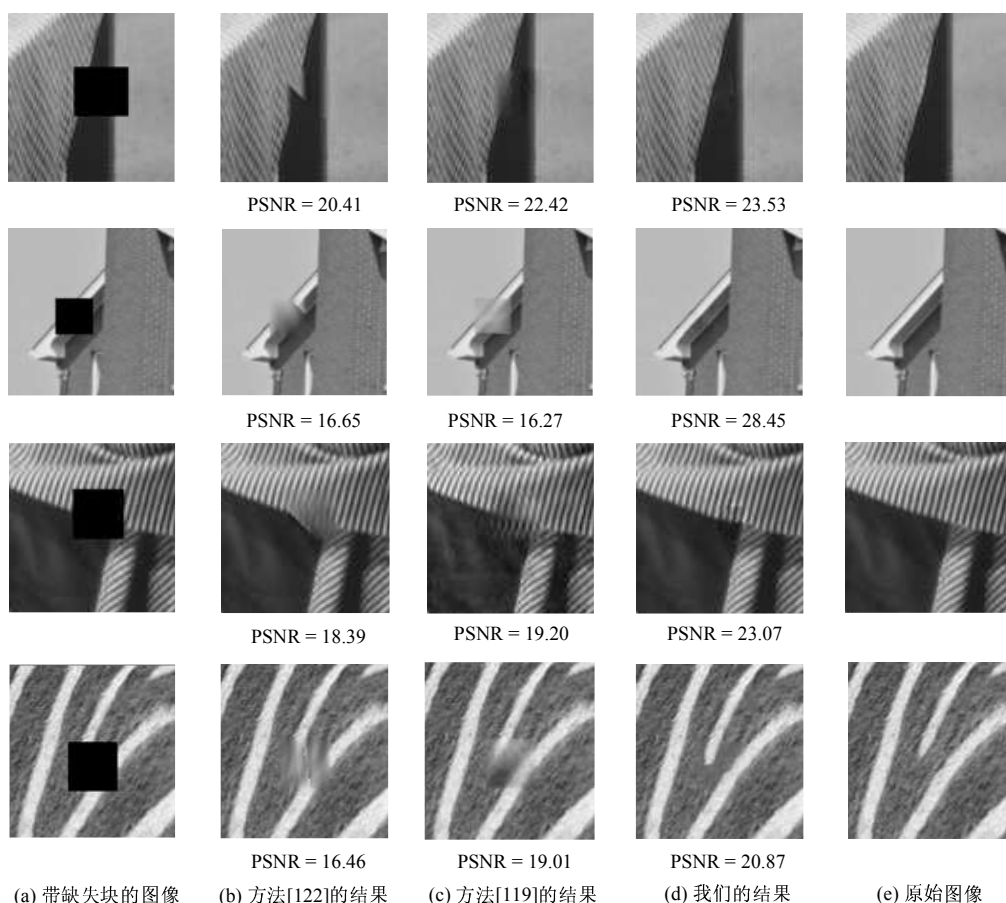


图 4-10 与基于稀疏性的图像填充算法比较。第一列为带有缺失图像块的输入图像，第二列到第四列图像显示Guleryuz的填充算法 [123]结果、Elad的填充算法 [120]结果和我们的算法结果，最后一列显示原始图像，填充性能用填充结果相对原始图像的PSNR值测度。

图4-10显示了算法应用于带有丢失块的图像填充问题。第一列为带有缺失图像块的输入图像，第二列到第四列图像显示Guleryuz的填充算法 [121–123]结果、Elad的填充算法 [120]结果和我们的算法结果，最后一列显示原始图像，填充性能用填充结果相对原始图像的PSNR值测度。正如文章 [121]所指出，Guleryuz的填充算法可以很好地填充较小图像块内容（不超过 16×16 ）。但是对于图4-10中的例子，丢失块区域较大，该算法不能精确填充。Elad的算法改进了Guleryuz的算法的填充结果，这是因为

¹<http://eeweb.poly.edu/%7Eonur/source.html>

²<http://www.greyc.ensicaen.fr/jfadili/demos/WaveRestore/downloads/mcalab/Home.html>

图像纹理和结构被分解且分别进行稀疏表达。但是对于复杂的结构，例如第二个例子中的结构没有被精确填充。

4.5.6 组合系数求和约束的必要性

为显示图像稀疏表达中组合系数求和约束(公式(4-12))的必要性，我们展示了一个例子以比较分别带有和不带有组合系数求和约束的图像填充结果。图4-11(a)和(b)分别显示了带有和不带有该约束的填充结果。由图4-11中(b)可见，缺失的岩石纹理被填充入很暗的纹理，即所填充的纹理的亮度与周围亮度不一致。而在约束稀疏线性组合系数求和为1之后，相当于对组合结果的亮度范围进行了约束，使其不会过大的超出例子图像块的亮度范围，因此填充结果更自然，如图4-11(a)。

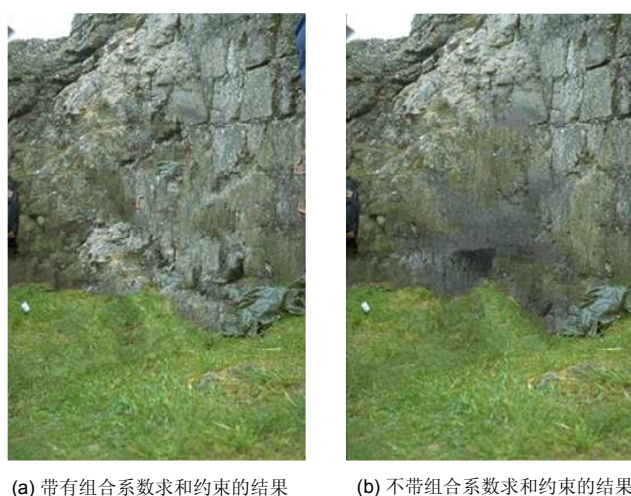


图 4-11 带有和不带有组合系数求和约束的图像填充结果比较

4.6 小结

本章提出了一种新的基于图像块稀疏性的填充算法，并成功应用于划痕/文字去除、物体移除和缺失块填充。该工作的创新之处在于：我们研究和发了一种新的自然图像稀疏性先验，即图像块稀疏性先验，并将该先验应用于图像填充算法设计中。该工作的研究思路来源于近年来人们对自然图像统计、稀疏表达的大量基础和适用性研究工作。

所提出的图像块稀疏性包括结构稀疏性和图像块稀疏表达两部分。结构稀疏性基于这样的先验：处于结构上的自然图像块与其相邻图像块的相似性的分布更为稀疏，因此结构稀疏性越强的图像块越可能位于图像的结构上。另一方面，图像块稀疏表达基于如下的先验：自然图像块可以由已知图像块的集合的稀疏线性组合进行表达。如上两种稀疏性被分别用于解决基于例子的图像填充算法中的两个关键问题：图像块选取和图像块填充。

实验表明：与经典基于例子的填充算法相比，我们的填充算法可以更稳健地区别结构和纹理，并且所填充的图像内容与周围内容具有更好的一致性；与以往基于稀疏性填充算法相比，我们的算法可以更好地填充大范围缺失内容，而且填充结果锐化度

高且更为自然。在下一步的研究工作中，我们将关注如何进一步发掘基于图像块的自然图像先验，进而应用于图像填充、超分辨率和纹理合成。

5 基于闪光先验和MRF先验建模的前景和背景分割算法

5.1 引言

图像分割是计算机视觉中的一个基本问题，研究如何将图像像素点进行重新聚合，以得到中层视觉信息表达，并应用于进一步的图像处理或理解。如何对单张图片进行自动的前景（距离相机近的感兴趣物体）和背景（前景物体背后的场景）分割是一个非常具有挑战性的问题。从数据挖掘角度来看，图像分割可以视为对象素特征的聚类、分类或半监督分类，目的是将像素点进行归类，判断每个像素点的类别信息，例如，对于前景/背景分割问题，即判别图像的像素点是属于前景类还是背景类。

实现前景/背景分割一般有两类基本方法。第一类方法是交互式图像分割算法 [92, 137, 138]。该算法允许用户通过交互式工具简要标记部分像素点的类别，然后通过已经标定的像素点学习其他未标定像素点的类别。该类算法采用基于图（Markov随机场）的半监督学习算法进行前景/背景分割。另一类算法融合更多的图像特征信息或者使用更多的图像（例如视频）进行分割。图像特征信息包括运动信息 [139–142]，立体视觉信息 [143]，附加光照信息 [144]，或者假设背景已知或者静止 [145]等。一般来说，利用多张图像的分割算法可以得到更好、更鲁棒的结果 [146, 147]。但是为了使分割算法更为实用，我们期望极小化所需要的图像数目，同时得到精确的图像分割结果。

本章，我们提出一种实用且高质量的前景/背景分割算法。该算法的输入为一对图像，即对同一场景，分别开启闪光灯和关闭闪光灯所拍摄的一对图像，我们称这样的图像为开闪光灯图像（flash image）和关闪光灯图像（flash-off image）。图 5-1显示了一个典型例子：自然场景由人（前景）和背景构成，且前景与背景具有一定的距离。分别关/开闪光灯拍摄一对图像，如图 5-1(a)(b)。注意：由于前后拍摄两幅图像具有一定的时间间隔，因此由于相机运动和场景运动的影响，两幅图像并不是逐点对齐的。我们的目标是以这对图像为输入实现自动的前景分割。对该场景进行单张图像的自动分割几乎是不可能的，这是因为前景/背景的颜色分布或者纹理等信息都无法完全区别前景和背景。但是利用如上得到的一对图像，我们可以实现精确的前景/背景分割。我们观察到：在闪光灯作用下，前景物体被充分闪光，亮度增大；而背景由于距离前景比较远，因此亮度受到闪光灯影响很小。这一观察的内在原因是：闪光灯的闪光能量随着距离的增大快速下降（与距离的平方成反比）。因此闪光灯信息给我们提供了一个重要的先验信息，可以区别前景与背景；另一方面，图像分割问题的重要先验信息是像素点分割标记（代表类别）具有光滑性，我们将用MRF(Markov random field)对光滑性进行建模。

于是，所提出的自动图像分割算法将主要应用如下两种重要的先验：

- 闪光先验：当前景/背景物体具有一定距离时，图像前景物体受到闪光灯的影响远远大于背景所受到的影响；
- 光滑性先验：图像像素的前景/背景标记具有光滑性，即相邻像素点具有相似的标记。

基于上述两种先验信息，我们将设计一种新的自动前景/背景分割算法——**闪光分割算法** (*flash cut*)。

闪光分割算法通过建模闪光先验信息和分割标记光滑性先验，并融合运动信息和颜色分布信息实现自动的前景/背景分割。我们通过全局颜色直方图对闪光先验信息进行建模，而通过Gibbs分布对分割标记光滑性先验信息进行建模。在Markov随机场框架下，我们将提出前景/背景分割能量模型，并通过Graph cut算法进行优化，而将模型优化结果作为分割结果。这一算法的优势是：对相机和物体的运动具有鲁棒性，且可以应用到具有树木/花草等复杂前景结构的图像分割问题中。因此，新算法模型实现了一个实用、简单的自动前景/背景分割技术，可实现利用带有闪光灯的家用相机进行高质量的前景提取。如图5-1(c)(d)所示，我们的算法实现了精确的前景/背景分割结果，并将原图背景替换为新的背景。



图5-1 相机分别关/开闪光灯连续拍摄的两张图像的前景/背景分割后替换背景的结果。(a) 未开闪光灯的照片。(b) 开启闪光灯的照片。(c)(d) 应用前景/背景分割算法进行分割并替换背景的结果。

5.1.1 相关工作

近年来，闪光灯相关技术是计算机视觉和计算机图形学研究中的一个重要方向。闪光灯技术已经被应用于生成高质量图像（例如图像增强）[31]、幻景消除和flash-exposure采样 [63]，以及图像去噪、细节转移、白平衡、红眼消除等图像处理问题 [32]。使用多闪光成像技术，利用闪光灯所产生的阴影还可以用于检测深度边缘 [148]和立体视觉计算 [28]。

与我们的工作最接近的算法是flash matting [29]算法，在该算法中，闪光信息被应用于产生高精度的图像软分割。但是，该算法要求输入的开/关闪光灯图像对是逐点对齐的，即相机和场景是静止的，而我们的算法对相机和物体的运动都具有鲁棒性，因此具有更强的实用性。

5.2 闪光分割算法

本节将建立闪光分割算法的能量模型。首先介绍闪光灯成像模型。

5.2.1 闪光成像模型

假定闪光灯为点光源且光源强度为 L ，空间中某物体表面上的一点受其辐射的能量强度为

$$E = L \cdot \rho \cdot r^{-2} \cdot \cos\theta, \quad (5-1)$$

其中, ρ 为给定闪光灯角度和观测角度下, 点所在表面的二向性反射分布函数 (Bidirectional Reflectance Distribution Function, BRDF), r 为点与闪光灯的距离, θ 为闪光方向和表面法向方向的夹角。可见, 闪光灯的强度随着距离 r 的增加迅速降低。

该工作中, 我们假设场景中背景层和前景层具有一定的距离。在该假设下, 前景的亮度会显著地受到闪光灯的影响, 而背景受到的影响会很小, 例如图 5-1(a)和(b)所示。由于闪光灯对前景/背景影响的显著不同, 为我们提供了很强的先验信息用于前景/背景分割。

5.2.2 分割能量模型

前景/背景分割问题可以建模为二类标记问题。给定两个输入图像: 开闪光灯图像 I^f 或者关闪光灯图像 I^{nf} , 目的是对图像中的任意一个像素点 p 进行标记: $x_p \in \{\text{背景}(= 0), \text{前景}(= 1)\}$ 。

我们用Markov随机场框架进行建模: 即将图像像素点作为结点 (Vertice), 并将邻接像素点的相邻关系作为结点之间的边 (Edge)。构建图 $G = (\mathcal{V}, \mathcal{E})$, 其中 \mathcal{V} 为结点集合, \mathcal{E} 为边集合。记每个节点为随机变量 X_p , 其取值 x_p 为1或者0, 那么 $X = \{X_p\}_{p \in \Omega_R}$ 构成随机场, 而 $x = \{x_p\}_{p \in \Omega_R}$ 为随机场 X 的观测, 进一步假设:

$$(1) P(x) > 0,$$

$$(2) P(x_p | x - \{x_p\}) = P(x_p | x_{N(p)}),$$

其中, $N(p)$ 为像素 p 的邻域。在该假设下, 随机场 X 为Markov随机场。依据Hammersley和Clifford定理, 随机场 X 满足Gibbs分布:

$$P(x) = \frac{1}{Z} \exp\left(-\frac{1}{T} E(x)\right), \quad (5-2)$$

其中, Z 为正规化参数, T 为温度参数, $E(x)$ 为定义在随机场上的能量函数。定义二阶Markov随机场, 那么 $E(x)$ 具有形式:

$$E(x) = \sum_p E_d(x_p) + \alpha \sum_{p < q} E_s(x_p, x_q), \quad (5-3)$$

其中, E_d 为定义在像素上能量项 (称为数据项), 而 $E_s(x_p, x_q)$ 为定义在相邻像素点上的能量项, 建模相邻像素标记的光滑性。前景层可以通过在图像分割标记空间 $\{x = \{x_p\}_{p \in \Omega_R}; x_p \in \{0, 1\}\}$ 中极大化Gibbs分布 $P(x)$, 即极小化能量函数(5-3)得到。参数 α 的作用在于平衡两项之间的影响, 在实验中, 我们设置该参数为30。

光滑项 $E_s(x_p, x_q)$ 用于约束相邻像素 p 和 q 的标记的一致性, 定义为

$$E_s(x_p, x_q) = |x_p - x_q| \cdot \exp(-\beta \|I_p - I_q\|^2), \quad (5-4)$$

其中, $\beta = (2 \langle \|I_p - I_q\|^2 \rangle)^{-1}$ [90], $\langle \cdot \rangle$ 为期望。

数据项 E_d 建模了前景上的闪光效应, 运动补偿的背景项以及像素的颜色似然值, 它由如下三项构成:

$$E_d(x_p) = \gamma_f E_f(x_p) + \gamma_m E_m(x_p) + E_c(x_p), \quad (5-5)$$

其中, γ_f 和 γ_m 在实验中设置为10。

- E_f 为前景项，建模了前景闪光先验信息：亮度受到闪光灯强烈影响的象素点更可能为前景。该项用图像的颜色分布直方图的变化进行建模，作为前景分割的一个重要的全局性线索。
- E_m 为背景项，建模了运动补偿的背景闪光先验信息：图像对可以很好匹配且亮度基本不受闪光灯影响的象素点更可能为背景。该能量项考虑了闪光和运动信息。
- E_c 为颜色项，分别建模了前景和背景颜色的分布，该分布通过混合高斯模型建立。

5.3 能量模型详细介绍

本节中，我们将分别详细介绍以上三个能量项的建模。

5.3.1 前景能量项

我们通过分析开闪光灯图像/关闪光灯图像对的颜色分布来建模闪光灯对前景的影响。由于采用颜色直方图进行建模，因此该项为全局性线索，对小的相机运动和物体运动具有很好的鲁棒性。

令 $H^f = \{h_k^f\}$ 和 $H^{nf} = \{h_k^{nf}\}$ 分别为开闪光灯图像/关闪光灯图像的RGB三维颜色直方图分布。 h_k^f 和 h_k^{nf} 分别为落在两直方图第 k 个bin中的象素个数。如果 $h_k^{nf} > h_k^f$ ，说明处于直方图 H^{nf} 中的某些象素受到闪光的影响而移到直方图 H^f 中具有更高亮度的bin中。因此，在关闪光灯图像中，处于直方图 H^{nf} 中的这种元素具有很高的可能性是前景元素。类似地，对于开闪光灯图像 I^f ，如果 $h_k^f > h_k^{nf}$ ，则意味着某些象素受到闪光影响而落入直方图 H^f 中第 k 个bin中。因此，在开闪光灯图像中，处于该类直方图中的象素具有更高可能是前景象素。我们将通过对图像对中的象素 p 定义闪光比率 (*flash ratio*) 来模型化这种观察：

$$r_p^f = \max\left\{\frac{h_{k_p}^f - h_{k_p}^{nf}}{h_{k_p}^f}, 0\right\}, \quad (5-6)$$

$$r_p^{nf} = \max\left\{\frac{h_{k_p}^{nf} - h_{k_p}^f}{h_{k_p}^{nf}}, 0\right\}. \quad (5-7)$$

其中， k_p 为象素 p 所在直方图bin的下标。闪光比率越大，说明象素 p 越有可能属于前景。图 5-2 显示了一个1维例子以说明闪光比率的意义。

图 5-3(b)显示了开/关闪光灯图像对的闪光比率，越亮代表闪光比率越高。由于直方图为全局性测度，因此闪光比率并非完全正确。例如，在关闪光灯图像中的地面部分（处于背景）的象素也具有较高的闪光比率，这是因为其上颜色与前景中人的衣物颜色很相似。两张图中黑色眼镜均具有较低的闪光比率是因为黑色物体由于反光性低而受到闪光影响很小。

基于闪光比率，我们定义前景能量项

$$E_f(x_p) = \begin{cases} 0 & , x_p = 1 \\ \frac{1}{1-\zeta}[\max\{r_p, \zeta\} - \zeta] & , x_p = 0 \end{cases}. \quad (5-8)$$

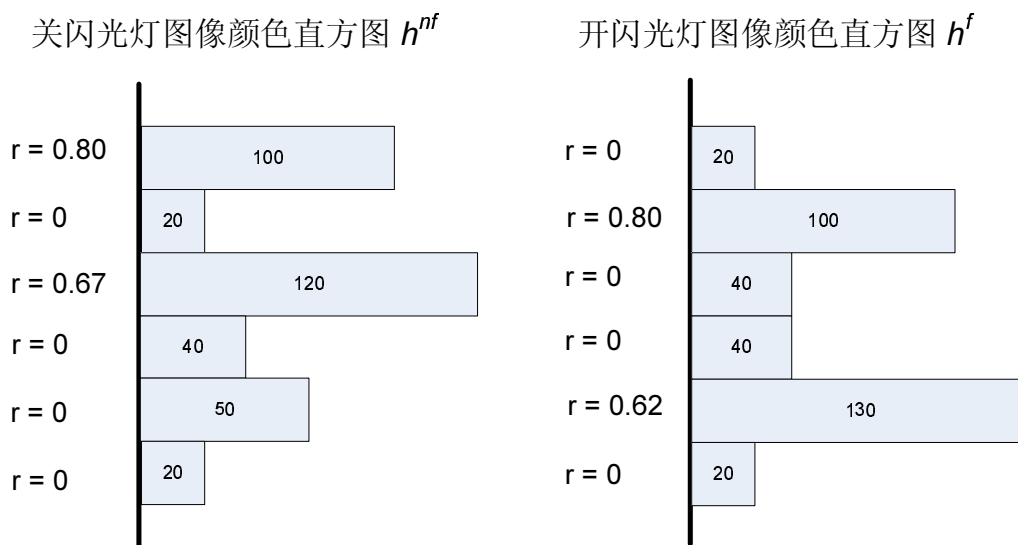


图 5-2 开/关闪光灯图像对的闪光比率示意图。以一维灰度直方图为例，对于每个直方图，越处于下方的bin亮度越大，而bin的长度反映处于该bin中的像素数目大小。直方图左端显示闪光比率。对于最上方bin，在关闪光灯图像直方图中频数为100，而在开闪光灯图像直方图中频数为20，因此闪光比率 $= (100-20)/100 = 0.80$ 。这说明：80个像素被闪光灯点亮并转移到了其它bin中。

其中， ζ 为鲁棒性参数，设置为0.2，该参数的重要性在于：如果 r_p 大于 ζ ，我们更可能标记像素 p 为前景，否则，我们没有特别的倾向性将该像素标记为前景或者背景。因此该能量项提供了一个比较保守的前景估计。

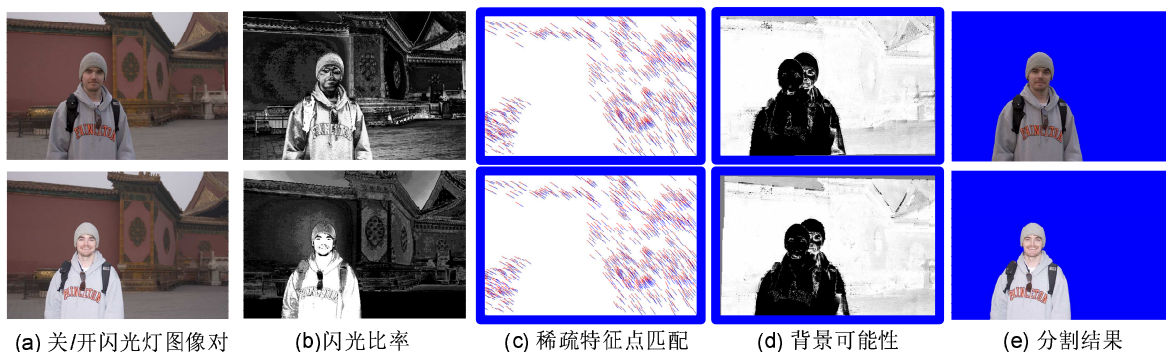


图 5-3 开/关闪光灯图像对分割的中间和最终结果。(a) 输入图像对。(b) 闪光比率。(c) SIFT特征点对应(上图)与背景SIFT特征点对应(下图)。(d) 背景可能性。(e) 分割结果。在(b)(d)中，像素越亮意味可能性越大。

5.3.2 运动补偿的背景能量项

假设我们已经得到了一对图像的稠密运动场 $\mathbf{m} = \{m(p)\}$ ，使得关闪光灯图像 I^{nf} 和开闪光灯图像 I^f 可以逐点准确匹配。那么， I^{nf} 中像素点 p 和其在 I^f 中的匹配点 $p' = m(p)$ 之间的亮度差为：

$$\Delta I_p = I_{m(p)}^f - I_p^{nf} = I_{p'}^f - I_p^{nf}. \quad (5-9)$$

我们定义该亮度差为**闪光亮度差(flash difference)**。由于我们假设用户在拍摄图像对时距离背景比较远且两张图像连续拍摄，因此背景亮度变化很小且空间一致。我们可以合理地将背景像素的闪光亮度差建模为高斯分布 $N(\Delta I_p|\mu, \sigma^2)$ ，均值为 μ 且方差为 σ^2 。因此，我们定义像素 p 属于背景的概率为：

$$p_b(x) = \exp(-\sigma_b(\Delta I_p - \mu)^2). \quad (5-10)$$

我们设定 $\sigma_b = \ln 2/(3\sigma)^2$ ，使得处于均值周围 $\pm 3\sigma$ 范围的闪光亮度差具有较高的可能性处于背景(可能性 ≥ 0.5)。我们称 $p_b(x)$ 为背景可能性，如图 5-3(d)为例。最终，能量项 $E_m(x_p)$ 定义为：

$$E_m(x_p) = \begin{cases} 2 \max\{p_b(x_p), 0.5\} - 1, & x_p = 1 \\ 0, & x_p = 0 \end{cases}. \quad (5-11)$$

利用该定义， $E_m(x_p)$ 被正规化到 $[0, 1]$ 之间，对于开闪光灯图像的能量项可以类似定义。

在上面定义中，我们假定参数 $\{\mu, \sigma^2\}$ 是已知的且运动流 \mathbf{m} 也是已知的。现在，我们描述如何估计这些参数。

1. 参数估计

我们通过分析图像对的特征点来对参数 $\{\mu, \sigma^2\}$ 进行估计。首先对两幅图像分别进行特征提取（SIFT特征提取 [149]），并用最近邻匹配找到两图像的匹配点集合：只有当特征点与其最匹配特征点的距离和次匹配特征点的距离比小于0.6时，特征点对为稳定匹配，且通过双向匹配找到每个特征点的唯一匹配点（构成单射关系）。图5-3(c)显示了图像对的匹配特征点，线段两端点分别为开/关闪光灯图像上匹配的特征点。

给定稳健的匹配特征点对，我们构造特征点相对于闪光亮度的分布（用直方图表达：测度具有不同闪光亮度的匹配特征点对数目）。在大多数情况下，仅仅有很少或者没有匹配特征点对处于前景区域，这是由于前景亮度/纹理受到闪光灯的影响而无法匹配，如图 5-3(c)所示。但在特殊情况下（例如前景特征受到闪光灯的影响也不是很大），所匹配的特征点对即来自前景也来自背景。例如，图 5-4(b)所示的分布图，由于所匹配的特征点对即来自前景也来自背景，因此匹配点对的闪光亮度具有多峰的特点。不管分布是否具为单峰，我们只关心第一个峰，该峰对应于背景匹配点对，

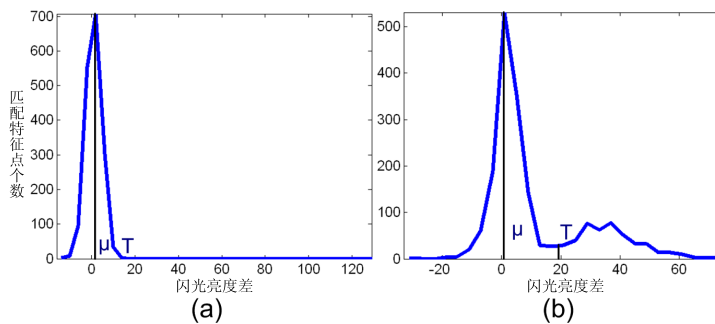


图 5-4 开/关闪光灯图像对匹配特征点的闪光亮度差直方图，X-坐标为闪光亮度差，Y-坐标为匹配点数目。(a) 具有单峰的分布。(b) 具有多峰的分布。参数 μ 可以通过第一个峰值估计，且 T 为 μ 右侧的第一个局部极小值。

闪光亮度的均值 μ 估计为直方图第一个局部极大值所对应的闪光亮度。闪光亮度的方差 σ^2 用闪光亮度低于 T 的匹配点对的闪光亮度进行估计，其中 T 为 μ 右侧第一个局部极小值所对应的闪光亮度，如图5-4。

2. 运动估计

由于只希望估计背景的运动，我们首先将闪光亮度大于 T 的匹配点对去除，然后用剩下的背景匹配点对估计背景运动场。例如，图5-3(c)和5-5(d)显示了背景特征点匹配。当背景为静止或距离相机足够远，例如图5-7(a)和(c)，我们可以用全局的运动模型（homography 或者 affine 运动模型）来建模背景运动，并可以直接通过匹配点对计算相应全局运动模型。但是，全局运动无法建模背景物体运动、径向扭曲的情况，例如图5-5中具有动态运动的背景。这些情况下，我们需要设计一种局部精细匹配算法进行图像匹配（即估计运动）。

以关闪光灯图像为例，通过对匹配特征点对的离散运动场的插值估计稠密背景运动场 \mathbf{m}^0 ，插值可采用自适应局部带权回归（Adaptively Locally Weighted Regression, ALWR）算法[149]。匹配点之间的颜色差为：

$$e_p = I_{m(p)}^f - I_p^{nf} - \mu_p. \quad (5-12)$$

其中， I^f 和 I^{nf} 分别为开/关闪光灯图像的颜色，而 μ_p 用来补偿闪光灯对颜色影响。显然，对于像素 p ，一个好的运动估计 $m(p)$ 应使得像素点 p 与其匹配像素点 p^{nf} 的颜色差尽可能小。

对于像素 p ，利用插值所估计的运动只提供一个比较好的初始估计，我们还需要精化对其运动的估计。该过程可通过如下的迭代算法实现Lucas-Kanade [150]。在第 k 次迭代，运动矫正项 $\Delta m^k(p)$ 为：

$$\Delta m^k(p) = -\left(\sum_{q \in w(p)} \nabla I_{m^k(q)}^f \nabla I_{m^k(q)}^{fT} \right)^{-1} \sum_{q \in w(p)} \nabla I_{m^k(q)}^f e_q^k.$$

其中， $w(p)$ 为像素点 p 周围 5×5 窗口。

图5-5(e)和(f)分别显示了用homography运动模型进行背景匹配时，我们的算法进行背景匹配后的背景可能性图(亮度越高意味背景可能性越大)。显然用我们的运动估计方法，背景可能性的估计更为精确。注意到，由于前景不能被很好地对齐，前景的背景可能性很小。

5.3.3 颜色项

前景和背景颜色分布通过高斯混合模型（Gaussian Mixture Models, GMMs）进行建模。以前景为例，我们取

$$p_c(I_p | x_p = 1) = \sum_{k=1}^K w_k^f \mathcal{N}(I_p | \mu_k^f, \Sigma_k^f), \quad (5-13)$$

其中 $\mathcal{N}(\cdot)$ 为高斯分布， $\{w_k^f, \mu_k^f, \Sigma_k^f\}$ 为高斯混合模型的第 k 个成分的权重、均值和协方差矩阵。 K 的典型取值为10。前景高斯混合模型通过EM(Expectation-maximization)算法在所有满足 $p_b(x) < 0.4$ 的像素点上进行学习。而背景颜色分布在满足 $p_b(x) > 0.6$ 的象

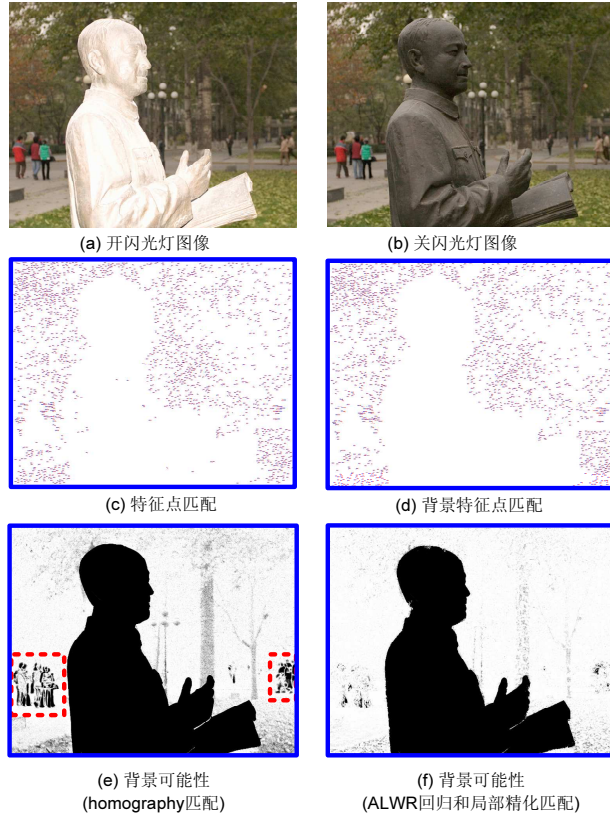


图 5-5 运动补偿后的背景可能性图。(a) 开闪光灯图像。(b) 关闪光灯图像。(c) 特征点对（线段两端点分别代表两图像匹配的特征点）。(d) 背景的特征点对应，即闪光亮度小于 T 的匹配点对。(e) 用homography运动模型进行背景匹配后的背景可能性图。由于矩形区域内有运动物体，因此其内的像素点匹配不精确。(f) 通过ALWR插值和局部精化匹配后的背景可能性图。

素点上学习。最终，颜色项 $E_c(x_p)$ 定义为：

$$E_c(x_p) = \begin{cases} -\log(p_c(I_p|x_p = 1)) & , x_p = 1 \\ -\log(p_c(I_p|x_p = 0)) & , x_p = 0 \end{cases} \quad (5-14)$$

5.3.4 模型优化

我们采用Graph cut算法 [102, 151]对前景/背景分割能量函数进行优化。该能量优化问题为基于图的标定 (Labelling) 问题：假定图 $G = (\mathcal{V}, \mathcal{E})$ 具有非负边权重，且有两个特殊的节点：source节点 s （代表前景标定1）和sink节点 t （代表背景标定0），图 G 的一个切 (cut) $C = (S, T)$ 将图节点 \mathcal{V} 分割为两个互不相交的集合 S 和 T ，使得属于 S 和 T 的节点分别属于前景和背景。优化目标为找到具有最小能量的切（从而为min-cut问题）。

图标定问题是公认的NP-难问题，直到D. Greig等人 [93]提出了Graph cut算法，该算法可以快速、有效地优化图标定能量函数。Graph cut算法通过迭代的 α -expansion move操作，不断更新具有更小能量的标定直到发现局部极小值。Graph cut算法可以快速地求解二值标定问题，被广泛的应用于图像恢复 [102]，交互式分割 [91]等。

下面我们证明所提出的前景/背景分割能量函数可以由Graph cut算法在多项式时间内优化。应用下述已知结果 [151]:

定理 5.1 [151] 设能量 E 为 n 个二值变量的函数, 即

$$E(x_1, x_2, \dots, x_n) = \sum_i E_d(x_i) + \sum_{p < q} E_s(x_p, x_q), \quad (5-15)$$

那么 E 可以由Graph cut算法在多项式时间内求得极小值的充分条件是:

$$E_s(0, 0) + E_s(1, 1) < E_s(0, 1) + E_s(1, 0). \quad (5-16)$$

其证明见 [151]。

下面我们证明所提出的前景/背景分割能量函数满足上述定理的条件, 因此可以由Graph cut算法在多项式时间内求得其极小值。

定理 5.2 前景/背景分割能量(5-3)可以由Graph cut算法在多项式时间内求得极小值。

证明 为证明能量(5-3)可以由Graph cut算法在多项式时间内求得极小值, 只需证明该能量满足关于 E_s 项的条件(5-16), 其中能量项 E_s 定义见(5-4), 事实上,

$$\begin{aligned} E_s(0, 0) &= E_s(1, 1) = 0, \\ E_s(1, 0) &= E_s(0, 1) = \exp(\beta \|I_p - I_q\|^2) > 0. \end{aligned}$$

因此,

$$E_s(0, 0) + E_s(1, 1) < E_s(0, 1) + E_s(1, 0)$$

一定满足, 所以定理得证。(证毕)

值得注意的是, Graph cut算法尽管还无法从理论上证明所收敛的极小值点是否为全局极小值点, 但可以证明所找到的局部极小值位于全局极小值的倍数范围内 [103]。

5.3.5 算法流程

总结以上讨论, 基于闪光信息和MRF建模的前景/背景分割算法可以概括为算法5.1。

5.4 实验结果

图 5-6显示了闪光分割算法和当前流行的一些分割算法的比较结果。图 5-6前两行为输入图象、闪光比率图、运动补偿的背景可能性图以及闪光分割的最终结果。为了说明闪光比率、背景可能性和颜色三者缺一不可, 图 5-6 (e)和(f)显示了只用闪光比率和颜色分布进行分割的结果和只用背景可能性和颜色分布进行分割的结果, 与融合三种信息的分割结果相比, 这些分割结果都不是很精确, 因此三种信息缺一不可。

算法5.1: 基于闪光信息和MRF建模的前景/背景分割算法

输入: 相机分别开关闪光灯拍摄的一对图像 I_f 和 I_{nf} 。

步骤1: 提取图像 I_f 和 I_{nf} 的SIFT关键点并进行关键点匹配;

步骤2: 计算图像 I_f 和 I_{nf} 的颜色分布, 并计算前景能量项;

步骤3: 对匹配关键点的亮度差进行分布统计, 得到背景匹配点, 然后根据匹配点估计背景运动模型, 并进行局部运动估计, 最终得到图像的逐点运动场;

步骤4: 根据图像对运动场计算背景能量项;

步骤5: 对背景可能性进行阈值化处理, 从而得到前景/背景的粗分割, 并分别对前景/背景的颜色进行高斯混合模型建模;

步骤6: 用Graph cut优化模型(5-3), 从而得到前景/背景分割结果;

输出: 图像对的前景/背景分割结果。

我们还与其它两种前景分割算法的结果进行了比较。这两种分割算法分别为GrabCut [138]算法和co-segmentation [152]算法。这两种算法也是目前为止最好的半自动/自动前景分割算法。图 5-6 (g)显示了GrabCut算法分割结果, 该算法为基于颜色分布的半自动分割算法。我们以图象外轮廓为初始分割, 由于前景/背景颜色具有混淆性, 因为该算法分割结果不精确。Co-segmentation是一种基于颜色直方图的全自动分割算法, 目的是通过一对图像, 分割出具有相似颜色分布的部分作为前景。图 5-6 (h)展示了应用该算法的分割结果, 由结果可以看到使用全局的颜色分布的区别是无法对细小的结构进行精确分割。

我们还将闪光分割算法应用到了一系列的室内/室外拍摄的开/关闪光灯图像对。图 5-7 显示了多个实验例子。图 5-7(a)显示了一个步行的人。基于运动的方法很难对该例子进行精确分割, 这是因为前后两张图像的运动非常大。图 5-7(b) 显示了具有运动背景的例子, 而图 5-7(c)具有非常细结构的前景物体的例子。显然GrabCut、co-segmentation和基于运动的算法都很难得到满意的结果。图 5-7(d)中, 前景为具有复杂结构的植物; 而在图 5-7(e)中, 背景为复杂结构的竹林, 由结果可见我们的闪光分割算法的优势在于可以分割出非常细小的结构。图 5-7(f) 为即有前景运动又有背景运动的例子。图 5-7(g)显示了另一个在室内拍摄的例子, 由结果可见我们的闪光分割算法的可以对相机和场景物体运动都具有很好的鲁棒性, 且对室内/室外拍摄的图像均可以进行有效分割。在最终的分割结果中, 我们使用coherent matting算法 [153]对边缘分割进行更为精确的分割, 使得替换背景的结果更真实。

5.5 小结

本章提出了一种使用闪光先验和光滑性先验的图像前景分割算法—闪光分割算法。通过融合闪光信息、运动和颜色信息, 闪光分割算法可以生成高质量的分割结果, 并且对物体运动和相机运动具有很好的鲁棒性, 因此该算法可以为具有带闪光灯相机的一般用户提供实用且高质量的分割结果。

由于将分割问题建模为二值标定问题, 即前景像素标记为1, 背景像素标记为0,

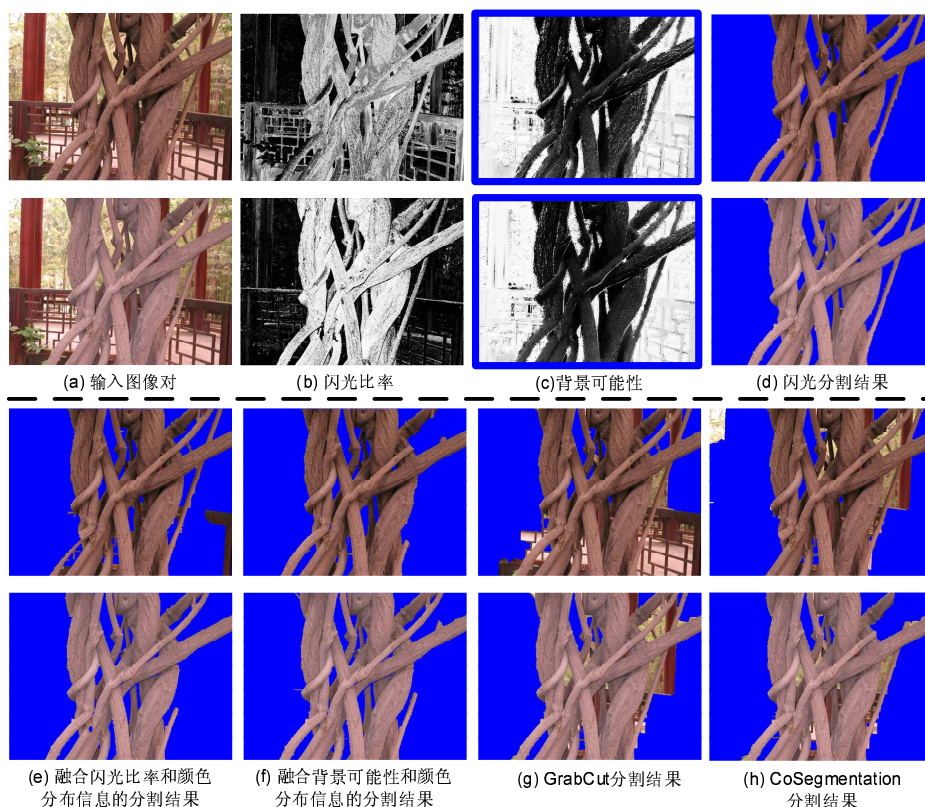


图 5-6 分割结果比较。(a-d) 输入图像对、闪光比率图、背景概率图与最终分割结果。(e) 只用闪光比率和颜色信息的分割结果。(f) 只用背景概率和颜色信息的分割结果。(g) GrabCut [137]结果。(h) CoSegmentation [151]结果。通过融合闪光比率图、背景概率和颜色信息，我们算法可以生成最精确的分割结果。

从而图像模型为： $I = \alpha F + (1 - \alpha)B$ 。其中， F, B 分别为前景和背景，而 α 为0或者1。由于采用非0即1的分割标记，因此并不能很好的分割软边缘（即 α 为 $[0, 1]$ 间的实数）的前景物体，例如长发软、皮毛和透明玻璃等。通过结合精细设计的Matting技术，我们的方法可能很好的处理这种情况；另外，同其他闪光相关的技术一样，我们的算法对大范围的前景阴影（即由于遮挡造成的闪光灯无法影响到的前景区域）比较敏感，但是由于标记的光滑性先验的作用，前景阴影区域也可以被正确标定，如图5-3、5-5、5-6和5-7(d)(f)所示。

下一步研究工作中，我们需要更为深入地研究闪光信息的建模方法。例如，因为像素点的颜色受到闪光灯的作用并不是任意的，而是某个通道（例如亮度通道）的分布受到显著影响，且影响程度应具有一定的规律性，我们应该应用这种规律性对闪光作用进行更有效的建模。另外，我们还希望将该算法应用到视频的前景分割中。

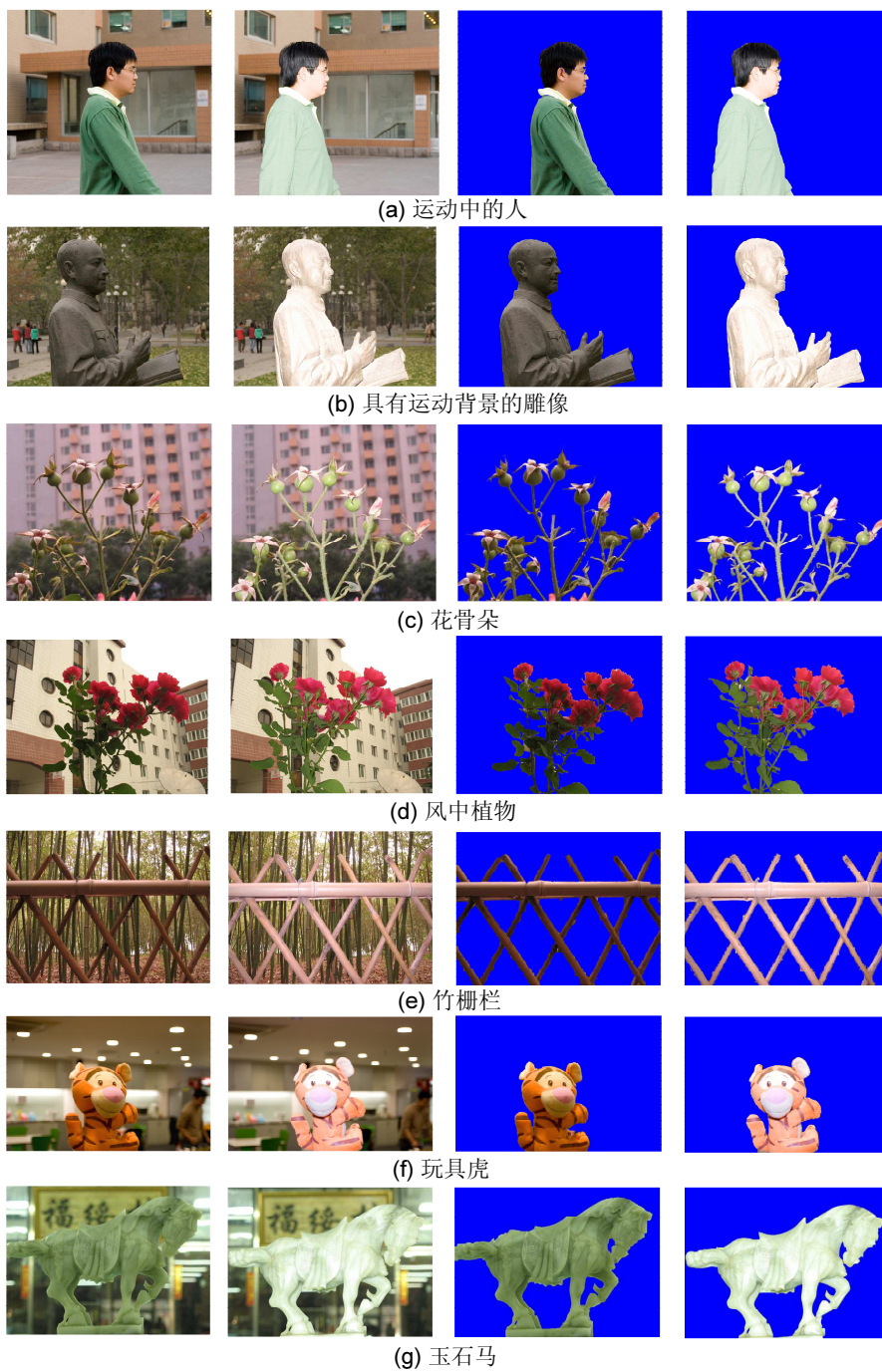


图 5-7 闪光分割结果。从左到右分别为：关闪光灯图像、开闪光灯图像、关闪光灯图像分割结果和开闪光灯图像分割结果。

6 结论与展望

6.1 本文主要贡献

本论文主要关注图像处理基本问题的建模与算法研究，系统研究了图像去噪、填充、超分辨率、增强和前景/背景分割五个基本问题。本论文从图像处理每个基本问题的特异性出发，以数学建模与分析为手段，通过提出图像的视觉先验模型，进而系统地发展了解决相应基本问题的优化模型及其优化算法。实验表明，这些模型与算法相对于当前的同类算法具有更好的图像处理效果。本论文的主要贡献如下。

第二章，提出了基于自然图像梯度轮廓先验的图像超分辨率与增强算法。通过建模自然图像梯度场在垂直于图像边缘方向上的梯度轮廓分布，提出了一种新的基于统计的自然图像先验—梯度轮廓先验。基于该先验，发展了解决图像超分辨率和增强问题的能量函数模型，并基于变分方法设计了优化算法。理论分析与应用实践表明：该模型和算法具有良好的数学性质，而且比已有算法具有更优良的图像处理效果。

第三章，提出了基于Markov随机场尺度选择模型的各向异性扩散去噪算法。通过对各向异性扩散去噪算法的尺度（即扩散停机时间）选择进行建模，提出了空间自适应的各向异性扩散滤波的尺度选择策略，并基于此发展了基于Markov随机场先验建模的尺度选择统计模型。结合该尺度选择模型，我们进而提出了新的基于各向异性扩散滤波的去噪算法。该算法也可以解释为各向异性扩散尺度空间的融合算法，有效解决了以往尺度选择策略为全局选取同一尺度的缺点。实验表明，所提出的基于Markov随机场尺度选择模型的各向异性扩散去噪算法是目前最好的去噪算法之一。

第四章，提出了基于图像块稀疏性先验的图像填充算法。通过定义图像块结构稀疏性，进而证明结构稀疏性是刻画图像重要结构的有用指标，结构稀疏性越大，则图像块越可能处于图像边缘；基于视觉编码的稀疏性先验，在假设待填充图像块可以由候选图像块集合的稀疏线性组合来表达的基础上，我们提出了图像块稀疏线性表达的优化模型和优化算法，并发展了基于图像块稀疏性先验的图像填充算法。相比以往的同类填充算法，所提出的新算法具有更好的填充性能。

第五章，提出了基于闪光信息和MRF先验建模的前景/背景分割算法。以开/关相机闪光灯设备所提供的图像对为输入，通过数学模型建模闪光设备所带来的光照变化的先验信息，提出了基于MRF先验的图像分割能量模型，并采用Graph cut算法进行优化。与已有的分割算法相比，所提出的算法可以得到更为精确的前景/背景分割结果。该模型的重要意义在于通过相机的闪光设备，提供了一种便捷、实用的前景提取模型和算法，可期待成为新一代相机的实用技术之一。

综上所述，本论文通过建模视觉先验信息，提出了解决图像处理五个基本问题的能量模型或统计模型，并设计了相应的优化算法。实验表明，所提出的模型与算法相对于已有的同类模型与算法具有更好的图像处理性能，因此具有更加广泛而直接的应用前景。

6.2 进一步的研究工作

在下一步的研究工作中，我们将继续深入地挖掘自然图像的内在规律性，并通过

视觉先验建模以更好地解决图像处理、分析与理解中的基本问题。特别地，我们将主要研究如下问题。

(1) 自然图像表达。本文已经在这方面做了如下的工作：自然图像梯度场轮廓先验和自然图像块稀疏性先验。这些研究工作的缺陷在于：自然图像梯度场轮廓先验仅描述了图像边缘的分布规律，但没有描述图像纹理的分布规律；另外，图像块稀疏性先验是针对图像块进行的稀疏建模和表达，缺乏对图像整体规律性的描述。因此我们将进一步研究：如何系统地同时描述图像的结构和纹理的分布规律，以真正实现对图像的完整表达，并应用该表达解决图像恢复、理解等问题。

(2) 基于物体的空间分布先验或形状先验的图像识别。本论文的研究工作主要集中在图像处理，而很少涉及到计算视觉的另一个本质问题，即如何进行图像识别或物体识别。因此，我们将进一步开展图像识别的研究工作，我们的研究思路是：基于描述物体的局部特征（例如，SIFT描述子等），找到描述并判别物体类别的特征基，使得局部特征在该基上具有更低维的表达，同时具有更高的分类判别性能。我们的研究重点是挖掘和建模物体的形状先验或空间组织结构先验，进而建立更为合理的特征基表达，以获得更高的物体识别能力。

(3) 视频数据的视觉先验建模。视频信息可以视为时空数据，生物视觉系统所采集到的数据本质上为视频数据。而目前的图像表达研究工作很少考虑时间信息，即很少建模视频信息的表达。下一步，我们也将研究和建模生物视觉系统对视频信息的处理机制，通过融合视频信息所提供的运动、深度等信息，发展视频数据的视觉先验模型，进而指导视频的分析与检索。

参考文献

- [1] Marr D. Vision: A computational investigation into the human representation and processing of visual information[M]. San Francisco: W. H. Freeman & Co Ltd, 1982.
- [2] Healey GE, Kondepudy R. Radiometric CCD camera calibration and noise estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(3): 267 - 276.
- [3] Bertalmio M, Sapiro G, Caselles V, et al. Image inpainting[J]. ACM Transaction on Graphics, 2000, 19(3): 417-424.
- [4] Hummel RA. Representations based on zero-crossings in scale-space[M]. New York: Courant Institute of Mathematical Sciences, New York University, 1986.
- [5] Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. Physica D, 1992, 60: 259-268.
- [6] Chan T, Vese LA. Active contours without edges[J]. IEEE Transactions on Image Processing, 2001, 10(2): 266-277.
- [7] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1990, 12(7): 629-639.
- [8] Black MJ, Sapiro G, Marimont DH, et al. Robust anisotropic diffusion[J]. IEEE Transaction on Image Processing, 1998, 7(3): 421-432.
- [9] Chan T, Marquina A, Mulet P. High-order total variation-based image restoration[J]. SIAM Journal on Scientific Computing, 2000, 22(2): 503 - 516.
- [10] Chan T, Osher S, Shen J. The digital TV filter and nonlinear denoising[J]. IEEE Transactions on Image Processing, 2001, 10(2): 231-241.
- [11] Huang JG, Mumford D. Statistics of natural images and models[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999, 540 - 547.
- [12] Levin A, Fergus R, Durand F, et al. Image and depth from a conventional camera and depth from a conventional camera with a coded aperture[J]. ACM Transactions on Graphics, 2007, 26(3): 1-9.
- [13] Roth S, Black MJ. Steerable random fields[C]. Proceedings of International Conference on Computer Vision, 2007.
- [14] Portilla J, Strela V, Wainwright MJ, et al. Image denoising using scale mixtures of Gaussians in the wavelet domain[J]. IEEE Transactions on Image Processing, 2003, 12(11): 1338-1351.
- [15] Zhu SC, Mumford D. Learning generic prior models for visual computation[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, 463-469.
- [16] Zhu SC, Mumford D. Prior learning and Gibbs reaction-diffusion[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, 19(11): 1236-1250.
- [17] S. Roth, M. J. Black. Fields of experts: a framework for learning image priors[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005, 860-867
- [18] Olshausen B, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images[J]. Nature, 1996, 381: 607-608.
- [19] Starck JL, Elad M, Donoho DL. Image decomposition via the combination of sparse representations and a variational approach[J]. IEEE Transactions on Image Processing, 2005, 14(10): 1570-1582.
- [20] Mairal J, Leordeanu M, Bach F, et al. Discriminative sparse image models for class-specific edge detection and image interpretation[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2008.
- [21] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms[C]. Proceedings of Advances in Neural Information Processing Systems Conference, 2006.
- [22] Donoho D. Compressed sensing[J]. IEEE Transaction on Information Theory, 2006, 52(4): 1289 - 1306.

- [23] Schass K, Rauhut H, Vandergheynst P. Compressed sensing and redundant dictionaries[J]. *IEEE Transaction on Information Theory*, 2008, 54(5): 2210–2219.
- [24] Maire J, Elad M, Sapiro G. Sparse representation for color image restoration[J]. *IEEE Transaction on Image Processing*, 2008, 17(1): 53–69.
- [25] Yang JC, Wright J, Ma Y, et al. Image super-resolution as sparse representation of raw image patches[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.
- [26] Liang L, Xiao R, Wen F, et al. Face alignment via component-based discriminative[C]. *Proceedings of European Conference Computer Vision*. 2008,
- [27] Burl M, Leung T, Perona P. Face localization via shape statistics[C]. *Proceedings of 1st International Workshop on Automatic Face and Gesture Recognition*, 1999, 154–159.
- [28] Feris R, Raskar R, Chen LB, et al. Discontinuity preserving stereo with small baseline multi-flash illumination[C]. *Proceedings of International Conference on Computer Vision*, 2005, 412–419.
- [29] Sun J, Li Y, Kang SB, et al. Flash matting[J]. *ACM Transaction on Graphics*, 2006, 25(3): 361–366.
- [30] Sun J, Sun J, Kang SB, et al. Flash cut: foreground extraction with flash and no-flash image pairs[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.
- [31] Eisemann E, Durand F. Flash photography enhancement via intrinsic relighting[J]. *ACM Transaction on Graphics*, 2004, 23(3): 673–678.
- [32] Petschnigg G, Agrawala M, Hoppe H, et al. Digital photography with flash and no-flash image pairs[J]. *ACM Transactions on Graphics*, 2004, 23(3): 664–672.
- [33] McGuire M, Matusik W, Pfister H, et al. Defocus video matting[J]. *ACM Transaction on Graphics*, 2005, 24(3): 567 – 576.
- [34] Lyu S, Simoncelli EP. Statistical modeling of images with fields of gaussian scale mixtures[C]. *Proceedings of Neural Information Processing Systems Conference*, 2006.
- [35] Tappen MF, Russell BC, Freeman WT. Exploiting the sparse derivative prior for super-resolution and image demosaicing[C]. *IEEE Workshop on Statistical and Computational Theories of Vision*, 2003.
- [36] Sun J, Zheng NN, Tao H, et al. Image hallucination with primal sketch priors[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003, 729–736.
- [37] Dai SY, Han M, Xu W, et al. Soft edge smoothness prior for alpha channel super resolution[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007.
- [38] Sun J, Sun J, Xu ZB, et al. Image super-resolution using gradient profile prior[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008, 2471–2478.
- [39] Levin A, Zomet A, Weiss Y. Learning how to inpaint from global image statistics[C]. *Proceedings of International Conference on Computer Vision*, 2003, 305–312.
- [40] Fergus R, Singh B, Hertzmann A, et al. Removing camera shake from a single photograph[J]. *ACM Transactions on Graphics*, 2006, 25(3): 787–794.
- [41] Levin A, Zomet A, Peleg S, et al. Seamless image stitching in the gradient domain[C]. *Proceedings of European Conference on Computer Vision*, 2005, 377–389
- [42] G. Hinton. Product of experts[C]. *Proceedings of International Conference on Artificial Neural Networks (ICANN)*, 1999, 1–6.
- [43] Hou HS, Andrews HC. Cubic splines for image interpolation and digital filtering[J]. *IEEE Transaction on Singnal Processing*, 1978, 26(6): 508–517.
- [44] Thevenaz P, Blu T, Unser M. *Handbook of Medical Imaging, Processing and Analysis*[M], Academic Press, San Diego, USA, 2000.
- [45] Kullback S. The Kullback-Leibler distance[J], *The American Statistician*, 1987, 41: 340 – 341.

-
- [46] Li X, Orchard MT. New edge-directed interpolation[J]. *IEEE Transaction on Image processing*, 2001, 10(10): 1521–1527.
- [47] Irani M, Peleg S. Motion analysis for image enhancement: resolution, occlusion and transparency[J]. *Journal of Visual Communication and Image Representation*, 1993, 4(4): 324–335.
- [48] Baker S, Kanade T. Limits on super-resolution and how to break them[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2002, 24(9): 1167–1183.
- [49] Lin ZC, Shum HY. Fundamental limits of reconstruction-based super-resolution algorithms under local translation[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2004, 26(1): 83–97.
- [50] Ben-Ezra M, Lin ZC, Wilburn B. Penrose pixels: super-resolution in the detector layout domain[C]. *Proceedings of International Conference on Computer Vision*, 2007
- [51] Freeman WT, Pasztor E, Carmichael O. Learning low-level vision[J]. *International Journal of Computer Vision*, 2000, 40(1): 25–47.
- [52] Freeman WT, Jones TR, Pasztor EC. Example-based super-resolution[J]. *IEEE Computer Graphics and Applications*, 2002, 22(2): 56–65.
- [53] Chang H, Yeung DY, Xiong Y. Super-resolution through neighbor embedding[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, 275–282.
- [54] Fattal R. Image upsampling via imposed edge statistics[J]. *ACM Transactions on Graphics*, 2007, 26(3): 1–8.
- [55] Liu C, Shum HY, Freeman WT. Face hallucination: theory and practice[J]. *International Journal of Computer Vision*, 2007, 75(1): 115–134.
- [56] Wang Q, Tang X, Shum HY. Patch based blind image super resolution[C]. *Proceedings of International Conference on Computer Vision*, 2005, 709–716.
- [57] Polesel A, Ramponi G, Mathews VJ. Image enhancement via adaptive unsharp masking[J]. *IEEE Transaction on Image processing*, 1999, 9(3): 505–510.
- [58] Rudin LI. Images, numerical analysis of singularities and shock filters[D]. Ph.D. Thesis, California Institute of Technology, 1987.
- [59] Guichard F, Morel JM. A note on two classical enhancement filters and their associated PDE's[J]. *International Journal of Computer Vision*, 2003, 52(2/3): 153–160.
- [60] Osher S, Rudin LI. Feature-oriented image enhancement using shock filters[J]. *SIAM Journal on Numerical Analysis*, 1990, 27(4): 919 - 940.
- [61] Schavemaker JGM, Reinders MJT, Gerbrands JJ, et al. Image sharpening by morphological filtering[J]. *Pattern Recognition*, 2000, 33(6): 997–1012.
- [62] Weickert J. Coherence-enhancing shock filters[J]. *Lecture Notes in Computer Science*, 2003, 2781: 1–8.
- [63] Agrawal A, Raskar R, Nayar SK, et al. Removing photography artifacts using gradient projection and flash-exposure sampling[J]. *ACM Transaction on Graphics*, 2005, 24(3): 828 – 835.
- [64] Agrawal A, Raskar R, Chellappa R. What is the range of surface reconstructions from a gradient field?[C]. *Proceedings of European Conference on Computer Vision*, 2006, 578–591.
- [65] Reinhard E, Ashikhmin M, Gooch B, et al. Color transfer between images[J]. *IEEE Computer Graphics and Applications*, 2001, 21(5): 34 – 41.
- [66] Tai YW, Jia J, Tang CK. Local color transfer via probabilistic segmentation by expectation-maximization[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, 747– 754.
- [67] Buades A, Coll B, Morel JM. A non-local algorithm for image denoising[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, 60–65.
- [68] Tomasi C, Manduchi R. Bilateral filtering for gray and color images[C]. *Proceedings of International Conference on Computer Vision*, 1998.

- [69] Sapiro G. Geometric partial differential equations and image analysis[M]. Cambridge University Press, 2001.
- [70] Tschumperle D. Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE's[J]. International Journal on Computer Vision, 2006, 68(1): 65–82.
- [71] Gilboa G, Sochen N, Zeevi YY. Variational denoising of partly-textured images by spatially varying constraints[J]. IEEE Transactions on Image Processing, 2006, 15(8): 2281–2289.
- [72] Donoho D. Denoising by soft-thresholding[J]. IEEE Transaction on Information Theory, 1995, 41(3): 613–627.
- [73] Witkin AP. Scale space filtering[C]. Proceedings of International Joint Conference on Artificial Intelligence, 1983, 1019–1021.
- [74] Lindeberg T. Edge detection and ridge detection with automatic scale selection[J]. International Journal of Computer Vision, 1998, 30(2): 117–154.
- [75] Elder JH, Zucker SW. Local scale control for edge detection and blur estimation[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1998, 20(7): 699–716.
- [76] Kennedy LM, Basut M. Scale space contours and localization property of a Gaussian derivative edge enhancement operator[C]. Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, 1997, 643–648.
- [77] Lindeberg T. Feature detection with automatic scale selection[J]. International Journal of Computer Vision, 1998, 30(2): 79–116.
- [78] Cheng HD, Li J. Fuzzy homogeneity and scale-space approach to color image segmentation[J]. Pattern Recognition, 2003, 36(7): 1545–1562.
- [79] Jolion JM, Megret R. Tracking scale-space blobs for video description[J]. IEEE Multimedia, 2002, 9(2): 34–43.
- [80] Leung Y, Zhang JS, Xu ZB. Clustering by scale-space filtering[J]. IEEE Trans. on Pattern Recognition and Machine Intelligence, 2000, 20(12): 1396–1410.
- [81] Lindeberg T. Scale-space theory in computer vision[M]. Kluwer Academic Publishers, 1994.
- [82] Babaud J, Witkin AP, Baudin M, et al. Uniqueness of the Gaussian kernel for scale-space filtering[J]. IEEE Transaction on Pattern Recognition and Machine Intelligence, 1986, 8(1): 26–33.
- [83] Chan T, Osher S, Shen J. Total variation denoising and enhancement of color images based on the CB and HSV color models[J]. Journal of Visual Communication and Image Representation, 2001, 12(4): 422–435.
- [84] Weickert J. Anisotropic diffusion in image processing[M]. Teubner-Verlag, 1998.
- [85] Dolcetta IC, Ferretti R. Optimal stopping time formulation of adaptive image filtering[J]. Applied Mathematics and Optimization, 2001, 43: 245–258.
- [86] Weickert J. Coherence-enhancing diffusion of colour images[J]. IEEE Transactions on Image Processing, 1999, 17(3): 201–212.
- [87] Solo V. Fast automatic stopping criterion for anisotropic diffusion[C]. Proceedings of IEEE International Conference on Acoustics, Speech and Signal, 2002, 1661–1664.
- [88] Mrazek P. Selection of optimal stopping time for nonlinear diffusion filtering[J]. International Journal of Computer Vision, 2003, 52(2/3): 189–203.
- [89] Papandreou G, Maragos P. A cross-validated statistical approach to scale selection for image denoising by nonlinear diffusion[C]. Proceedings of the Conference on Computer Vision and Pattern Recognition, 2005, 625–630.
- [90] Blake A, Rother C, Brown M, et al. Interactive image segmentation using an adaptive GMMRF model[C]. Proceedings of European Conference on Computer Vision, 2004, 428–441.
- [91] Boykov Y, Kolmogorov V. Computing geodesics and minimal surfaces via graph cuts[C]. Proceedings of International Conference on Computer Vision, 2003.
- [92] Li Y, Sun J, Tang CK, et al. Lazy snapping[J]. ACM Transaction on Graphics, 2004, 23(3): 303–308.

-
- [93] Agarwala A, Dontcheva M, Agrawala M, et al. Interactive digital photomontage[J]. *ACM Transaction on Graphics*, 2004, 23(3): 292–300.
- [94] Sun J, Zheng NN, Shum HY. Stereo matching using belief propagation[J]. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 2003, 25(7): 1–14.
- [95] Kwatra V, Schodl A, Essa I, et al. Graphcut textures: image and video synthesis using graph cuts[J]. *ACM Transactions on Graphics*, 2003, 22(3): 277–286.
- [96] You YL, Xu WY, Tannenbaum A, et al. Behavioral analysis of anisotropic diffusion in image processing[J]. *IEEE Transactions on Image Processing*, 1996, 5(11): 1539–1553.
- [97] Aubert G, Kornprobst P. *Mathematical problems in image processing - partial differential equations and the calculus of variations*[M]. New York: Springer-Verlag, 2002.
- [98] Zitnick CL, Jovic N, Kang SB. Consistent segmentation for optical flow estimation[C]. *Proceedings of the Tenth IEEE International Conference on Computer Vision*, 2005, 1308–1315.
- [99] Liu C, Freeman WT, Szeliski R, et al. Automatic estimation and removal of noise from a single image[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2008, 30(2): 299–314.
- [100] Li SZ. *Markov random field modeling in image analysis*[M]. Tokyo:Springer-Verlag, 2001.
- [101] Modestino JW, Zhang J. A Markov random field model-based approach to image interpretation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 1992, 14(6): 606–615.
- [102] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2004, 26(9): 1124–1137.
- [103] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2001, 23(11): 1222–1239.
- [104] Felzenszwalb PF, Huttenlocher DP. Efficient belief propagation for early vision[J]. *International Journal of Computer Vision*, 2006, 70(1): 41–54.
- [105] Murphy K, Weiss Y, Jordan M. Loopy belief propagation for approximate inference: an empirical study[C]. *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, 1999, 467–475.
- [106] Wainwright MJ, Jaakkola TS, Willsky AS. MAP estimation via agreement on (hyper)trees: message-passing and linear-programming approaches[J]. *IEEE Transactions on Information Theory*, 2005, 51(11): 3697–3717.
- [107] Kolmogorov V. Convergent tree-reweighted message passing for energy minimization[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2006, 28(10): 1568–1583.
- [108] Szeliski R, Zabih R, Scharstein D, et al. A comparative study of energy minimization methods for Markov random fields[C]. *Proceedings of European Conference on Computer Vision*, 2006, 16–29.
- [109] DeStefano A, White PR, Collis WB. Training methods for image noise level estimation on wavelet components[J]. *EURASIP Journal on Applied Signal Processing*, 2004, 16: 2400–2407.
- [110] Simoncelli E. Statistical models for images: compression restoration and synthesis[C]. *Proceedings of Asilomar Conference on Signals, Systems and Computers*, 1997, 673–678.
- [111] Bertalmio M, Bertozzi AL, Sapiro G. Navier-Stokes, fluid dynamics, and image and video inpainting[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, 417–424.
- [112] Ballester C, Bertalmio M, Caselles V, et al. Filling-in by joint interpolation of vector fields and gray levels[J]. *IEEE Transactions on Image Processing*, 2001, 10(8): 1200–1211.
- [113] Efros A, Leung T. Texture synthesis by non-parametric sampling[C]. *Proceedings of International Conference on Computer Vision*, 1999, 1033–1038.
- [114] Bertalmio M, Vese L, Sapiro G, et al. Simultaneous structure and texture image inpainting[J]. *IEEE Transactions on Image Processing*, 2003, 12(8): 882–889.
- [115] Criminisi A, Perez P, Toyama K. Object removal by exemplar-based image inpainting[C]. *Proceedings of International Conference on Computer Vision*, 2003, 721–728.

- [116] Wu J, Ruan Q. Object removal by cross isophotes exemplar-based image inpainting[C]. Proceedings of International Conference on Pattern Recognition, 2006, 810–813.
- [117] Wong A, Orchard J. A nonlocal-means approach to exemplar-based inpainting[C]. Proceedings of IEEE International Conference on Image Processing, 2008.
- [118] Tziritas G, Komodakis N. Image completion using efficient belief propagation via priority scheduling and dynamic pruning[J]. IEEE Transactions on Image Processing, 2007, 16(11): 2649–2661.
- [119] Jia JY, Tang CK. Image repairing: robust image synthesis by adaptive nd tensor voting[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, 643–650.
- [120] Elad M, Starck JL, Querre P, et al. Simultaneous cartoon and texture image inpainting using morphological component analysis[J]. Applied and Computational Harmonic Analysis, 2005, 19: 340 – 358.
- [121] Guleryuz OG. Nonlinear approximation based image recovery using adaptive sparse reconstructions[C]. Proceedings of IEEE International Conference on Image Processing, 2003.
- [122] Guleryuz OG. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part I: theory[J]. IEEE Transactions on Image Processing, 2006, 15(3): 539 – 554.
- [123] Guleryuz OG. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part II: adaptive algorithms[J]. IEEE Transactions on Image Processing, 2006, 15(3): 555 – 571.
- [124] Criminisi A, Perez P, Toyama K. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200–1212.
- [125] Li YP, Huttenlocher D. Sparse long-range random field and its application to image denoising[C]. Proceedings of the European Conference of Computer Vision, 2008.
- [126] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries[J]. IEEE Transaction on Image Processing, 2006, 15(12): 3736 – 3745.
- [127] Mairal J, Bach F, Ponce J, et al. Discriminative learned dictionary for local image analysis[C]. Proceedings of European Conference on Computer Vision, 2008.
- [128] Winn J, Criminisi A, Cristianini N. Object categorization by learned universal visual dictionary[C]. Proceedings of International Conference on Computer Vision, 2005.
- [129] Peyre G. Non-negative sparse modeling of textures[C]. Proceedings of Scale Space and Variational Methods in Computer Vision, 2007.
- [130] Chang H, Yeung DY, Xiong YM. Super-resolution through neighbor embedding[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004.
- [131] Mallat S, Zhang Z. Matching pursuit in a time-frequency dictionary[J]. IEEE Transactions on Signal Processing, 1993, 41: 3397 – 3415.
- [132] Pati YC, Rezaiifar R, Krishnaprasad PS. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition[C]. Proceedings of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, 1993.
- [133] Chen SS, Donoho DL, Saunders MA. Atomic decomposition by basis pursuit[J]. SIAM Review, 2001, 43(1): 129 – 159.
- [134] Tibshirani R. Regression shrinkage and selection via the Lasso[J]. Journal of Royal Statistical Society B, 1996, 58(1): 267–288.
- [135] Figueras i Ventura RM, Vandergheynst P, Frossard P, et al. Color image scalable coding with matching pursuit[C]. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004.
- [136] Gribonval R, Fvotte C, Vincent E. Performance measurement in blind audio source separation[J]. IEEE Transaction on Speech, Audio, Language Processing, 2006, 14(4): 1462–1469.
- [137] Boykov Y, Jolly MP. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images[C]. Proceedings of International Conference on Computer Vision, 2001.

-
- [138] Rother C, Blake A, Kolmogorov V. Grabcut - interactive foreground extraction using iterated graph cuts[J]. *ACM Transaction on Graphics*, 2004, 23(3): 309–314.
- [139] Bhat P, Zheng KC, Snavely N, et al. Piecewise image registration in the presence of multiple large motions[C]. *Proceedings of Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, 2491–2497.
- [140] Torr PHS, Szeliski R, Anandan P. An integrated bayesian approach to layer extraction from image sequences[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2001, 23(3): 297–303.
- [141] Wills J, Agarwal S, Belongie S. What went where[C]. *Proceedings of Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003, 37–44.
- [142] Xiao JJ, Shah M. Motion layer extraction in the presence of occlusion using graph cut[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, 972–979.
- [143] Kolmogorov V, Criminisi A, Blake A, et al. Probabilistic fusion of stereo with color and contrast for bi-layer segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(9): 1480–1492.
- [144] Bolle RM, Connell JH, Haas N, et al. Object imaging system[P]. U. S. Patent 5,631,976. 1994.
- [145] Sun J, Zhang W, Tang X, Shum HY. Background cut[C]. *Proceedings of European Conference on Computer Vision*, 2006, 628–641.
- [146] Apostoloff N, Fitzgibbon A. Bayesian video matting using learnt image priors[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, 407–414.
- [147] Chuang YY, Agarwala A, Curless B, et al. Video matting of complex scenes[J]. *International Journal of Computer Vision*, 2002, 21(3): 243–248.
- [148] Raskar R, Tan K, Feris R, et al. Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging[J]. *ACM Transactions on Graphics*, 2004, 23(3): 673–678.
- [149] Lowe DG. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91–110.
- [150] Lucas B, Kanade T. An iterative image registration technique with application to stereo vision[C]. *Proceedings of International Joint Conference Artificial Intelligence*, 1981, 674–679.
- [151] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts?[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(2)
- [152] Rother C, Kolmogorov V, Minka T, et al. Cosegmentation of image pairs by histogram matching-Incorporating a Global Constraint into MRFs[C]. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, 993–1000.
- [153] Shum HY, Sun J, Yamazaki S, et al. Pop-up light field: an interactive image-based modeling and rendering system[J]. *ACM Transaction of Graphics*, 2004, 23(2): 143–162.
- [154] Kramer HP, Bruckner JB. Iterations of a non-linear transformation for enhancement of digital images[J]. *Pattern Recognition*, 1975, 7(1-2):53–58.

致谢

本文是在我的导师徐宗本教授的悉心指导和帮助下完成的。在攻读博士学位期间，他对我的研究和生活都给予了无微不至的关怀与支持，本文的选题、写作与修改均凝聚着恩师的心血。徐老师严谨的治学态度、广博而精髓的学术造诣、勤奋努力的拼搏精神和对学生的生活与前途的关怀，都对我产生了深刻的影响，值得我一生去学习。今后，我只会更加努力、勤奋地学习与工作，以不辜负徐老师对我的关怀与培养。

特别感谢微软亚洲研究院研究员孙剑博士、微软全球副总裁沈向洋博士和微软Redmond研究院主任研究员Kang Singbing博士。在微软亚洲研究院访问研究期间，他们在科研和论文写作方面所给予的有力指导和帮助使我受益匪浅，他们勤奋专注的科研精神和注重实践的治学态度深深地影响着我。

感谢陈白丽老师和张继洁老师，她们在生活和科研上所给予的热情关怀与帮助，使我能够更好地开展研究工作。

感谢所有的师兄（弟姐妹）和在微软共同实习的同事，与他们在科研上广泛而有益的讨论推动着我不断前进。

最后，特别感谢我的父母与妻子，他们所给予我的无私的支持、付出与家庭的温暖是我勤奋工作与不断进步的动力。

攻读学位期间取得的研究成果

- [1] **Sun J**, Xu ZB. Scale selection for anisotropic diffusion filter by Markov random field model[J]. Pattern Recognition, Accepted, 2009 (SCI, EI 源期刊).
- [2] **Sun J**, Sun J, Xu ZB, Shum HY. Image super-resolution using gradient profile prior[C]. In Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2008, 1: 2471-2478 (EI: 083911592126).
- [3] **Sun J**, Sun J, Kang SB, Xu ZB, Tang X, Shum HY. Flash cut: foreground extraction with flash and no-flash image pairs[C]. In Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2007, 1: 843-850 (EI: 074110860595).
- [4] **Sun J**, Xu ZB. An edge preserving regularization model for image restoration based on Hopfield neural network[C]. Lecture notes in computer science, 2006, 3972:563-568 (SCI: BET84, EI: 062910011239).
- [5] **孙剑**, 徐宗本. 计算机视觉中的尺度空间方法[J]. 工程数学学报, 2005, 22 (6): 951-962.
- [6] 谷雨,徐宗本,**孙剑**,郑锦辉. 基于PCA与ICA特征提取的入侵检测集成分类系统[J]. 计算机研究与发展, 2006, 43(04): 633-638.
- [7] 谷雨,郑锦辉,**孙剑**,徐宗本,基于独立成分分析和支持向量机的入侵检测方法[J]. 西安交通大学学报, 2006, 39(8): 876-879 (EI: 05369347382).
- [8] Xu ZB, **Sun J**. Image inpainting by patch propagation using sparsity of natural image patches[J]. Submitted to IEEE Transactions on Image Processing, 2009.
- [9] **Sun J**, Sun J, Xu ZB, Shum HY. Gradient profile prior and its applications in image super-resolution and enhancement[J]. Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence (in revision), 2009.

学位论文独创性声明

本人声明，所呈交的学位论文系在导师指导下本人独立完成的研究成果。文中依法引用他人的成果，均已做出明确标注或得到许可。论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果。

本人如违反上述声明，愿意承担以下责任和后果：

1. 交回学校授予的学位证书；
2. 学校可在相关媒体上对作者本人的行为进行通报；
3. 本人按照学校规定的方式，对因不当取得学位给学校造成的名誉损害，进行公开道歉。
4. 本人负责因论文成果不实产生的法律纠纷。

论文作者签名： 日期： 年 月 日

学位论文知识产权权属声明

本人在导师指导下所完成的论文及相关的职务作品，知识产权归属学校。学校享有以任何方式发表、复制、公开阅览、借阅以及申请专利等权利。本人离校后发表或使用学位论文或与该论文直接相关的学术论文或成果时，署名单位仍然为西安交通大学。

论文作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

（本声明的版权归西安交通大学所有，未经许可，任何单位及任何个人不得擅自使用）