

图像协同关联性约束的研究与应用

Research and Application of Image Correspondence Constraint

一级学科: 计算机科学与技术
学科专业: 计算机应用技术
作者姓名: 付华柱
指导教师: 操晓春 研究员

计算机科学与技术学院
二零一三年四月二十五日

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得天津大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名: 签字日期: 年 月 日

学位论文版权使用授权书

本学位论文作者完全了解天津大学有关保留、使用学位论文的规定。特授权天津大学可以将学位论文的全部或部分内容编入有关数据库进行检索，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构递交论文的复印件和磁盘。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名: 导师签名:
签字日期: 年 月 日 签字日期: 年 月 日

摘要

目标间的协同关联性约束研究是计算机视觉领域中的热点问题之一。其主要指在单张图像处理受到制约或瓶颈时，通过引入其他图像或参考源，分析相互间关系，构建全局的关联性约束，以此来辅助提升图像处理的效果。但是现有的协同关联性约束算法大多数集中于多张图像之间的近似物体关联约束，而且算法构造复杂、计算量大。在本论文中，我们将深入研究协同关联性约束的核心思想，理解其构成形式。一方面，进一步挖掘多图像间的关联性约束，提高多张图片间的协同处理效果。另一方面，我们还将借鉴协同关联分割的思想，将其协同约束的思想拓展到单幅图像的应用中，通过构建单张图像不同区域间的关联，以及单图像相关源间的关联，提高图像的处理效果。本文的主要研究内容包括：

1) 多张图像间的关联性约束：我们通过聚类构建多张图像之间的协同关联性约束。基于图像生物视觉显著性原理和多图像间目标关联性约束，提出一种基于聚类方法的多图片间的协同显著性目标检测算法。协同显著性主要指多图像中的重复出现的同一或近似的视觉显著性物体。我们的算法通过聚类构建多图像之间的物体全局关联结构，依据三种自底向上的协同显著性测度，包括特征对比度、中心偏移度、以及图间分布度，对图像进行聚类级别的协同显著性测量，最终将各测度结果进行融合，提取出多张图片中的协同显著性目标。我们的协同显著性物体检测算法避免了大量繁重的学习训练过程，具有简单高效的特点。该算法不仅在多张图片的协同显著性物体检测中取得优异结果，而且在单张图片的视觉显著性物体检测中，也取得了良好的测试结果。此外，我们还深入挖掘协同显著性物体检测算法的相关应用，提出四种典型应用：协同分割、鲁棒的图像测距、弱监督学习、以及视频前景提取等，进一步彰显了协同显著性检测在图像处理中的应用潜力。

2) 单张图像不同区域间的关联性约束：我们挖掘单张图像中物体自身的几何结构约束，对图像的不同区域间建立协同联系。自然图片中的物体往往满足一定的几何结构，这使得各个区域之间可以构建协同关联性约束。我们利用物体自身的这种几何结构约束，引入协同分割思想，提出一种图像内关联区域之间的几何结构约束分割算法。为了反映不同区域之间的协同关联性约束，我们首先利用物体自身的几何结构约束建立一个像素级别的稠密几何结构映射矩阵，该矩阵在描述前景物体的位置分布的同时，也反映了物体自身的几何结构约束。然后我们将这种几何结构矩阵引入到基于图的能量函数模型中，提出新的基于几何结构约束的图分割模型，该模型满足子模性，可

以通过图分割算法优化求解。除此之外，我们还将该几何结构约束模型拓展到基于组件的分割框架中，以满足弱几何约束、以及复杂几何约束的物体分割需求。实验表明，这种不同区域之间的协同关联性约束提供了一种高级别的图像语义关联，相对于仅仅使用低级别特征的分割算法，我们的算法取得了更好的分割效果。

3) 单图像相关源间的关联性约束：除了单张图片不同区域间的关联以外，我们还发现，同图片的不同处理源之间也可以协同关联性约束。通过研究这些不同处理源之间的关联性约束，建立单幅图像与镜头畸变、以及单幅图像与其不同操作处理层之间的协同关联性，我们提出一种基于图像畸变度的广角图像伪造区域盲检测分割算法。镜头的径向几何畸变在绝大多数图像处理中通常被认为是需要消除的不良性质，但我们提出该几何畸变反映了镜头的内部结构属性，而且对图像提供一个全局的几何约束。因此我们将这种镜头径向几何畸变性质作为度量指标，引入到图像篡改盲检测取证算法中。首先，提出了一种镜头径向畸变映射模型，在该模型中，空间直线将在映射半球上投影成一个大切圆。其次，依据该几何约束，设计了两种底层测度用于获取目标图像的伪造取证分布图，检测图像中伪造物体的位置。最后，通过构造基于图分割的能量函数，将目标图像与其伪造取证分布图关联起来，作为单图像的两个处理层，构建协同关联，提取出像素级别的目标取证呈现。

通过研究我们发现，协同关联性约束不仅存在于多图像之间，而且也可以存在于单张图像之中，通过图像的不同结构区域、以及相关源之间体现出来。这种协同关联性约束可以突破现有单张图像的局限，提高单/多张图像处理的结果，提供新的结构语义描述，以及提供有效的参考验证准则。

关键词： 关联性约束，显著性检测，图像协同分割，协同显著性检测，前景分割，图像取证

ABSTRACT

Correspondence constraint is the hot issue in computer vision. Most existing correspondence constraint methods focus on the similar objects in multiple images, and they are often computationally demanding. In this paper, we study the theory of correspondence constraint, demonstrate potential usages of the correspondence constraint. On one hand, we discover the correspondence constraint between the multiple images, and promote its performance. On other hand, we extend the idea of correspondence constraint into the single image processing, and generate the relation corresponding between the multiple regions and processing sources of the single image. Our paper includes:

1) Correspondence constraint between the multiple images: we employ the clustering to generate the correspondence constraint between the multiple images. We introduce a new cluster-based algorithm for co-saliency detection, which is based on laws of the visually salient stimuli and correspondence constraint. Global correspondence between the multiple images is implicitly learned during the clustering process. Three visual attention cues are devised to effectively measure the cluster saliency. The advantage of our method is mostly bottom-up without heavy learning, and has the property of being simple, general, efficient, and effective. Experimental results demonstrate the advantages of the proposed method over the competing co-saliency methods. Our method on single image also outperforms most the state-of-the-art methods. Furthermore, we apply the co-saliency method on four applications to demonstrate the potential usages of the co-saliency map.

2) Correspondence constraint between the multiple regions in the single image: we employ the geometry structure to build correspondence constraint between the multiple regions in the single image. We take the geometry structure constraint into the foreground extraction, and propose a novel geometry constraint segmentation method for extracting the foreground. Firstly, the geometry foreground map is used to represent the geometry structure of the image, which includes the geometry matching magnitude and the foreground location prior. Then, the geometry constraint model is built by introducing this geometry structure into the graph-based segmentation function. Finally, the segmentation result is obtained via graph cut. Moreover, our geometry constraint segmentation is also extended to the weak geometry object under a part-based framework. Experiments demonstrate that the high-level property of geometry constraint significantly improves the low-level segmentation results.

3) Correspondence constraint between the multiple processing sources of the single image: we build the correspondence constraint between the single image and lens distortion/multiple processing layers, and propose a novel forensic method for detecting the forgery object. We employ the radial distortion as the intrinsic property of the lens, which could offer a global constraint. A modified spherical projection model is adopted, which is equivalent to the other captured rays-based models of the fisheye lens with only one free parameter. In this model, the straight world line is projected into a great circle on the viewing sphere, which provides a unique geometric constraint. Two saliency measure cues are provided to compute the untrustworthy likelihoods of the candidate lines. Finally, a fake saliency map is obtained according to the untrustworthy likelihood to segment the fake region.

Above all, we find out that the correspondence constraint is not only valid on the multiple images, but appears in the single image processing, which could be provided by multiple regions and processing sources. This correspondence constraint breaks the limited of single image, which improves the performance of image processing, provides the new semantics structure, and offers the effective reference.

KEY WORDS: Correspondence Constraint, saliency detection, co-segmentation, co-saliency detection, foreground segmentation, image forensics

目 录

摘要	I
ABSTRACT	III
目 录	V
第一章 绪论	1
1.1 研究意义与相关背景	1
1.1.1 主要相关背景	2
1.2 本文的主要研究内容与贡献	8
1.3 本文的结构	9
第二章 多图像间的关联性约束	11
2.1 相关背景工作	11
2.2 协同显著性目标检测方法	13
2.2.1 基于聚类结构的全局关联	14
2.2.2 聚类级的显著性测度	14
2.2.3 显著性测度的融合	18
2.3 实验与分析	19
2.3.1 单图片的显著性目标检测实验	19
2.3.2 协同显著性目标检测实验	21
2.3.3 运行速度测试	23
2.3.4 相关讨论	25
2.4 协同显著性检测的应用	27
2.4.1 协同分割应用	27
2.4.2 鲁棒的图像测距	30

2.4.3 弱监督学习应用	32
2.4.4 视频前景提取	32
2.5 本章总结	33
 第三章 单张图像不同区域间的关联性约束	37
3.1 相关背景工作	37
3.2 几何结构映射矩阵	38
3.2.1 前景先验分布图	39
3.2.2 几何约束强度分布图	40
3.2.3 几何结构强度因子	41
3.3 几何结构约束模型	42
3.3.1 嵌入几何约束的数据惩罚项	43
3.3.2 几何约束平滑项	44
3.3.3 基于几何结构约束的前景分割能量函数	44
3.4 基于组件的扩展框架	45
3.5 实验与分析	46
3.5.1 几何结构约束模型分解测试	46
3.5.2 分割准确度比较	48
3.5.3 数据库的批量实验结果	48
3.5.4 与嵌入几何结构的level-set 算法比较	51
3.5.5 算法鲁棒性讨论	52
3.6 本章总结	53
 第四章 单图像相关源间的关联性约束	55
4.1 相关背景工作	55
4.2 镜头径向畸变映射模型	57
4.2.1 反反射模型	57
4.2.2 径向畸变映射模型的相关性质	58
4.3 基于直线畸变几何约束的取证测度	59
4.3.1 直线畸变几何约束	59

目 录

4.3.2 体积取证测度	61
4.3.3 距离取证测度	61
4.3.4 测度取证融合	62
4.4 伪造取证分布图	62
4.4.1 取证非置信度检测	63
4.4.2 取证分布图生成	64
4.4.3 检测结果分割呈现	65
4.5 实验结果分析	66
4.5.1 径向畸变映射模型实验	66
4.5.2 模拟实验	66
4.5.3 真实图像测试	70
4.5.4 伪造取证分布图的应用	71
4.5.5 相关失败情况	73
4.6 本章总结	76
第五章 总结与展望	79
5.1 总结	79
5.2 展望	80
参考文献	81
发表论文和参加科研情况说明	90
致 谢	92

第一章 绪论

“所谓联系是指事物内部诸要素之间以及事物之间的相互影响、相互制约和相互作用。”

——唯物辩证法

1.1 研究意义与相关背景

目标间的协同关联性研究是计算机视觉领域中的热点问题之一。其主要指在单张图像处理受到制约或瓶颈时，通过引入其他图像或参考源，分析相互间关系，构建全局的关联性约束，通过相互合作、协同制约等形式，提升各自的图像处理的结果。常见的关联性约束主要体现在多张图片之间，可以基于任意的图像处理操作。

一个典型的协同关联性应用就是图像分割问题^[1-4]。图像分割是计算机视觉领域的基础方向，其可以广泛应用于目标识别，场景理解，三维重建，图像检索等领域。其中对图像前景目标的提取是图像分割的一种常见处理，其主要是通过对图像进行二值处理，以区分图像中前景目标和背景区域。通常的图像分割算法主要依靠像素颜色、纹理、边缘等低级别（low-level）的像素信息，缺少完整的物体级别（object-level）信息，以及相关的语义特征信息，因此分割精度有限。同时，由于遮挡或复杂背景的干扰，容易造成信息缺失。此外，由于图像中的物体通常存在前/背景的二义性，往往需要介入人工交互选取来实现，无法实现自动化。近年来，图像协同分割（co-segmentation）的概念被提出^[5-11]，图像协同分割将构建多张图像间的全局关联性约束，将像素级别的不同测量度引入到一个全局约束函数中，该全局约束可以作为一个高级别的物体/语义关联，而提供全局协同约束，在此基础上，协同分割算法可以有效地提高多张图片中近似物体的分割精度，可以获得缺失信息的互补，同时，去除前景物体的二义性。

可见，在引入协同关联性约束后，图像处理效果可以得到有效提升，而且减少了图像前景的语义二义性。但是，现有的协同关联性约束算法大多数集中于多张图像之间的近似物体关联，同时，关联结构约束构造复杂、计算量大。因此，在本文中我们深入研究基于协同关联性约束的图像分割思想，并将这种多目标的协同关联约束推广到其他相关应用中。首先，研究多张图片之间的协同关联性约束，通过基础聚类方式，挖掘多图像间的关联性约束

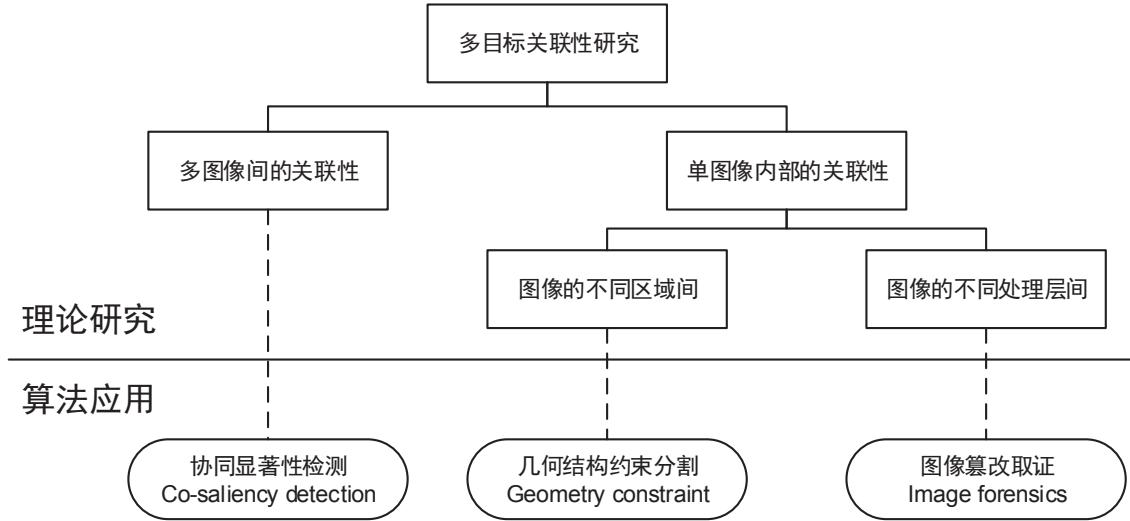


图 1-1 本文的主要研究内容与结构框架。

目标约束。其次，本文还将借鉴协同关联分割的思想，但不局限于多张图片的协同分割中，将其协同约束的思想拓展到单幅图像的应用中，通过构建单张图像不同区域间的关联，以及单图像相关源间的关联，提高图像的处理效果。除此之外，本文将通过具体应用来研究相关理论基础，我们主要涉及三方面研究，如图1-1所示：1) 多张图像间的关联性约束：我们通过聚类构建多图像之间的全局物体关联，依据图像生物视觉显著性原理和多图像间目标关联性约束，提出基于聚类方法的基于多张图片的协同显著性目标检测分割算法。2) 单图像不同区域间的关联性约束：我们发现单幅图像中的不同区域之间也可以近似的作为一种协同分割应用。因此我们利用单张图像中物体自身的几何结构约束，对图像的不同区域间建立协同联系，分割图像中满足几何结构约束的前景物体。3) 单图像相关源间的关联性约束：同一图像的不同处理源之间存在着关联性约束。我们受协同分割算法的启发，研究这些不同处理源之间的关联性约束，建立图像与拍摄镜头、以及图像与其不同操作处理层之间的协同关联性约束，并提出一种基于图像畸变度的广角图像伪造区域盲检测分割算法。

1.1.1 主要相关背景

在本小节中，我们将主要展示协同关联性约束的相关研究背景。由于在每个单章中，我们还将更加详细地介绍每个技术点的相关背景工作，因此，为了避免重复，在这里我们仅仅介绍一些粗略的相关研究背景。

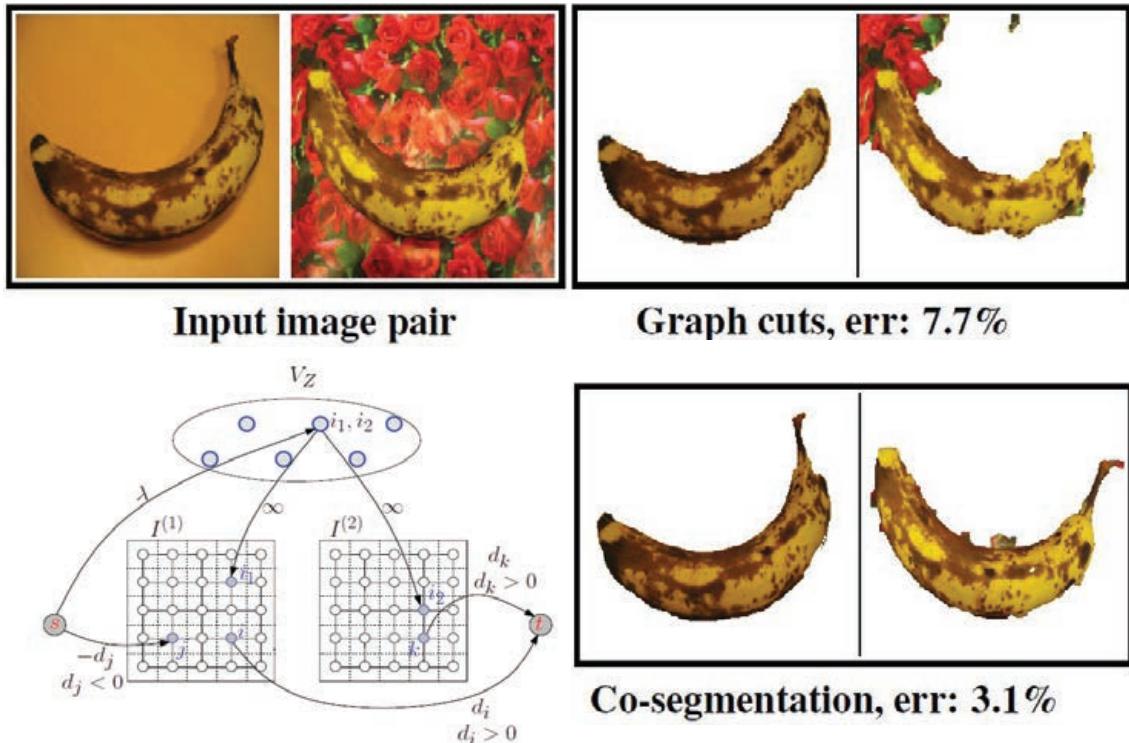


图 1-2 本图引自文献^[6], 左上为输入图像对, 左下为构建全局关联性约束图, 右上为单幅图像的分割, 右下为图像协同分割。可见基于多张图像的协同关联性约束, 图像协同分割相对于单张图像分割, 可以明显的提高分割精度, 减少二义性。

1.1.1.1 多图像的协同分割思想

如图1-2所示, 协同分割是从两幅或多幅图像中, 分割出相似的前景目标区域或者物体。最早由文献^[5]引入该概念, 并迅速引起关注。图像的协同分割主要分为两种方式, 一是根据传统的马科夫随机场(MRF)定义, 引入全局约束项, 定义为:

$$\min E(L_1, L_2) = E^{image}(L_1, L_2) + E^{global}(h_1, h_2), \quad (1-1)$$

其中, $E^{image}(L_1, L_2)$ 为传统的图分割能量方程^[12-14], $E^{global}(h_1, h_2)$ 为全局约束项, 该全局约束项通常是通过两幅图像间的直方图分割近似度来设定, 不同方法间的区别主要在于: L1 norm 模型^[5], L2 norm 模型^[15], 奖励模型^[6], 修改的Boykov-Jolly 模型^[7], 以及L12 norm 模型^[6]等几种不同的全局约束项。这种方法可以取得极高的精确性, 但缺点是首先, 对于图分割来说, 需要预先手动设定初始的源点值, 其次, 这些模型都局限于两幅图像, 很难拓展到多幅图像中。相对的, 另一种基于多组图像的协同分割方法逐渐被提出, 其主要有以下几种: 文献^[8]提出的基于少量人机交互操作的半监督协同分割算法。文献^[16]提出的基于分离聚类的协同分割算法。文献^[17]提出的基于同向异构

扩散的协同分割算法。文献^[9] 基于特征矩阵秩约束的尺度不变协同分割算法。以及文献^[10] 提出的基于物体检测级别的协同分割算法。这些算法的共同性是将协同约束作为一个全局性特征的弱约束条件，而不是直方图的距离匹配，因此可以有效的拓展到多张图片中。

这些协同分割算法都获得了良好的分割效果，其主要得益于优秀的物体级别（object-level）的全局约束。因为单张图片所能提供的信息量较少，容易被遮挡或者复杂背景干扰，造成信息缺失，此外，单张图片只能提供低级别的像素信息。相对的，当引入多张图片形成一个图像子集后，在子集中的近似物体便可以提供更多的高级别属性，而且还可以获得缺失信息的互补。基于此，协同处理往往可以获得比单幅图片更好的分割效果。

1.1.1.2 协同显著性目标检测



图 1-3 给一组图片（第一行），现有的显著性目标检测算法（例如^[18]）容易被复杂的背景所干扰，而且缺乏多图片间的关联信息（第二行）。相对于单张图片的显著性目标检测，我们的协同显著性目标检测算法可以准确的提取出多张图片间的重复出现的近似目标（第三行），保持多图像间的关联性。

我们第一个研究内容是多图像间的关联性约束。多图像间的关联性约束是一种常见的协同约束方式，当多张图片包含相同或相似物体时，物体在图像彼此之间就构成了关联性约束。我们将基于图像生物视觉显著性原理与多图像间目标关联性约束相结合，提出一种基于聚类方法的多图片间的协同显著性目标检测算法（Cluster-based Co-saliency Detection Method）。

显著性检测通常被认为是一种有效的预处理过程，其依据人类生物视觉特性规则快速定位目标^[18-22]。大多数现有显著性检测算法都侧重于从单张图片中检测显著目标^[23-25]。近年来，多张图像间的协同关联作用受到越来越多的重视，基于此协同显著性目标的概念也随之被提出。协同显著性目标检

测首先被定义为在一组近似图像中，挖掘不同物体^[26]。然而，这种定义需要这组近似图像是在短时间内拍摄，这种约束条件限制了检测算法的应用。另一种被广为接受的协同显著性目标概念被定义为在一组图片中，提取重复出现的近似的显著性目标^[27-29]。在这种定义下的协同显著性目标检测具有更防范的应用场景，例如，协同分割（cosegmentation）^[5,6,10]、近似目标检测（common pattern discovery）^[30,31]、目标协同识别（object co-recognition）^[32,33]、图像检索（image retrieval）^[34]、以及图像简报生成（image summaries）^[8,35]等。在本文中，我们将按照后一种的协同显著性目标定义，提出基于聚类方法的基于多张图片的协同显著性目标检测算法。

图1-3 展示了一组多图像协同显著性目标检测的例子。其中单张图片的显著性检测算法^[18]（第二行）提取每一张的显著区域，但是容易被复杂背景干扰。例如，前两张图片中的观众席就被错误的标准为显著目标。此外，单张图片图像还丢失了多图像间的相关性。相对地，我们的协同显著性目标将多图的重复出现性作为一种约束性质，挖掘了多组图片中重复出现的显著性目标，例如图1-3 中的第三列所示的红色运动员。

1.1.1.3 几何结构约束的前景分割



图 1-4 自然界中的物体自身存在大量的几何结构，如对称、旋转、重复、平移等。

我们第二个主要研究内容是单张图像不同区域间的关联性约束。自然图片中的物体往往满足一定的几何结构，这些几何结构可以使图像中的不同区域构成一定的结构约束。我们将利用物体自身的这种几何结构约束，引入协同分割思想，提出一种图像内关联区域之间的几何结构约束分割算法（Geometry Constraint for Foreground Extraction）。

前景分割是计算机视觉的基础问题之一，它的目的是将图片分成前景和背景两个独立的部分^[36-38]。大多数已有算法都是构建一个图函数，并进行二

值分割求解，从而分割出前景物体^[1,3,39,40]。但这些算法都仅仅考虑构建前/背景像素之间的关系，而忽略物体自身的结构属性，例如对称性、重复性等。自然图片中的物体往往满足一定的几何约束，例如对称、旋转、复制等几何结构（如图1-4所示），这些几何结构将图片按照一定的结构约束，划分成多个关联区域，各个区域之间通过几何结构约束相互联系。这种几何结构约束的优势在于它反应了物体内在的形状属性，简单、普及，而且不需要训练学习。尽管几何结构属性被提出了很长时间^[41-45]，但如何将这种高级别的几何特征融入到低级别的像素分割还需进一步研究。在本文中，我们将主要讨论提取这种物体自身的几何结构信息，并将其引入到前景分割算法中去。我们提出一种单幅图像中前景物体几何结构映射矩阵，该矩阵可以有效地反应前景物体的几何结构分布情况，并提供一个稠密的像素级别的图像对称结构分布图。同时，将这种几何结构矩阵引入到传统的马尔科夫随机场模型（Markov Random Fields, MRF）中，来提高图像前景的分割精度。

近年来已有一些方法将几何约束概念引入到图像处理中^[46-49]。例如，文献^[50]通过映射约束（transformation constraints）将目标划分为结构化子区域。但该方法并没有涉及前景提取。文献^[48]将对称关联矩阵（affinity matrix）被引入到区域增长算法中（region growing algorithm），但该方法仅仅涉及一般的图像分割问题，对于几何结构在前景提取中的作用，并没有深入展开。文献^[49]则是基于对称边缘约束，其将物体的对称形状作为动态形状先验加入到level-set分割函数中。但该算法仅能针对对称结构的物体，而无法推广到普通情况。相对而言，我们的算法与这些已有算法不同，区别在于：首先，我们的算法侧重于提取前景物体，而不是特殊几何结构的物体。几何结构约束仅仅作为一个约束项加入到分割函数中。第二，我们的几何结构映射矩阵是基于全局几何结构和局部最优匹配而获得的，具有更高的鲁棒性。第三，我们的几何结构约束模型满足子模性，可以通过图分割算法求取最优解。

1.1.1.4 基于镜头畸变度的图像伪造取证

我们第三个研究内容是单图像相关源间的关联性约束。数字图像往往存在许多相关源，例如拍摄设备、拍摄镜头等硬件，以及在图像处理过程中的中间环节，例如预处理图像层、检测结果分布图等。我们将主要研究这些不同处理源之间的关联性约束，建立单幅图像与镜头畸变、以及单幅图像与其不同操作处理层之间的协同关联性，并提出一种基于图像畸变度的广角图像伪造区域盲检测算法。

通常来说，镜头畸变（distortion）是计算机视觉处理中的典型问题^[51-55]。大量的研究工作主要针对如何消除或者较弱镜头畸变所带来的不良影响^[56-59]，很少有工作把镜头畸变看做是一个重要的图像内在属性。而实际上，对于存在明显畸变的广角镜头（extreme wide-angle lens）或者全景镜头（omnidirectional lens）来说，镜头畸变可以提供一种非常有效地图像全局几何约束。在本文中，我们将展示如何利用镜头畸变作为测度量，测量广角镜头

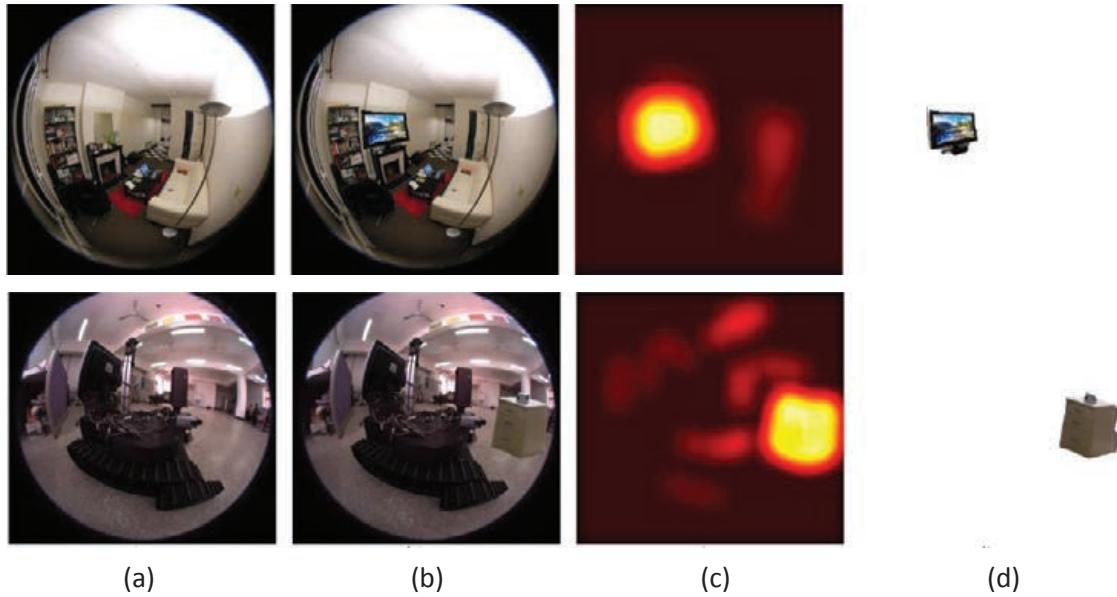


图 1-5 检测伪造的鱼眼镜头图像。(a) 原始图像。(b) 伪造的图像。(c) 我们算法检测的伪造取证分布图。(d) 像素级别的取证呈现。

图像中的伪造物体。

由于超广的视角 (field of view, FOV) 优势, 广角镜头通常被用来作为室内外的监控设备, 而监控视频常被用来作为重要事件的司法取证。因此, 如何确定监控图像/视频的真伪度便成为重要的鉴定核心之一。此外, 由于技术的发展, 大量价廉高质量的广角镜头大量出现, 越来越多的人出于兴趣原因使用广角镜头来采集图像/视频^[60,61]。如其他数字图像一样, 广角镜头图像同样很容易被大量功能强劲的图像处理软件进行编辑和篡改, 这也容易导致图像的可信度测量问题^[62]。图1-5 (b) 展示了两个伪造的鱼眼镜头图像。伪造的前景物体被粘贴复制到原始图像 (图1-5 (a)) 中。

图像内容取证技术指不需要对原始数据进行预处理, 而直接对目标进行取证分析操作^[63], 更适合于对未知来源的多媒体进行真实性检测, 如网络环境中图像/视频/音频真实性区分、电子商务中的身份认证和资格认证、以及司法取证中图像/视频/音频证据的鉴定等应用, 具有重要的科学价值和社会价值。图像内容取证技术按照所使用的约束, 可以粗略的分为五种类型: 1) 基于基本像素 (pixel-based)^[64], 主要针对复制、重采样、剪切等操作的图像编辑检测。2) 基于存储格式 (format-based)^[65,66], 主要检测图像有损压缩中的编辑检测。3) 基于相机格式 (camera-based)^[67-71], 主要基于相机镜头、传感器、以及后处理芯片的图像检测。4) 基于物理规则 (physically-based)^[72,73], 主要对真实环境中的光线、映射等物理规则进行建模。5) 基于几何约束 (geometric-based)^[74,75], 通过图像成像映射规则, 构建视角几何约束, 以此来进行图像检测。我们的取证技术属于第三类和第五类混合, 即通过广角

镜头，构建几何约束条件，以此来进行图像内容的取证工作。如前文所述，广角镜头的畸变是一把双刃剑。一方面，普通镜头的几何约束不再适用于广角镜头；但另一方面，镜头级别又提供了一些新的约束，这些约束条件可以被用做自底向上的测度，来进行图像取证工作。在本文中，我们主要针对物体所包含的直线部分进行检测。因为真实环境中的物体成像，绝大多数包含直线部分，而空间直线正好提供了一个准确的畸变约束。此外，我们还提供了一种伪造取证分布图，以此来以此来提供像素级别的取证内容呈现。图1-5(c-d)展示了我们算法的伪造取证分布图，以及相应的分割结果。

1.2 本文的主要研究内容与贡献

在本论文中，我们将深入研究协同分割的核心思想，理解其构成形式，一方面，进一步挖掘多张图片中的协同关联约束，提高多张图片间的协同处理效果，另一方面，我们将这种协同分割思想其拓展到单张图像的处理应用中，挖掘单张图像不同区域间、相关源间的协同关联约束，提高相应单张图像的处理效果。主要研究内容与贡献包括：

- **多图像间的关联性约束：** 我们提出一种基于聚类方法的多图片间的协同显著性目标检测算法。我们的算法通过聚类构建多图像之间的物体全局关联结构，并设计自底向上的协同显著性测度，检测多张图片中的协同显著性目标。我们提出的协同显著性物体检测避免了大量繁重的学习训练过程，具有简单高效的特点。我们的主要贡献是：
 - 我们提出一种新型的基于聚类方法的多图片间的协同显著性目标检测算法，通过聚类构建多张图片间的协同性关联约束。
 - 基于聚类结构，提出了三种自底向上的协同显著性测度，包括特征对比度，中心偏移度，以及图间分布度，对图像进行聚类级别的协同显著性测量。
 - 我们的算法不仅在多张图片的协同显著性物体检测中取得优异结果，而且在单张图片的视觉显著性物体检测中，也取得了良好的测试结果。
 - 我们还将深入挖掘协同显著性物体检测算法的相关应用，提出四种典型应用：协同分割、鲁棒的图像测距、弱监督学习、以及视频前景提取等，进一步证明了协同显著性物体检测在图像处理中的应用潜力。
- **单图像不同区域间的关联性约束：** 自然图片中的物体往往满足一定的几何约束，基于这种几何结构将图片划分成多个的区域，这些区域之间可以构建关联性约束。我们将利用这种单张图像中的几何结构约束，引入协同分割思想，提出一种单幅图像中前景物体对称约束分割

方法。我们的主要贡献是：

- 提出一种稠密的图像几何结构映射矩阵，该矩阵在描述前景物体的位置分布的同时，也反映了物体自身的几何结构约束。
- 将这种几何结构矩阵引入到基于图的能量函数模型中，提出新的基于几何结构约束的图分割模型。我们的模型满足子模性，可以通过图分割算法优化进行求解。
- 我们还将该几何结构约束模型拓展到基于组件的分割框架中，以满足弱几何约束、以及复杂几何约束的物体分割需求。
- **单图像相关源间的关联性约束：**我们提出依据镜头径向几何畸变性质作为图像内部属性，来进行图像篡改盲检测。镜头的径向几何畸变在绝大多数图像处理中通常被认为是需要消除的不良性质，但我们提出该几何畸变反映了镜头的内部结构属性，而且对图像提供一个全局的几何约束。我们将镜头径向几何畸变性质作为度量指标引入到图像篡改盲检测取证算法中。我们还提出一种图像伪造度显著性分布图，并在原始图像和分布图之间建立协同联系，协同原始图像，分割单幅图像中的伪造区域。我们的主要贡献是：
 - 提出一种新型的基于广角镜头畸变约束的图像取证算法。算法通过图像内容与拍摄镜头性质构成关联约束，以此来检测图像中的伪造直线。
 - 设计了一种广角镜头的径向畸变模型，该模型简化了传统的广角镜头模型。在该模型下，空间直线在投影半球面将投影为一个切圆，从而提供镜头的内在几何约束。
 - 算法提供了两个自底向上的图像内容取证测度：体积取证测度和距离取证测度。通过取证测度，可以有效的检测图像中的伪造直线。
 - 我们还生成一个伪造区域显著性分布图。该分布图可以作为目标图像的一个处理层，通过与原始目标图像进行协同关联约束，分割出检测物体。

1.3 本文的结构

本文围绕协同关联性约束研究展开讨论，其具体组织结构如下：

第一章介绍本文的研究背景、研究内容和主要贡献等。

第二章研究多张图像间的关联性约束，以及基于此的多图片间的协同显著性目标检测算法。

第三章研究单张图像不同区域间的关联性约束，以及基于此的图像内几何结构关联区域之间的约束分割算法。

第四章研究单张图像相关源间的关联性约束，以及基于此的图像与镜头关联的基于畸变测度的图像伪造取证算法。

第五章对本文工作进行总结，并对协同关联性约束的未来工作进行展望。

第二章 多图像间的关联性约束

本章将主要研究多图像间的关联性约束。多图像间的关联性约束是一种常见的协同约束方式，当多张图片包含相同或相似物体时，物体在图像彼此之间就构成了关联性约束。但是大多数多图像协同约束要求图像的背景各不相同，否则容易构成前景的二义性。在本章，我们将基于图像生物视觉显著性原理与多图像间目标关联性约束相结合，提出一种基于聚类方法的多图片间的协同显著性目标检测分割算法（Cluster-based Co-saliency Detection Method）。协同显著性（Co-saliency）主要指多图像中的重复出现的同一或近似的视觉显著性物体。协同显著性物体检测是一个具有潜力但尚未深入开发的新研究领域。我们提出的基于聚类的协同显著性物体检测算法通过聚类构建多图像之间的物体全局关联结构，依据三种自底向上的协同显著性测度（saliency cue），包括特征对比度（Contrast cue）、中心偏移度（Spatial cue）、以及图间分布度（Corresponding cue），对图像进行聚类级别的协同显著性测量，最终将各测度结果进行融合，提取出多张图片中的协同显著性目标。我们的协同显著性物体检测算法避免了大量繁重的学习训练过程，具有简单高效的特点。该算法不仅在多张图片的协同显著性物体检测中取得优异结果，而且在单张图片的视觉显著性物体检测中，也取得了良好的测试结果。此外，在本章的最后，我们还深入挖掘协同显著性物体检测算法的相关应用，提出四种典型应用：协同分割（co-segmentation）、鲁棒的图像测距（robust image distance）、弱监督学习（weakly supervised learning）、以及视频前景提取（video foreground detection）等，进一步彰显了协同显著性检测在图像处理中的应用潜力。

2.1 相关背景工作

将多张图片关联作为一种附加约束用来检测协同显著性目标，已经被证明是一种行之有效的解决方案^[27-29]。文献^[27]提出了一种针对图像对的基于近似重复区域的协同显著性目标算法，其利用显著性滤波器测试图像中的所有像素的显著性反应，并利用相关熵（Kullback-Leibler divergence）来增强图像对中的显著性区域。文献^[28]则利用一个复杂的多层次图结构（co-multilayer graph），获取图像的全局关联信息，并且将单张图片显著性结果和多张图片显著性结果进行线性组合，获取最后的协同显著性目标。然而，这两种方

法^[27,28]仅能针对一组图像对进行操作，而无法拓展到多幅图像中。文献^[29]认为协同显著性目标是单张图像显著性目标和其在多张图像中出现频率的相乘。但是该方法仅把协同显著性目标检测作为协同分割的前期处理，并没有深入展示协同显著性目标的特性和优势。此外，文献^[29]的方法还存在算法复杂度高、运行效率低下的缺点。相对而言，我们提出的基于聚类的多图像协同显著性目标检测算法具有结构简单、运行速度快等优点。同时，在大量标准数据集上的试验结果显示，我们的算法相对于已有的算法更加快捷高效，获得了令人满意的检测结果。

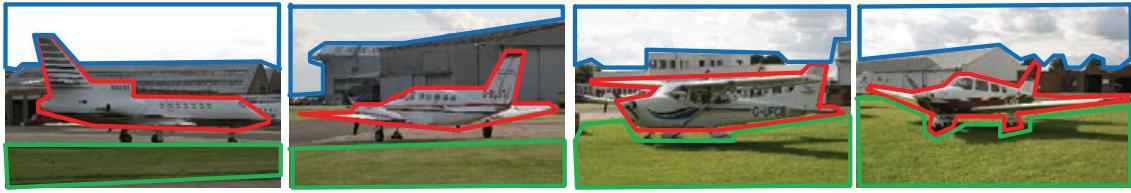


图 2-1 三幅图中，人类视觉机制很容易确定飞机（红色框）是四张图片中的近似物体。但是由于天空（蓝色框）、草地（绿色框）等也是重复出现，而且所占面积较大，这就使得自动协同分割算法把它们错认为前景而提取。

另一个与协同显著性目标检测相关的研究领域是‘协同分割’，其目的是分割出多组图片中的近似物体^[5,6,76]。相对于协同分割，我们的协同显著性检测主要基于低级别（low-level）的先验概念，更多的偏向于人类视觉关注机制。此外，协同分割还存在三点不足：首先，图像中的近似但非显著的背景会产生关联，从而影响无监督（unsupervised）情况下的协同分割^[9,17]。图2-1所示，对于人类视觉机制来说，很容易确定飞机是多张图片中的近似前景物体，但是由于天空、草地等背景也在两幅图像中重复出现，而且所占面积较大，这就使得自动协同分割算法常把它们错认为前景而提取。其次，一些协同分割算法^[6,15]需要借助人工输入来指导系统区分前背景，从而无法做到自动化。第三，协同分割算法通常计算量大，尤其是对于大图像集合，很难做到快速分割。而在实际需求中，很多应用，例如图像尺度调节（image retargeting）^[77,78]、目标定位与识别（object location and recognition）^[79]等，仅仅需要快速粗略的定位多图像中的重复物体。相对于协同分割来说，我们的协同显著性目标检测算法可以自动的检测重复出现的显著物体。同时由于算法的简单高效，我们的算法可以作为一些高级别图像理解工作的预处理。更重要的是，我们在大量的协同分割数据库上测试我们的协同显著性目标检测算法发现，我们的算法尽管没有经过复杂的学习训练过程，但其仍然取得了相当良好的测试结果，毫不逊色于大多数已有的协同分割算法。

2.2 协同显著性目标检测方法

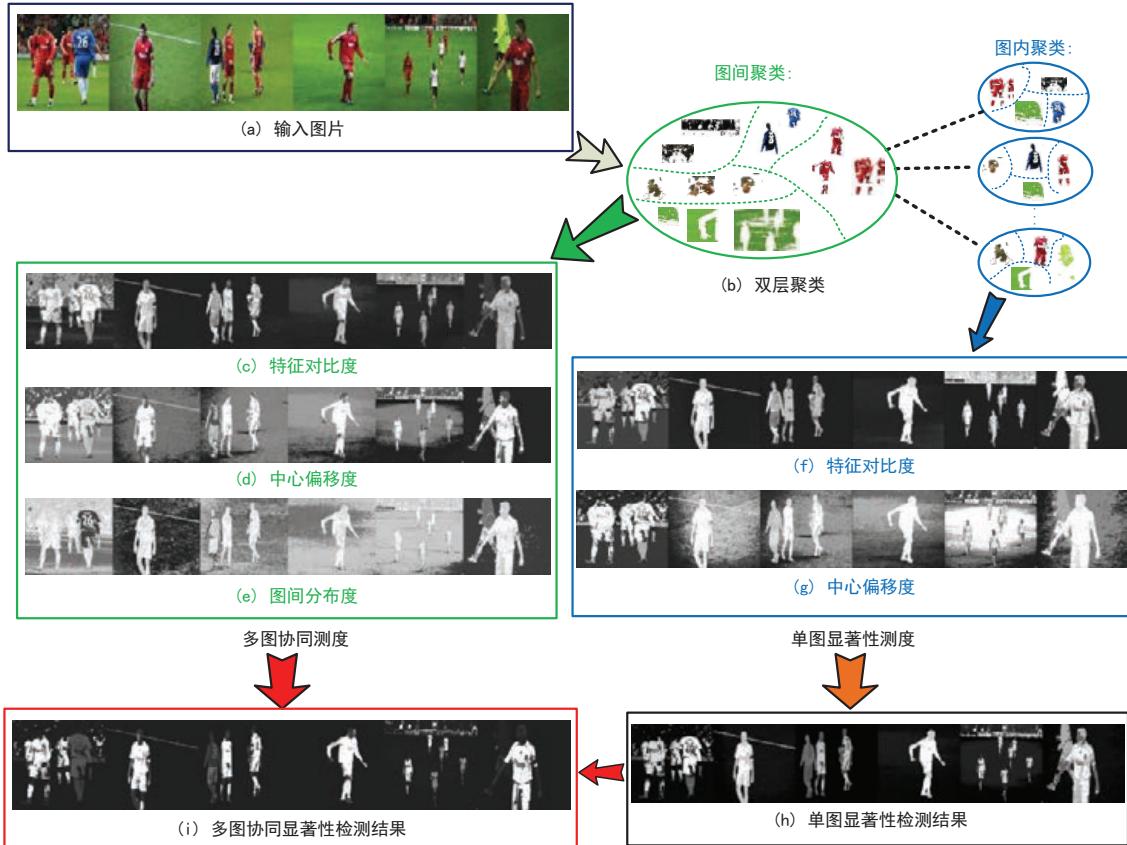


图 2-2 基于聚类的协同显著性目标检测算法流程图。(a) 输入图像集。(b) 对输入图像进行图像内和图像间聚类。特征对比度(c, f)、中心偏移度(d, g)、以及图间分布度(e)是我们设计的三种自底向上的协同显著性测度。(h) 单张图像的视觉显著性检测结果。(i) 多张图片的协同显著性检测结果。

协同显著性目标检测算法虽然有很大应用潜力，但并没有受到很大的关注，而且已有的算法也没有取得很好的检测效果^[27,29]。在此，我们首先定义协同显著性目标算法应该具备的基本特性：

- (图内显著性) 协同显著性目标应该满足单幅图片中的视觉显著性准则，其相对于图像背景应该具有可分辨性。
- (图间显著性) 协同显著性目标在多张图像间，应该具有高近似性，因此在一定程度上，反映出全局的重复性、出现频率高等特征。
- 协同显著性目标检测作为其他应用的前期预处理，其还应该具有移植性好、计算速度快等特点。

根据以上需求，我们提出了一种基于聚类的多张图片协同显著性目标检测算法。图2-2展示了我们的算法流程图：在输入一组图片后，我们的算法首先进

行一次双层聚类 (two-layer clustering)。第一层聚类发生在每幅图像内部 (图内聚类)，第二层聚类发生在全部图像中 (图间聚类)；然后，我们计算每一类的显著性测度 (saliency cue)，显著性测度将测量每一类的独特性 (图内/图间)、中心距离 (图内/图间)、以及分布度 (图间)；最后，基于这些类级别的显著性测度，我们计算每一个像素点的协同显著性值，生成最终的协同显著性目标分布图。

2.2.1 基于聚类结构的全局关联

基于聚类的全局关联方法是受单幅图像中的全局对比度 (global-contrast)^[18,80-82] 方法的启发。全局对比度方法将像素的特征划分为直方图 (histogram) 形式，以此来测量特征对比的差异。显著性物体的特征与图像中其他特征会具有明显的区别。但是这种基于直方图的方法在区间 (bin) 边缘是不连续的，容易导致硬划分。相对来说，我们使用聚类来划分区间，避免了实际区间的断层，获取了更内聚 (cohesive) 的全局约束。此外，在多图像间的聚类还提供了全局关联性。在我们的方法中，我们对聚类所使用的方法不进行特殊选择，在这里使用Kmean聚类。

在聚类过程中有两个需要注意的问题：1) 如何确定聚类数目；2) 聚类错误的像素点 (例如背景点与前景点聚类到一起) 是否会影响到显著性检测的结果。过少的聚类数目会导致同类别的像素具有相同的显著性测度值，从而缺乏足够的显著性层次。为了避免‘离散’的聚类，我们使用一个概率框架去柔性为每一个像素分配显著值。另外，基于这个概率框架，聚类数目并不影响最终的显著性检测结果。

2.2.2 聚类级的显著性测度

在本小节，我们主要介绍三个基于聚类的显著性测度用来测量聚类级别的显著性。头两种测度被称为特征对比度和中心偏移度，它们最早是被单张图像的显著性检测算法所引入。我们将这两种显著性测度拓展到我们的基于聚类的框架下，并且用它们去检测单张以及多张图像的显著性测度。我们同样引入一种新的测度—图间分布度，去描述多张图片间的近似物体出现频率。我们基于聚类的方法不同于基于像素的检测方法，显著性测度主要应用于聚类级别 (cluster-level)。

符号注释： $\{p_i^j\}_{i=1}^{N_j}$ 表示图像 I^j 的第 i 个像素，其中 N_j 表示图像的维度 (lattice)。 $\{z_i^j\}_{i=1}^{N_j}$ 表示图像 I^j 中像素 p_i^j 的归一化位置 (normalized location)。给 M 张图片 $\{I^j\}_{j=1}^M$ ，我们获取 K 个类¹ $\{C^k\}_{k=1}^K$ 。我们用一个D维向量 $\{\mu^k\}_{k=1}^K$ 代表类，其

¹在实际中，我们将引用两个聚类数目， K_1 和 K_2 分别表示图内聚类数目和图间聚类数目。但是在基于聚类级别的描述中，我们暂不区分它们，统一使用 K 来代表聚类数目。

中 μ^k 代表聚类 C^k 的类中心 (cluster center)。函数方程 $b: \mathbb{R}^2 \rightarrow \{1 \dots K\}$ 代表像素 p_i^j 与聚类 $b(p_i^j)$ 的映射关系。

2.2.2.1 特征对比度测度

特征对比度测度 (Contrast cue) 主要测量单张/多张图片中的视觉特征独特性 (uniqueness)。在单张图片显著性物体检测中，对比度是一种被广泛使用的测度^[18,19,24]，一个主要原因是对比度测度，尤其是颜色、纹理对比度，非常符合人类的视觉接收域感知。同样的，在多张图片中，特征对比度依旧有效。其和单张对比度面向像素操作不同，我们的特征对比度主要作用于聚类级别，从而提供一个全局的关联性对比约束。类 C^k 的特征对比度被定义为其特征向量与其它类的对比：

$$w^c(k) = \sum_{i=1, i \neq k}^K \left(\frac{n^i}{N} \|\mu^k - \mu^i\|_2 \right), \quad (2-1)$$

其中 L_2 模被用来计算特征向量距离， n^i 表示类 C^i 所包含的像素数目， N 代表所有图片的像素数目和，这个定义使得更大的类起到更大的作用。此公式(2-1)近似于文献^[18] 中的基于直方图对比度 (Histogram-based Contrast)。但这里有三个主要的区别：第一，文献^[18] 计算显著性测度通过一个简化的直方图，而我们使用了聚类，因此我们的方法具有更好的内聚性。第二，在我们的方法中，对比度测度仅仅作为三个协同显著性测度的其中一个。图2-7 展示了在单张图片显著性检测中，我们的方法与文献^[18] 的检测效果区别。

特征对比度测度可以作用于单张图片和多张图片中，如图2-2(c) 和(f) 所示。特征对比度的优势在于特征相对稀少的类将更加显著，例如图2-2 中的运动员。但是，特征对比度很容易受到复杂背景的干扰 (例如图2-2 中的观众)。此外，它也没有反应多张图像中的重复特征。

2.2.2.2 中心偏移度测度

在人类视觉注意力系统中，靠近图像中心的物体总能收到更多的注意力关注^[82-84]，而且这也符合摄影学原理。当物体远离图像中心时，物体的注意力也逐渐下降，该现象被称为 ‘中心偏移准则’ (central bias rule)。我们将这种单张图片中的视觉规则拓展到多张图片中，用来测量目标类在多张图片中的全局位置分布 (spatial distribution)，我们称之为中心偏移度测度 (Spatial cue)。类 C^k 的中心偏移度测度 $w^s(k)$ 被定义为：

$$w^s(k) = \frac{1}{n^k} \sum_{j=1}^M \sum_{i=1}^{N_j} \left[\mathcal{N}(\|z_i^j - o^j\|^2 | 0, \sigma^2) \cdot \delta[b(p_i^j) - C^k] \right], \quad (2-2)$$

其中 $\delta(\cdot)$ 表示 Kronecker delta 方程， o^j 表示图像 I^j 的中心位置，高斯核 (Gaussian kernel) $\mathcal{N}(\cdot)$ 用来计算像素点 z_i^j 与图像中心 o^j 的欧氏距离 (Euclidean distance)，方差 (variance) σ^2 设定为图像的归一化半径。归一化参数 (normalization coefficient) n^k 代表聚类数目 C^k 。相对于单张图片中的中心偏移准则，我们的

中心偏移度测度描述了聚类级别的先验位置信息，其提供了一个多张图片的全局中心偏移约束。

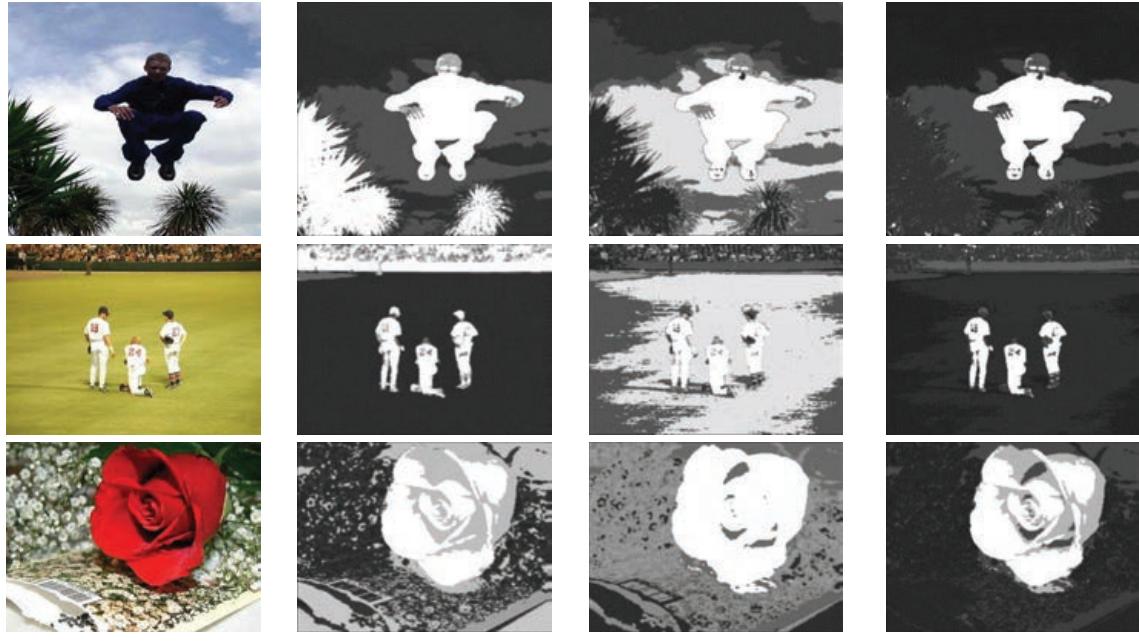


图 2-3 我们的单张图片显著性目标检测效果。(a) 输入图片。(b) 特征对比度检测，其可以区分大多数显著性目标。(c) 中心偏移度检测，其可以区分在图像四周的高纹理的背景。(d) 我们的最终单张图片显著性目标检测结果，其融合了特征对比度和中心偏移度，并取得了理想的显著性目标分布图。

与特征对比度一样，中心偏移度测度也可以同时作用于单张和多张图片中。如图2-2 (d) 和图2-2 (g) 所示，位于中心的红色运动员相对于蓝色运动员具有更高的测度值。图2-3 进一步展示了特征对比度和中心偏移度的区别，其中特征对比度可以检测出大部分显著性目标，相对而言中心偏移度更适合去除高纹理复杂背景，尤其是远离图像中心的干扰性背景。一方面，中心偏移度可以弥补特征对比度的缺陷，抑制复杂背景的干扰（例如图2-3 第一行的大树，以及第二行的观众）。但另一方面，位于图像中心的背景（例如图2-3 第二列的操场）有可能造成中心偏移度的测量误差。在结合了两个测度后，我们的单幅图像显著性目标检测取得了令人满意的结果，如图2-3 (d)。

2.2.2.3 图间分布度测度

不同于特征对比度和中心偏移度测度，我们的第三个测度，图片分布测度 (Corresponding cue)，主要用于测量一个类别在多幅图像间 (Bin) 的分布情况。重复率 (repetitiveness) 描述了一个物体在多张图片中反复出现的频率，是一种非常重要的重复物体的全局属性。而在我们的基于聚类的框架中，图间聚类本身就反映了多张图片中的全局关联性。图2-4 展示了一个聚类分布

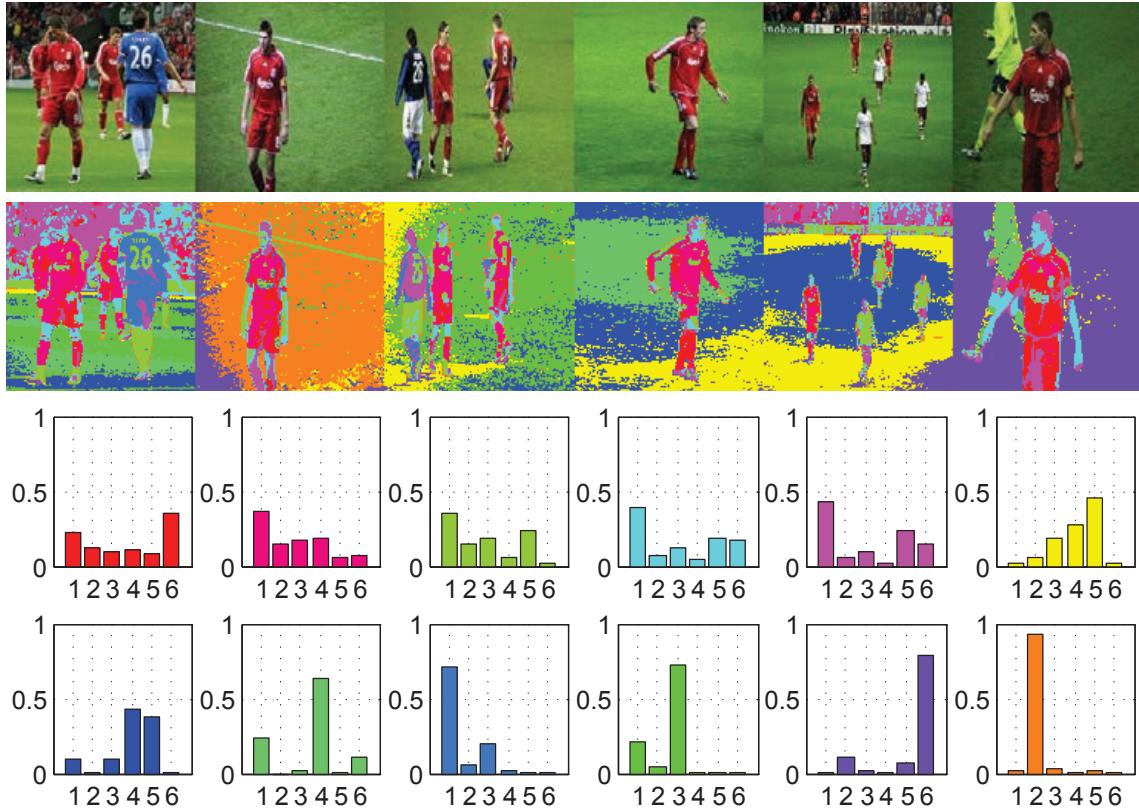


图 2-4 图间分布度测度展示。顶行：输入图片。中间行：图间聚类结果（类数目为 $K = 12$ ）。底行：每一个聚类对应的 M 区间划分的直方图。其中这12个类按照它们的方差排序，颜色与中行的聚类结果相对应。 M 代表图像数目，本例中为6。

的例子，其中重复物体（红色运动员）在每一张图片中都反复出现。基于此，我们可以使用类的方差来粗略描述某一聚类在多张图片中出现分布情况。

首先，通过一个 M 区间划分的直方图 $\hat{\mathbf{q}}^k = \{\hat{q}_j^k\}_{j=1}^M$ 来描述聚类 C^k 在 M 张图像中的分布度：

$$\hat{q}_j^k = \frac{1}{n^k} \sum_{i=1}^{N_j} \delta[b(p_i^j) - C^k], \quad j = 1 \dots M. \quad (2-3)$$

其中， n^k 是聚类 C^k 所包含的像素数目，并且附加归一化条件 $\sum_{j=1}^M \hat{q}_j^k = 1$ 。然后，我们的区间分布度测度 $w^d(k)$ 定义为：

$$w^d(k) = \frac{1}{\text{var}(\hat{\mathbf{q}}^k) + 1}, \quad (2-4)$$

其中 $\text{var}(\hat{\mathbf{q}}^k)$ 代表了聚类 C^k 的直方图 $\hat{\mathbf{q}}^k$ 的方差。某一类的图间分布度越高，则代表其包含的像素越均匀的分布在多张图片之中。

图2-2 (e) 展示了图间分布度测度的检测结果，其中红色运动员频繁的出现在多张图片中，相对于蓝色运动员其具有更高的分布度。然后，近似的背景同样会拥有较高的区间分布度，不过由于其他两个测度的综合作用，这

些误判的背景会在最终的协同显著性目标检测结果分布图中被消除。

2.2.3 显著性测度的融合



图 2-5 协同显著性分布图融合比较：相加（上图）与相乘（下图）。可见，相乘可以得到较好的视觉效果，其抑制了噪声点。

至此，我们可以得到三种自底向上协同显著性测度量²。每一种显著性测度独立使用时，都有其优缺点。常见的测度值融合方法是线性相加^[19,24]以及相乘^[85]。图2-5展示了相加和相乘融合的区别。相比来说，相乘融合更适合抑制噪声点，而相加可以取得更好的查全率（Recall）。对于显著性目标检测来说，查准率（Precision）要比查全率更加重要^[81]。在本算法中，我们同样偏重查准率，因此我们使用相乘去融合显著性测度值。

在融合测度值之前，我们先使用标准的高斯（standard Gaussian）归一化（normalize）所有测度分布图。然后，类 k 的聚类级别协同显著性测度概率 $p(k)$ 定义为：

$$p(C^k) = \prod_i w_i(k), \quad (2-5)$$

其中 $w_i(k)$ 代表显著性测度。

此时计算出的聚类级别协同显著性值是离散分配的，即在同一类中的像素点具有相同的协同显著性测度值，从而缺乏区分性。因此，我们需要进一步平滑处理每一个像素点。类 C^k 中的像素 x 的显著性似然值（saliency likelihood）满足高斯分布 \mathcal{N} （Gaussian distribution）：

$$p(x | C^k) = \mathcal{N}(\|v_x, \mu^k\|_2 | 0, \sigma_k^2), \quad (2-6)$$

²实际上，整体算法将输出七个测度值：三个多图像的测度值，以及两个单图像输出的测度值。

其中 v_x 像素 x 的特征向量, 高斯方差 σ_k 直接使用类 C^k 的方差。因此, 最后显著性的边缘分布 $p(x)$ (marginal saliency probability) 可以通过所有聚类的联合分布 $p(C^k)p(x|C^k)$ 计算得到:

$$p(x) = \sum_{k=1}^K p(x, C^k) = \sum_{k=1}^K p(x | C^k)p(C^k). \quad (2-7)$$

最终得到了像素级别的协同显著性分布图, 例如图2-2 (i) 所示。

2.3 实验与分析

实验主要首先从两个方面测试我们的算法: 单张图片显著性目标检测、以及多张图片的协同显著性目标检测。实验将在三个公开数据库中测试比较我们的算法与其他已有算法的性能。其次, 我们将进一步分析算法中各个显著性测度的作用、算法运行效率、以及聚类的参数调节等相关问题。在实验中, 我们使用CIE lab 颜色特征和Gaber 滤波器^[86] 作为像素的特征描述。Gaber 滤波器包含8个方向 (orientations) 响应, 带宽 (bandwidth) 设置为1, 使用1个尺度 (scale)。我们将融合Gaber 滤波器的8个方向响应, 计算最终的强度值分布图 (magnitude map) 作为像素得纹理特征 (texture) 向量。在我们的算法中, 使用Kmean 来进行聚类, 图内 (单张图片) 聚类数目设置为 $K_1 = 6$, 图间 (多张图片) 聚类数目设置为 $K_2 = \min\{3M, 20\}$, 其中 M 代表图像数目。

2.3.1 单图片的显著性目标检测实验

首先我们进行单张图片的显著性目标检测实验。我们使用公开的MSRA1000 显著性检测数据库^[81], 该数据库是最大的显著性目标监测数据库之一, 包含1000张图片, 并提供像素级别的目标人工标注 (pixel-level ground truth)。我们比较我们的单张图片显著性目标检测算法 (SS) 与其他五个已有的检测算法: Spatiotemporal Cues^[80] (SC)、Frequency-tuned saliency^[81] (FT)、Spectral residual^[23] (SR)、以及文献^[18] 中的Region-based Contrast (RC) 和Histogram-based Contrast (HC)。图2.6a) 展示了在MSRA1000 显著性图像数据库中的查准率/查全率曲线 (Precision/Recall) 实验结果, 其中F 测度为:

$$\text{F-measure} = \frac{(1 + \beta^2)\text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}}, \quad (2-8)$$

其中, 我们按照文献^[18,81] 中设置 $\beta^2 = 0.3$, 使查准率略重要于查全率。此外, 我们还测试了算法中的特征对比度 (CoC) 和中心偏移度 (SpC), 如图2.6a) 中等个虚线。我们的特征对比度测度形式上与HC^[18] 近似, 区别在于我们使用聚类替代了直方图。因此, 我们的特征对比度结果 ($F = 0.755$) 也近似于HC^[18] 的结果 ($F = 0.751$)。实验中我还发现中心偏移度的结果要优于特征

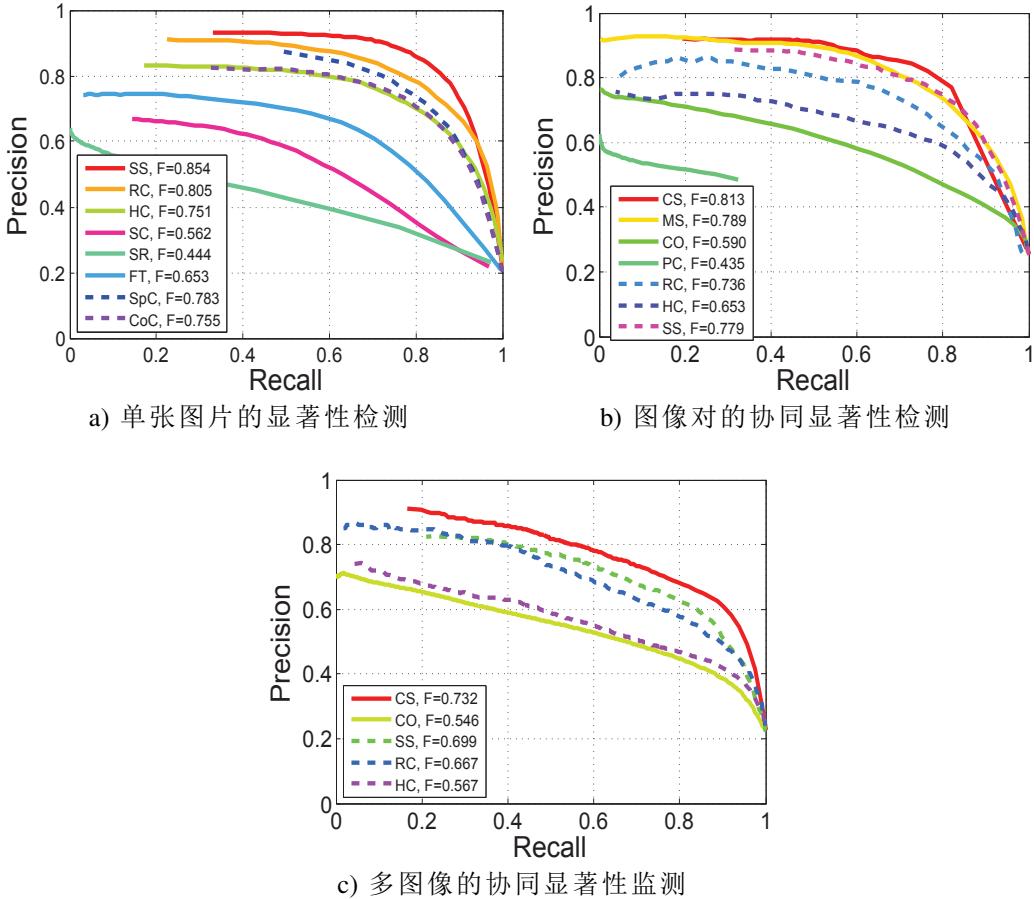


图 2-6 在三个公开测试数据库中算法运行的查准率/查全率 (Precision/Recall) 曲线比较实验结果。SS 表示我们算法的单张显著性目标检测 (single image saliency) 结果, CS 代表我们的多张图片协同显著性目标检测 (co-saliency) 结果。(a) 在 MSRA1000 显著性图像数据库中的查准率/查全率曲线, 我们的算法与已有的单张显著性目标检测算法比较, 包括: RC^[18], HC^[18], SC^[80], SR^[23], FT^[81]。同时还展示我们算法的特征对比度测度 (CoC) 与中心偏移度测度 (SpC) 曲线。(b) 在协同显著性图像对数据库中的查准率/查全率曲线, 我们的算法与另外三种相关算法比较: CO^[29], MS^[28] and PC^[27]。(c) 在 iCoseg 协同分割数据库中的查准率/查全率曲线。

对比度, 其中一个原因就是绝大多数是自然拍摄的图片都满足中心偏移准则。这一现象也被最近的其他算法^[82,83,87]所证实。虽然单独的中心偏移度测度结果并没有优于RC^[18], 但是我们的整体单图显著性目标检测算法在融合了特征对比度和中心偏移度后, 取得了优于RC^[18]的测试结果, 我们的F-measure为0.854, 高出RC^[18] ($F = 0.805$) 将近5%。

图2-7展示了不同的显著性目标检测算法的检测结果。对于前两张图片, 特征对比度测度要优于中心偏移度测度。然而, 中心偏移度依然通过一个较高的测度值检测到显著性目标, 将其与背景分离开。当图片具有复杂纹理的

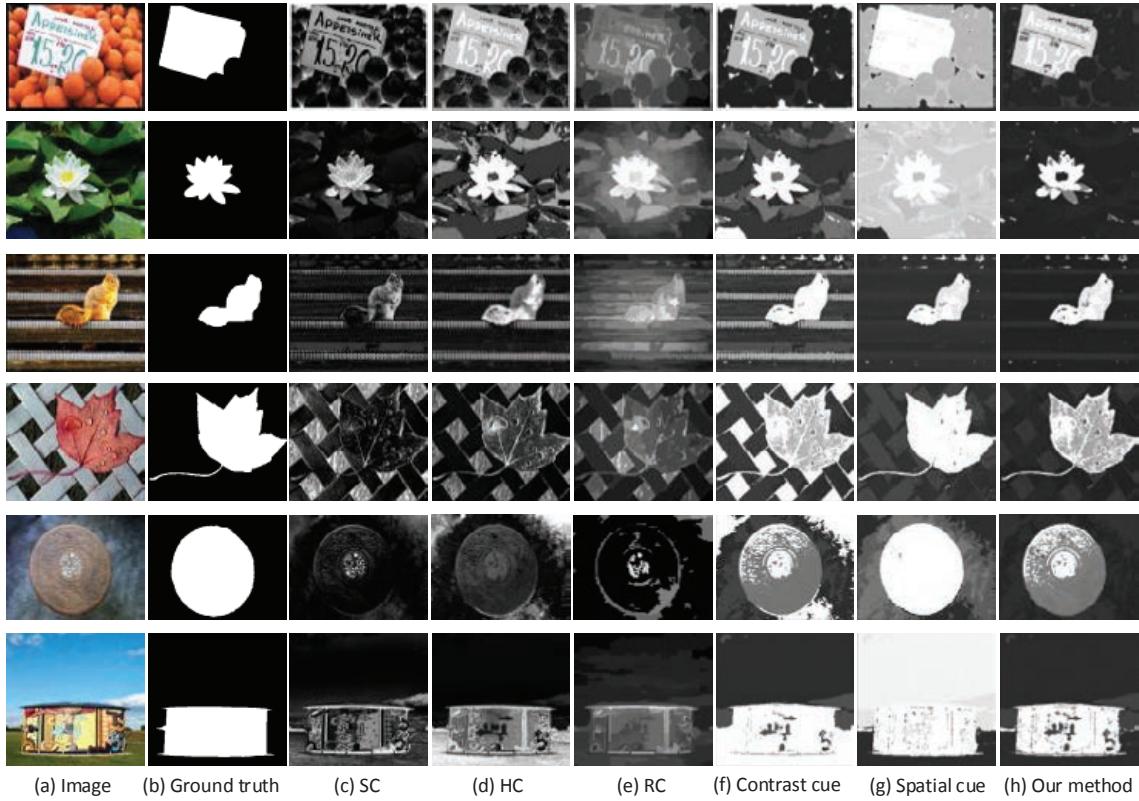


图 2-7 在MSRA1000 显著性图像数据库中的检测结果。(a) 输入图像。(b) 标定图像。显著性检测结果: (c) SC^[80], (d) HC^[18], (e) RC^[18], (f) 我们的特征对比度测度, (g) 我们的中心偏移度测度, (h) 我们的最终单张图像显著性目标检测结果。

背景或者低对比度时,例如图2-7 中的第三到第五张图片,大多数已有算法无法取得理想的效果。因为它们仅采用低级别的特征对比,很容易被背景中的高对比度噪声(第三、四张图片)以及低对比度的前景(第五张图片)所干扰。相对而言,中心偏移准则,其基于空间位置(spatial layout)而不是特征对比,更适合于这些情况。我们的算法在融合了特征对比度和中心偏移度后,更具有鲁棒性,取得了最佳的显著性检测结果。在图2-7 最后一张图例中,由于草地的高对比度干扰,特征对比度测度无法准确的定位显著性目标。而中心偏移度测度由于天空的干扰也产生了偏差,即图中的显著性物体低于天空的测度值。然后我们的算法融合了两种测度的优势,准确的检测出正确的视觉显著性目标。

2.3.2 协同显著性目标检测实验

大多数协同显著性目标检测算法都针对一组图片进行检测,实验将比较我们的协同显著性目标检测算法(CS)与其他三种已存的协同显著性算法: co-saliency^[29] (CO)、preattentive co-saliency^[27] (PC)、以及multi-image

saliency^[28] (MS)。测试数据库采用协同显著性图像对数据库^[28] (Co-saliency Pairs dataset)，其包含210张图片（105对）。每对图片包含一个或多个近似目标以及不同的背景。图2.6b) 展示了该数据库上的运行结果。作为对比，我们同样提供了单张图片显著性检测算法的结果：文献^[18]的RC、HC、以及我们的单张检测算法（SS），如图中的虚线所示，与单张检测结果近似，我们的算法仍然在在单张图片检测中取得了最好的结果。此外，我们的单张检测算法（SS）结果高于协同显著检测算法PC^[27] 和CO^[29]，接近于MS^[28] 的结果。主要的一个原因是在协同显著性图像对数据库中包含的前景大多与背景明显区别，这就降低了第二张图片的贡献，即部分降低了协同信息的作用。但我们依然可以看出，在包含了协同信息后，我们算法的检测结果明显高于其他算法SS ($F = 0.779$) 和MS ($F = 0.789$)，接近 $F = 0.813$ 。

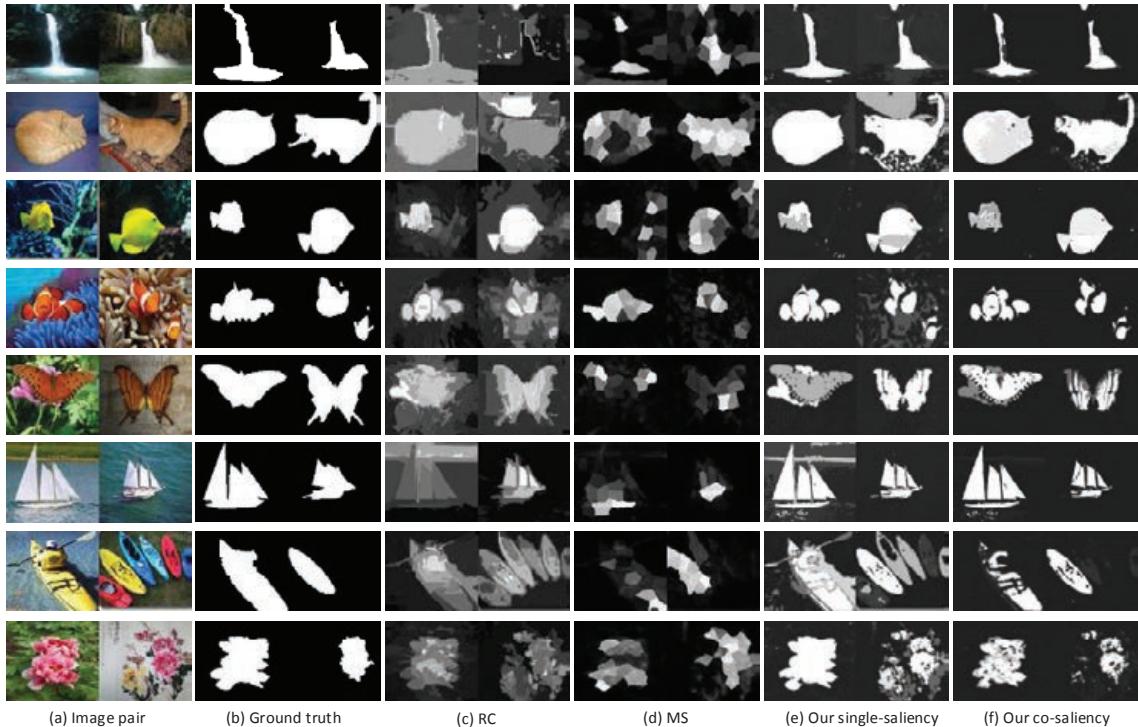


图 2-8 在协同显著性图像对数据库中的检测结果。(a) 输入图像对。(b) 标定图像。显著性检测结果：(c) SRC^[18]，(d) MS^[28]，(e) 我们的单张显著性检测，以及 (f) 我们的协同显著性检测。

图2-8展示了在协同显著性图像对数据库中的检测结果。总体来说，我们的方法取得了准确的检测结果。在方法RC^[18] 中，非显著性的高纹理背景没有被消除，例如图2-8 中的前两行图像。对于方法MS^[28]，大物体很难被检测到，例如图2-8 中的第二行以及第五行图例。一个可能原因是，方法MS^[28] 仅仅图像块（patch）的局部对比度信息，而缺少全局信息。因此，大物体的中心图像块在于其周围的对比度相比，并没有任何显著性特征，因此无法被检

测到。相对而言，我们的算法通过将大物体聚类解决了此类问题，获取的更好的检测结果。此外，对于复杂背景，例如图2-8 中的第三行和第六行图例，同样会影响到方法RC^[18] 和MS^[28] 的效果。由于中心偏移度的作用，我们的算法针对复杂背景具有良好的鲁棒性，获得了最好的结果。图2-8 中最后两个图例展示了协同显著性目标检测与单张显著性检测的区别。单张图像的显著性检测算法侧重于挖掘单张图像内的所有显著性目标，而协同显著性检测算法则可以以同一个全局的关联性，提取出多张图片中重复出现的目标，例如图例中的黄色的皮艇与红色牡丹花。

最后，我们在CMU Cornell iCoseg 协同分割数据库中测试我们的协同显著检测算法在多张图片中的效果 (image number $\gg 2$)。iCoseg 协同分割数据库是最大的公开发行数据库之一，其包含643 张图片，分为38 个类别。因为协同显著性检测算法MS^[28] 和PC^[27] 无法作用于多张图像，因此我们仅比较方法CO^[29]。如上一组实验类似，我们依旧提供单张图片检测的结果作为参考。图2.6c) 展示了这些方法的查准率/查全率曲线。我们的方法取得了最佳的检测结果 ($F = 0.732$)。由于iCoseg 数据库主要针对协同分割应用而不是显著性检测算法评估，图像中的部分物体无法提供自底向上的显著性特征属性。因此所有方法的检测结果都略低于其在协同显著性图像对数据库中的检测结果。图2-9 展示了在iCoseg 协同分割图像数据库中的检测结果图。

2.3.3 运行速度测试

我们的协同显著性目标检测算法主要基于自底向上的显著性测度，而不需要大量繁重的学习训练过程。同时，基于聚类的算法结构相对于像素级别的操作，更加节省空间与运行时间。实验平台为Dual Core 2.8 GHz 处理器，4GB 内存，算法代码编译环境为matlab。我们以iCoseg 数据库为基础，在同一类别图像中随机挑选一定数目的图像作为输入，所有图像统一尺度放缩到 128×128 。对于每一个图像数目，我们运行20 次程序，并统计平均时间，以及其方差。图2-10 展示了在iCoseg 数据库上的处理不同数目图像的运行时间统计，其中红色曲线维算法整体运行时间 (Entire method)，蓝色曲线为聚类算法运行时间 (K-means)。显然，运行时间随着图像数目的增加而增长。当图像多于五张时，聚类时间 (K-means) 将占用算法的大多数时间。通常来说，我们的算法处理一对 128×128 的图像耗时3秒左右，而处理四张图像，大概耗时4 秒。相对其他算法，协同检测算法^[29] 需要大概50 秒处理四幅图像，而算法^[28] 则需要450 秒处理一对图像。我们的算法极大地提高了运行速度，而且取得了更好的检测结果。

与我们的算法相近似的是协同分割算法，其目的是从多图像中提取出像素级别的准确目标分割。但是，大多数协同分割算法都需要大量的运算时间。例如算法^[9] 需要30 秒计算两幅图像，而算法^[76] 则更需要4 到9 个小时去

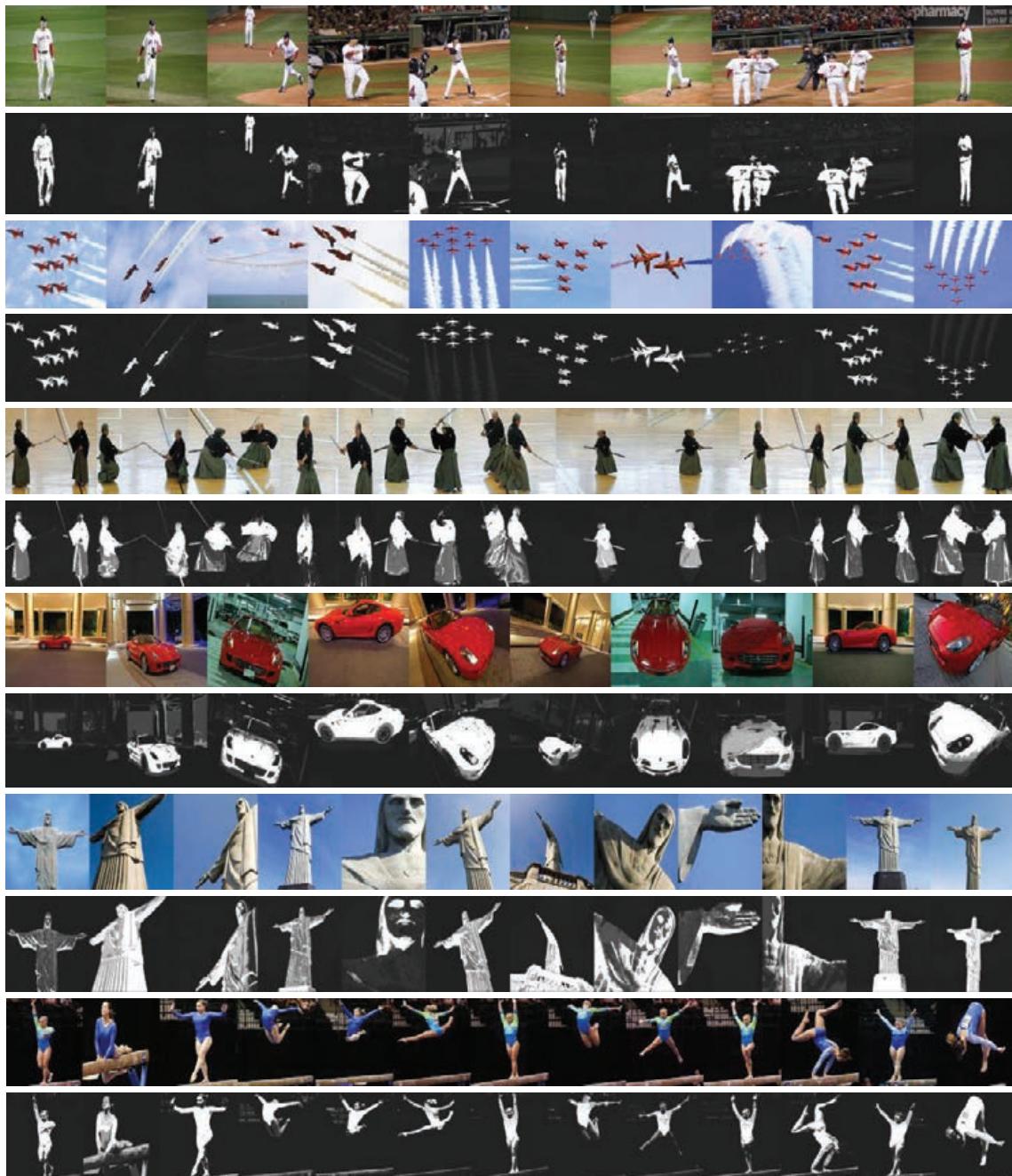


图 2-9 在 iCoseg 协同分割图像数据库中的检测结果，我们的协同显著性目标检测算法提供了一个准确的多张图片间的重复目标蒙版（mask）。

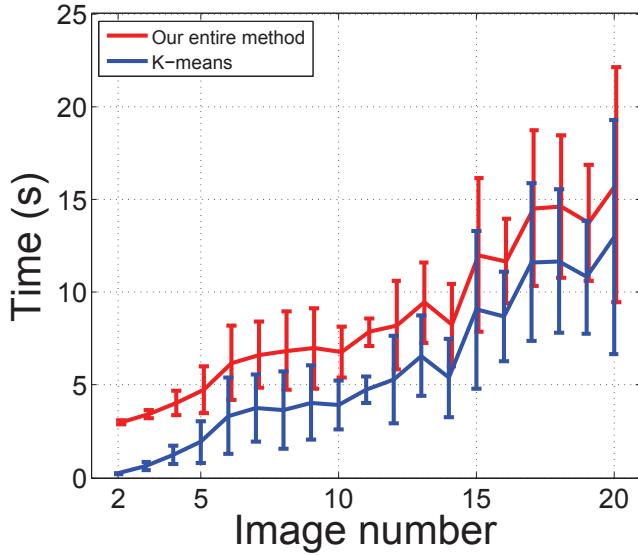


图 2-10 在iCoseg 数据库上的处理不同数目图像的运行时间统计，图像尺寸为 128×128 ，图间聚类数目为 $K = \min\{3M, 20\}$ ，其中 M 代表图像数目。图中红色曲线维算法整体运行时间 (Rntire method)，蓝色曲线为聚类算法运行时间 (K-means)

处理30幅图像。相对来说，我们的协同显著性检测算法处理300张图片仅需要22分钟，并且提供相当可观的准确度。更多的我们算法与协同分割的细节比较将在第2.4.1节讨论。

2.3.4 相关讨论

本节将讨论我们算法所涉及的三种相关因素：各显著性测度作用、据类数目分析、以及一些退化情况的讨论。

2.3.4.1 各显著性测度的作用分析

我们的使用了三种自底向上的显著性测度去测量协同显著性目标。为了评估每一种显著性测度所起的作用，我们将其分别在iCoseg 数据库上进行测试：协同显著性检测 (CS)、单张图片协同性检测 (SS)、单张图片中的特征对比度 (SS-CoC)、中心偏移度 (SS-SpC)、以及多张图片中的特征对比度 (CS-CoC)、中心偏移度 (CS-SpC)、和图间分布度 (CS-Corr)。图2-11展示了实验结果，从中我们可以归纳出四个结论：第一点，相同显著性测度在多张图像中的检测结果要优于其在单张图像中相应的检测结果。因为多张图片所提供的全局关联可以有效地对显著性检测结果起到约束作用，使其更加准确的在众多候选目标中筛选出最具视觉显著性的目标。第二点，特征对比度测度在iCoseg 数据库中，其单张图像的检测结果与多张图片的检测结果近似。这是由于在iCoseg 数据库中同一类的物体图片都呈现相似的背景，这就

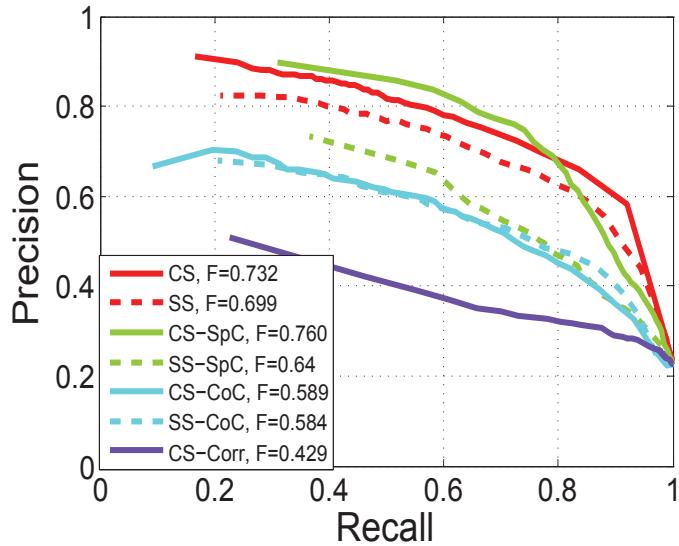


图 2-11 不同显著性测度在 iCoseg 数据库上的查准率/查全率分析。

导致特征对比度在多张图片中所提供的全局约束相对较弱。第三点，中心偏移度测度 (CS-SpC) 在数据库中取得了最好的结果，其效果甚至高于完整的协同显著性检测算法 (CS)。这是由于在 iCoseg 数据库中大部分前景物体全都集中在图像的中心区域，这使得中心偏移度起到了最大的作用。但是，这种位置先验情况并不是总是存在，尤其是在实际图像中，一个明显的例子就是图2-7最后一行。第四点，图间分布度测度 (CS-Corr) 本身并没有取得较好的结果。这是意料之中的，因为图间分布度的侧重点在于提取目标在多张图片间的关联度，而不是分离出显著性目标。简单来说，图间分布度的作用是用来在候选区域中辅助区分最重复出现的显著性目标。因此，弥漫在所有图像中的背景噪声，也有可能被图间分布度错误提取。

2.3.4.2 聚类数目分析

由于公式(2-7)软分配各像素显著性值的作用，聚类数目不再是我们的主要因素。在这里，我们仅仅测试不同的图间聚类数目在 iCoseg 数据库中的运行结果。图2-12展示了不同聚类数目在 iCoseg 数据库上的测试结果。我们发现在聚类数目较小的时候，检测结果随着数目增多而增加。但当聚类数目超过15时，效果提升将变得趋于平滑，但大量的聚类数目会导致运行时间的上升。因此，在实际中，我们选择一个较为宽松的聚类数目上限 $K = \min\{3M, 20\}$ ，其中 M 代表图像的数目。

2.3.4.3 退化情况分析

和其他大多数算法一样，我们的多图像协同显著性目标检测算法也存在一定的退化现象。第一种退化现象是目标物体由多部分组成，如图2-13 (a)

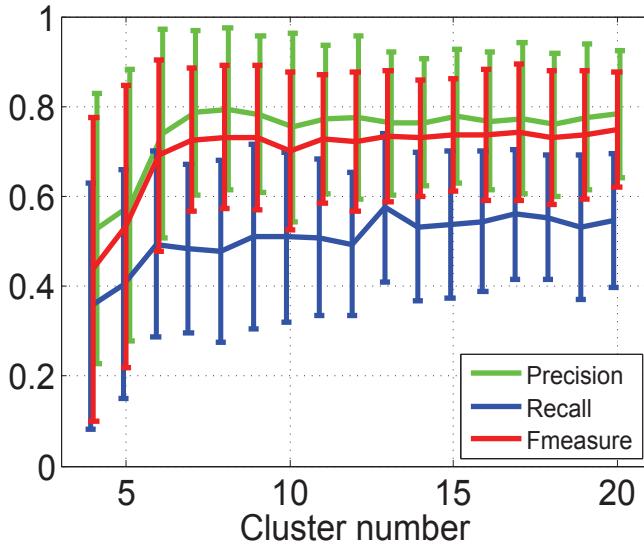


图 2-12 不同据类数目在iCoseg 数据库上的查准率/查全率/F因子分析。

和 (c) 所示。我们的算法容易陷入局部最优，即仅能检测显著性目标的最显著部分，而不是全部。造成这种情况的主要原因是由于我们的算法仅仅采用自底向上的特征测度，而没有大规模学习训练过程，因此无法提供一个物体级别 (object-level) 的特征表述。此外，我们的算法也容易受聚类影响，即非显著性的背景与显著性前景存在近似的特征。图2-13 (b) 展示了这种情况，其中小孩的衣服与背景中房子的墙颜色近似。这种情况也容易发生在动物的保护色情况，例如图2-13 (d-e) 所示。

2.4 协同显著性检测的应用

近几年中，单张图像的显著性目标检测算法受到了极大的重视，大量的图像处理应用被相继提出。然而，协同显著性目标检测算法却未受广泛重视，缺乏应用支撑。在本节，我们将讨论四种潜在应用：协同分割 (co-segmentation)、鲁棒的图像测距 (robust image distance)、弱监督学习 (weakly supervised learning)、以及视频前景提取 (video foreground detection) 等，充分挖掘协同显著性目标检测的应用潜力。

2.4.1 协同分割应用

第一个直接关联的应用就是协同分割。大多数算法都将协同分割问题转化为优化求解一个分割能量函数^[5-7,9]。分割能量函数主要包括图像内的能量项 (energy term) 以及全局的约束项 (constraint term)。但复杂的分割能量函

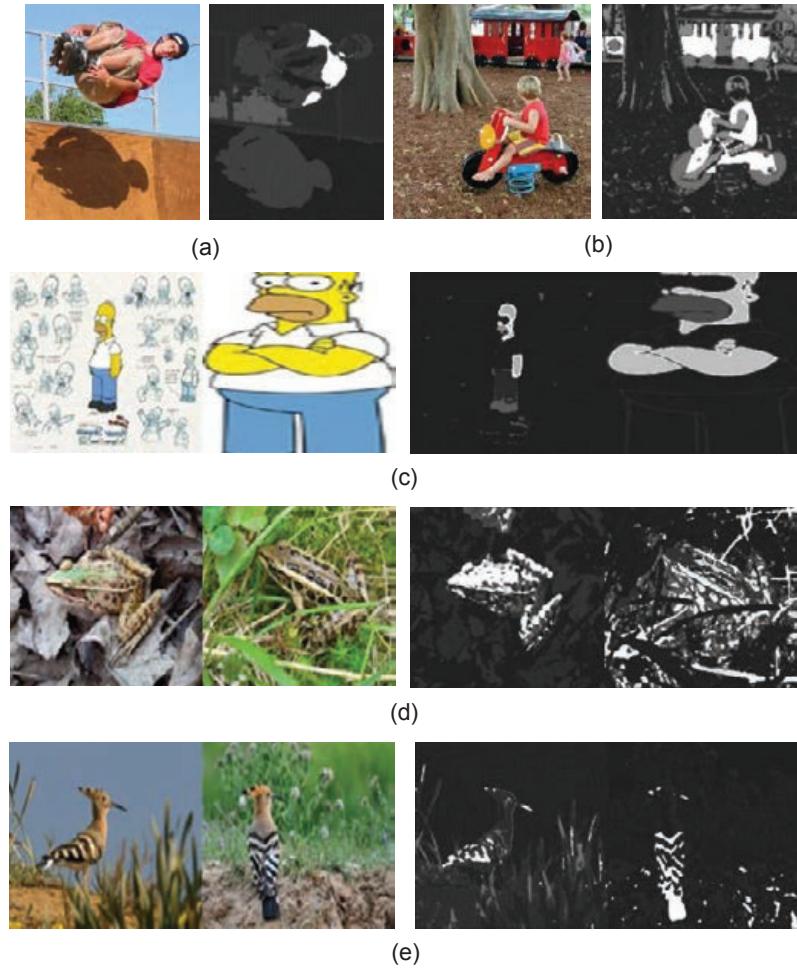


图 2-13 我们的算法一些退化情况。(a-b) 为单张图像显著性检测的退化情况, (c-e) 为多张图像协同显著性检测的退化情况。

数往往增加运算复杂度,降低效率。更重要的是,协同分割算法侧重于从一个小图像集中分割“近似物体”(similarly looking object),而容易导致错分“近似背景”,无法实现完全的自动化系统。常见去除二义性的办法是加入手工标注笔刷(strokes)^[8,39],或者是手工标注框(bounding boxes)^[3,88]。相对而言,协同显著性检测算法可以直接提供了一个反映近似物体的初始化模板(mask),取代人工输入。

在实验中,我们通过马尔科夫随机场(Markov random field)去进行前背景图像的二值分割。能量函数 E 定义为:

$$E = \sum_{i \in I} p(x_i) + \sum_{\{p,q\} \in N} V_{p,q}(x_p, x_q), \quad (2-9)$$

其中, $\mathbf{x} = \{x_i | i \in I\}$ 代表图像 I 的二值标签结果(binary-valued label), $p(x_i)$ 代表像素*i*的协同显著性测度值。 $V_{p,q}$ 是平滑惩罚函数(smoothness penalty),它

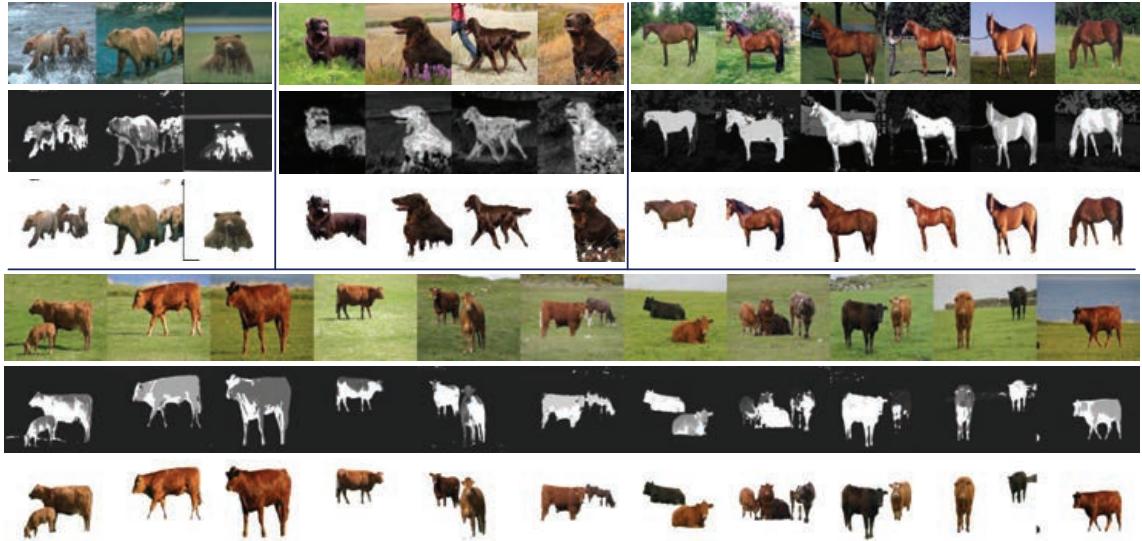


图 2-14 基于我们的协同显著性检测的图像分割结果。自上到下：输入图片，我们的协同显著性检测结果，以及分割结果。

使邻接像素尽可能分配到相同的标签， \mathcal{N} 代表邻接像素集。权重 $V_{p,q}$ 被定义为^[40]：

$$V_{p,q} = \lambda \exp(-\beta \|z_p - z_q\|^2), \quad (2-10)$$

其中， z_p 代表像素 p 的RGB 颜色值， $\beta = (2 \langle (z_p - z_q)^2 \rangle)^{-1}$ ， $\langle \cdot \rangle$ 代表图像期望， λ 代表参数。图2-14 展示了基于我们的协同显著性检测的图像分割结果。图像数量为分别为3、4、6、以及11。总体上说，我们的协同显著性目标检测算法可以得到一个精确的前景目标模板。

表2-1 展示了我们的算法在MSRC 数据库中^[89] 与其他方法^[29,76] 的比较。我们计算其准确度，即正确分类的前景背景像素数目与图像所有数目的比值。表中最后一列 ‘Avg.’ 代表所有分类的平均成绩。我们的算法超过方法^[76] 将近8%。算法^[29] 使用协同显著性检测作为先验知识，然后构建全局能量函数去分割图像，从而获得了最高的分割准确率（Avg. = 85.95%）。我们的算法略逊于^[29]，获得了Avg. = 83.55% 的成绩。主要原因是我们的分割算法，公式(2-9)，并没有使用全局的约束项。对于一些物体类别来说，例如cat、face、和bike，同一类的物体在不同图片中具有极大的光线和特征变化，这使得我们的聚类算法很难把他们聚到同一个类中，从而导致结果下降。但值得注意的是，算法^[29] 的全局约束函数将耗费大量的运行时间，对四张分别率为 128×128 的图像，算法^[29] 需要近40秒去处理。而我们的算法仅需要5秒钟就可以获得近似的分割效果，极大地提高了运行效率。

表 2-1 在MSRC 数据库上的分割准确度 (%)。

Class	Plane	Cow	Cat	Face	Bike	Avg.
方法 ^[76]	73.8	81.6	74.4	84.3	63.3	75.48
方法 ^[29]	87.66	91.36	86.68	87.27	76.76	85.95
我们的方法	86.75	90.39	83.53	84.83	72.28	83.55

2.4.2 鲁棒的图像测距

第二个协同显著性目标检测的应用是鲁棒的图像测距。图像视觉距离计算是一个基础问题，它被广泛应用于多个相关领域：图像检索重排序 (reranking of image retrieval)^[90,91]、以及图像内容聚类 (image content clustering)^[92,93]。近年来，基于物体内容的图像对测距被证明有助于提高基于全局特征的图像检索结果^[10,34,94,95]。例如，文献^[5]使用了协同分割函数能量作为图像对的距离测度。然而，这种方法受限于它的分割测度，无法拓展到更常见的其他视觉特征中去。因此，我们提出了一种更加高效与普遍的基于协同显著性检测的鲁棒性图像测距方法。两张图片 $I_{\{1,2\}}$ 的任一图像距离测度用 $D(I_1, I_2)$ 表示。输入一组图像对，首先通过协同显著性算法对每幅图片进行分割，分为协同显著性前景 I^f 和背景 I^b ，然后我们定义一个显著性比重权值 r^f ：

$$r^f = \text{Size}(I_1^f + I_2^f) \times \text{Mean}(I_1^f + I_2^f), \quad (2-11)$$

其中， $\text{Size}(\cdot)$ 和 $\text{Mean}(\cdot)$ 代表前景 I^f 的像素数目与协同显著性测度值的平均值。相同的，背景也可以定义一个显著性比重权值 r^b 。最终，我们的基于协同显著性目标检测的鲁棒性图像测距 $D'(I_1, I_2)$ 被定义为：

$$D'(I_1, I_2) = \frac{r^f}{r^f + r^b} D(I_1^f, I_2^f) + \frac{r^b}{r^f + r^b} D(I_1^b, I_2^b), \quad (2-12)$$

其中

$$I_i = I_i^f + I_i^b, \quad i = 1, 2.$$

简单而言，我们的算法将图像的协同前景 (co-saliency region) 与背景分开，使前景在图像距离的计算中起到更大的比重。

图2-15 展示了传统图像距离与我们的鲁棒性图像测距的比较。我们的方法不局限在于特定的图像距离测量方式，在本实验中，我们使用颜色直方图的卡方分布距离 (Chi-squared)。在传统的全局图像距离计算中，两幅不相关的图像容易被近似的背景所干扰，从而导致错误的计算距离。相对而言，我们的鲁棒性图像测距则集中于协同显著性目标，而减弱了背景干扰。相比较于其他基于协同分割的距离计算方法^[5,10]，我们的框架具有更高的运行效率、容易移植，而且任何距离计算方式都可以应用到我们的框架中，非常符合实



图 2-15 基于协同显著性目标检测的鲁棒性图像测距结果。中列的图像是输入图像，左侧为正确的匹配图像，包含近似的物体，右侧为不相关的图片，但具有近似的背景。通过协同显著性目标的辅助，我们的方法降低了正确匹配图像的距离，提升了错误匹配对的距离。

际中图像测距的需求。

2.4.3 弱监督学习应用

常见的计算机学习方法主要分为全监督学习(supervised learning)与无监督学习(unsupervised learning)两种。全监督学习需要人工标注(label)像素级别(pixel-level)的图像训练库,这种方法耗时费力,当数据训练库很大的时候,尤其是海量数据的情况下,几乎无法获得像素级别的精确标注数据。相对而言,无监督学习可以自发地挖掘数据的相关内容,但是在无任何交互约束的情况下,而且容易造成数据语义的二义性,导致学习效果低下,很难获取令人满意的学习效果。近年来,随着机器学习方法的深入发展,弱监督学习的概念被逐渐提出^[79,96-98]。弱监督学习(weakly supervised learning),介于全监督学习与无监督学习之间,其主要通过提供图像/视频级别的内容标注,无需明确标注物体的准确位置,而直接进行相关训练学习处理。这种方法极大减轻了全监督中人工精准标注的工作量,同时又提供足够的标注信息,消除无监督学习中的语义二义性,从而保证了训练学习的准确性和鲁棒性。我们的协同显著性检测可以用来作为弱监督学习的一种有效地解决方案。通过协同显著性目标检测,配合前景分割,可以得到一个近似的全监督标定训练图集,这就使得很多全监督分类算法,例如Auto-context^[36],无需大量的人工标注。图2-16展示了我们的协同显著性检测分布图配合Auto-context^[36]的分类检测结果。首先图像经过弱监督标注进行筛选,然后我们的协同显著性目标检测算法输出协同目标二值图,作为全监督训练的训练图集。最后使用Auto-context获得目标类别的分类器,分类器将更适合于检测复杂图片中的目标,如图2-16 (d) 所示。

2.4.4 视频前景提取

最后一个应用是视频前景提取。视频可以看做是一组有序的图像集,其中的前景物体可以作为显著性目标被检测^[99-101]。视频的显著性目标依旧满足特征对比度属性^[24,80,102],以及中心偏移度。同时前景也将满足图间分布度,因为前景会反复出现在视频中。因此我们的协同显著性目标检测可以直接的应用到视频前景检测中。图2-17展示了基于我们的协同显著性目标检测的视频前景提取结果,我们的算法准确的提取了前景目标。此外,作为多张图片的全局检测,我们的方法对于噪声帧(outlier)具有很高的鲁棒性,例如图2-17第二个例子中的第一帧和最后一帧,前景(鲸鱼)并没有出现。注意到,我们的方法仅仅使用了颜色和纹理特征去聚类,以及提取协同显著性目标。而其它的时空(spatio-temporal)特征,例如光流(optical flow)等,也可以很容易的移植到我们的方法中,进一步提高视频前景提取的效果。

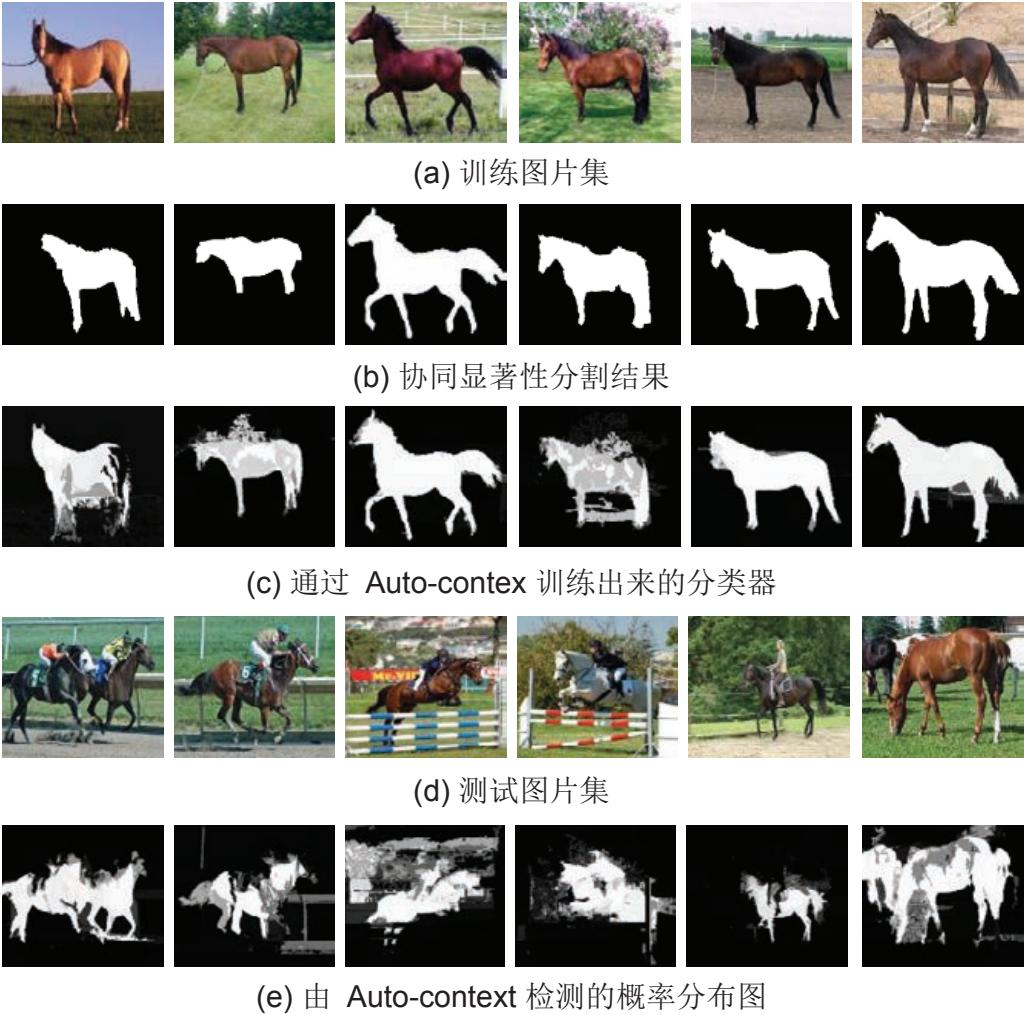


图 2-16 我们的协同显著性检测分布图配合Auto-context^[36]的分类检测结果。(a) 通过弱监督标注的训练图片。(b) 我们的协同显著性检测分布图被用来作为全监督标注图像。(c) 通过Auto-context训练的分类器。(d) 复杂的测试图片。(e) 通过Auto-context分类器的检测概率图。

2.5 本章总结

在本章中，我们研究了多图像间的关联性约束问题，并提出了一种基于聚类的协同显著性目标检测算法。我们通过聚类，构建全局关联约束，避免了大量繁重的训练过程。并提出三种显著性测度去进行协同显著性目标的检测，其中特征对比度和中心偏移度可以广泛的应用于大量单张/多张图片显著性检测数据库中，图间分布度则挖掘物体在多张图片中的出现频率，提供全局约束。我们的算法在单张和多张图片数据库中均取得了理想的结果。我们的算法高效、简洁、精准，作为一个快速全自动的预处理方法，可广泛的应

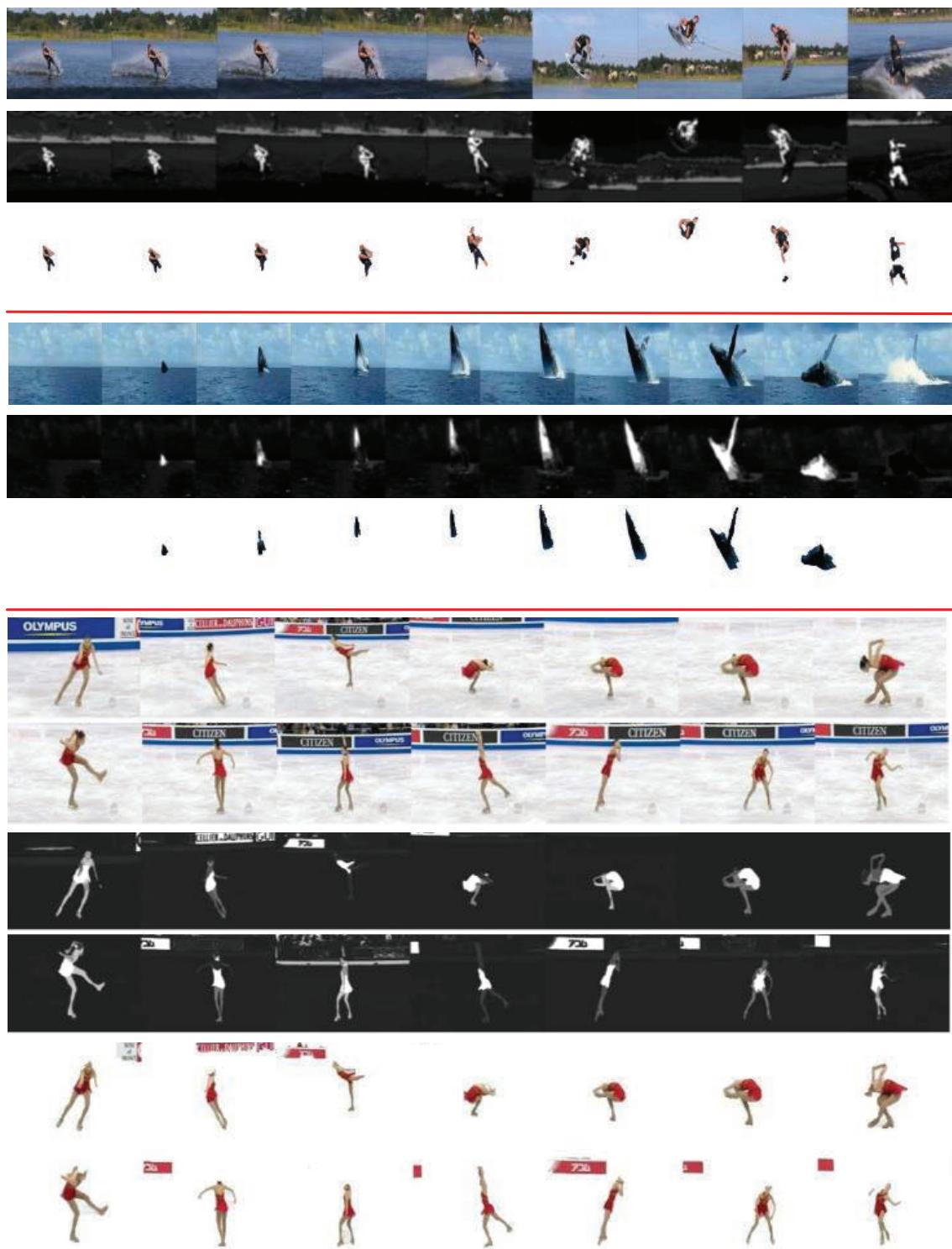


图 2-17 基于我们的协同显著性目标检测的视频前景提取结果。从上而下分别为：输入视频序列，协同显著性目标检测结果，视频前景提取结果。

用到其他相关图像处理应用中。

研究显示，多图像间的关联性约束可以有效地降低图像目标的语义二义性。同时，通过图像间的协同关联作用，可以明显的促进图像自身的处理准确性和鲁棒性。

第三章 单张图像不同区域间的关联性约束

本章将主要研究单张图像不同区域间的关联性约束。自然图片中的物体往往满足一定的几何结构，例如自身全局及局部的对称结构（Symmetry）、重复区域间的仿射结构（Affine）等，这些几何结构可以使图像中的不同区域构成一定的结构约束，各个区域之间构建协同关联矩阵，以此来相互约束，辅助提升处理效果。在本章中，我们将利用物体自身的这种几何结构约束，引入协同分割思想，提出一种图像内关联区域之间的几何结构约束分割算法（Geometry Constraint for Foreground Extraction）。为了反映不同区域之间的协同关联性约束，我们首先利用物体自身的几何结构约束建立一个像素级别的稠密几何结构映射矩阵（Geometry foreground map），该矩阵在描述前景物体的位置分布的同时，也反映了物体自身的几何结构约束。然后我们将这种几何结构矩阵引入到基于图的能量函数模型中，提出新的基于几何结构约束的图分割模型（Geometry constraint model），该模型满足子模性（Submodular），可以通过图分割算法优化求解（Graph cut）。除此之外，我们还将该几何结构约束模型拓展到基于组件的分割框架中（part-based framework），以满足弱几何约束、以及复杂几何约束的物体分割需求。

3.1 相关背景工作

前景提取：前景目标通常被定义为与背景有明显区分的特殊物体^[37,103,104]。已存的方法大多通过用户交互去减少前景由于复杂特征以及弱边缘（weak boundaries）所导致的二义性^[3,39,40]。近年来，出现了大量算法基于目标性（objectness）^[105–108]、视觉显著性（saliency）^[18,24,81]等技术自动提取前景。例如，文献^[109]通过产生大量候选分割子集，然后对其进行完整性排序，最终输出排名最高的分割前景。而文献^[103]则使用了人类视觉显著性准则作为依据，提取初始的前景目标，再进行分割。在我们的算法中，我们同样使用视觉显著性准则生成前景先验分布图，以此来提供前景的初始位置信息。我们将前景先验分布图与几何结构融合生成最终的几何结构映射矩阵，以此来描述图像前景的几何结构分布。

几何结构检测：对称检测的研究已经发展了很多年^[110,111]，比如：文献^[112]中，通过对称直方图投票来筛选所有可能的对称轴。文献^[113]提出一种夹角关联度（angular correlation）用来测量图像中的特征点夹角，这种方法

可以快速的检测出旋转与镜面对称结构。文献^[114]则提出一种基于对称增长 (geometry-growing) 的方法，去检测对称特征结构。在文献^[115]中，MIFT 特征 (mirror SIFT, MIFT) 被用来检测对称点特征。然而，这些方法只能提供一些稀疏的不充足的匹配点对，无法提供足够的像素对级别的 (pixel-wise) 分割支持。通常来说，几何变换单应矩阵 (transformation homography) 可以通过少数几个匹配点对计算出来。几何变换单应矩阵是一个线性的 (linear) 非奇异的 (non-singular) 3×3 矩阵 \mathbf{H}_T ^[116]。我们用 Λ 代表图像像素的数目，对每一个像素点 $u \in \Lambda$ ，都存在一个几何结构匹配的对应点坐标 $\mathbf{x}_{u'}$ ：

$$\mathbf{x}_{u'} = \mathbf{H}_T \mathbf{x}_u, \quad (3-1)$$

其中， $\mathbf{x}_u = (x_u, y_u, 1)^T$ 代表像素点 u 的齐次坐标。显然，几何变换单应矩阵提供了一个全局的几何约束，但其缺乏鲁棒性，尤其是对复杂/弱几何结构的物体来说。因此，在我们的方法中，我们将提供一个稠密的几何结构映射去替代稀疏的几何特征点对。我们的几何结构映射以全局的几何变换单应矩阵为基础，同时满足局部的最优化特征匹配。

图分割函数：前景分割问题通常可以被格式化为一个全局图优化问题，其可以通过图分割 (graph cuts) 进行优化求解^[14,117-119]。图分割能量函数可以定义为^[12,40]：

$$E(L) = \sum_{u \in \Lambda} D_u(l_u) + \sum_{(u,v) \in \mathcal{N}} V_{u,v}(l_u, l_v), \quad (3-2)$$

其中， $L = \{l_u | u \in \Lambda\}$ 代表了图像的二值标签 (binary-valued label)，对于硬分割来说 $l_u \in \{0, 1\}$ 。 D_u 被称为数据惩罚项 (data penalty term)， $V_{u,v}$ 是平滑项 (smoothness penalty term)，它使邻接像素尽可能分配到相同的标签， \mathcal{N} 代表邻接像素集。在图像分割中，数据惩罚项 D_u 用来独立的评估像素点与前景/背景特征的紧密度，而平滑项 $V_{u,v}$ 则保证区域邻接像素点的紧密性。这个能量函数形式又被称为马尔科夫随机场 (Markov Random Fields, MRF) 模型^[13]。我们的方法将几何结构约束嵌入到这个能量方程中，以此来提高目标的分割精度。同时，我们的新模型依旧满足子模型，可以通过图分割来求解。

3.2 几何结构映射矩阵

图3-1 给出了我们的基于几何结构约束的前景分割算法流程图。首先根据视觉显著性目标检测算法，还可以得到一个前景先验分布图 (foreground prior map)。。其次，通过几何结构检测得到几何变换单应矩阵。在我们的方法中，不限制特定的几何结构检测算法，这里我们使用方法^[110]为例。根据几何变换单应矩阵和局部特征优化，我们可以得到一个几何约束强度分布图 (geometry magnitude map)。然后，我们将两个分布图结合，得到几何结构映射矩阵 (geometry foreground matrix)。该矩阵一方面提供了基于全局几何结

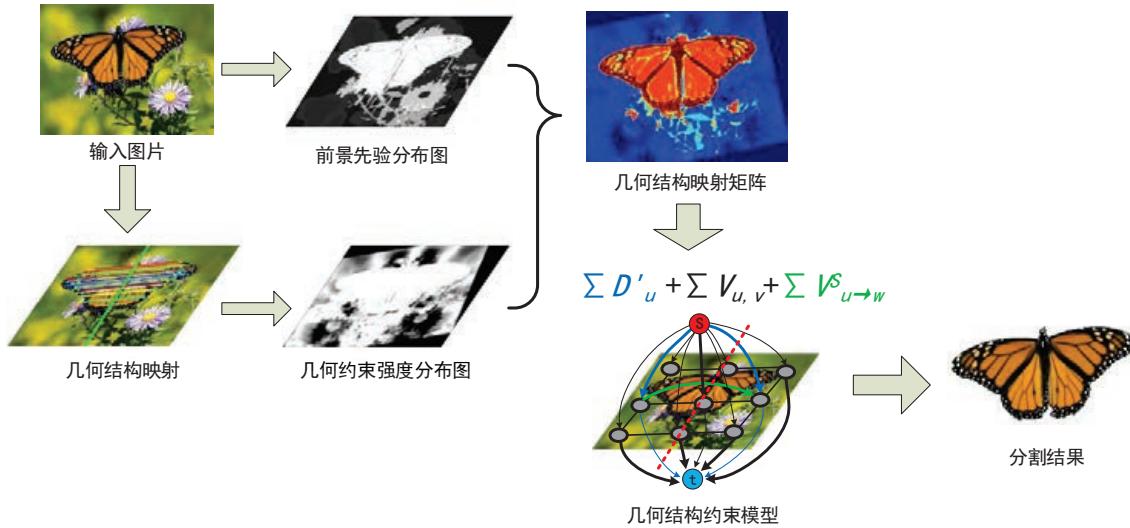


图 3-1 基于几何结构约束的前景分割流程图。我们首先生成两个基本分布图：前景先验分布图通过视觉显著性特征，提供前景目标的初始化位置。几何约束强度分布图提供全局的几何匹配约束。然后将两个分布图进行融合，生成几何结构映射矩阵，该矩阵提供了一个稠密的几何结构匹配关系。最后，我们将几何结构映射矩阵嵌入到图像分割函数中，生成几何结构约束模型，并通过图分割，提取出最终的前景物体。

构约束的稠密像素匹配结果，另一方面也提供了前景目标的初始位置信息。最终，我们将几何结构映射矩阵融合到图像分割能量函数中，得到几何结构约束模型，通过图分割优化，提取出前景目标物体。下面我们将详细介绍每一个步骤。在本文中，我们主要基于对称几何结构去介绍我们的算法，而不失一般性。一方面是由于对称结构是最简单直接的几何约束形式，其广泛存在于自然真实物体中。另一方面是由于绝大多数物体都可以通过对称轴变换（bilateral reflection with the medial axis transform）来进行表示^[120]。

3.2.1 前景先验分布图

第一个基础分布图是前景先验分布图。我们使用视觉显著性准则，来生成前景位置。显著性检测是基于人类生物视觉注意力机制出发，可以通过底层特征相应生成^[18,19,121]。我们不限定视觉显著性检测的方法，在本文中我们使用我们的基于聚类的单张图像视觉显著性检测算法。详细的算法细节已经在第2章描述，我们直接使用其单张图片的检索结果作为输出，生成前景先验分布图S。图3-2 (b) 展示了一些前景先验分布图的例子。由于视觉显著性目标检测仅仅依赖于低级别自底向上测度，导致其无法应对复杂场景以及高纹理背景。但前景先验分布图的主要作用是提供一个前景目标的初始位置，尽可能的抑制背景，而无需精确地分割出前景目标。

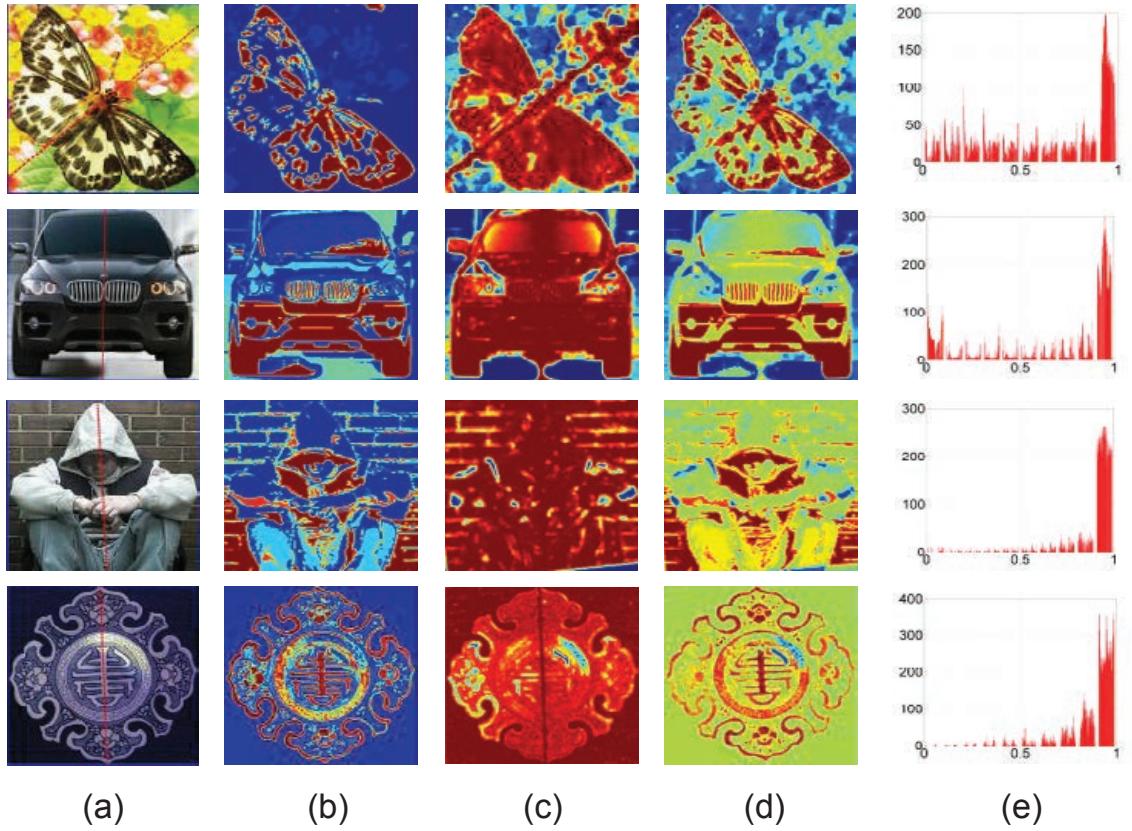


图 3-2 (a) 四种典型的对称前景。(b) 前景先验分布图。(c) 几何约束强度分布图。(d) 几何结构映射矩阵。(e) 几何强度约束分布的直方图。

3.2.2 几何约束强度分布图

第二个基础分布图是几何约束强度分布图。如同第3.1所述，单纯的图像几何结构可以通过全局的几何变换单应矩阵（homography） \mathbf{H}_T 得到。不过这种单应矩阵仅能提供一个全局的约束，缺少据不匹配的鲁棒性。因此我们提出一种几何约束强度分布图，加入区域特征匹配来优化全局几何结构约束。对于任意图像点 u ，首先基于单应矩阵变换公式（3-1），求取其对应的全局几何关联参考点（geometric anchor） u' 。然后计算几何关联参考点 u' 的领域中，每一个像素点 v' 与原始像素点 u 的特征匹配强度 $m(v')$ ：

$$m(v') = \frac{-\|z_u - z_{v'}\|_2}{\mathcal{G}[dis(\mathbf{x}_{u'}, \mathbf{x}_{v'}) | 0, \sigma^2]}, \quad (3-3)$$

其中， $\mathbf{x}_u = (x_u, y_u, 1)^T$ 代表了像素 u 的齐次坐标， z_u 代表了像素 u 的RGB特征值，L2 norm 被用来计算特征距离。我们使用高斯核（Gaussian kernel） \mathcal{G} 去作为欧氏距离（Euclidean distance） $dis()$ 的惩罚函数，系数 σ 设置为玲姐半径。可见我们的特征匹配强度公式（3-3）融合了两个方面的信息：全局几何结构信息、以及局部特征匹配信息。换句话说，公式（3-3）可以看做是一个寻找最

佳几何匹配对的代价函数 (cost function)，其包含与几何关联参考点 u' 的空间距离，以及与像素点 u 的特征匹配度。

最终，最佳几何匹配点 w 可以通过计算几何关联参考点 u' 的领域中，特征匹配强度最大值求得：

$$u \rightarrow w = \arg \max_{w \in N(u')} \{m(w)\}, \quad (3-4)$$

其中， $N(u')$ 几何关联参考点 u' 的领域。注意几何关联参考点 u' 仅仅代表原始像素点 u 的初试几何对应位置，而通过公式 (3-4) 计算的最佳几何匹配点 w 则包含了局部的特征匹配优化，其针对弱几何关系的情况更具有鲁棒性。

图3-2 (c) 展示了几何约束强度分布图 $\mathbf{M} = [m(u)]$ 的例子。显然几何约束强度分布图提供了一个更加稠密的点对点几何结构匹配关系。但是可以注意到对于背景也满足几何关系的情况，几何约束强度分布图则不加以区分，例如图3-2 (c) 的最后两例。

3.2.3 几何结构强度因子

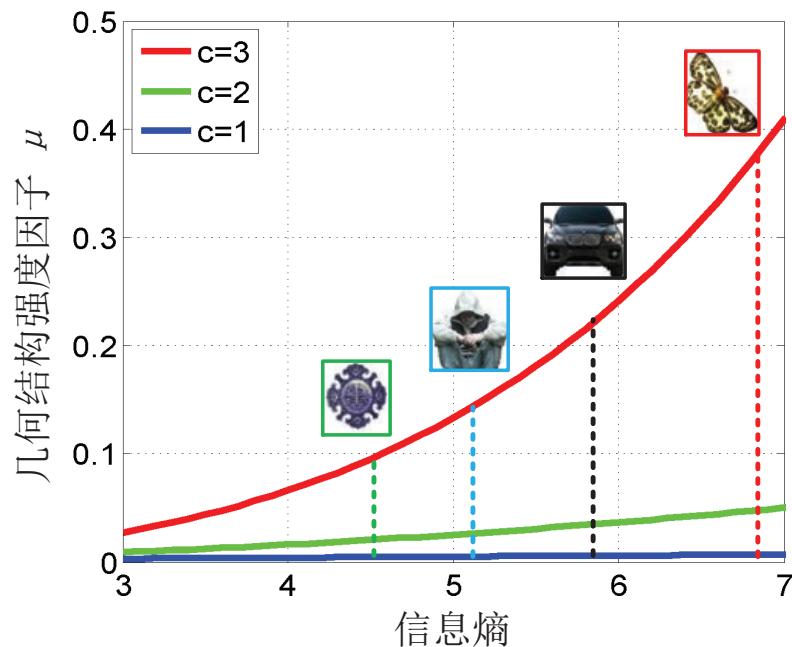


图 3-3 几何结构强度因子 μ 与图像信息熵 $E_n(\mathbf{M})$ 的函数关系曲线。

对于每一个分布图来说，单独使用时都会存在不足，因此我们将其叠加，形成一个初始的几何结构映射矩阵：

$$\mathbf{H}'_s = \frac{\mathbf{S} + \mathbf{M}}{2}, \quad (3-5)$$

其中， \mathbf{S} 和 \mathbf{M} 分别代表前景先验分布图和几何约束强度分布图。

如前文所述，如果背景很难区分，而且也存在一定的对称性时，我们的方法将退化。图3-2 (a) 展示了四种典型的背景强弱情况：蝴蝶图例中的前景具有明显的对称结构，而大量背景没有对称性；汽车图例则是对称前景占据了绝大部分空间；人物图例展示了弱对称前景和对称背景同时存在；最后的标志图例展示了前景和背景都具有强对称性。我们希望几何结构映射矩阵可以有效的区分这四种情况。图3-2 (d) 则展示了由公式 (3-5) 计算的几何结构映射矩阵。我们发现该矩阵依然很难描述背景也满足几何约束的情况。例如：当非几何背景出现的时候（如前两张图例），几何结构映射矩阵可以很好地反映出前景情况，区分开前景/背景。但当背景复杂且满足几何结构时（后两行图例），几何结构映射矩阵容易混淆前景/背景，如果依然使用几何结构约束的话，反而会影响前景分割的结果。因此，我们希望当出现这种退化现象时，几何结构的约束可以被减弱，甚至消除，因此我们将导入一个新的变量，几何结构强度因子 (quantity parameter)，用来评估图像的几何结构情况。

在实际情况中，我们发现几何约束强度分布图更容易受到背景干扰。我们统计几何约束强度分布图的直方图统计，如图3-3 所示，发现非几何约束背景的直方图比几何约束的背景具有更平滑的直方图。因此我们简单的引入信息熵去描述这个全局结构，提出几何结构强度因子 μ 的概念：

$$\mu = \exp \left\{ \frac{E_n^c(\mathbf{M})}{a} \right\} - b, \quad (3-6)$$

其中， $E_n(\mathbf{M})$ 代表了几何约束强度分布图 \mathbf{M} 的图像信息熵，其包含256 个bin，值域为[0, 8]。超参数 (hyperparameters) (a, b, c) 控制着几何结构强度因子 μ 的分布情况，并保证其归一化到(0, 1) 区间内。在实验中，我们设置 $a = 1000$, $b = 1$ 。图3-3 展示了几何结构强度因子 μ 与图像信息熵 $E_n(\mathbf{M})$ 的函数关系曲线。我们同样标出了不同超参数 $c = 1f2f3$ 的曲线作为对比，可见当 $c = 3$ 时，具有最佳的梯度变化。

最后，我们得到加入几何结构强度因子 μ 的几何结构映射矩阵 (Geometry foreground map) \mathbf{H}_s 为：

$$\mathbf{H}_s = 2\mu \frac{\mathbf{S} + \mathbf{M}}{2} + 1 - \mu. \quad (3-7)$$

其中，几何结构映射矩阵的值域被限制在 $(1 - \mu, 1 + \mu)$ 内。通常来说，几何结构强度因子 μ 控制着几何结构映射矩阵的总体强度，即约束分割函数的强度。如果图像背景也呈现很强的几何结构，那么信息熵 $E_n(\mathbf{M})$ 将导致一个较小的几何结构强度因子 μ ，有助于降低几何结构约束对前景分割的影响。

3.3 几何结构约束模型

在本节，我们将介绍如何将几何结构映射矩阵嵌入到传统的分割能量函

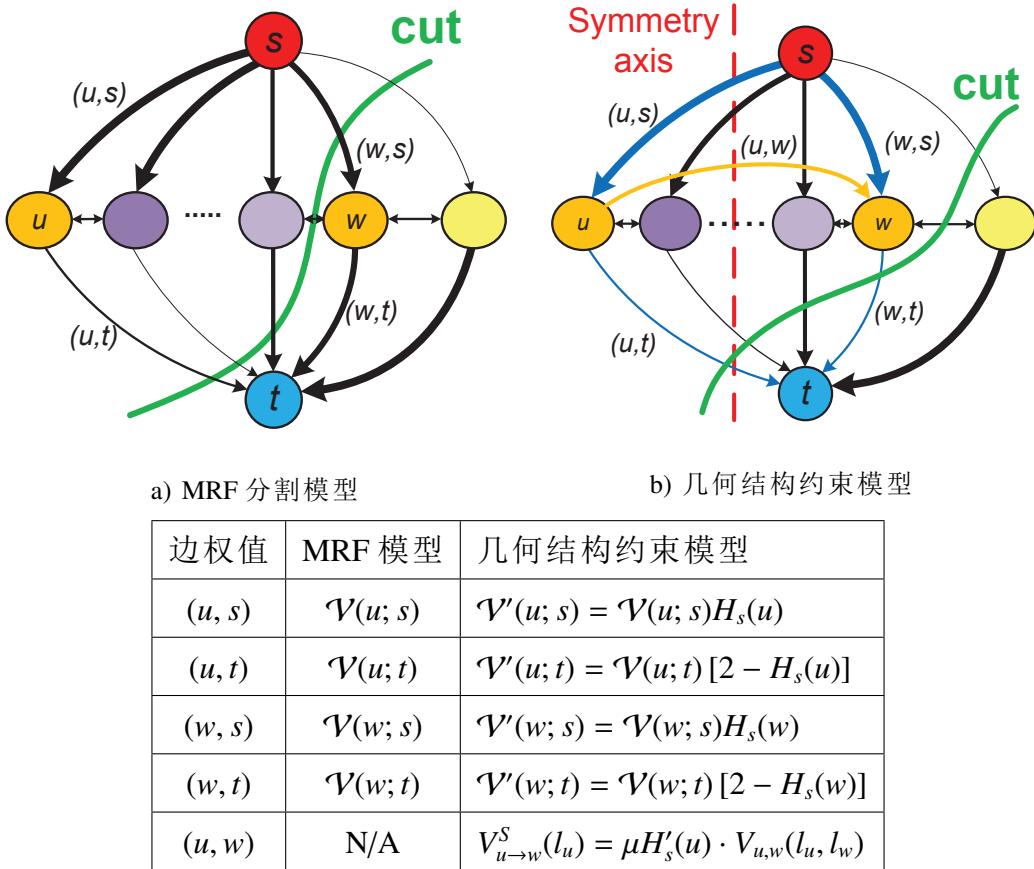


图 3-4 二值 s/t 分割图示意图。在图中节点 u 拥有几何结构匹配点 w 。红色的虚线代表一个集合约束关系，线的粗细代表的边权值的大小。在 (b) 中新加入的黄色边代表了几何平滑项。列表中显示了每一条边再加入几何结构约束后的变化。

数，本文以以MRF 公式 (3-2) 为例。从而生成新的前景分割能量函数，我们称其为“几何结构约束模型”(geometry constraint model)。

3.3.1 嵌入几何约束的数据惩罚项

在传统的MRF 函数中 (公式 (3-2))，数据惩罚项 (data term) D_u 是基于前景/背景种子点而决定的。前景/背景用 s 和 t 表示，数据惩罚项常写为：

$$D_u(u) = \mathcal{V}(u; l_u), \quad \text{with } l_u = \{s, t\}, \quad (3-8)$$

其中， $\mathcal{V}(\cdot)$ 代表了像素点 u 与前景/背景的近似分布度。

在我们的框架中，我们使用几何结构映射矩阵作为前景的均衡参数。嵌

入几何约束的数据惩罚项 D'_u 可表述为:

$$D'_u(u) = \begin{cases} \mathcal{V}(u; l_u)H_s(u) & , \text{if } l_u = s \\ \mathcal{V}(u; l_u)[2 - H_s(u)] & , \text{if } l_u = t \end{cases}, \quad (3-9)$$

其中, $H_s(u)$ 代表了像素(节点) u 对应的几何结构映射矩阵值。在我们的嵌入几何约束的数据惩罚项中, 满足几何结构的像素点被鼓励分类为前景, 而不满足几何约束的点更有可能被分割为背景。对于非几何结构的前景和满足几何约束的背景, 我们的数据惩罚项有可能误分割它们。但是由于几何结构强度因子的作用, 当背景也是对称时, 因子趋于0, 因此嵌入几何约束的数据惩罚项将退化成传统的MRF 分割能量函数。

3.3.2 几何约束平滑项

在平滑项中, 我们使用一个近似传统MRF 平滑项的边权值, 连接几何匹配点对:

$$\sum_{\{u,v\} \in \mathcal{N}} V_{u,v}(l_u, l_v) + \sum_{u \in \Lambda} V_{u \rightarrow w}^S(l_u), \quad (3-10)$$

其中, $u \rightarrow w$ 代表像素 w 是像素 u 的最佳几何约束匹配点。平滑项 $V_{u,v}$ 是公式 (3-2) 中的传统MRF 平滑项。几何约束平滑项 $V_{u \rightarrow w}^S$ 将鼓励满足几何约束匹配的点对分配成同样的类中, 其具体方式定义为:

$$V_{u \rightarrow w}^S(l_u) = \mu H'_s(u) \cdot V_{u,w}(l_u, l_w), \quad (3-11)$$

其中, $H'_s(u)$ 代表公式 (3-5) 中的初始的几何结构映射矩阵, 而 μ 代表几何结构强度因子。不同于公式 (3-7) 中的几何结构映射矩阵, 其值域范围为 $(1 - \mu, 1 + \mu)$, 在几何约束平滑项中的 $\mu H'_s(u)$ 其值域被限制在 $[0, \mu]$ 中。其含义是, 当几何结构映射矩阵值高的时候, 几何约束平滑项将起到一个更大的作用。而对于非几何结构的像素点, 几何约束平滑项接近于0。

3.3.3 基于几何结构约束的前景分割能量函数

最终, 我们的几何结构约束模型定义为:

$$\mathbf{E}_s(L) = \sum_{u \in \Lambda} D'_u(l_u) + \sum_{\{u,v\} \in \mathcal{N}} V_{u,v}(l_u, l_v) + \sum_{u \in \Lambda} V_{u \rightarrow w}^S(l_u), \quad (3-12)$$

同时, 前景分割结果可以通过全局最小优化得到:

$$\mathbf{L}^* = \arg \min \mathbf{E}_s(L). \quad (3-13)$$

图3-4 (a) 展示了传统的MRF 分割模型, 而图3-4 (b) 则展示了我们的几何结构约束模型。其中, 节点 u 拥有几何结构匹配点 w 。红色的虚线代表一个集合约束关系, 线的粗细代表的边权值的大小。相对于图3-4 (a), 蓝色的边表示与传统的MRF 分割模型相对应的嵌入几何约束的数据惩罚项(边),

而黄色的边代表新加入的几何约束平滑项。值得注意的是，我们的几何结构映射矩阵 \mathbf{H}_s 是单一映射，因此几何约束平滑项是从节点 u 到节点 w 的有向边。

性质：公式 (3-12) 定义的能量分割函数 $\mathbf{E}_s(L)$ 满足子模型 (submodular)，因此可以通过图分割 (graph cut) 进行优化求解。

由于几何结构强度因子 $\mu \in (0, 1)$ 的作用，几何结构映射矩阵 \mathbf{H}_s 的值域被限制在 $[1 - \mu, 1 + \mu]$ 。因此几何结构约束模型中，在公式 (3-9) 嵌入参数 $H_s(u)$ 和 $(2 - H_s(u))$ 后，并没有改变数据惩罚项的符号，因此该项是正则的 (regular)。在公式 (3-12) 中的几何约束平滑项 $\sum V_{u \rightarrow w}^S$ 可以被看做是一个特殊的平滑边，连接两个满足几何结构约束的节点。因此，公式 (3-12) 定义的能量分割函数 $\mathbf{E}_s(L)$ 满足子模型 (submodular)，可以通过图分割优化求解。

3.4 基于组件的扩展框架

现在，我们将提出的基于几何结构约束的前景分割算法扩展到更加一般的情况，我们分两种情况去分析：

弱几何结构物体：如前文所述，我们的几何结构约束模型仅仅提供一个弱几何约束(weak geometry constraint)。我们的模型鼓励满足几何约束的区域 (像素) 分配到前景，同时抑制非几何约束的区域 (像素) 分配到背景，而不是强制分配。尤其是在几何结构强度因子的作用下，对于退化的非几何结构图像，我们的方法将会退化成普通的MRF模型，从而减弱甚至去除几何结构约束。

复杂/非几何结构物体：通常来说，大多数物体都可以通过骨架中轴分解成多个组件部分^[120,122]。例如，对于非线性的对称几何变换，弯曲对称模型 (curved symmetry pattern)^[45]，可以通过整合多个局部对称来拟合。同样，无明显几何结构的物体也可以通过部件分解，分成多个满足局部几何结构的组件^[123]。受到基于组件框架的启发，我们可以首先通过局部几何结构检测或中轴检测 (part-geometry detection or medial axis detection)^[120] 获取局部组件的结构结构，如图3-5所示。然后对每一个几何结构区域 $\{R_k\}_{k=1}^K$ ，将拓展几何约束算法应用到局部部件中，计算其相应的局部几何强度分布图 $\{M_k(u)\}_{k=1}^K$ ，其中 K 代表组件的数目。最后通过组件融合，再形成全局的几何强度分布图 M ，融合的几何强度分布图为：

$$M(u) = \begin{cases} \max\{M_k(u)\} & , \text{if } u \in \{R_k\}_{k=1}^K \\ 0 & , \text{otherwise} \end{cases} . \quad (3-14)$$

最后，通过公式 (3-7) 得到最终的几何映射矩阵，进行前景分割。

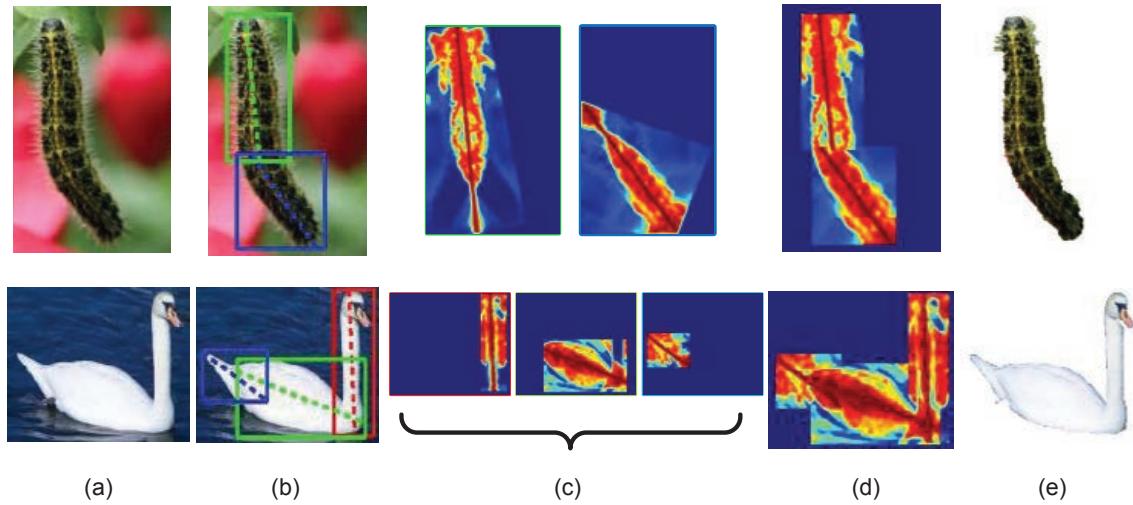


图 3-5 对于复杂/弱几何约束物体的基于组件的扩展分割流程图。(a) 输入图片。(b) 局部几何结构检测结果, 图中每一个局部组件都被标出, 虚线代表几何结构(对称轴)。(c) 每个局部组件对应的几何强度分布图。(d) 通过融合局部组件的全局几何强度分布图。(e) 最终的分割结果。

3.5 实验与分析

在本节中, 我们主要测试三个方面的内容: 1) 测试几何结构约束模型中每个部分的作用; 2) 与传统的MRF模型以及其他前景分割算法^[49,103,109]进行比较; 3) 讨论几何结构模型的鲁棒性。

实验主要针对基于组件的扩展框架进行, 因为单一个的几何结构约束可以看做是组件模型的特例。我们的算法运行设置为: 首先通过全局几何结构检测算法^[110]检测物体的全局对称轴。如果检测失败, 则继续使用局部的组件几何结构检测算法^[120]检测组件结构。然后通过公式 (3-7) 计算几何结构映射矩阵, 生成几何结构约束模型, 并通过最大流最小割算法^[12]求得分割结果。

3.5.1 几何结构约束模型分解测试

我们首先测试几何结构约束模型每个部分的作用。我们的几何结构约束模型主要包括两部分: 嵌入几何约束的数据惩罚项, 以及几何约束平滑项。图3-6展示了不同部分的分割结果, 其中图3-6 (b) 是生成的几何结构映射矩阵。图3-6 (c) 展示了未加任何几何约束的传统MRF模型分割效果, 我们标注为 $D_v + V_{u,v}$ 。可以看到当背景和前景特征近似时, 容易造成分欠分割(例如汽车阴影、花瓣等部位)或者过分割(例如车灯、蝴蝶翅膀等部位)。图3-6

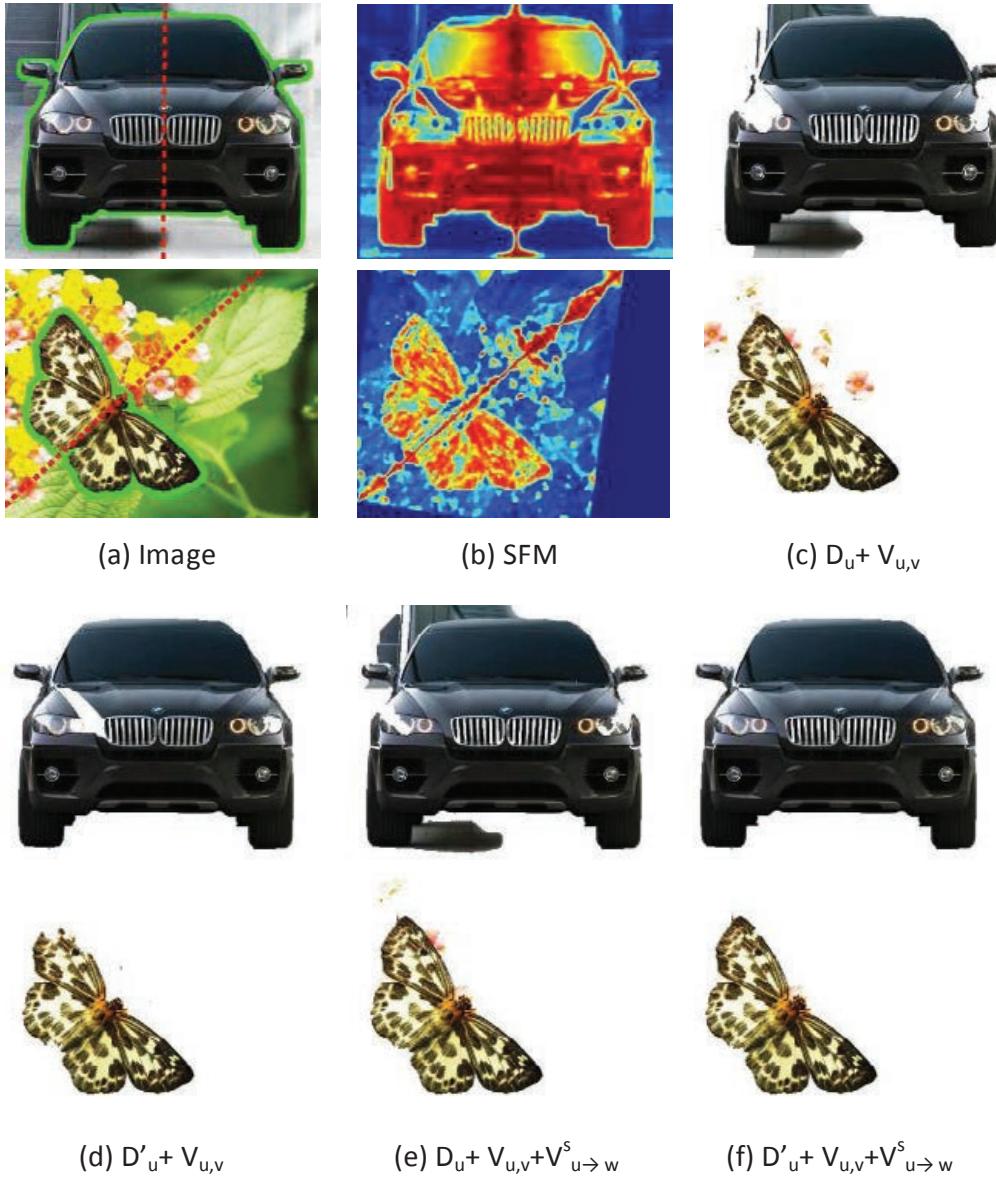


图 3-6 几何结构约束模型中不同部分的分割结果。(a) 输入图片，其中红色虚线为几何对称轴，绿色线为ground truth。(b) 几何结构映射矩阵。(c) 传统MRF模型分割结果。(d) 几何结构加入到数据惩罚项的分割结果。(e) 加入几何约束平滑项的分割结果。(f) 完整的几何结构约束模型分割结果。

(d) 展示了在数据惩罚项嵌入几何结构约束后的分割效果，表示为 $D'_u + V_{u,v}$ 。可以看到欠分割部分基本被消除，因为那些部分都不满足几何结构。图3-6(e)展示了仅加入几何约束平滑项的前景分割结果，我们的几何约束平滑项仅仅增强满足几何匹配的像素对分配同样的标签(label)，例如车灯部分、和蝴蝶翅膀。但对于不满足几何约束的部分，不起作用，例如汽车阴影等。当

所有约束全都融合后($D'_u + V_{u,v} + V_{u \rightarrow w}^S$)，我们的完整几何结构约束模型取得了最好的结果，如图3-6 (f) 所示。

3.5.2 分割准确度比较

由于几何结构约束的加入，我们的方法可以得到一个稳定的分割结果，而无需迭代。我们采用传统的MRF模型作为基准分割结果(baseline)。在实验中，我们使用高斯混合模型(Gaussian Mixture Model, GMM)来作为前景/背景种子点的数据惩罚测度^[40]。传统MRF模型的数据惩罚项被定义为^[3]:

$$D_u = \log p(z_u|GMM_b) - \log p(z_u|GMM_f), \quad (3-15)$$

其中， z_u 代表了像素 u 的描述特征(appearance)， GMM_f/GMM_b 代表前景/背景的高斯混合模型。我们同样改进这个框架，加入几何结构约束:

$$D'_u = (2 - H_s(u)) \log p(z_u|GMM_b) - H_s(u) \log p(z_u|GMM_f), \quad (3-16)$$

其中， H_s 为几何结构映射矩阵。传统MRF模型的数据平滑项为:

$$V_{u,v} = \frac{\lambda \exp(-\beta \|z_u - z_v\|^2)}{dis(u, v)}, \quad (3-17)$$

其中， $dis(\cdot)$ 代表相邻像素的欧氏距离， $\beta = (2 \langle (z_u - z_v)^2 \rangle)^{-1}$ ， $\langle \cdot \rangle$ 代表全图像的期望值(expectation)， λ 是调节系数。注意到我们的实验主要侧重高级别的几何结构约束是否可以增强传统的低级别的分割结果。因此实验中的MRF模型不同于常见算法(例如^[3,40])的配置，我们仅使用一个标注窗口(bounding box)，来作为初始的数据权重。

图3-7展示了三种几何结构图像。第一种(图3-7 (a))是简单的几何结构物体，其可以通过全局的几何映射矩阵来描述。相对于传统的MRF模型，我们的几何结构约束模型抑制的非几何结构化的背景，取得了明显的效果提升。图3-7 (b)展示了无全局几何结构的前景。基于局部几何结构检测^[120]，我们可以计算出全局的几何映射矩阵，从而完成分割。图3-7 (c)展示了第三种情况，即有些目标可以被全局和局部结构同时表示，例如旋转、平移等结构。可以看到，我们的算法无论是全局还是局部，都优于传统的MRF模型的分割结果。同时注意到在本例中，全局约束的分割结果要好于局部组件的分割结果。其一种可能的原因是由于图像背景中的绿叶，也存在局部几何结构，因此影响了局部前景的提取(如图3-7 (c) 中局部几何结构约束的SFM图中的黄箭头所示)。

3.5.3 数据库的批量实验结果

为了大量测试分割结果，我们使用了两个标准数据库去评估我们的算法与传统MRF模型的区别。我们同样附加了另外两种前景分割算法^[103,109]。其

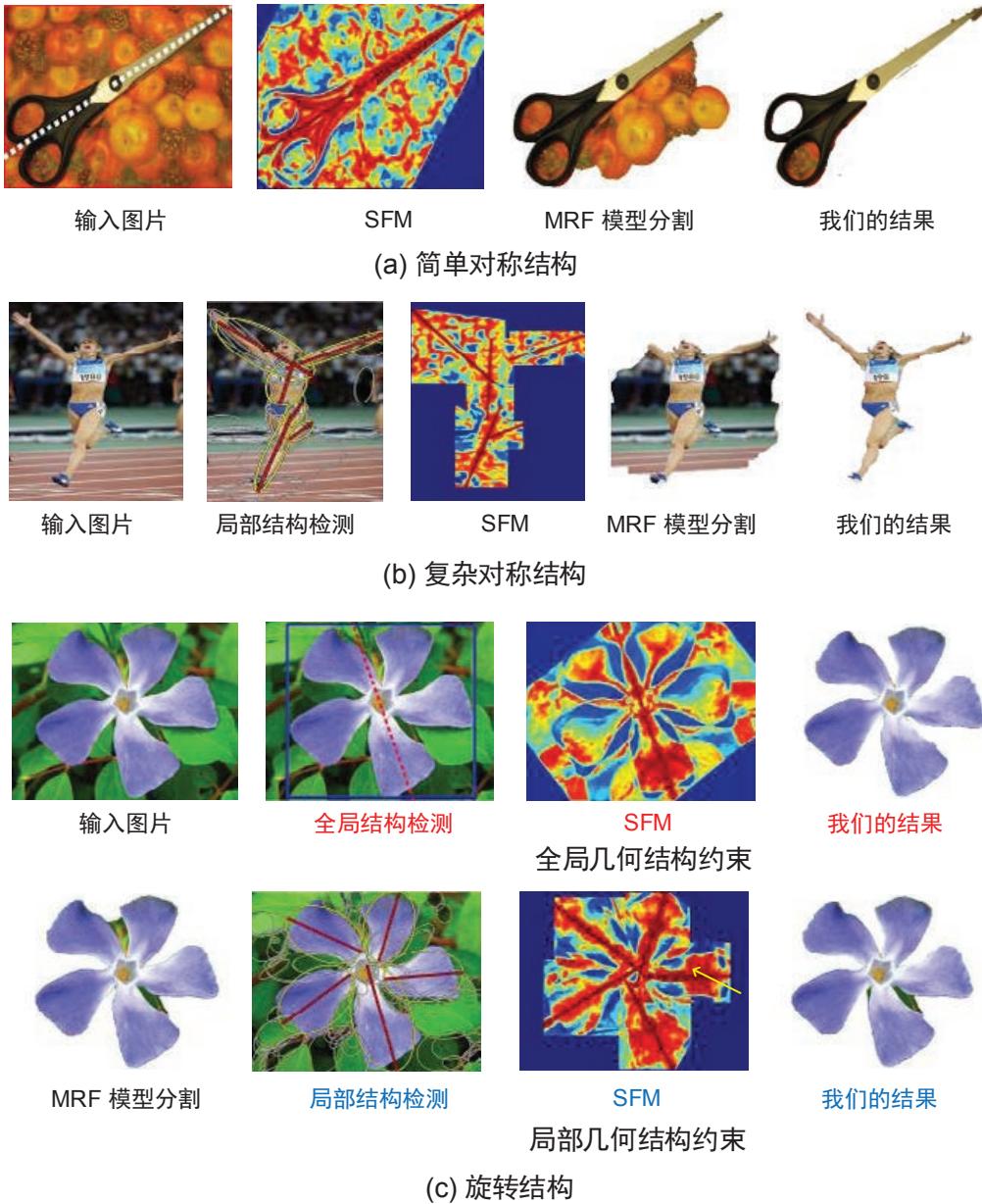


图 3-7 对于不同几何结构的分割效果，其中 SFM 代表几何结构映射矩阵。(a) 简单的对称几何结构，其满足全局对称性。(b) 复杂结构物体通过组件分解成局部几何约束。(c) 同时满足全局几何结构和组件结构的物体，可以用两种方法分别各自分割。

中，算法^[109]将生成一组前景分割候选集合，并依据训练分类器去排序所有候选结果，将得分最高的结果作为输出。文献^[103]则是采用视觉显著性关注点(fixation point)去生成前景的初始标定，然后使用极坐标边缘优化来分割前景物体。为了比较，我们定义groundtruth的中心为视觉显著性关注点，作为方法^[103]的初始输入。此外，我们统计所有算法的分割错误率^[3,40]，作为前景

表 3-1 在两个数据库上的分割结果比较

	MRF 模型	方法 ^[103]	方法 ^[109]	几何结构约束模型		
				$V_{u \rightarrow w}^S$	D'_u	$D'_u + V_{u \rightarrow w}^S$
Leeds 蝴蝶数据库(%)	13.73	12.84	14.78	12.56	10.74	8.83
Grabcut 数据库(%)	11.39	10.58	12.40	12.14	9.81	9.21

分割效果的评价指标。



图 3-8 蝴蝶数据库的分割结果。首行是输入图像，中间行是我们的几何结构映射矩阵，最底行是我们的分割结果。

第一个测试数据库是Leeds 蝴蝶数据库^[124] (Butterfly Dataset)，其一共包含832张图像，并提供每张的像素级标注 (groundtruth)。蝴蝶是自然界中最具有几何结构 (对称结构) 的物体之一。我们在此数据库上测试几何结构对分割的提升。图3-8 展示了分割结果，其中中间行是我们的几何结构映射矩阵，最底行是我们的分割结果。我们的几何结构映射矩阵反映了目标的几何结构，并提供一个准确的分割模板 (mask)。此外，对于图像中的小物体或者复杂背景，例如图3-8 中后两列图例，我们的几何结构映射矩阵容易造成一些噪声。但对于我们的几何结构约束分割模型来说，几何约束仅仅提供一个弱限制，我们的最终分割结果完全去除了这些噪声干扰。

表3-1的第一行展示了算法在蝴蝶数据库的分割结果，自左向右分别是：MRF模型，方法^[103]，方法^[109]，仅使用几何约束平滑项 ($V_{u \rightarrow w}^S$)，仅使用几何约束数据惩罚项 (D'_u)，完整的几何结构约束模型 ($D'_u + V_{u \rightarrow w}^S$)。实验结果表

明，加入几何结构约束后，分割结果均取得了提升一定程度的提升。尤其是我们的完整的几何结构约束模型与传统的MRF模型相比，具有将近5%的效果提升。分割模型的各部分实验结果与第3.5.1节的实验结果一致。注意到方法^[109] 的分割结果较差，这主要是因为方法^[109] 主要基于常用数据训练的函数，最佳的分割结果可能排名靠前（top-rank），但不是排名第一（first rank）。而实验中，主要将排名第一的结果作为输出比较。

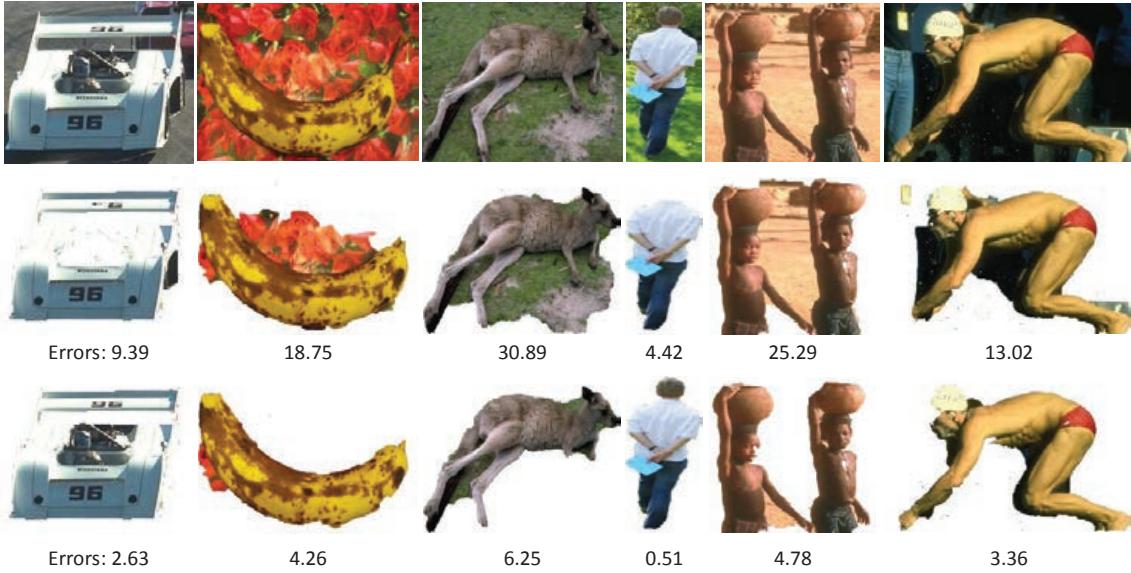


图 3-9 我们的分割结果（第三行）与传统的MRF 模型（第二行）结果比较，其中错误率标注在相应结果的下方。我们的几何结构约束可以明显的提升分割效果。

我们实验中的第二个数据库将采用GrabCut 数据库^[3]。该数据库被广泛使用，包含大量真实环境中的复杂/非几何结构前景物体。图3-9 展示了部分实验结果，其中分割误差率被标注在相应结果下面。在几何结构约束下，我们的算法取得了更好的分割结果。表3-1 的第二行展示了分割结果。我们的算法取得了9.21% 的分割效果，优于传统MRF 模型的分割结果（11.39%）。值得注意的是仅使用几何约束平滑项 ($V_{u \rightarrow w}^S$) 的结果比传统MRF 模型有所下降。有可能的一个原因是在无数据惩罚项的帮助下，几何约束平滑项有可能使得对称匹配点对出现过分割，

3.5.4 与嵌入几何结构的level-set 算法比较

文献^[49] 提出了一种将几何结构约束嵌入到level-set 分割模型中的算法，其使用一种基于level-set 的弧线测度技术去获取满足几何结构的边线部分，并提供一个动态的形状先验信息。如第3.1 节所述，我们的方法不同于文献^[49]。最重要的是我们的方法侧重于“前景分割”，而不是“几何物体分割”。因此，

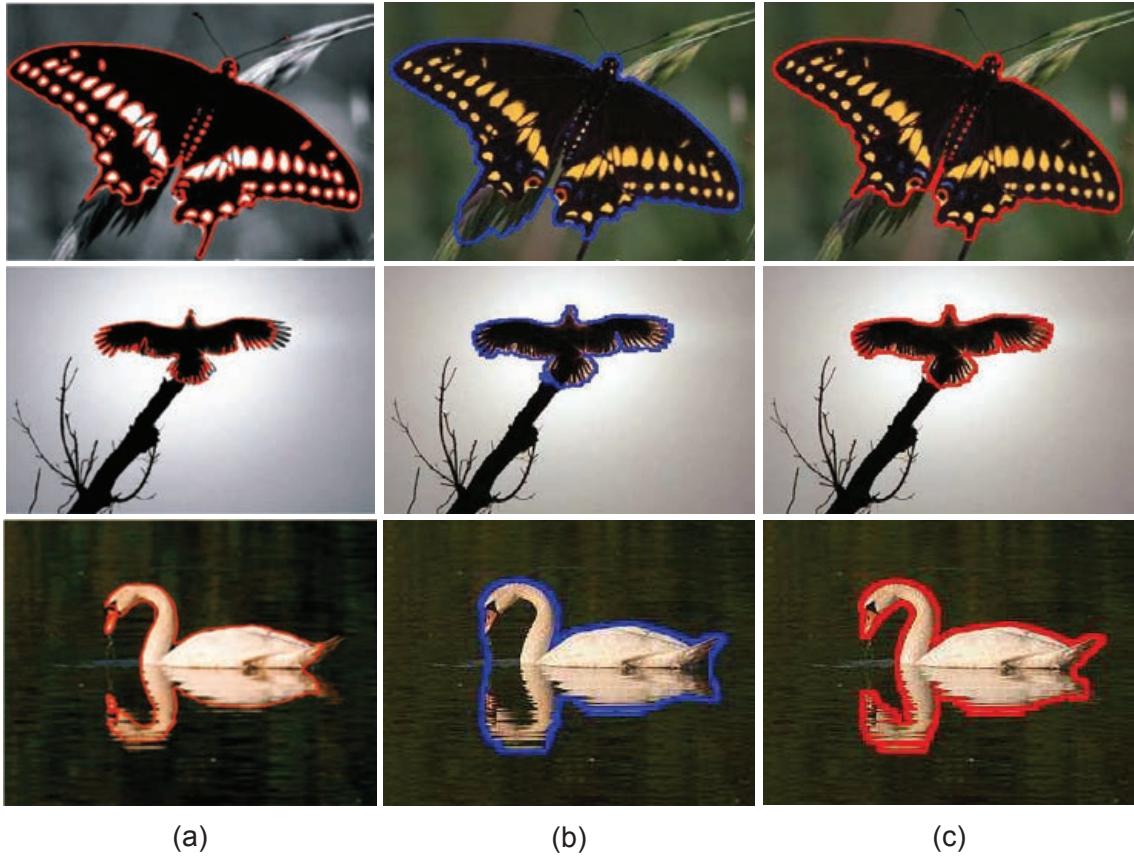


图 3-10 我们的算法与同样嵌入几何结构约束的level-set 算法比较^[49]。(a) 算法^[49]的分割结果。(b) 传统MRF 模型的分割算法。(c) 我们的几何结构约束模型的分割结果。

几何结构约束仅仅被用来作为几何/非几何物体的鼓励/惩罚系数。由于弱约束的原因，我们的算法同样可以适用于复杂/非几何结构的物体分割，这一点与文献^[49]明显不同。图3-10展示了我们的算法与文献^[49]的比较，可见我们的算法获取了更好的分割效果。例如在第一个图例中，我们的算法准确的提取出蝴蝶翅膀的花纹，而方法^[49]出现了过分割。第二章图片中，我们的算法准确的分割了鸟的翅膀，同样避免了在方法^[49]中出现的过分割。

3.5.5 算法鲁棒性讨论

在我们的方法中，几何结构检测被作为前期输入。通过公式 (3-4) 取得的几何结构约束模型对几何结构误差具有一定的鲁棒性。以对称几何结构为例，图3-11展示了一些非对称物体被错误地检测对称轴后的分割结果。可以看到由于几何结构强度因子 μ 的作用，由公式 (3-7) 计算的几何结构映射矩阵的强度将被抑制，因此几何结构约束强度降低。同时，这种弱几何约束依

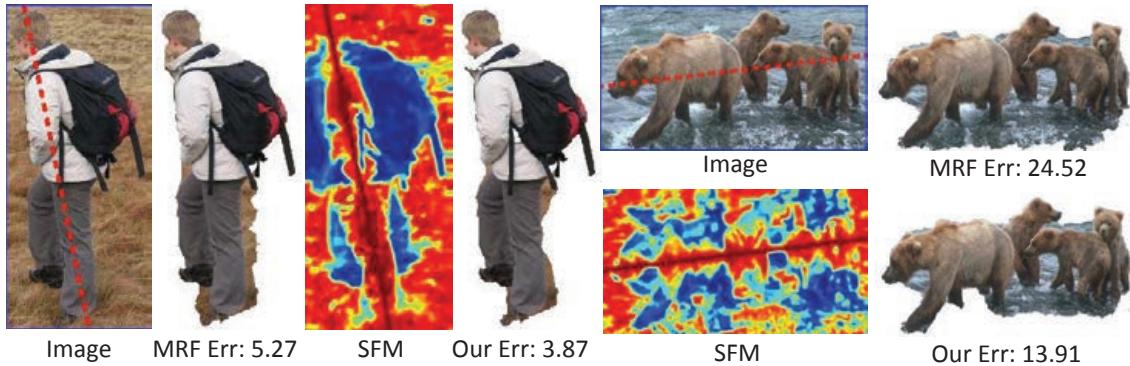


图 3-11 错误的几何结构检测将导致几何结构约束模型出现大量误匹配。红色虚线代表检测错误的对称轴, SFM 代表相应的几何结构映射矩阵。尽管几何约束被抑制了, 我们的弱约束模型依旧取得了比传统MRF 要好的分割效果, 尤其是是针对非几何结构的背景。

旧对非几何结构的背景产生作用, 相对于传统的MRF 模型提高分割效果。

3.6 本章总结

在本章中, 我们主要研究单张图像不同区域间的关联性约束, 并提出了一种基于几何结构约束的前景分割算法。我们通过四个方面体现几何结构信息: 几何结构约束分布图、几何结构强度因子、几何结构映射矩阵、以及几何结构约束模型。而且这种几何结构信息是物体自身内在属性的体现, 无需大量繁重的学习训练过程。实验表明, 我们的算法在引入高级别的几何结构信息后, 相对于传统的分割模型, 可以明显的提升低级别像素分割的准确度。

研究显示, 单张图像不同区域间的关联性约束是一种有效的效果提升方法。满足一定的结构关联的区域/目标间构建的关联结构, 体现了一种全局约束条件, 可以明显的提升目标分割的准确性和鲁棒性。

第四章 单图像相关源间的关联性约束

本章将主要研究单图像相关源间的关联性约束。数字图像往往存在许多相关源，例如拍摄设备、拍摄镜头等硬件，以及在图像处理过程中的中间环节，例如预处理图像层、检测结果分布图等。我们将主要研究这些不同处理源之间的关联性约束，建立单幅图像与镜头畸变、以及单幅图像与其不同操作处理层之间的协同关联性，并提出一种基于图像畸变度的广角图像伪造区域盲检测分割算法。镜头的径向几何畸变在绝大多数图像处理中通常被认为是需要消除的不良性质，但同时，该几何畸变也反映了镜头的内部结构属性，在图像拍摄过程中，被如实的记录反应在图像中。因此，镜头畸变可以看做是对图像提供了一个全局的几何约束。对此，我们将这种镜头径向几何畸变性质作为度量指标，引入到图像篡改盲检测算法中。首先我们提出了一种镜头径向畸变映射模型（radial distortion projection model），在该模型中，空间直线将在映射半球（viewing sphere）上投影成一个大切圆（great circle）。其次，依据该几何约束，设计了两种底层测度用于获取目标图像的伪造取证分布图（fake saliency map）。最后，通过构造基于图分割的能量函数，将目标图像与其伪造取证分布图关联起来，提取出像素级别的目标取证呈现。

4.1 相关背景工作

图4-1展示了小孔成像镜头与广角镜头的区别，其中 \mathbf{P}_w 是空间中的3D点， \mathbf{p}' 是图像成像面中基于小孔成像模型的投影点， \mathbf{p} 则是在径向畸变作用下的投影点。 f 代表焦距， z 代表光轴。 θ 和 ϕ 分别代表入射角和方位角。与投影点 \mathbf{p}' 相比广角镜头投影点 \mathbf{p} 具有明显的径向畸变。理想的小孔成像模型可以被描述为：

$$r(\theta) = f \tan(\theta), \quad (4-1)$$

其中， θ 代表入射光线 $\mathbf{P}_w\mathbf{p}'$ 与主光轴的入射角， r 代表投影点与主光轴点的径向距离。而相对来说，广角镜头存在着严重的径向畸变，因此其成像模型通常有以下几种^[125-127]：

$$\begin{aligned} r_1(\theta) &= f_1 \theta && (\text{Equidistance}), \\ r_2(\theta) &= 2f_2 \sin(\theta/2) && (\text{Equisolid angle}), \\ r_3(\theta) &= f_3 \sin(\theta) && (\text{Orthogonal}), \\ r_4(\theta) &= 2f_4 \tan(\theta/2) && (\text{Stereographic}), \end{aligned} \quad (4-2)$$

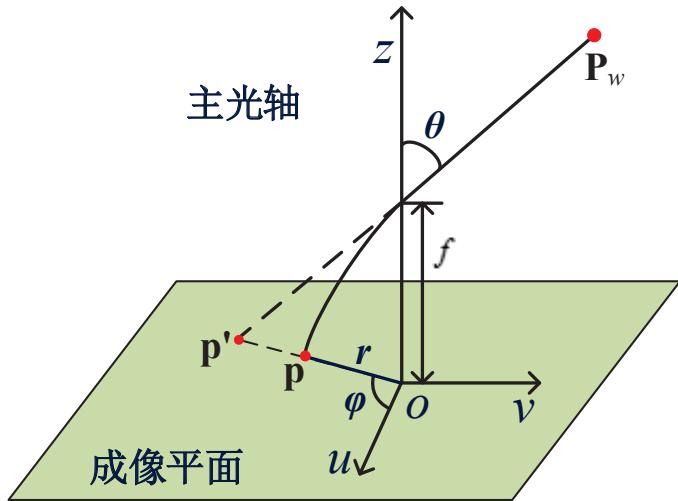


图 4-1 小孔成像镜头与广角镜头的区别, \mathbf{P}_w 是空间中的3D点, \mathbf{p}' 是图像成像面中基于小孔成像模型的投影点, \mathbf{p} 则是在径向畸变作用下的投影点。

其中, 径向畸变模型 $r(\theta) = fg(\theta)$ 也被称为发射线模型 (captured rays-based model)。一种典型的广角镜头就是鱼眼镜头 (fisheye lens), 其可以同时获取整个半球面的场景信息 (semi-spherical view)。由于接近360度的超广视角, 鱼眼镜头通常被用来研究径向畸变^[57-59]。在本章中, 我们也以鱼眼镜头为主, 其不失一般性, 而且很容易可以推广到其他广角镜头的情况。

将公式 (4-2) 的径向畸变模型规划到统一模型 (Unified model) 中, 将非常适合于图像处理分析, 尤其是在未标定 (uncalibrated) 的图像中, 我们无法获取畸变模型。在文献^[128] 中, 一种针对所有反射式镜头 (central catadioptric systems) 的统一模型被提出。在该模型中, 所有的反射式镜头都是同构的 (isomorphic), 其通过一个正交与投影平面的投影中心, 构建球面与投影平面的投影映射关系。在文献^[129] 中, 作者进一步拓展了该模型, 通过附加一个二次曲面映射, 使其适用于鱼眼镜头。在文献^[130] 中, 通过半球映射的反射镜头的统一模型被证明是等价于具有径向畸变的发射线模型, 因此该统一模型可以被直接用来作为鱼眼镜头的成像模型。然而, 这些统一模型具有两个不足: 首先, 这些统一模型包含多个参数, 这些参数无法在伪造图像中被准确标定; 其次, 这些模型无法处理在入射角接近 $\theta = \pi/2$ 的情况。因此, 我们简化了统一成像模型, 生成一种径向畸变映射模型 (Radial Distortion Projection model)。我们的模型仅使用一个参数和更简单的形式, 而且等价于公式 (4-2) 中的所有发射线模型。

广角镜头的畸变提供了一个非常重要的几何约束, 例如空间中的直线将在映射半球 (viewing sphere) 上投影成一个大圆 (great circle)^[131,132]。基于该约束, 我们将引入两种自底向上的测度, 使用它们去评测候选直线的取证置信度 (untrustworthy likelihoods)。在我们的方法中, 候选直线通过半自动化

检测提取：首先，通过边缘检测自动提取出候选弧线；然后，每一个候选直线被人工选取。手工选取候选直线是因为现有的算法无法在识别畸变情况下的空间直线。我们使用三个点去描述候选直线：两个端点、以及中心点。

4.2 镜头径向畸变映射模型

在本节，我们将介绍一种基于统一模型^[128,130]改进的镜头径向畸变映射模型（Radial Distortion Projection, RDP）。径向畸变映射模型描述了一个从三维空间到二维畸变图像点的映射关系。如图4-2所示，模型包含一个单位半球为基础，称之为映射半球（viewing sphere）。虚拟光轴中心**c**和坐标原点**o**分别称为光轴中心（optical center）和图像主投影中心（principal projection center）。光轴中心**c**坐标为 $[0, 0, -f_v]^T$ ，其中 $f_v > 0$ 代表虚拟镜头焦距（virtual focal length）。图像的成像投影过程可以被描述为：

- 步骤一：对于空间坐标系中的任意点 $\mathbf{P}_w = [X_w, Y_w, Z_w]^T$ ，其半球映射点 $\mathbf{P}_s = [x_s, y_s, z_s]^T$ 可以表示为：

$$\mathbf{P}_s = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} = \frac{\mathbf{P}_w}{\|\mathbf{P}_w\|} = \begin{bmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{bmatrix}, \quad (4-3)$$

其中， $\|\mathbf{P}_w\| = \sqrt{X_w^2 + Y_w^2 + Z_w^2}$ 和 (ϕ, θ) 分别代表半球映射点 \mathbf{P}_s 的二维球坐标， ϕ 代表方位角， θ 代表入射角。

- 步骤二：半球映射点 \mathbf{P}_s 再投影到归一化的图像成像平面，形成投影点 $\mathbf{p} = [u_p, v_p, 1]^T$ 。该映射过程是一个基于虚拟光轴中心**c**的映射过程，类似于小孔成像模型。投影点的齐次坐标系为：

$$\mathbf{p} = \mathbf{K}_s \mathbf{P}_s, \quad (4-4)$$

其中， \mathbf{K}_s 代表相机内参，具有形式： $\mathbf{K}_s = \begin{bmatrix} f_v & & \\ & f_v & \\ & & 1 \end{bmatrix}$ ，其中 f_v 为虚拟焦距。

4.2.1 反映射模型

在实际中，图像点坐标往往是唯一的输入参数，因此反映射模型（back-project formulation）更加具有实用价值。 $\mathbf{p} = [u, v, 1]^T$ 成像平面上的映射点，其中坐标原点设置为图像中心**o**。点 \mathbf{p} 的径向半径 r 可以通过 $r = \sqrt{u^2 + v^2}$ 计算。为了简单描述，我们将坐标点 \mathbf{p} 与其径向半径 r 结合，组成中间向

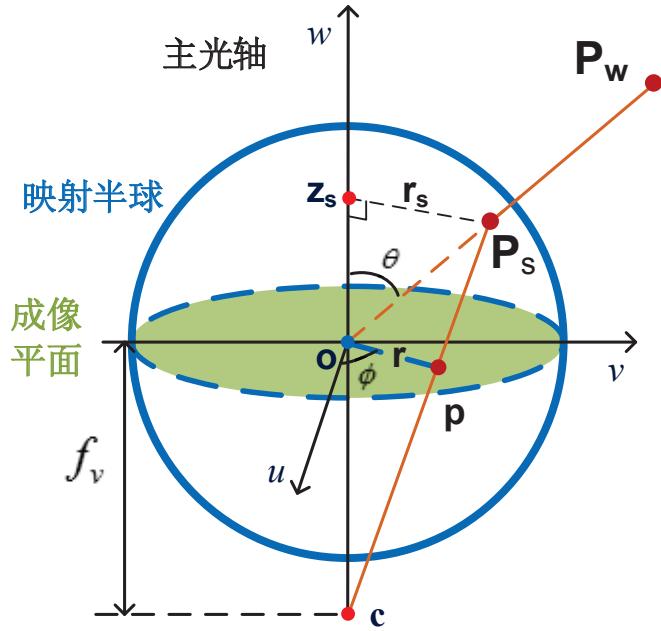


图 4-2 广角镜头的径向畸变映射模型。其中, \mathbf{P}_w 代表空间中的点, \mathbf{P}_s 代表映射半球上的半球映射点 (spherical projection point), \mathbf{p} 代表图像映射平面上的最终投影点。

量 $\tilde{\mathbf{p}} = [u, v, r]^T$ 。我们的径向畸变映射模型的反射模型可以表示为:

$$\mathbf{P}_s = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} = \begin{bmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{bmatrix} = \frac{1}{r} \begin{bmatrix} \sin(\theta) & & \\ & \sin(\theta) & \\ & & \cos(\theta) \end{bmatrix} \begin{bmatrix} u \\ v \\ r \end{bmatrix} = \frac{1}{r} \mathbf{K}_b \tilde{\mathbf{p}}, \quad (4-5)$$

其中, $\mathbf{P}_s = [x_s, y_s, z_s]^T$ 代表映射半球上的单位点向量, \mathbf{K}_b 被定义为反射矩阵。 $\phi = \arctan(v/u)$ 代表方位角, θ 代表入射角, 其可以通过下式计算:

$$\theta = \begin{cases} 2 \arctan(r), & \text{if } f_v = 1 \\ 2 \arctan\left(\frac{f_v - \sqrt{f_v^2 + r^2 - f_v^2 r^2}}{r(f_v - 1)}\right), & \text{if } f_v \neq 1 \end{cases}. \quad (4-6)$$

通过公式 (4-5), 图像坐标点可以被反射到映射半球上的单位半球坐标中。

4.2.2 径向畸变映射模型的相关性质

我们的镜头径向畸变映射模型具有一些优秀的属性:

1) 我们的径向畸变映射模型可以仅通过一个参数 f_v 等价于公式 (4-2) 中的其他发射线模型。径向畸变映射模型可以变成 $r = f_v g(\theta)$, 其通过近似三角形 $\triangle \mathbf{ocp} \simeq \triangle \mathbf{cz}_s \mathbf{P}_s$, 有:

$$\frac{r}{\sin \theta} = \frac{f_v}{f_v + \cos \theta} \Rightarrow r = \frac{f_v \sin \theta}{f_v + \cos \theta}. \quad (4-7)$$

因此，径向畸变映射模型与发射线模型的关系可以近似为：

$$\hat{f}_v = \arg \min_{f_v} \sum_{\theta \in [0, \pi/2]} \left\| f_g g(\theta) - \frac{f_v \sin \theta}{f_v + \cos \theta} \right\|^2, \quad (4-8)$$

其中， $f_g g(\theta)$ 代表公式 (4-2) 中的发射线模型。

2) 我们的径向畸变映射模型减少了传统统一模型^[128,130]中的映射平面距离参数，这使得径向畸变映射模型更侧重于畸变约束。此外，我们在入射角为 $\theta = \pi/2$ 时依旧有效，这与传统的统一模型^[129,130]具有很大的优化。在径向畸变映射模型中，入射角为 $\theta = \pi/2$ 的点将投影到图像的边缘，映射在一个半径为 $r = 1$ 的水平圆上。

3) 由于径向畸变映射模型中的映射半球投影过程，实际空间中的直线将在映射半球上投影为一个大切圆 (great circle)，进一步在成像平面上投影为一个二次曲线 (conic)。映射半球上的投影切圆可以提供一个几何约束，使得我们无需对图像进行矫正，仅在映射半球上就可以进行图像内容取证，同时，避免了尺度变换所带来的误差。为了展示尺度的变化，我们假设畸变图像被矫正成正确的平面中，例如平行于成像平面的矫正图像 (corrected image)，如图4-8。很明显，具有较大入射角 θ 的点在矫正图像中，将被投影到无穷远处 (infinite position)。因此，在成像平面中的具有大入射角的短线段也将被投影成矫正图像中的无穷长直线。这种尺度变化，很容易导致测量误差，尤其是当入射角 θ 接近 $\pi/2$ 时。相对来说，在映射半球上归一化尺度，可以有效的避免尺度变化带来的测量误差。

4) 如传统的统一模型一样，我们的径向畸变映射模型也同样可以适用于反射式镜头 (catadioptric lens)，通过设置虚拟焦距为 $f < 0$ 。此外，我们的径向畸变映射模型还可以适用于其它广角镜头。一个简单的计算虚拟焦距 f_v 的公式为：

$$f_v = r_w \frac{\cos \theta_w}{\sin \theta_w - r_w}, \quad (4-9)$$

其中， r_w 代表图像维度， θ_w 代表广角镜头的视场角。

4.3 基于直线畸变几何约束的取证测度

在本节中，我们首先描述径向畸变映射模型中的直线畸变几何约束，然后提出两种自底向上的测度，体积取证测度和距离取证测度，去判断图像直线的可靠性。

4.3.1 直线畸变几何约束

在我们的径向畸变映射模型中，实际空间中的直线在映射半球上投影

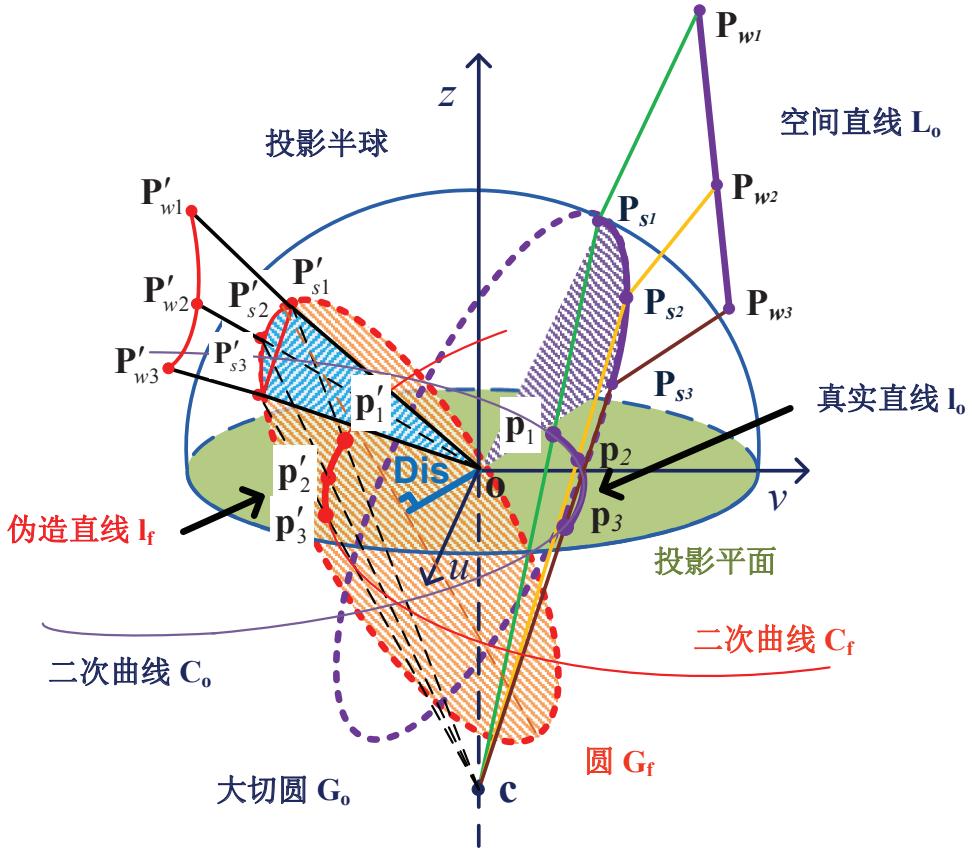


图 4-3 径向畸变映射模型的直线畸变几何约束。实际空间中的直线在映射半球上投影为一个大切圆，然后在成像平面上投影为一个二次曲线。紫色线代表直线的投影过程（真实直线），红色线代表非直线的投影过程（伪造直线）。蓝色的阴影代表体积取证测度，变量Dis 代表距离取证测度。

为一个大切圆，然后在成像平面上投影为一个二次曲线。图4-3 展示了这个映射过程。其中紫色线 L_o 代表空间中直线 $P_{w1}P_{w2}P_{w3}$ 的投影。二次曲线 C_o 可以由空间直线的投影点 (p_1, p_2, p_3) 拟合而成。投影点 (p_1, p_2, p_3) 组成了候选直线 l_o 的两个端点以及一个中心点。并且反投影到映射半球上，分别形成点 (P_{s1}, P_{s2}, P_{s3}) 。半球映射点 (P_{s1}, P_{s2}, P_{s3}) 定义了空间直线 L_o 与图像中心 \mathbf{o} 在投影半球上的大切圆 G_o 。相对来说，伪造直线上的点 (p'_1, p'_2, p'_3) 虽然也可以在成像平面中投影为一个二次曲线段 C_f 。但在映射半球上得反射射点 $(P'_{s1}, P'_{s2}, P'_{s3})$ 所确定的切圆 G_f 不是大切圆。因为伪造的直线将不再满足这个几何约束条件。换句话说，由伪造线生成的切圆确定一个切面，该切面将不再穿过映射半球的中心点，如图4-3 中的红色虚线圆所示。由于该直线约束的存在，我们可以提出两种自底向上的测度，体积取证测度和距离取证测度，用来检测图像中的伪造直线。

4.3.2 体积取证测度

基于几何约束，如果成像平面上的三个投影点($\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$)是来源于空间直线的话，那么它们在映射半球上的反射投影点($\mathbf{P}_{s_1}, \mathbf{P}_{s_2}, \mathbf{P}_{s_3}$)将满足几何约束：

$$Volume_{\Delta} = \det(\mathbf{P}_{s_1} \mathbf{P}_{s_2} \mathbf{P}_{s_3}) = 0, \quad (4-10)$$

其中， $Volume_{\Delta}$ 代表由映射半球中心点 \mathbf{o} 以及大切圆 \mathbf{G}_o 上的点($\mathbf{P}_{s_1}, \mathbf{P}_{s_2}, \mathbf{P}_{s_3}$)组成的四面体的体积，如图4-3 中的紫色阴影所示。通常来说，伪造的直线无法满足这个几何约束，如图4-3 中蓝色阴影所示。因此，我们可以将体积作为一个自底向上测度，即体积取证测度 w^V ：

$$w^V = \frac{1}{\|\mathbf{l}_f\|^3} \left\| \det(\mathbf{P}'_{s_1} \mathbf{P}'_{s_2} \mathbf{P}'_{s_3}) \right\|, \quad (4-11)$$

其中， $\|\mathbf{l}_f\|$ 代表伪造直线 \mathbf{l}_f 在图像中的长度。图4-4 (b) 展示了每一个候选直线的在不同的虚拟焦距下的几何测度值。可以看到真实直线的体积取证测度接近于0，而伪造直线的体积取证测度则远离于0。同时在实验中，我们也发现，当伪造直线远离图像中心的时候，其几何测度无法将其鉴别出来，例如图4-4 (b) 中最右边的候选直线。注意到，当虚拟焦距为 $f = 0$ 时，我们的径向畸变映射模型将退化，所有的畸变值降为0，这是由于在第4.2.1 节的步骤二中的映射变换将无法使用，公式 (4-6) 中的像素点入射角将同质化。

4.3.3 距离取证测度

如同体积取证测度一样，由映射半球面上的三个投影点 \mathbf{P}_{s_i} 所确定的切面，与中心点 \mathbf{o} 之间的距离也可以作为一个取证测度。映射半球上的三个点 \mathbf{P}_i 确定一个切面 Π ，空间中任意一点与平面可以组成矩阵 $\mathbf{M} = [\mathbf{P}, \mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3]$ 。对于平面 Π 上的任意点，可以存在 $\det(\mathbf{M}) = 0$ ，我们可以计算平面参数^[116]：

$$\Pi = (D_{234}, -D_{134}, D_{124}, -D_{123})^T, \quad (4-12)$$

其中， D_{jkl} 由矩阵 $[\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3]$ 的第 jkl 行组成。在矩阵中， \mathbf{P}_i 通过其次坐标 $\mathbf{P}_i = [x_i, y_i, z_i, 1]^T$ 描述，因此矩阵 $[\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3]$ 是一个 4×3 的矩阵。平面 Π 与空间中任意一点 $\mathbf{P}_0 = [x_0, y_0, z_0]^T$ 的空间距离可以计算为：

$$Dis = \frac{|x_0 D_{234} - y_0 D_{134} + z_0 D_{124} - D_{123}|}{\sqrt{D_{234}^2 - D_{134}^2 + D_{124}^2}}. \quad (4-13)$$

如上文描述，空间中直线将在映射半球上形成一个大切圆 \mathbf{G}_L ，因此映射半球中心 \mathbf{o} 将穿过平面，即距离为 $Dis = 0$ ，如图4-3 中的紫色虚线。相对的，伪造直线的反射点无法生成大切圆，导致距离 $Dis > 0$ ，如图4-3 红色阴影线。因此我们可以将切面与中心的距离定义为一个取证测度，即距离取证测

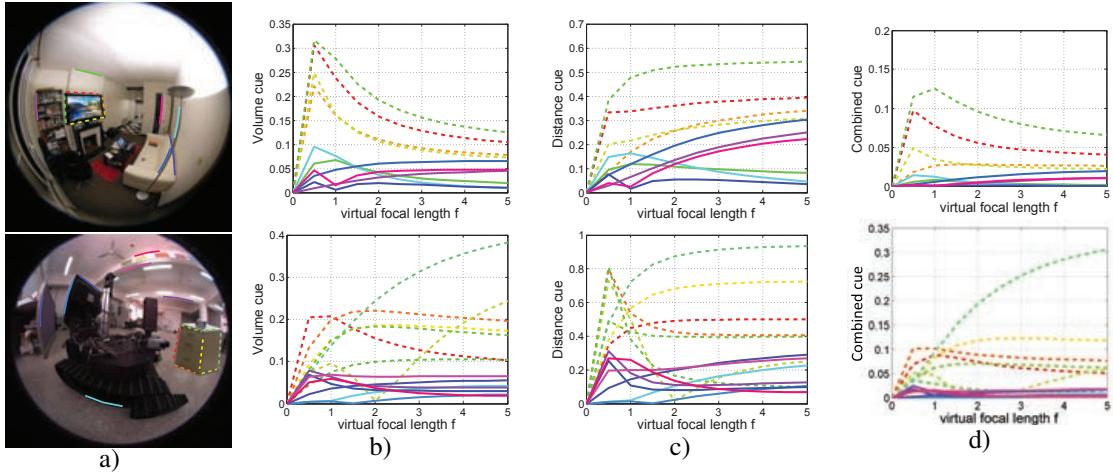


图4-4 候选直线在不同虚拟焦距下对应的畸变测度。(a)为输入图像，其中候选直线用不同的颜色标记出来。伪造的直线使用虚线，真实的直线使用实线标记。(b-d)展示了体积取证测度 w^V 、距离取证测度 w^D 、以及综合测度 w^C 在不同的虚拟焦距下的曲线图。

度 w^D :

$$w^D = \frac{|D_{123}|}{\sqrt{D_{234}^2 + D_{134}^2 + D_{124}^2}}. \quad (4-14)$$

由于映射半球的中心点**o**位于坐标原点 $[0, 0, 0]^T$ 上，因此公式(4-13)中分子的前三项将被移除。图4-4 (c)展示了候选直线在不同的虚拟焦距中的距离取证测度值。相比与体积取证测度，距离取证测度更加适合于远离图像中心的候选直线。但是在实验中我们也发现，距离取证测度容易受噪声干扰，而且相对于几何取证测度具有较低的准确率。

4.3.4 测度取证融合

为了获取正好的取证结果，我们将两种取证测度进行融合，生成一种简单的融合取证测度 w^C :

$$w^C = \log(w^V w^D + 1), \quad (4-15)$$

图4-4 (d)展示了混合测度针对不同的虚拟焦距的检测曲线，其中可以看到伪造直线与真实直线被明显区分开了。

4.4 伪造取证分布图

由于伪造的物体存在噪声以及遮挡等情况，使得伪造物体的分割提取

是一个常见问题。在本节中，我们引入伪造取证分布图的概念去分割检测结果。首先，取证非置信度被用来生成初始的伪造取证候选点，然后生成伪造取证分布图，其描述伪造物体的位置与区域。最终，伪造物体通过分割函数分割，展现像素级别的检测结果呈现。我们的伪造取证分布图并不局限任何伪造度检测算法，在本节中，我们将继续延续基于畸变测度的检测方法。

4.4.1 取证非置信度检测

通常来说，阈值（Thresholding）是一种常见的判别指标，然而，其容易导致阈值附近分布的候选直线的误判。因此，我们使用一种自适应伪造值来进行候选直线判断，我们称之为取证非置信度（untrustworthy likelihood）。首先，我们将生成一个模拟测试数据集，其中包含40000个模拟真实直线，以及40000个虚拟模拟伪造直线。我们将在后面的实验中（第4.5.2节）再一次详细介绍该数据集。我们统计每个候选直线的取证测度值分布直方图，如图4-5所示。我们使用 $w_i, i = 1 \dots N$ 代表取证测度的bin，其中 N 代表bin的总数目（如图4-5中， $N = 10$ ）。由于在实际测量中，真是直线的取证测度值基本都低于0.01，因此我们将统计直方图的范围限制在[0, 0.01]以内。同时，对于取证测度值 $w^C > 0.01$ 将全部放置于最后一个bin w_N 中。基于统计直方图，畸变取证测度的分布曲线可以计算为：

$$p(x|w_i) = \frac{N_O(i)}{N_O(i) + N_F(i)}. \quad (4-16)$$

其中， $p(x|w_i)$ 代表 x 基于 w_i 的先验概率值， $N_O(i)$ ($N_F(i)$) 代表真实/伪造直线在bin i 中的数目。图4-5中的绿色曲线描述了取证测度的分布情况。

为了简便起见，我们使用高斯核（Gaussian kernel）去拟合这个分布曲线，

$$G(x|\alpha, \beta) = \alpha \exp\left(-\frac{x^2}{2\beta^2}\right) \quad (4-17)$$

其中， α 可以通过数据集中的 $\hat{\alpha} = p(x|w_i = 0)$ 推导出。参数 β 可以通过最小二乘拟（least square error）合计算出来：

$$\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^N (p(x|w_i) - G(x|\hat{\alpha}, \beta))^2. \quad (4-18)$$

在确定了参数 $\hat{\alpha}$ 和 $\hat{\beta}$ 后，我们就可以判断任何待测曲线的预测值。在图4-5中的黑色曲线是通过参数 $\hat{\alpha} = 0.89$ 和 $\hat{\beta} = 0.2$ 拟合而成。

最后，我们可以定义取证非置信度 U 为：

$$U = 1 - \hat{\alpha} \exp\left(-\frac{w^2}{2\hat{\beta}^2}\right), \quad (4-19)$$

其中， w 代表畸变取证测度。很明显，伪造直线的取证非置信度要高于真实图像。我们使用取证非置信度来获取初始的取证分布图，计算每一个候选直

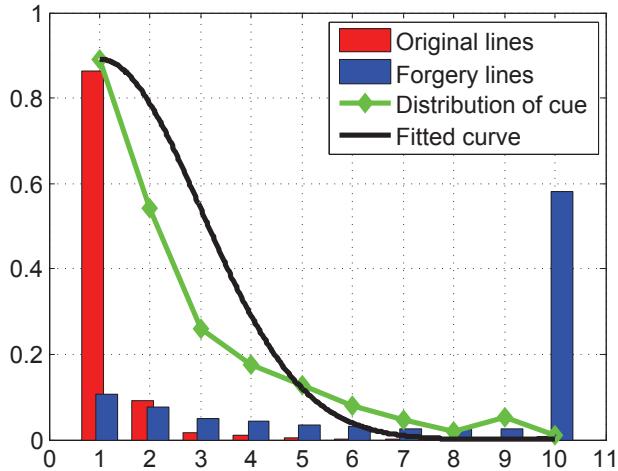


图 4-5 在虚拟数据库中的取证测度统计直方图，其中绿色曲线代表取证非置信度 bin，黑色曲线代表公式 (4-18) 的拟合结果。x 轴代表区间数目，y 轴代表分布频率。红色/蓝色柱状条代表总体的真实/伪造直线的相应数目比例。

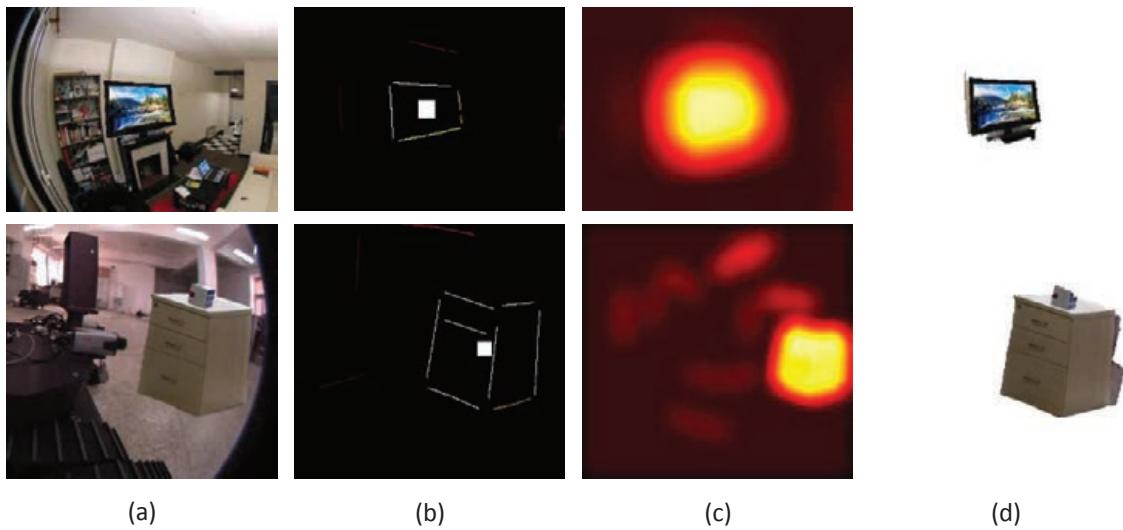


图 4-6 伪造取证分布图的生成过程。(a) 图1-5 (b) 放大的伪造区域。(b) 初始的候选直线检测结果，热度图描述了相应的取证非置信度。(c) 对应的伪造取证分布图。(d) 最终的分割结果呈现。

线的伪造取证度，如图4-6 (b) 所示。

4.4.2 取证分布图生成

我们的伪造性检测目标是获取像素级别的检测结果呈现。已存的分割方法大多数是基于图像的颜色、纹理、边缘来分割物体^[133,134]。可这些方法并

不适合我们的伪造取证工作，主要因为两个原因：第一，检测到的伪造直线不一定是物体的边缘；第二，我们仅能获取待测直线的非置信度测量值，而无法判断直线的哪一侧是伪造区域。然而，从统计的观点来看，位于伪造直线中间的区域更有可能是伪造物体。基于此点，我们将借助视觉显著性的观点，视觉显著性主要指生物视觉注意力机制中第一眼观测目标^[18,19,24]。所不同的是，我们侧重于伪造物体的检测。

对于图像 \mathbf{I} 来说， $\mathbf{L}^* = [l_i]$ 代表候选直线集合， U_i 代表候选直线 l_i 的取证非置信度。为了更好的覆盖伪造物体，我们同时计算了高非置信度值的中心区域作为基点。使用 $\mathbf{G}(\mathbf{c}_g, r_g)$ 代表基点中心，其中 \mathbf{c}_g 代表基点中心位置， r_g 代表基点中心半径。基点中心可计算为：

$$\mathbf{c}_g = \frac{1}{\|\mathbf{L}^*\|} \sum_{l_i \in \mathbf{L}^*} l_i U_i. \quad (4-20)$$

图4-6 (b) 展示了候选直线与基点中心所构成的初始取证分布图，其中热图值代表了相应的取证非置信度测量值。然后对于图像中的每一个点 $x_k \in \mathbf{I}$ ，我们可以计算相应的取证权重值 S_k ：

$$S_k = \frac{1}{\|\mathbf{L}^* \cap \mathbf{G}\|} \sum_{x_i \in \mathbf{L}^* \cap \mathbf{G}} U_i \exp\left(-\frac{\text{dis}(x_k, x_i)}{\gamma^2}\right), \quad (4-21)$$

其中， \mathbf{G} 代表通过公式 (4-20) 计算的基点中心， $\text{dis}(x_k, x_j)$ 代表像素点 x_k 和 x_j 之间的欧氏距离。 γ 代表尺度参数，在实验中我们设置为 $\text{Radius}/40$ ， Radius 代表图像的尺度半径。图4-6 (c) 展示了我们的取证分布图 $\mathbf{S} = [S_k]$ 。

4.4.3 检测结果分割呈现

在本节中，我们将使用二值分割函数去提取伪造物体。我们将分割问题规划为一个最大流/最小割的图分割问题。分割能量函数 E 定义为：

$$E(x_i) = \sum_{i \in I} S_i(x_i) + \sum_{\{p, q\} \in \mathcal{N}} V_{p,q}(x_p, x_q), \quad (4-22)$$

其中， $x = \{x_i | i \in I\}$ 代表图像 I 的二值分割label，对于二值分割来说 $x_i \in \{0, 1\}$ ，0 代表原始数据，1 代表伪造数据。 S_i 代表像素 i 通过公式 (4-21) 计算的取证权重值。 $V_{p,q}$ 代表平滑项 (smoothness penalty)， \mathcal{N} 代表邻接像素集合。平滑项的权重 $V_{p,q}$ 定义为^[3]：

$$V_{p,q} = \frac{\lambda \exp(-\beta \|z_p - z_q\|^2)}{\text{dis}(p, q)}, \quad (4-23)$$

其中， $\text{dis}(,)$ 代表图像中任意两个像素点之间的欧氏距离。 $\beta = (2 \langle (z_p - z_q)^2 \rangle)^{-1}$ ，其中 z_x 代表像素 x 的RGB 颜色值。 $\langle \cdot \rangle$ 代表图像期望， λ 代表参数。

最终分割结果可以通过最小化能量函数求解:

$$\mathbf{x}^* = \arg \min E(x_i). \quad (4-24)$$

该函数被称为马尔科夫随机场 (Markov Random Fields)，其满足子模性，可以通过图分割 (graph cuts) 算法求解^[13,117]。图4-6 (d) 展示了使用我们的分割函数提取的伪造物体呈现。

4.5 实验结果分析

我们主要分三部分进行实验。首先评估我们的径向畸变映射模型，然后在模拟数据和真实图像上测试我们的算法。在模拟数据测试中，我们主要测试各个取证测度的作用，以及其对应噪声的鲁棒性。在真实图像测试中，我们将在模拟图像和视觉图像中进行算法评测。最后，我们将讨论取证分布图在其他图像检测取证算法中的应用。除此之外，我们还将进一步谈论算法的退化情况。除非特别注释，在实验中，我们默认图像的几何中心即为光轴中心，在径向畸变映射模型中的虚拟焦距设为 $f = 1$ ，测试图像的分辨率为 800×800 。

4.5.1 径向畸变映射模型实验

如第4.2.2 节所述，我们的径向畸变映射模型等价于其它的发射线，模型。在实验中我们假设镜头的视场角为 $\pi/2$ ，投影点映射到归一化的成像平面中，因此参数 f_g 可以通过 $f_g = 1/g(\pi/2)$ 计算。图4-7 展示了我们的径向畸变映射模型与不同的传统发射线模型的关联，可以看到通过不同的虚拟焦距 f_v ，我们的径向畸变映射模型可以很好地拟合其他传统发射线模型。图4-8 展示了一些基于我们的径向畸变映射模型的图像矫正示例。注意到其中最后两幅图像是拍摄于广角镜头，可见我们的畸变模型并不受视场角的影响。

4.5.2 模拟实验

在模拟实验中，我们评估我们方法中的各个测度针对不同因素的影响，这些因素包括：图像投影中心偏移、模型虚拟焦距、以及噪声鲁棒性。模拟数据的生成方式为：首先随机生成一组空间直线。对于单位半径畸变图像 ($R = 1$)，空间直线的长度限制在 $[0.2, 0.8]$ 中，以及方位角范围为 $[0, \pi]$ 。然后，将这些模拟直线通过传统的统一映射模型，投影到成像平面。这些直线可以作为真实直线数据。将这些直线再进行随机移动，形成伪造直线数据。此外，针对公式 (4-2) 中的每一种传统映射模型，我们都生成10000 条真实直线和10000 条伪造直线，因此，数据库共包括40000 条真实直线和40000 条伪造

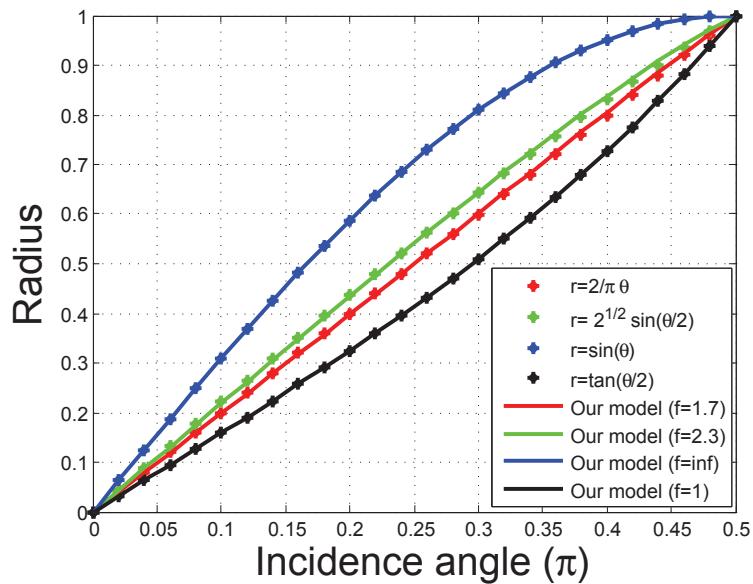


图 4-7 我们的径向畸变映射模型与传统的发射线模型的关联。星点代表公式 (4-2) 中的发射线模型曲线，实线曲线代表我们的径向畸变映射模型在不同的虚拟焦距下的拟合曲线 f_v 。



图 4-8 基于我们的径向畸变映射模型的图像矫正，其中第一行维畸变图像，第二行为矫正后的图像。

直线。

4.5.2.1 图像投影中心偏移影响

对于未标定的图像来说，畸变中心通常被默认为畸变图像的几何中心点，在实际操作中，这种默认并不影响准确度^[57]。在模拟实验中，我们首先评估图像投影中心偏移对算法中不同的取证测度的影响。我们通过加入高斯噪声，来模拟图像中心偏移情况。高斯噪声被设置为零均值(zero-mean)，以及方差 δ 从0变化到120像素。图4-9 (a-c) 分别展示了体积取证测度、距离取证测度、以及融合测度针对不同的图像投影中心偏移(fisheye center bias) 的影响。在实验中，误差bar的标准差被划分为两部分，大于均值与小于均值。红色的曲线代表真实直线的测试结果，蓝色曲线代表伪造直线的测试结果。注意到真实直线的测度并没有接近于0，这主要是因为模拟直线中包含四种广角镜头模型，而实验中，我们的虚拟焦距是设定为1的，因此导致的了误差存在。虽然各种测度值随着图像投影中心偏移的增大而增加，但它们依然保持了真实直线与伪造直线之间的区分度。

为了更好的评估各个取证测度的检测能力，我们提出了一种新的测量指标区分测度Score，其定义为：

$$Score = \frac{|mean(w_f) - mean(w_o)|}{mean(w_f)}, \quad (4-25)$$

其中， $mean(w)$ 代表取证测度 w 的均值， w_f 和 w_o 分别代表伪造直线与真实直线的取证测度。区分测度Score 的值越高，意味着该测度中伪造直线与真是直线的差距越大。图4-9 (d) 展示了各个测度针对图像投影中心偏移的区分测度，其中混合测度取得了最好的分区测度结果。我们发现区分测度的曲线并没有随着图像投影中心偏移的增大而线性下降。其主要的原因是区分测度描述的是真实直线与伪造直线取证测度之间的区分程度。

4.5.2.2 虚拟焦距影响

另一个径向畸变映射模型的参数是虚拟焦距 f 。在第4.2.2 节中，我们证明了径向畸变模型仅通过虚拟焦距 f ，就可以等价于其他传统的统一成像模型。对于图像盲检测来说，拍摄镜头是未知的，因此通常需要固定一个虚拟焦距。因此，在这个实验中我们将评测不同的虚拟焦距值对取证测度结果的影响。图4-10 (a-c) 展示了实验结果。由于当虚拟焦距 $f > 6$ 时，结果曲线就趋于稳定了，所以我们重点分析值域在 $f \in [0, 6]$ 的范围。如第4.3.2 节所述，当虚拟焦距为 $f = 0$ 时，所有的测度值都等于零。实验显示针对不同的虚拟焦距，我们的这些自底向上测度都可以从候选直线中，清楚的区分伪造的直线。对于真实直线数据集，距离取证测度的波动大于其他两种测度，但是仍然保持在0.2 以下。图4-10 (d) 展示了区分测度Score的结果。实验中，混合测度依旧保持最好的实验结果，而且受虚拟焦距的影响很小，稳定维持在0.97 以上，尤其是在虚拟焦距为 $f = 1$ 时，其取得了最高成绩Score = 0.99。因此，

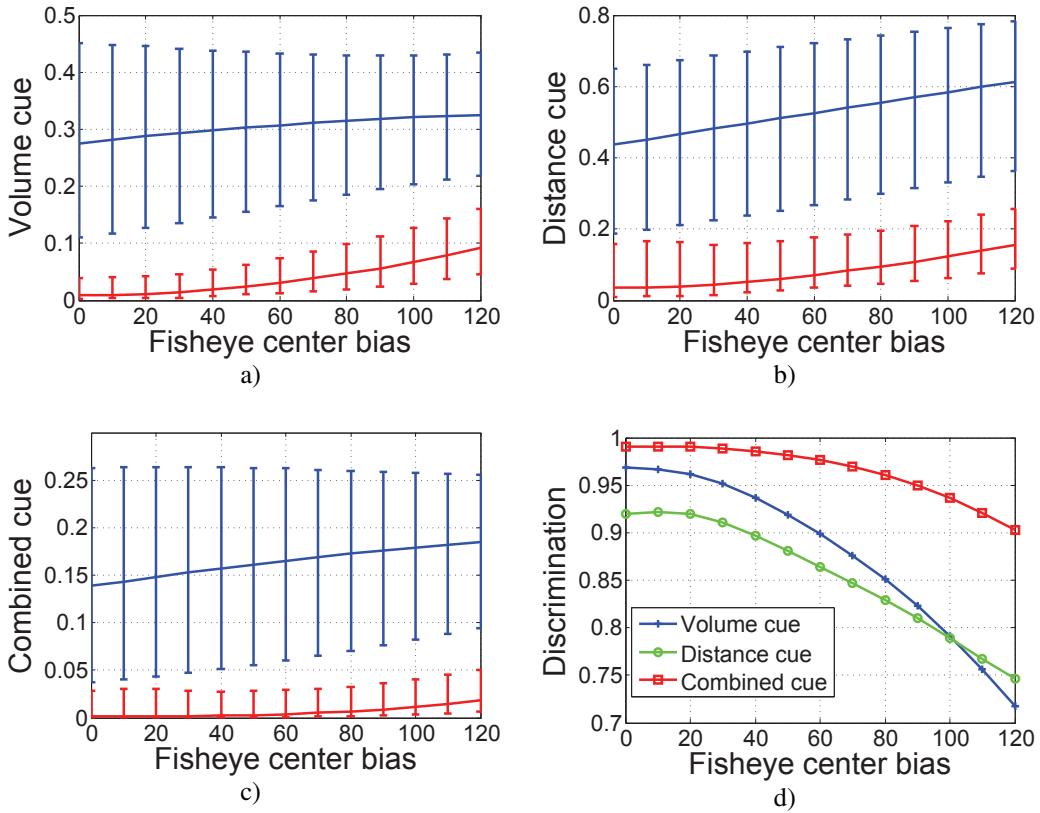


图 4-9 (a-c) 体积取证测度、距离取证测度、以及融合测度针对不同的图像投影中心偏移 (fisheye center bias) 的影响，其中红色的曲线代表真实直线的测试结果，蓝色曲线代表伪造直线的测试结果。(d) 每一个取证测度针对图像投影中心偏移的区分测度 Score，其中蓝色、绿色、红色曲线分别代表体积取证测度、距离取证测度、以及融合测度。

对于通常的情况，我们都将虚拟焦距设定为 $f = 1$ 。

4.5.2.3 噪声鲁棒性

在最后一个模拟测试中，我们将测试各种取证测度对噪声的鲁棒性。在我们的方法中，候选直线通过三个点来表达，因此我们加入高斯噪声来测试鲁棒性，噪声程度范围从0到4像素。图4-11 (a) 展示了体积取证测度对噪声的鲁棒性，其中随着噪声程度的增加，真实直线的变化由0.0085提升到0.1204。伪造直线与真实直线的在少于1.5个噪声以下，保持足够的区分度。图4-11 (b) 展示了距离取证测度对噪声的鲁棒性，显然其真实直线的测度值受噪声的干扰更大。当噪声大于1个像素时，距离取证测度下的真实直线和伪造直线将存在一定的混淆。可见，距离取证测度对噪声的鲁棒性较差。图4-11 (c) 展示了混合取证测度对噪声的鲁棒性，受距离取证测度的影响，在噪声程度为 $\delta = 1$ 时，也出现了混淆。图4-11 (d) 展示了取证测度针对不同

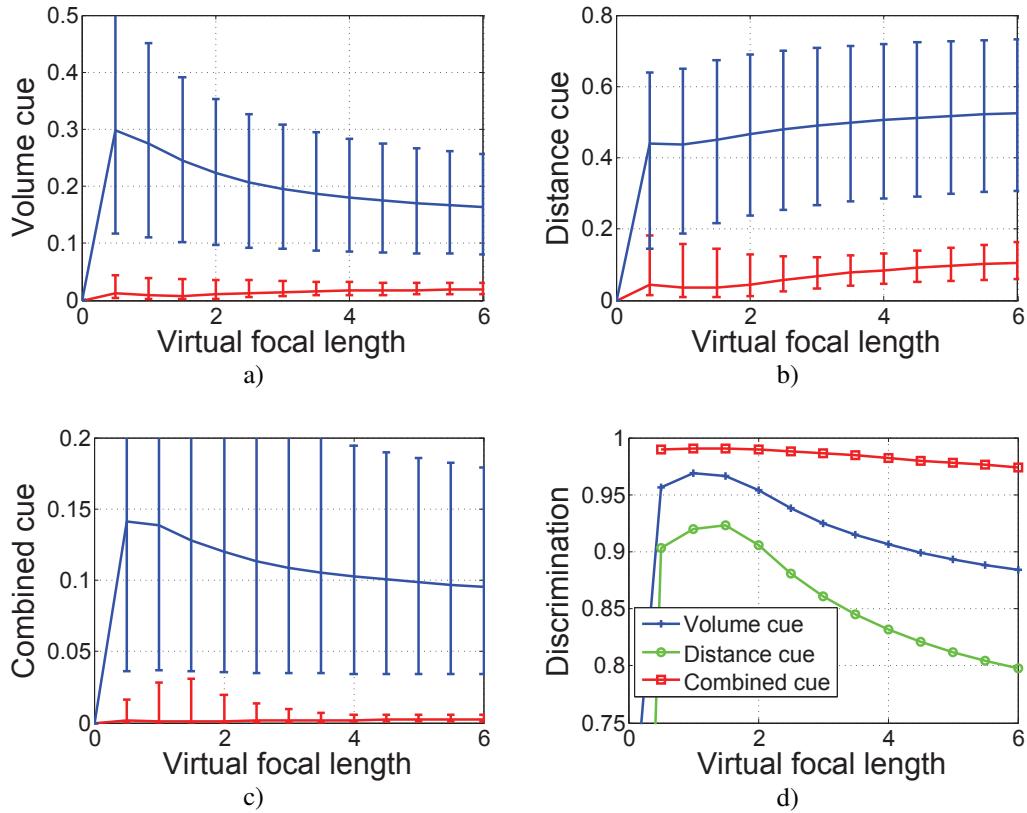


图 4-10 (a-c) 体积取证测度、距离取证测度、以及融合测度针对不同的虚拟焦距 (virtual focal length) 的影响, 其中红色的曲线代表真实直线的测试结果, 蓝色曲线代表伪造直线的测试结果。(d) 每一个取证测度针对虚拟焦距的区分测度 Score, 其中蓝色、绿色、红色曲线分别代表体积取证测度、距离取证测度、以及融合测度。

噪声程度的区分测度 Score, 其中混合取证测度依旧保持了最佳的区分结果。

4.5.3 真实图像测试

在实验中, 我们将建立一个真实图像测试数据库。图4-12 展示了图像的伪造编辑流程。首选从网络上收集原始的真实图像, 然后将伪造目标复制到原始图像中。在图像篡改过程中, 还随机的加入一些尺度变换、旋转变换、颜色均衡等图像处理操作。我们使用11张原始图片, 配合13个目标物体进行随机生成伪造的图像数据。数据库一共包含100 张图片, 标记出1582 条候选直线, 其中包括648 条伪造直线。

在图像的取证过程中, 我们分别输出生成候选直线取证测度 (图4-12 (e))、取证分布图 (图4-12 (f))、以及最终的取证结果呈现 (图4-12 (g))。其中在候选直线取证测度判定中, 我们将混合测度值为 $w^C > 0.02$ 的候选直线判定为伪造直线, 即对应的取证非置信度测量值约为 $U \approx 0.45$ 。图4-13 展示

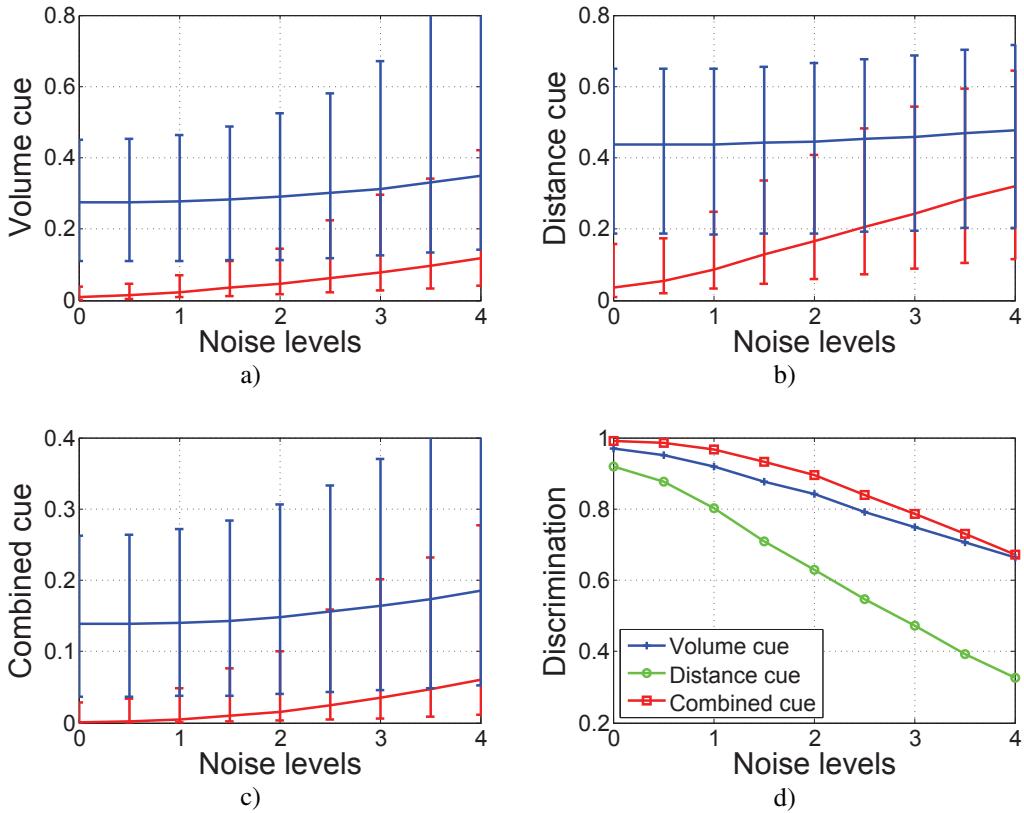


图 4-11 (a-c) 体积取证测度、距离取证测度、以及融合测度针对不同的噪声程度 (noise levels) 的鲁棒性，其中红色的曲线代表真实直线的测试结果，蓝色曲线代表伪造直线的测试结果。(d) 每一个取证测度针对不同噪声程度的区分测度 Score，其中蓝色、绿色、红色曲线分别代表体积取证测度、距离取证测度、以及融合测度。

了在真实图像测试库上的查准率/查全率曲线，可见三种取证测度的结果非常接近，混合测度在F评测指标上略微高出其它两种测度，到达0.77。

在图4-12 (e) 中可以看到一些伪造直线被漏检测，如在伪造物体上的绿色通道。相对来说，在图4-12 (f) 中的取证分布图取得了更好的检测效果，大多伪造物体被准确的检测定位出来。再经过能量方程 (4-22) 的分割，检测结果被很好地呈现出来，如图4-12 (g) 所示。图4-14 展示了更多的伪造图像内容取证结果。

4.5.4 伪造取证分布图的应用

如第4.4.2 节描述，我们的伪造取证分布图不仅仅局限于基于径向畸变的图像取证，同时也可以应用于其他取证方法。在本实验中，我们将其应用在另外两种图像取证方法中，以提高其取证分割精度。第一个方法是基于平面单映矩阵约束的伪造取证方法^[75]，该方法通过构建多视角中的平面单映矩阵

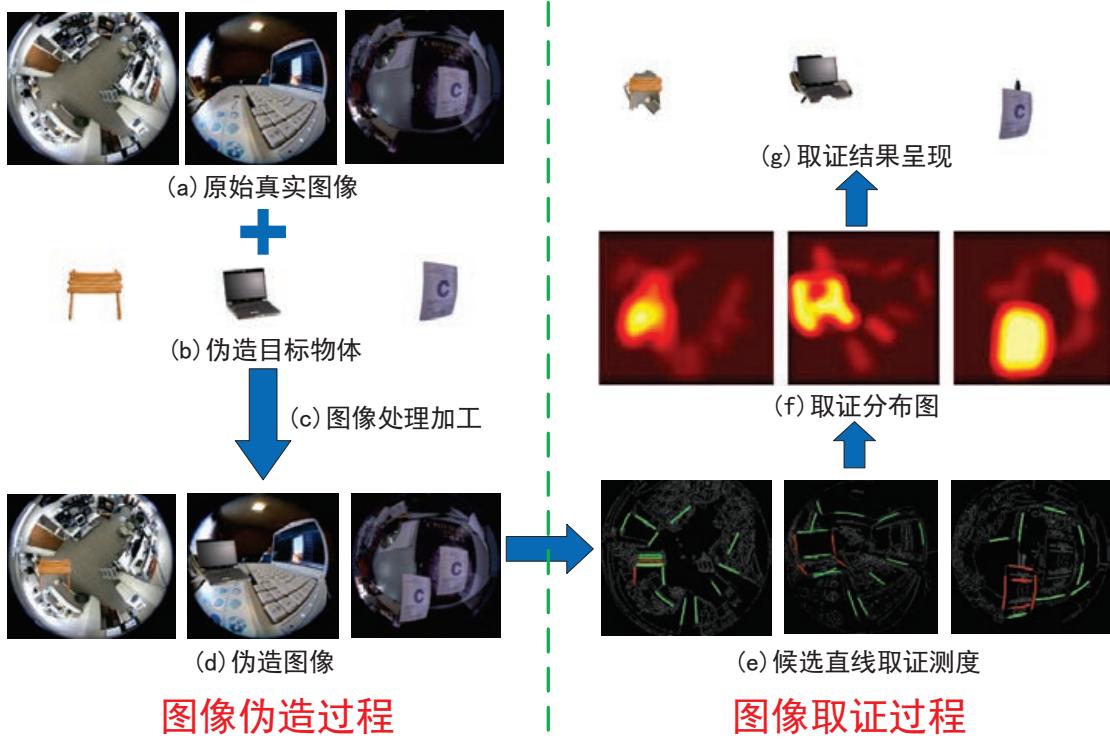


图 4-12 基于真实图片的取证算法测试流程图。在图像伪造过程中，使用原始真实图像（a）融合伪造的目标物体（b）生成伪造图像（d）。同时，在图像融合过程中，进行尺度变换、旋转变换、颜色均衡等图像处理加工，使伪造图像更加自然。在图像取证过程中，我们将生成候选直线取证测度（e）、取证分布图（f）、以及最终的取证结果呈现（g）。其中，在候选直线取证测度（e）中，混合取证测度值 $w^C > 0.02$ 的直线用红色标出作为初始检测结果。

约束，定位伪造区域，并通过在线特征选取（online feature/parameter selection），进行结果分割。在这里，我们依旧保持平面单映矩阵约束，并以此来构建伪造取证分布图的初始种子点 U ，然后使用基于取证分布图的能量函数（4-22）取代在线特征选取，进行检测结果的目标分割。图4-15展示了分割结果。可以看出原始的方法^[75]容易导致欠分割（under-segmentation），如前两个图例，以及过分割（over-segmentation），如后两个图例。而我们的方法则取得了较好的结果。

另一个常见的伪造方法是区域复制^[135-137]，其通过简单的粘贴复制将目标编辑到图像中的其他位置。我们首先使用SIFT 点配合RANSAC 算法去生成伪造取证分布图的初始种子点 U ，然后使用伪造取证分布图去提供分割模板，分割取证结果。图4-16展示了基于SIFT 复制伪造检测的图像取证结果，我们的算法可以得到一个像素级别的取证结果呈现。

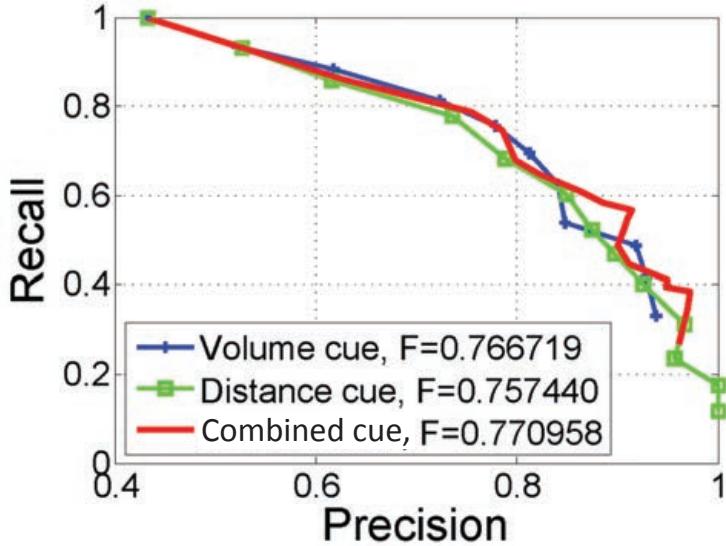


图 4-13 在真实图像测试库上的查准率/查全率曲线，其中蓝色、绿色、红色曲线分别代表体积取证测度、距离取证测度、以及融合测度。

4.5.5 相关失败情况

本节我们将讨论一些失败的情况。第一种失败情况是伪造直线的漏检。因为我们是基于几何约束来进行检测取证，当伪造的直线满足该几何约束时，算法将失效。而且一个典型的情况即是伪造直线穿过图像中心。图4-17 (a) 展示了这种情况，伪造物体上的某个直线，满足几何约束，导致无法被直接检测出来。但是，通过其它部分的伪造直线检测，我们依旧可以通过伪造取证分布图定位伪造物体。

另一种失败情况存在于分割阶段。由于我们的算法主要是基于伪造检测种子点，例如本文中的检测直线，以及上一节中的SIFT 点。当前景目标过大时，我们的伪造取证分布图很难覆盖物体，因此导致分割失败，如图4-17 (b) 所示。我们算法仅能提取出伪造直线附近的区域，而无法分割出整个物体。

最后一种退化现象是弱镜头畸变 (weakly distortion)。因为我们的方法是基于几何畸变约束，当直线接近镜头中心时，尤其是穿过镜头中心，其径向畸变将减弱，导致我们的算法退化。为了更好地描述这种情况，我们进行一个模拟直线评估实验。模拟直线评估实验设置方式如图4-18 (a) 所示，畸变图像中的虚拟伪造直线 L 具有两个参数：直线长度 l 、以及直线与图像中心 o 的垂直间距 r 。我们设定直线的无方向角以及任何径向畸变，这样，只有当直线穿过图像中心时，直线才有可能为真，其他情况，直线肯定为伪造的。当垂直间距 r 增大时，镜头畸变越强烈，导致直线的取证测度将远离真实直线的取证测度，即远离零值。直线长度 l 和垂直间距 r 的范围分别设

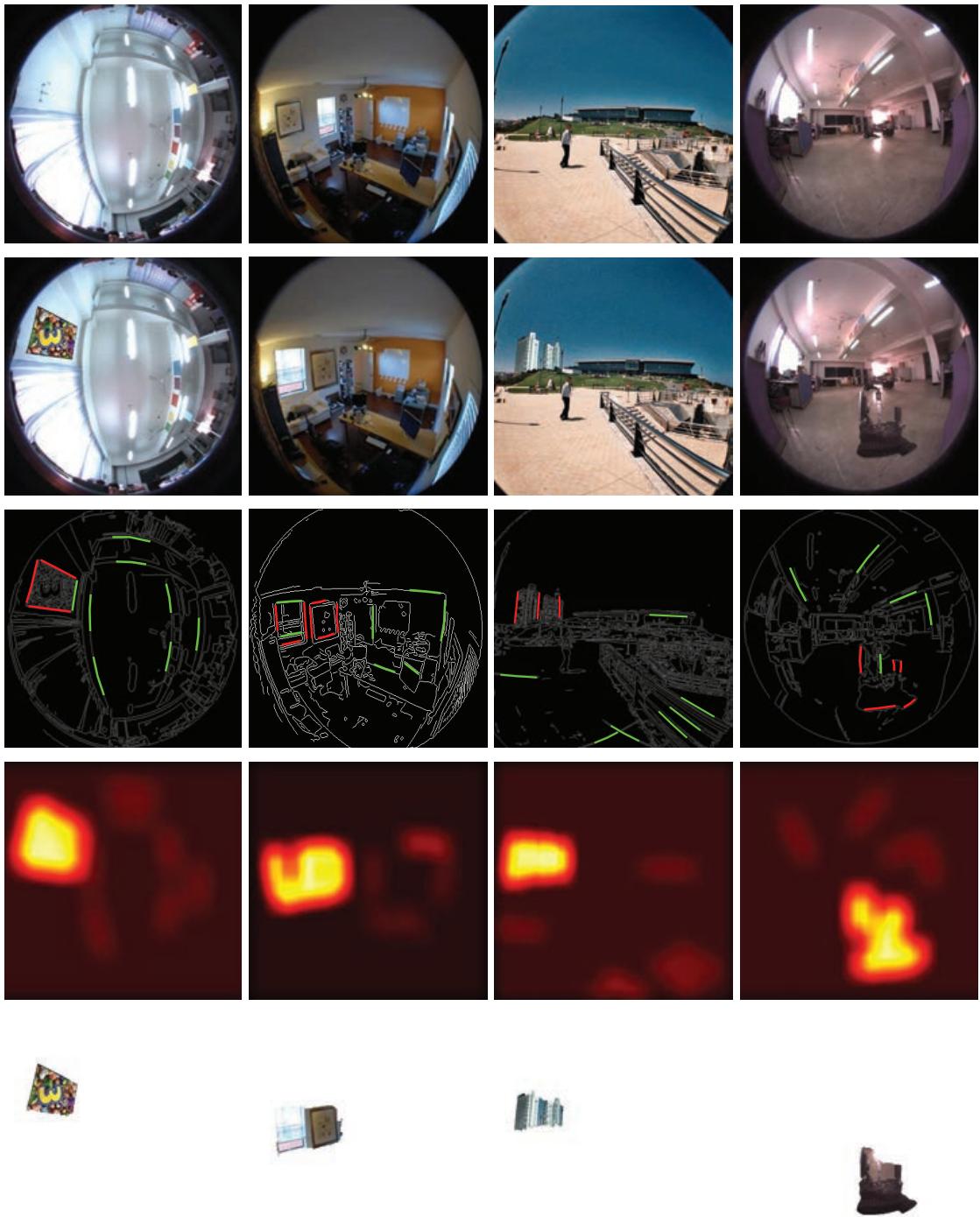


图 4-14 伪造图像内容取证结果。自上到下：原始真实图像，伪造的图像，候选直线取证测度，取证分布图，以及最终的取证结果呈现。其中在候选直线取证测度中，我们将混合测度值为 $w^C > 0.02$ 的候选直线判定为伪造，用红色线标记。可以看到虽然候选直线直接进行取证测度判定存在一个的漏检，但经过取证分布图后，可以准确的定位到物体级别。

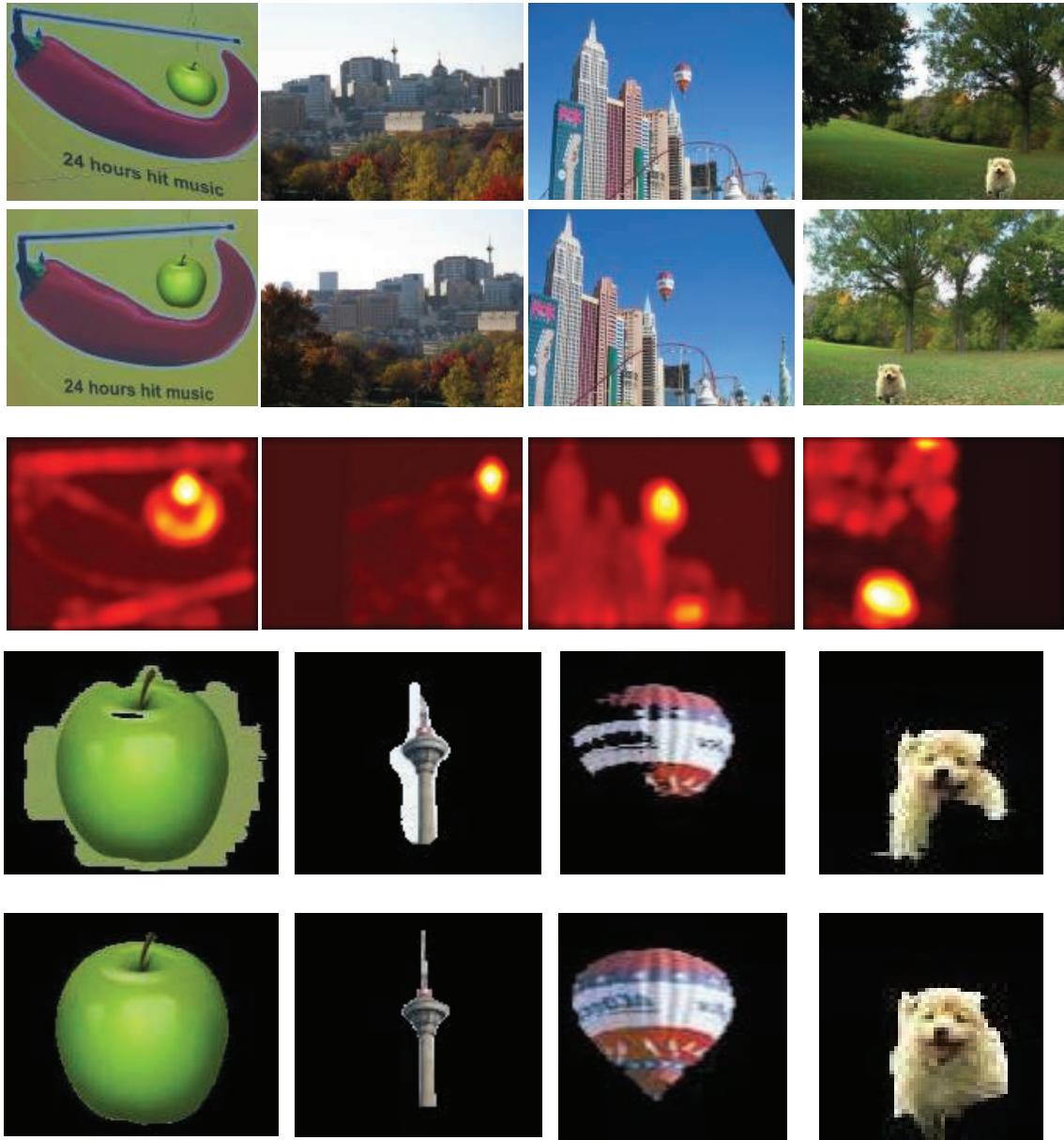


图 4-15 伪造取证分布图的应用：基于多视角的平面单映矩阵约束伪造取证。自上到下分别是：输入图像对、基于平面单映矩阵约束的伪造取证分布图、使用原始文献方法^[75]得到的分割结果、以及我们的方法得到的分割结果。

定在 $[0.1, 1]$ 和 $[0, 1]$ 之间。图 4-18 (b) 展示了体积取证测度的影响。由于公式 (4-11) 中 $\|I_f\|^3$ 分量的影响，体积取证测度几乎不受直线长度 l 的影响。但其也不与垂直间距 r 成正比，当体积取证测度保持增长一直到垂直间距 $r = 0.5$ ，然后随着垂直间距的增长而下降。不同于体积取证测度，距离取证测度几乎随着垂直间距 r 的增大而增长，如图 4-18 (c) 所示。图 4-18 (d) 展示了混合测度的结果，混合测度避免了体积取证测度当垂直间距 r 过大时的退化，同时

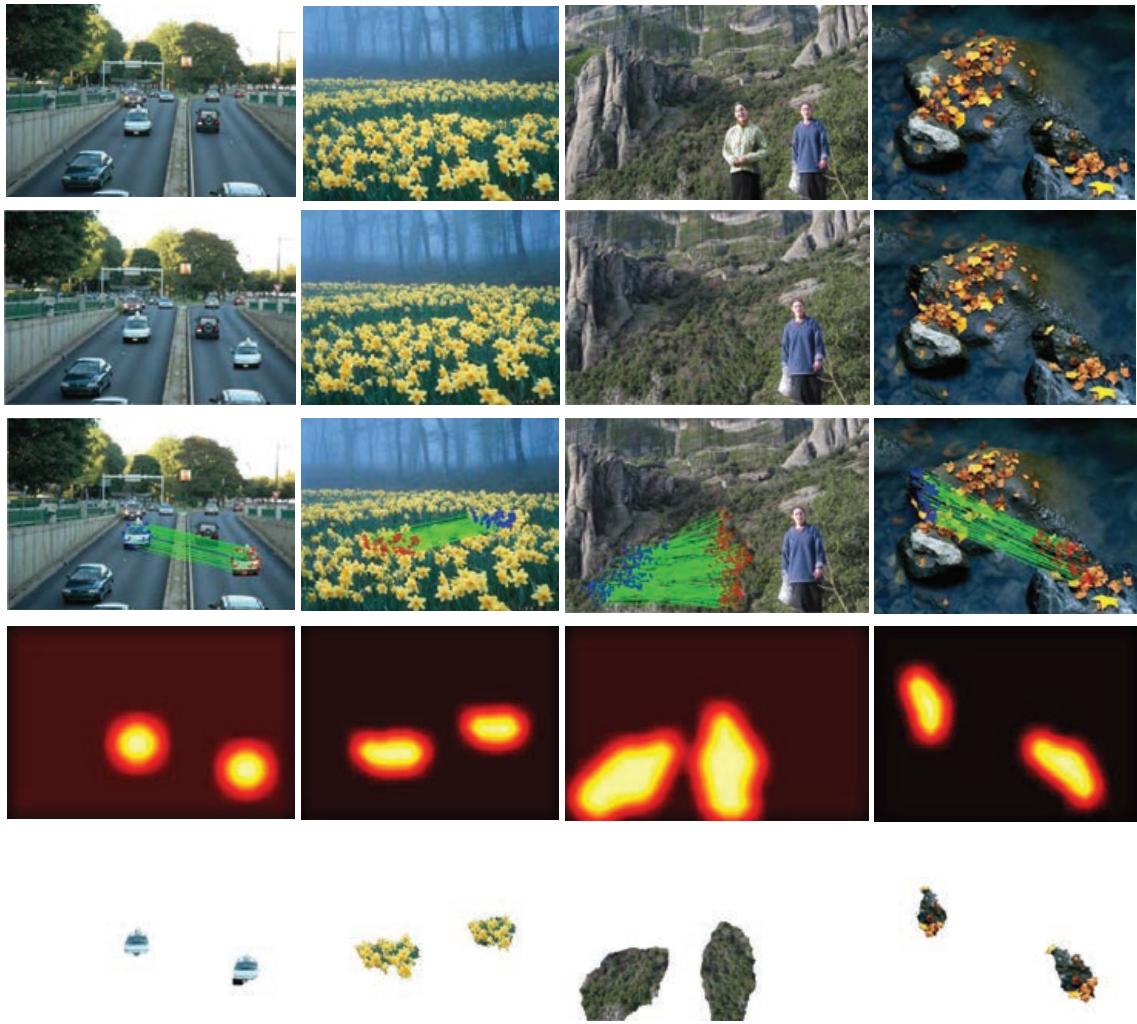


图 4-16 伪造取证分布图的应用：基于SIFT 复制伪造检测的图像取证。自上到下分别是：原始图片、伪造的图片、SIFT 特征匹配结果、相应的伪造取证分布图、以及最后的分割结果。

相比距离取证测度具有更好的噪声鲁棒性与检测准确性。

4.6 本章总结

在本章中，我们主要研究单图像相关源间的关联性约束。我们建立单幅图像与镜头畸变、以及单幅图像与其不同操作处理层之间的协同关联性，并提出一种基于图像畸变度的广角图像伪造区域盲检测算法。我们基于镜头径向畸变映射模型，推导出一种镜头畸变几何约束，并基于此提出两种自底向上的伪造取证测度。此外，我们使用伪造取证分布图进行取证结果的定位，并进行像素级别的分割。

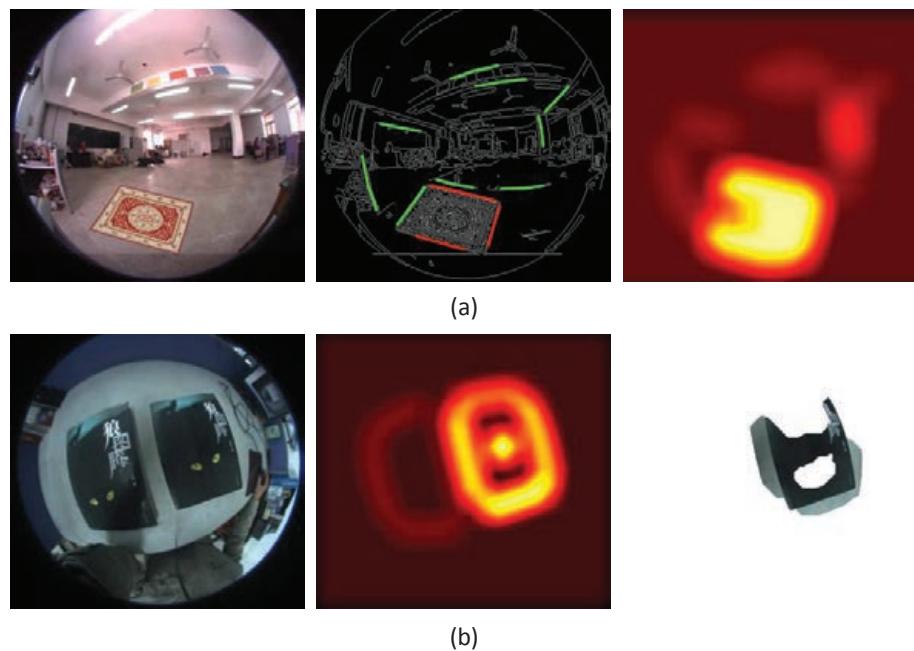


图 4-17 检测方法的失败情况。(a) 伪造直线的漏检，其中伪造物体上的部分直线没有被混合测度直接检测出来。(b) 物体分割失败，由于物体过大，导致欠分割。

研究显示，单图像相关源间的关联性约束可以为图像操作提供新的自身参考约束，使得一些需要参考源的算法，可以独立单幅图像完成，在一定程度上放宽了部分图像处理算法的适用性。

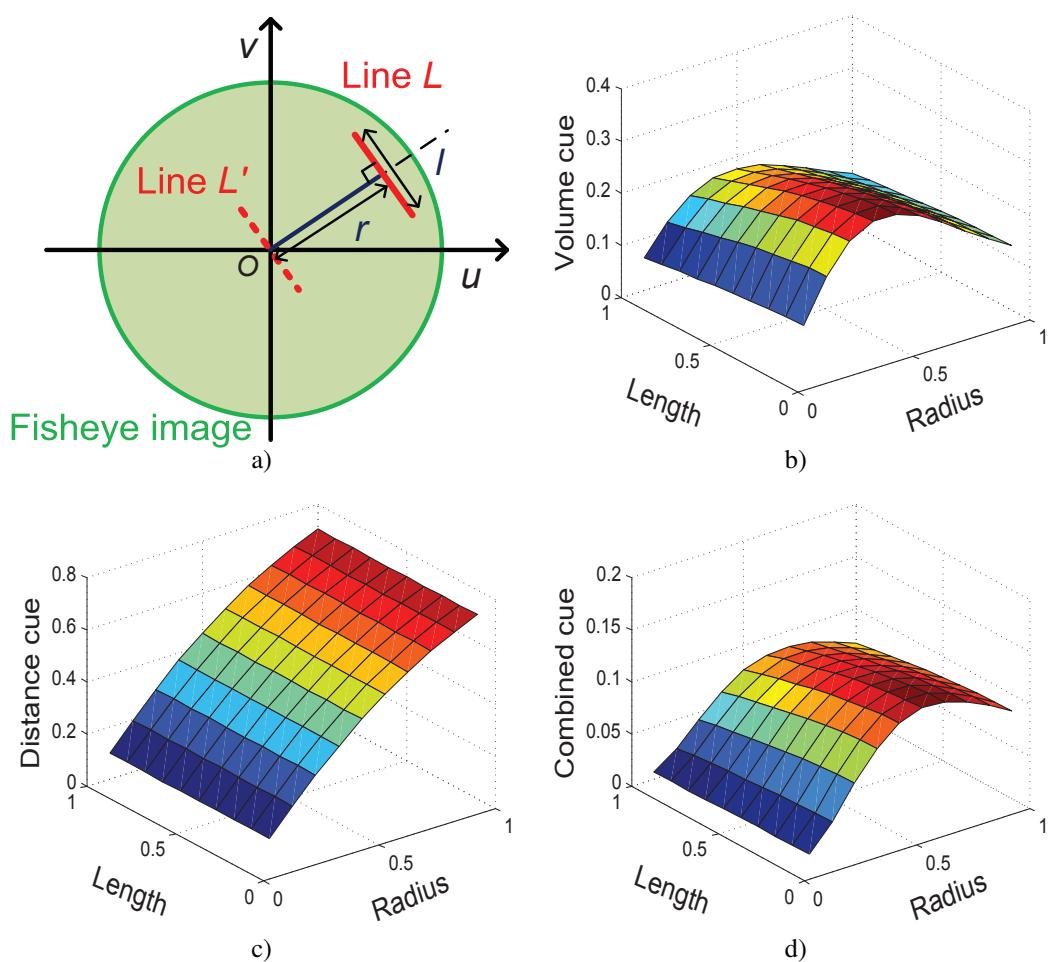


图 4-18 (a) 模拟直线评估实验设置。(b-d) 各个畸变取证测度对不同的直线长度 l 以及垂直间距 r 的影响。

第五章 总结与展望

5.1 总结

在本文中我们深入研究图像协同关联性约束思想，并将这种多目标的协同关联约束推广到其他相关应用中。我们进行了三个方面研究与拓展：多张图像间的关联性约束、单张图像不同区域间的关联性约束、以及单图像相关源间的关联性约束。我们以三个具体应用作为平台，研究协同合作的相关算法，其主要涉及：

1) 多张图像间的关联性约束：我们通过聚类，构建多张图像之间的协同关联性约束。通过引入图像生物视觉显著性原理，配合多图像间的目标关联性约束，提出多张图像的基于聚类的协同显著性目标检测分割算法。我们发现基于底层特征描述的聚类操作，可以有效地构建图像之间的弱关联结构。相对于基于复杂图结构的多图像关联算法，我们的聚类算法更加简单、高效，而且可以米面复杂的训练学习过程。我们将显著性测度直接应用于聚类级别上，可以看做是对多图像的协同关联结构进行直接操作，大大减少了传统算法的基于像素级别的操作算法复杂度。而且实验结果也表明，经过聚类算法，多图像的协同关联性约束被较好的体现出来，这是一种简单、高效的关联算法。

2) 单张图像不同区域间的关联性约束：我们挖掘单张图像中物体自身的几何结构约束，对图像的不同区域间建立协同联系。我们发现单幅图像中的不同区域之间也可以近似的作为一种协同分割应用。而不同区域之间的关联性约束则是通过几何结构约束所体现。为了反映不同区域之间的协同关联性约束，我们设计了一种几何结构映射矩阵，这种稠密的映射关系矩阵可以有效的提供一个像素级别的协同关联性约束。在这个基础上，我们使用基于几何结构约束的图分割模型去进行前景分割。实验表明，这种不同区域之间的协同关联性约束提供了一种高级别的图像关联，相对于仅仅使用低级别的分割算法，我们的算法取得了更好的分割效果。因此可以看到，在单张图片中引入协同关联性约束，也可以有效的提高相应的单张图像分割效果。

3) 单图像相关源间的关联性约束：除了单张图片不同区域间的关联以外，我们还发现，同图片的不同处理源之间也可以协同关联性约束。例如相机镜头与所拍摄的图片、以及图像与不同的处理层。因此，我们通过研究这些不同处理源之间的关联性约束，挖掘单幅图像与镜头、以及单幅图像与其

不同操作处理层之间的关系。发现不同处理源之间可以构建协同关联性约束，这种关联性约束可以提供一个非常有效的图像内容参考。基于此，我们提出了种基于图像畸变度的广角图像伪造区域盲检测分割算法。在镜头与其所拍摄的图像间，建立验证关系。在图像处理层与原始图像间，建立协同分割关系。实验表明，这种相关源之间的协同关联性约束近似于在不同图像间反应同样的目标区域，可以提供了一种有效的参考验证约束。

通过研究我们发现，协同关联性约束不仅存在于多图像之间，而且也可以存在于单张图像之中，通过图像的不同结构区域、以及相关源之间体现出来。这种协同关联性约束可以突破现有单张图像的局限，提高单/多张图像处理的结果，提供新的结构语义描述，以及提供有效的参考验证准则。

5.2 展望

尽管我们对协同关联性约束进行了深入研究与拓展，但鉴于时间与能力上的限制，研究仍存在很多的缺陷与不足。这些问题，可以作为以后未来工作的目标与方向，主要包括：

1) 在多张图像间的关联性约束中，虽然聚类可以有效地构建多张图像之间的协同关联性约束，但聚类所使用的特征仍然是基于颜色、纹理等底层特征，缺乏足够的语义关系。当聚类失败或者目标过于复杂时，方法有可能退化。因此，在以后的工作中，我们将加入高级别的语义描述，以期更好的解决多图像关联问题。此外，我们检测算法虽然可以处理多张图片，但仍然无法适应海量数据的需求（百万级别的图像）。如何在大规模数据的基础上，快速有效地建立图像间的关联，也是未来工作的重心之一。

2) 在单张图像不同区域间的关联性约束中，虽然我们证明了在单张图片中，可以通过进入几何结构约束，将图像的不同区域间建立协同联系。但依然要求目标满足一定的几何结构约束，因此方法具有局限性。未来，我们将进一步拓展单张图像不同区域间的关联，发展更加普遍的结构关系，例如邻接关系、物体级别的语义关系、以及空间结构等。这样可以更加拓展单张图像不同区域间的关联性约束的适用范围。

3) 在单图像相关源间的关联性约束中，我们的基于图像畸变度的广角图像伪造区域检测算法，虽然可以有效地检测出图像的伪造物体。但需要镜头满足一定的畸变度，我们的算法适用性。在实际中，图像多源之间的关系多种多样，我们将挖掘更多的关联约束方法，例如光照约束、相机噪声约束、图像格式约束等，进一步挖掘更多的单图像相关源间的关联性约束。

参考文献

- [1] Shi J, Malik J. Normalized cuts and image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8):888–905.
- [2] Tu Z, Zhu S. Image Segmentation by Data-Driven Markov Chain Monte Carlo[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5):657–673.
- [3] Rother C, Kolmogorov V, Blake A. “GrabCut”: interactive foreground extraction using iterated graph cuts[J]. ACM Transactions on Graphics, 2004, 23(3):309–314.
- [4] Grady L. Random Walks for Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(11):1768–1783.
- [5] Rother C, Minka T, Blake A, et al. Cosegmentation of Image Pairs by Histogram Matching - Incorporating a Global Constraint into MRFs[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2006, 1:993–1000.
- [6] Hochbaum D S, Singh V. An efficient algorithm for Co-segmentation[C]. International Conference on Computer Vision (ICCV), 2009:269–276.
- [7] Vicente S, Kolmogorov V, Rother C. Cosegmentation Revisited: Models and Optimization[C]. European Conference on Computer Vision (ECCV), 2010:465–479.
- [8] Batra D, Kowdle A, Parikh D, et al. Interactively Co-segmentating Topically Related Images with Intelligent Scribble Guidance[J]. International Journal of Computer Vision, 2011, 93(3):273–292.
- [9] Mukherjee L, Singh V, Peng J. Scale invariant cosegmentation for image groups[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011:1881–1888.
- [10] Vicente S, Rother C, Kolmogorov V. Object cosegmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011:2217–2224.
- [11] Kim G, Xing E. On multiple foreground cosegmentation[C]. International Conference on Computer Vision (ICCV), 2012:837–844.
- [12] 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.
- [13] Kolmogorov V, Zabin R. What energy functions can be minimized via graph cuts?[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(2):147–159.

-
- [14] Szeliski R, Zabih R, Scharstein D, et al. A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(6):1068–1080.
 - [15] Mukherjee L, Singh V, Dyer C. Half-integrality based algorithms for cosegmentation of images[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009:2028–2035.
 - [16] Fan J, Wu Y, Dai S. Discriminative spatial attention for robust tracking[C]. *European Conference on Computer Vision (ECCV)*, 2010:480–493.
 - [17] Kim G, Xing E, Fei-Fei L, et al. Distributed Cosegmentation via Submodular Optimization on Anisotropic Diffusion[C]. *International Conference on Computer Vision (ICCV)*, 2011:169–176.
 - [18] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011:409–416.
 - [19] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11):1254–1259.
 - [20] Lee W, Huang T, Yeh S, et al. Learning-Based Prediction of Visual Attention for Video Signals[J]. *IEEE Transactions on Image Processing*, 2011, 20(11):3028–3038.
 - [21] Toet A. Computational versus Psychophysical Bottom-Up Image Saliency: A Comparative Evaluation Study[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(11):2131–2146.
 - [22] Valenti R, Sebe N, Gevers T. What Are You Looking at? Improving Visual Gaze Estimation by Saliency[J]. *International Journal of Computer Vision*, 2012, 98(3):324–334.
 - [23] Hou X, Zhang L. Saliency Detection: A Spectral Residual Approach[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007:1–8.
 - [24] Liu T, Yuan Z, Sun J, et al. Learning to Detect a Salient Object[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(2):353–367.
 - [25] Lang C, Liu G, Yu J, et al. Saliency Detection by Multitask Sparsity Pursuit[J]. *IEEE Transactions on Image Processing*, 2012, 21(3):1327–1338.
 - [26] Jacobs D, Goldman D, Shechtman E. Cosaliency: where people look when comparing images[C]. *ACM symposium on User interface software and technology*, 2010:219–228.
 - [27] Chen H. Preattentive co-saliency detection[C]. *International Conference on Image Processing (ICIP)*, 2010:1117–1120.
 - [28] Li H, Ngan K. A Co-saliency Model of Image Pairs[J]. *IEEE Transactions on Image Processing*, 2011, 20(12):3365–3375.

参考文献

- [29] Chang K, Liu T, Lai S. From co-saliency to co-segmentation: An efficient and fully unsupervised energy minimization model[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011:2129–2136.
- [30] Tan H, Ngo C. Common pattern discovery using earth mover’s distance and local flow maximization[C]. International Conference on Computer Vision (ICCV), 2005:1222–1229.
- [31] Yuan J, Wu Y. Spatial Random Partition for Common Visual Pattern Discovery[C]. International Conference on Computer Vision (ICCV), 2007:1–8.
- [32] Toshev A, Shi J, Daniilidis K. Image Matching via Saliency Region Correspondences[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007:1–8.
- [33] Cho M, Shin Y, Lee K. Co-recognition of Image Pairs by Data-Driven Monte Carlo Image Exploration[C]. European Conference on Computer Vision (ECCV), 2008:144–157.
- [34] Yang L, Geng B, Cai Y, et al. Object Retrieval Using Visual Query Context[J]. IEEE Transactions on Multimedia, 2011, 13(6):1295–1307.
- [35] Goferman S, Tal A, ZelnikManor L. Puzzle-like Collage.[J]. Computer Graphic Forum, 2010, 29(2):459–468.
- [36] Tu Z, Bai X. Auto-Context and Its Application to High-Level Vision Tasks and 3D Brain Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(10):1744–1757.
- [37] Kumar M P, Torr P H S, Zisserman A. OBJCUT: Efficient Segmentation Using Top-Down and Bottom-Up Cues[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(3):530–545.
- [38] Maire M, Yu S, Perona P. Object detection and segmentation from joint embedding of parts and pixels[C]. International Conference on Computer Vision (ICCV), 2011:2142–2149.
- [39] Li Y, Sun J, Tang C, et al. Lazy snapping[J]. ACM Transactions on Graphics, 2004, 23(3):303–308.
- [40] Blake A, Rother C, Brown M, et al. Interactive image segmentation using an adaptive GMMRF model[C]. European Conference on Computer Vision (ECCV), 2004:428–441.
- [41] Yang A, Huang K, Rao S, et al. Symmetry-based 3-D reconstruction from perspective images[J]. Computer Vision and Image Understanding, 2005, 99(2):210–240.
- [42] Liu Y, Hel-Or H, Kaplan C, et al. Computational Symmetry in Computer Vision and Computer Graphics[J]. Foundations and Trends® in Computer Graphics and Vision, 2010, 5(1-2):1–195.
- [43] Jiang N, Tan P, Cheong L. Multi-view repetitive structure detection[C]. International Conference on Computer Vision (ICCV), 2011:535–542.
- [44] Zhao P, Quan L. Translation symmetry detection in a fronto-parallel view[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011:1009–1016.

- [45] Lee S, Liu Y. Curved Glide-Reflection Symmetry Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(2):266–278.
- [46] Yang A, Rao S, Huang K, et al. Geometric segmentation of perspective images based on symmetry groups[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2003:1251–1258.
- [47] Park H, Martin G, Bhalerao A. Structural Texture Segmentation using Affine Symmetry[C]. IEEE International Conference on Image Processing, 2007:49–52.
- [48] Sun Y, Bhanu B. Symmetry integrated region-based image segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009:826–831.
- [49] RiklinRaviv T, Sochen N, Kiryati N. On Symmetry, Perspectivity, and Level-Set-Based Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(8):1458–1471.
- [50] Elor Y, Shaked D, Bruckstein A. Crazy-Cuts: From Theory to App[J]. The Mathematical Intelligencer, 2012, 34:50–55.
- [51] Devernay F, Faugeras O. Straight lines have to be straight[J]. Machine Vision and Applications, 2001, 13(1):14–24.
- [52] Claus D, Fitzgibbon A. A rational function lens distortion model for general cameras[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005, 1:213–219.
- [53] Kannala J, Brandt S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(8):1335–1340.
- [54] Liu H, Javed O, Taylor G, et al. Omni-Directional Surveillance for Unmanned Water Vehicles[C]. The Eighth International Workshop on Visual Surveillance, 2008.
- [55] Fu H, Cao Z, Cao X. Embedded omni-vision navigator based on multi-object tracking[J]. Machine Vision and Applications, 2011, 22(2):349–358.
- [56] Fitzgibbon A W. Simultaneous linear estimation of multiple view geometry and lens distortion[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2001, 1:125–132.
- [57] Hartley R, Kang S B. Parameter-Free Radial Distortion Correction with Center of Distortion Estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(8):1309–1321.
- [58] Tardif J, Sturm P, Trudeau M, et al. Calibration of Cameras with Radially Symmetric Distortion[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(9):1552–1566.
- [59] Kukelova Z, Pajdla T. A Minimal Solution to Radial Distortion Autocalibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12):2410–2422.

参考文献

- [60] Mičušík B, Pajdla T. Estimation of omnidirectional camera model from epipolar geometry[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2003, 1:485–490.
- [61] Wei J, Li C, Hu S, et al. Fisheye Video Correction[J]. IEEE Transactions on Visualization and Computer Graphics, To appear.
- [62] Lyu S, Farid H. How Realistic is Photorealistic?[J]. IEEE Transactions on Signal Processing, 2005, 53(2):845–850.
- [63] Farid H. Image forgery detection[J]. IEEE Signal Processing Magazine, 2009, 26(2):16–25.
- [64] Popescu A, Farid H. Exposing Digital Forgeries by Detecting Traces of Re-sampling[J]. IEEE Transactions on Signal Processing, 2005, 53(2):758–767.
- [65] Fan Z, de Queiroz R L. Identification of bitmap compression history: JPEG detection and quantizer estimation[J]. IEEE Transactions on Image Processing, 2003, 12(2):230–235.
- [66] Luo W, Huang J, Qiu G. JPEG Error Analysis and Its Applications to Digital Image Forensics[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(3):480–491.
- [67] Johnson M, Farid H. Exposing digital forgeries through chromatic aberration[C]. Proceedings of the 8th ACM workshop on Multimedia and security, 2006:48–55.
- [68] Popescu A, Farid H. Exposing Digital Forgeries in Color Filter Array Interpolated Images[J]. IEEE Transactions on Signal Processing, 2005, 53(10):3948–3959.
- [69] Swaminathan A, Wu M, Liu K. Nonintrusive component forensics of visual sensors using output images[J]. IEEE Transactions on Information Forensics and Security, 2007, 2(1):91–106.
- [70] Cao H, Kot A. Accurate Detection of Demosaicing Regularity for Digital Image Forensics[J]. IEEE Transactions on Information Forensics and Security, 2009, 4(4):899–910.
- [71] Swaminathan A, Wu M, Liu K. Digital image forensics via intrinsic fingerprints[J]. IEEE Transactions on Information Forensics and Security, 2008, 3(1):101–117.
- [72] Johnson M, Farid H. Exposing Digital Forgeries in Complex Lighting Environments[J]. IEEE Transactions on Information Forensics and Security, 2007, 2(3):450–461.
- [73] Liu Q, Cao X, Deng C, et al. Identifying Image Composites Through Shadow Matte Consistency[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3):1111–1122.
- [74] Johnson M, Farid H. Metric Measurements on a Plane from a Single Image[R].[S.l.]: Department of Computer Science, Dartmouth College, 2006.
- [75] Zhang H, Guo X, Cao X. Water Reflection Detection Using a Flip Invariant Shape Detector[C]. International Conference on Pattern Recognition (ICPR), 2010:633–636.
- [76] Joulin A, Bach F, Ponce J. Discriminative clustering for image co-segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010:1943–1950.
- [77] Wu H, Wang Y, Feng K, et al. Resizing by symmetry-summarization[J]. ACM Transactions on Graphics, 2010, 29(159):1–10.

- [78] Fang Y, Chen Z, Lin W, et al. Saliency Detection in the Compressed Domain for Adaptive Image Retargeting[J]. *IEEE Transactions on Image Processing*, 2012, 21(9):3888–3901.
- [79] Nguyen M H, Torresani L, de la Torre F, et al. Weakly supervised discriminative localization and classification: a joint learning process[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009:1925–1932.
- [80] Zhai Y, Shah M. Visual attention detection in video sequences using spatiotemporal cues[C]. *ACM Multimedia*, 2006:815–824.
- [81] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009:1597–1604.
- [82] Duan L, Wu C, Miao J, et al. Visual saliency detection by spatially weighted dissimilarity[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011:473–480.
- [83] Tatler B W. The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions[J]. *Journal of Vision*, 2007, 7(14):1–17.
- [84] Judd T, Ehinger K, Durand F, et al. Learning to predict where humans look[C]. *International Conference on Computer Vision (ICCV)*, 2009:2106–2113.
- [85] Lang C, Nguyen T, Katti H, et al. Depth Matters: Influence of Depth Cues on Visual Saliency[M]//*European Conference on Computer Vision (ECCV)*, 2012:101–115.
- [86] Feichtinger H, Strohmer T. Gabor analysis and algorithms: Theory and applications[M].[S.l.]: Birkhauser, 1998.
- [87] Liu L, Chen R, Wolf L, et al. Optimizing Photo Composition[J]. *Computer Graphic Forum*, 2010, 29(2):469–478.
- [88] Lempitsky V, Kohli P, Rother C, et al. Image segmentation with a bounding box prior[C]. *International Conference on Computer Vision (ICCV)*, 2009:277–284.
- [89] Shotton J, Winn J, Rother C, et al. TextonBoost for Image Understanding: Multi-Class Object Recognition and Segmentation by Jointly Modeling Texture, Layout, and Context[J]. *International Journal of Computer Vision*, 2009, 81(1):2–23.
- [90] Zhang R, Zhang Z. Effective Image Retrieval Based on Hidden Concept Discovery in Image Database[J]. *IEEE Transactions on Image Processing*, 2007, 16(2):562–572.
- [91] Deselaers T, Ferrari V. Global and efficient self-similarity for object classification and detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010:1633–1640.
- [92] Goldberger J, Gordon S, Greenspan H. Unsupervised image-set clustering using an information theoretic framework[J]. *IEEE Transactions on Image Processing*, 2006, 15(2):449–458.
- [93] Blaschko M, Lampert C. Correlational spectral clustering[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008:1–8.

参考文献

- [94] Jiang W, Er G, Dai Q, et al. Similarity-based online feature selection in content-based image retrieval[J]. *IEEE Transactions on Image Processing*, 2006, 15(3):702–712.
- [95] Wang P, Wang J, Zeng G, et al. Salient Object Detection for Searched Web Images via Global Saliency[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012:3194–3201.
- [96] Kim G, Torralba A. Unsupervised Detection of Regions of Interest Using Iterative Link Analysis[C]. *The Conference on Neural Information Processing Systems*, 2009:961–969.
- [97] Deselaers T, Alexe B, Ferrari V. Localizing objects while learning their appearance[C]. *European Conference on Computer Vision (ECCV)*, 2010:452–466.
- [98] Zhu J, Wu J, Wei Y, et al. Unsupervised Object Class Discovery via Saliency-Guided Multiple Class Learning[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012:3218–3225.
- [99] Yang M, Yuan J, Wu Y. Spatial selection for attentional visual tracking[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007:1–8.
- [100] Ken F, Miyazato K, Kimura A, et al. Saliency-based video segmentation with graph cuts and sequentially updated priors[C]. *IEEE International Conference on Multimedia and Expo*, 2009:638–641.
- [101] Lee Y, Kim J, Grauman K. Key-segments for video object segmentation[C]. *International Conference on Computer Vision (ICCV)*, 2011:1995–2002.
- [102] Rahtu E, Kannala J, Salo M, et al. Segmenting Salient Objects from Images and Videos[C]. *European Conference on Computer Vision (ECCV)*, 2010:366–379.
- [103] Mishra A, Aloimonos Y, Cheong L, et al. Active Visual Segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(4):639–653.
- [104] Mishra A, Shrivastava A, Aloimonos Y. Segmenting “simple” objects using RGB-D[C]. *IEEE International Conference on Robotics and Automation (ICRA)*, 2012:4406–4413.
- [105] Carreira J, Sminchisescu C. Constrained parametric min-cuts for automatic object segmentation[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010:3241–3248.
- [106] Feng J, Wei Y, Tao L, et al. Salient object detection by composition[C]. *International Conference on Computer Vision (ICCV)*, 2011:1028–1035.
- [107] Alexe B, Deselaers T, Ferrari V. Measuring the Objectness of Image Windows[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11):2189–2202.
- [108] Kuettel D, Ferrari V. Figure-ground segmentation by transferring window masks[C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012:558–565.
- [109] Endres I, Hoiem D. Category Independent Object Proposals[C]. *European Conference on Computer Vision (ECCV)*, 2010:575–588.

- [110] Loy G, Eklundh J. Detecting Symmetry and Symmetric Constellations of Features[C]. European Conference on Computer Vision (ECCV), 2006:508–521.
- [111] Lee S, Liu Y. Skewed Rotation Symmetry Group Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9):1659–1672.
- [112] Prasad V, Yegnanarayana B. Finding axes of symmetry from potential fields[J]. IEEE Transactions on Image Processing, 2004, 13(12):1559–1566.
- [113] Keller Y, Shkolnisky Y. A Signal Processing Approach to Symmetry Detection[J]. IEEE Transactions on Image Processing, 2006, 15(8):2198–2207.
- [114] Cho M, Lee K M. Bilateral Symmetry Detection via Symmetry-Growing[C]. British Machine Vision Conference, 2009:1–11.
- [115] Guo X, Cao X. MIFT: A framework for feature descriptors to be mirror reflection invariant[J]. Image and Vision Computing, 2012, 30(8):546–556.
- [116] Hartley R I, Zisserman A. Multiple View Geometry in Computer Vision[M].[S.l.]: Cambridge University Press, 2004.
- [117] 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.
- [118] Felzenszwalb P, Huttenlocher D. Efficient Graph-Based Image Segmentation[J]. International Journal of Computer Vision, 2004, 59(2):167–181.
- [119] Szummer M, Kohli P, Hoiem D. Learning CRFs Using Graph Cuts[C]. European Conference on Computer Vision (ECCV), 2008:582–595.
- [120] Levinstein A, Dickinson S, Sminchisescu C. Multiscale symmetric part detection and grouping[C]. International Conference on Computer Vision (ICCV), 2009:2162–2169.
- [121] Hou X, Zhang L. Dynamic Visual Attention: Searching for coding length increments[C]. The Conference on Neural Information Processing Systems, 2008:681–688.
- [122] Blum H. A transformation for extracting new descriptors of shape[J]. Models for the perception of speech and visual form, 1967, 19(5):362–380.
- [123] Bai X, Wang X, Latecki L, et al. Active skeleton for non-rigid object detection[C]. International Conference on Computer Vision (ICCV), 2009:575–582.
- [124] Wang J, Markert K, Everingham M. Learning Models for Object Recognition from Natural Language Descriptions[C]. British Machine Vision Conference, 2009:1–11.
- [125] Benosman R, Kang S B. Panoramic Vision: Sensors, Theory, and Applications[M].[S.l.]: Springer Verlag, 2001.
- [126] Geyer C, Daniilidis K. Catadioptric Projective Geometry[J]. International Journal of Computer Vision, 2001, 45(3):223–243.

参考文献

- [127] Hughes C, Denny P, Glavin M, et al. Equidistant Fish-Eye Calibration and Rectification by Vanishing Point Extraction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(12):2289–2296.
- [128] Geyer C, Daniilidis K. A Unifying Theory for Central Panoramic Systems and Practical Implications[C]. European Conference on Computer Vision (ECCV), 2000:445–461.
- [129] Ying X, Hu Z. Can We Consider Central Catadioptric Cameras and Fisheye Cameras within a Unified Imaging Model?[C]. European Conference on Computer Vision (ECCV), 2004, 3021:442–455.
- [130] Courbon J, Mezouar Y, Eck L, et al. A generic fisheye camera model for robotic applications[C]. International Conference on Intelligent Robots and Systems, 2007:1683–1688.
- [131] Ying X, Hu Z. Catadioptric camera calibration using geometric invariants[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(10):1260–1271.
- [132] Ying X, Hu Z, Zha H. Fisheye Lenses Calibration Using Straight-Line Spherical Perspective Projection Constraint[C]. Asian Conference on Computer Vision (ACCV), 2006:61–70.
- [133] Ferrari V, Fevrier L, Jurie F, et al. Groups of Adjacent Contour Segments for Object Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(1):36–51.
- [134] Arbelaez P, Maire M, Fowlkes C, et al. From Contours to Regions: An Empirical Evaluation[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009:2294–2301.
- [135] Pan X, Lyu S. Region Duplication Detection Using Image Feature Matching[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(4):857–867.
- [136] Zhang C, Guo X, Cao X. Duplication Localization and Segmentation[C]. Advances in Multimedia Information Processing, 2010, 6297:578–589.
- [137] Amerini I, Ballan L, Caldelli R, et al. A SIFT-Based Forensic Method for Copy/Move Attack Detection and Transformation Recovery[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3):1099–1110.

发表论文和参加科研情况说明

(一) 博士期间的期刊论文

- [1] **Huazhu Fu**, Xiaochun Cao, Zhuowen Tu, “Cluster-based Co-saliency Detection”, *IEEE Transactions on Image Processing (TIP)*, 2013. (In press)
- [2] **Huazhu Fu**, Xiaochun Cao, Zhuowen Tu, Dongdai Lin, “Symmetry Constraint for Foreground Extraction”, *IEEE Transactions on Cybernetics*, 2013. (In press)
- [3] **Huazhu Fu**, Xiaochun Cao, “Forgery Authentication in Extreme Wide-angle Lens Using Distortion Cue and Fake Saliency Map”, *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 7, no. 4, pp. 1301-1314, 2012.
- [4] **Huazhu Fu**, Xiaochun Cao, Dai Tang, “Topology Preserved Regular Superpixels and Supervoxels”, *Submitted to IEEE Transactions on Multimedia (TMM)*. (Invited paper for ICME 2012)

(二) 博士期间的会议论文

- [1] Xiaochun Cao, Zhiqiang Tao, Wei Feng, Bao Zhang, **Huazhu Fu**, Xuewei Li, “Saliency map fusion based on rank-one constraint”, in *10th IEEE International Conference on Multimedia and Expo (ICME)*, 2013, San Jose, California, USA
- [2] Dai Tang, **Huazhu Fu**, Xiaochun Cao, “Topology preserved regular superpixel”, in *9th IEEE International Conference on Multimedia and Expo (ICME)*, 2012, pp. 765-768.
- [3] Ling Du, Xiaochun Cao, Muhua Zhang, **Huazhu Fu**, “Blind Robust Watermarking Mechanism based on Maxima Curvature of 3D Motion Data”, in *14th Information Hiding Conference (IH)*, 2012, pp. 110-124.

(三) 博士期间申请的发明专利

- [1] 操晓春, 李思远, 付华柱. 医学图像中对称器官的分割方法: 中国, 201210124019.8, 申请日: 2012-04-25.

致 谢

“为什么要读博？”在读博的时间里，很多人问过我这个问题。我想一万个苦逼博士，会有一万个不同的答案：为了拿到一个理想的学位；为了取得高薪的敲门砖；为了不离开喜欢的校园生活；为了提高自己的科研能力；为了完成自己坚持的信念；为了满足自己对学术的一点点贪婪。无论是什么，我们都在寻求，以期接近，心里的那份固执。很幸运，我付出了，收获了，并且无憾。我想我还会任性地坚持下去，直到有一天，我走累了，坐在路边，开始向所谓的现实妥协。但是，即使到那时，我也会说，这条路也许不是那么金光熠熠，但是我过得很快乐。

说起致谢，首先要感谢我的父母。说来愧疚，在别人孩子都在为票子、车子、房子而奋斗的时候，我的脑子里却只有算法、程序、实验、论文。感谢你们对我的支持，容忍我的固执，为我提供生活的一切，让我按照心意去走喜欢的路。

其次，要感谢我的博士导师，操晓春老师。师者，传道授业解惑也。博士生所学的不仅仅是学术上的知识，更多的是一种对待科研的态度。感谢操老师这些年对我的指导。从他的身上，我看到了对科研的狂热执着、对治学的严谨认真、以及对工作的精益求精，这些都令我印象深刻，也必将对我今后的研究和工作产生深远影响。

同时，我还要感谢UCLA 的屠卓文老师，以及微软亚洲研究院的张霖涛老师。在微软亚洲研究院访问实习期间，他们给予我很多的指导和帮助，让我了解并学会如何进行高水平的科研。

当然，我还要感谢在实验室一起学习和工作过的兄弟姐妹们，郭晓杰、张华、韦星星、杜玲、熊越、张宝、赵汉东、刘奇光、张晓静、张雪娟、李璟洁、唐黛等等，对不起，实验室人员太多，咱就不一一列出了:) 感谢你们陪我度过的时光，让我的博士生活色彩缤纷。

“宠辱不惊，闲看庭前花开花落；去留无意，漫随天外云卷云舒。”
——以此，献给我博士的三年时光。

付华柱
2013年4月于天津大学