

北京交通大学

硕士学位论文

基于高斯混合模型的图像分割的研究

Study on Image Segmentation Based on Gaussian Mixture Model

作者：欧垚江

导师：于双元

北京交通大学

2015 年 6 月



学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通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，提供阅览服务，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送交论文的复印件和磁盘。学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

(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：欧秋江

导师签名：

签字日期：2015年6月11日

签字日期：2015年6月11日

学校代码: 10004

密级: 公开

北京交通大学

硕士学位论文

基于高斯混合模型的图像分割的研究

Study on Image Segmentation Based on Gaussian Mixture Model

作者姓名: 欧垚江

学 号: 12120441

导师姓名: 于二元

职 称: 副教授

学位类别: 工学

学位级别: 硕士

学科专业: 计算机科学与技术

研究方向: 并行与分布式计算

北京交通大学

2015 年 6 月

致谢

本论文的工作是在我的导师于双元老师的悉心指导下完成的，于双元老师严谨的治学态度、渊博的专业知识以及科学的工作方法给了我极大的帮助和影响。于双元老师不仅悉心指导我完成了实验室的科研工作，在平时的学习和生活中也给予我极大的关心和帮助，在此衷心感谢三年来于双元老师对我的关心和指导。

感谢北京交通大学给我提供深造研究生的机会，使我能够继续学习和提高。三年的时光中我接受了浓郁的学术氛围熏陶在专业知识方面更上了一个台阶，并且在学术期刊上发表了论文。这些都离不开我的导师于双元老师的支持和学校给予的帮助。

感谢我的父亲和母亲，是他们从小教育我树立学习意识并一直支持我攻读学业，他们对我生活上和精神上极大的支持，我将以实际行动报答你们的养育之恩。

摘要

图像分割是计算机视觉中的重要内容,广泛的应用在图像识别体系中。经过几十年的研究,关于图像分割许多学者提出了各种方法和理论。其中基于模型的图像分割技术是分割方法中的一个热点,而高斯混合模型(GMM)又是图像分割模型中的一个有效且著名的模型。基于高斯混合模型的聚类分析方法可以解决存在不确定性分类以及具有复杂内容的图像分割的问题。可是基于高斯混合模型的先验概率分布没有引入邻域像素之间的空间关系,从而导致了使用该方法进行图像分割得到的结果对于噪声的干扰比较敏感。因此,空间可变有限混合模型 SVFMM 首次提出了引入邻域像素之间的空间关系到混合模型的先验概率分布中,使得模型具有较好抗噪性。之后也有学者提出了许多对 SVFMM 模型的改进,以提高抗噪性。为了克服 GMM 对噪声的敏感问题,需要对 GMM 的先验概率分布进行空间关系约束,有两种途径可以进行先验概率的约束。

本文首先对基于高斯混合模型的图像分割方法进行了研究,介绍马尔科夫随机场以及将马尔科夫随机场引入到高斯混合模型的先验概率分布中去。介绍了两种进行参数估计的方法,期望最大化(EM)算法和梯度下降算法。并对高斯混合模型进行图像分割的优点以及缺点进行阐述,为本文深入研究提供基础。

其次研究了高斯混合模型中进行先验概率约束的第一种途径的一种方法,一种拓展的高斯混合模型。基于相邻的像素的先验概率分布更可能是趋于相同或者相似的思想,提出了一种利用像素在邻域空间内的均值取代像素本身的值的方法构造先验概率分布,将邻域权重函数的均值作用到先验概率分布中,提高模型的抗噪性,通过梯度下降算法进行模型参数的估计,并给出了算法实现的伪代码。

最后将高斯核函数及拓展的高斯混合模型中每个像素属于每个类别的权重函数一起对像素的先验概率进行约束,提出了一种基于高斯核函数的先验概率约束的高斯混合模型。分别利用该模型及 GMM 对人工合成图像进行了实验,结果表明该模型的抗噪效果较好。

关键词: 图像分割; 高斯混合模型; 马尔科夫随机场; 梯度下降算法

ABSTRACT

Image segmentation is an important part of computer vision and widely used in image recognition system. After decades of research, many scholars have proposed a variety of methods and theories about image segmentation. Image segmentation technique based on models is a hot spot in these methods and techniques, Gaussian Mixture Model (GMM) is an effective and well-known model in these model-based techniques. Cluster analysis methods based on Gaussian Mixture Model can resolve complex content and uncertainty classification questions about image segmentation. But prior probability distribution based on Gaussian Mixture Model without introducing spatial relationships between neighboring pixels, resulting in segmentation results which got from Gaussian Mixture Model for image segmentation are more sensitive to noise. Therefore, Spatially Variant Finite Mixture Model (SVFMM) firstly proposed the introduction of spatial relationships between neighboring pixels to prior probability distribution of mixture model, so that mixture model has better noise immunity. To improve mixture model's noise immunity many scholars improved the SVFMM. In order to overcome the problem of noise sensitive to GMM, there are two ways to make spatial relationship constraints prior probability distribution on the GMM

Firstly, study on image segmentation method based on Gaussian Mixture Model in this paper, as well as introduce the Markov Random Field (MRF) and introduction of Markov Random Field to prior probability distribution of Gaussian Mixture Model. And then introduce two methods for parameter estimation, Expectation-Maximization (EM) algorithm and Gradient Descent Algorithm. Also describe advantages and disadvantages of Gaussian Mixture Model for image segmentation, And these can be used as the basis of the further study.

Secondly, we study a method which is an extension of the Gaussian Mixture Model, it is a method of the first kind of the prior probability constraint in the Gaussian Mixture Model. We proposed a method of constructs prior probability distribution replacing each pixel value in an image with the average value of its neighbors(including itself), based on the thought that the distribution of prior probability of adjacent pixels are more likely to be tended to be the same or similar. The mean of the neighborhood of the weight function is applied to the prior probability distribution, improve mixture model's noise immunity, and use the Gradient Descent Algorithm to estimate parameters, and gives the

pseudo code of the algorithm.

Finally, we propose a Gaussian Mixture Model based on Gaussian kernel function's prior probability, the model get the Gaussian kernel function and weight function that each pixel belonging to each category of the extension of the Gaussian Mixture Model together constraint the priori probability. The model and GMM were used to experiment with the synthetic images, and the results showed that the model had better noise immunity.

KEYWORDS: Image Segmentation; Gaussian Mixture Model; Markov Random Field; Gradient Descent Algorithm

目录

摘要	III
ABSTRACT.....	IV
1 引言	1
1.1 图像分割及其研究背景及意义	1
1.2 图像分割的研究现状	3
1.3 论文的主要研究工作	7
1.4 论文的结构安排	8
2 基于模型的图像分割	9
2.1 基础知识	9
2.1.1 高斯分布	9
2.1.2 高斯分布参数估计的极大似然法	10
2.2 高斯混合模型	11
2.3 基于马尔科夫随机场的高斯混合模型简介	13
2.3.1 马尔科夫随机场基本理论	13
2.3.2 基于马尔科夫随机场的高斯混合模型	15
2.4 模型的参数估计方法	19
2.4.1 EM 算法	19
2.4.2 梯度下降算法	20
2.5 本章小结	21
3 高斯混合模型的图像分割及其先验概率约束	23
3.1 扩展的高斯混合模型	23
3.2 基于空间邻域关系的高斯混合模型	24
3.3 参数学习	27
3.4 本章小结	28
4 基于高斯核函数的先验概率约束的高斯混合模型	30
4.1 高斯混合模型及图像分割	30
4.2 基于高斯核函数的先验概率约束的高斯混合模型	31
4.3 实验结果及分析	33
4.4 本章小结	35

5 总结与下一步工作

5.1 总结

5.2 下一步工作

参考文献

附录

作者简介

独创性声明

学位论文数据集

36

36

36

38

41

43

44

45

1 引言

1.1 图像分割及其研究背景及意义

人类主要通过视觉和听觉这两种途径来感知外部世界, 据统计, 人类靠通过视觉接收到的信息占总接收信息比例的 60%, 通过听觉接收到的信息占 20%, 其他的(比如嗅觉, 触觉等)占了 20%。而图像的信息在视觉所获得的信息中具有重要地位, “百闻不如一见”、“耳听为虚, 眼见为实”等, 这些都说明了视觉, 图像信息在传播信息的过程中的重要性, 因此, 对于图像的信息处理和提取就显得更为重要了。人类虽然在认识和改造世界的过程中具有一定的主观能动性, 但是自身的能力却是有限的, 一般情况下我们人类需要借助某种机器或者工具来协助或者代替我们完成某些工作。利用外部的工具可以快速、有效的完成大量复杂的外部场景的识别、定位以及追踪, 这是科技进步的表现, 设计出具有人类的视觉特征的智能系统是具有重要意义的。在这种需求下, 于上世纪五十年代出现了一门新的学科——计算机视觉, 或称之为机器视觉、图像处理和理解。

计算机视觉是指用计算机实现人的视觉功能——对客观世界三维场景的感知、识别和理解^[1]。一副静态图像, 它反映的信息是一维的, 计算机视觉研究的是, 可以让机器像人一样, 具有人类的某些能力, 可以从一维静态的图像中, 通过计算分析得出景物的二维图像, 进而实现人类的视觉所能观测到的三维世界中的属性特征和真实结构, 这样计算机才是真正的具有智能性, 这样才可以在复杂的场景环境中完成分析、识别及跟踪等任务。一般二维图像是由许多像素组成, 这些像素具有一定的灰度并且反映了场景的深度信息, 以像素灰度作为单位给出的图像的信息量是非常庞大的, 并不适合对这些图像直接进行分析和处理。而为了能够得到使人们满意的图像的分析结果或者较为理想的结果, 计算机视觉系统主要的工作就是: 从庞大的图像信息数据里分离出或突出感兴趣图像的数据, 利用这些经处理过的数据进一步去构造出某种较为理想的图像的表示形式, 使它们能够有效的用于图像的分析及理解, 这些就是图像预处理、特征检测及特征提取所要研究的问题。

图像分割作为图像处理中的一项关键技术和重要环节, 其本身也是一道经典难题, 图像分割发展到今天还是没有找到一种通用的方法进行分割, 也还没有制定出一个通用的标准去判定分割算法的好坏。对于一副图像, 人们经常只是对该图像中的某些部分感兴趣, 比如一副遥感图像, 若从军事的角度来看可能只是对

导弹基地、机场或者兵工厂等这些军事目标感兴趣；若从环境生态的角度来看，可能只是对湿地、森林或者草地等目标感兴趣。而这些感兴趣的部分一般只是对应图像中具有特殊性质的、特定的多个或者单个区域，这些感兴趣的部分我们称之为目标，目标之外的其他部分我们称之为背景。为分析和辨别目标我们就需要把这些多个或者单个目标从待分割图像中划分出来，这就是对图像分割的一种简单的理解。图像分割需要研究的问题之一就是怎么去确定感兴趣的目标。

从集合的角度来看，图像分割可以定义如下^[2]：

设集合 R 表示整个图像空间， R 依据选定的一致性准则 P 划分为互相不重叠的非空子区域或非空子集： $\{R_1, R_2, \dots, R_n\}$ ，这些子区域(子集)须得满足下面的条件：

- (1) $R = \bigcup_{i=1}^n R_i$ ；
- (2) 对于任意的 i 与 j ，当 $i \neq j$ 时， $R_i \cap R_j = \emptyset$ ；
- (3) 对 $i=1, 2, \dots, N$ ， $P(R_i) = \text{True}$ ；
- (4) 对 $i \neq j$ ，有 $P(R_i \cup R_j) = \text{False}$ ；
- (5) 对 $i=1, 2, \dots, N$ ， R_i 是连通的区域。

其中 \emptyset 表示空集， $P(R_i)$ 是集合 R_i 中所有元素的逻辑谓词。条件(1)指的是待分割图像的分割结果的所有子区域的并集或者总和应该能够包含图像中的全部像素，换句话说就是图像分割应该将每个像素都分割到某一个子集里面去。条件(2)指的是待分割图像的分割结果的所有子集是互相不重叠的，也就是说分割的结果中某个像素不可以同时属于多个子集。条件(3)指的是分割出来的每个子集或子区域都具有独特的特征，也就是说分割结果中同一子集的像素都具有某些相同的特征。而条件(4)与条件(3)相反，说的是分割结果中不同的子区域中的像素应该是具有不同的特征的，且没有公共元素。条件(5)说的是分割结果中某一个分割出的子集或者子区域应该是一个连通组元。

图像分割在现实中已经得到了广泛的应用，在文档图像处理，工业自动化，生产过程监控，线产品检测，遥感和生物医学图像分析，保安监视，以及军事，农业，体育，磁共振成像，合成孔径雷达图像，目标检测等方面都得到了应用。在这些应用里面，图像分割的结果精确与否将会直接影响到后续的各种处理。例如在医学图像领域中图像分割就扮演着重要的角色^{[3]-[5]}，精确的医学图像分割结果会给诊断的医生提供更加可靠且准确的信息，从而导致误诊情况减少。总的来说，在各种图像应用中，对图像进行目标提取、目标测量等等都离不开图像分割。但是由于图像的重叠强度、低对比度以及噪声的扰动使得图像的自动分割仍然是一个具有挑战性的研究课题^[6]。

尽管对于图像分割的研究已经有五十多年的历史，但是还没有形成一个快速、通用的分割方法。随着对图像分割的逐渐深入，人们发现图像分割问题是一个信息不全的不适定性的问题：图像分割，其目的是为了理解，可分割却是依赖于理解后所得的知识。而分割的不适定性会对问题求解形成很大的困难，到目前为止虽然已经提出的分割算法有很多，却并没有一种分割理论可以通用，也没有一种图像的分割算法可以适用于所有情况，每一种分割的方法都有其优点，同时也有其局限性。目前图像分割算法所存在的主要问题有：欠分割，即分割后的区域包含不一致的区域；过分割，即分割出来的区域不完整；分割性能的不稳定；分割所需要的时间较长等。有些分割的方法没有考虑到像素之间的空间信息导致对于含有噪声的图像的分割会存在误分，分割边界不清等。因此，如何去完善图像分割仍然还是计算机视觉所研究的一个热点。

1.2 图像分割的研究现状

当前，对于图像分割的方法主要分为两大类：基于边界的图像分割方法以及基于区域的图像分割方法^[7]。基于边界的图像分割方法是去试图寻找到图像中目标对象的边缘；基于区域的图像分割方法则是去试图将一副图像划分成为连通的区域。

(1) 基于边界的图像分割方法

基于边界的图像分割方法一般是通过改变边界曲线，从而优化某个目标函数，以实现最后的分割。边界曲线可使用图像的边缘作为初值去拟合样条曲线。图像中不同的区域之间总会存在边缘，边缘是颜色值或者灰度等属性不连续的结果，一副图像中的边缘含有异常丰富的信息。为获取一副图像的边缘，许多学者相继提出了多种边缘检测的方法，例如 LOG 算子^[8]、Sobel 算子^[9]、Canny 算子^[10]等等。但是由于这类微分算子对于噪声非常敏感，其检测到的边缘很可能不是边缘，因此需要通过形态学、平滑等操作去除噪声点、空洞等假的边缘。此外，为得到封闭、单像素的边缘，常常需要进行连接、细化以及跟踪等操作来形成目标对象的边界，但是这些操作的计算量比较大，处理起来也比较慢。在过去的这几十年来，人们相继提出了很多的基于边界的图像分割方法^{[11][12]}，这些方法都存在一些问题：在图像分割之前需要指定其目标对象的近似的边缘，而且目标函数的优化容易陷入局部最优。

(2) 基于区域的图像分割方法

基于区域的图像分割方法是试图将一副图像分成多个区域或类别，该区域或类别是一致的，根据形成的区域的不同，这种方法又可以分为：阈值分割方法、

区域生长法、分水岭分割方法、聚类分割方法等等。

1) 阈值分割方法对于一副待分割的图像的前景和背景的差别较为明显的时候,其分割的效果比较好。阈值是一个在图像分割的时候区分前景和背景的门限值,阈值分界了不同的区域,在实际的图像处理系统中经常会需要用到阈值处理。阈值分割分为两步:(a)选择确定合适的阈值;(b)根据选择的阈值将待分割的图像的每个像素归为对应的类别。然而对于阈值的选择一直都是阈值分割技术的难题,如果阈值的选择过低或者过高,那么图像分割出来的区域就会相应的扩大或者缩小。Prewitt 等人提出的直方图阈值法^[13]是一种常见的选择阈值的方法,这种方法根据统计图像的局部或者整体的灰度信息,从而把图像分割成为几个目标区域或类别,简单来说就是若灰度直方图所呈现出来的是双峰状态的时候,那么就选择双峰之间的低谷(谷底)对应的灰度值来作为阈值。该方法较为简单,但是在实际的应用中,图像的灰度直方图却往往比较粗糙,很难确定极大值和极小值,另外,若是属于单峰直方图或者双峰之间的谷底比较平坦的时候,这种分割方法就不适用了。

而对于彩色图像来说,其不止只是有灰度这个属性,因此,这种方法对于彩色图像的处理与灰度图像的处理有较大的不同。一般多数的方法是对待分割的彩色图像的每个分量都分别采用直方图阈值方法,之后再选出最佳的阈值。尽管阈值分割方法文献很多,其分割的方法也较为简单,但是该方法依据的仅仅只是像素的灰度值或二维灰度值,其对像素的空间信息位置以及边缘信息等等未做考虑,因此,这类方法对于一些分割质量要求不太高的应用较为有效,而对于多阈值组成的或者是含有噪声的图像的分割的效果就不是很好了。最大类间方差法是较为常用的阈值分割方法,它是通过最大化类间方差来选择一个全局的最优的阈值。除了最大类间方差法之外还有最小误差法、矩量保持法、简单统计法、模糊集法以及基于熵的阈值分割法等等。

2) 区域生长法^[14]是根据同一物体的区域内像素的相似性来聚集像素点的方法。从初始的区域开始,将具有相同性质的相邻的像素或者区域合并到邻近区域中来,进而逐步地增长区域,直到不需要合并为止。可以用平均灰度值、颜色以及纹理等信息去度量区域内的像素的相似性。区域生长法可以用来分割自然景物等较为复杂的图像,是一种较为普遍的方法。基于区域的方法假设在同一个区域中的像素应该在视觉上具有相似的特征,例如颜色特征、灰度特征以及纹理特征等。用区域生长法进行图像分割有两个方面的难点:第一个是对种子点选择,选择不同的种子点会得到不同的分割结果;第二个是区域生长过程中设置相似性准则,如何自动且有效的进行种子点的选择,从而得到精确的分割结果,这是一个难题。区域生长法有优点也有缺陷,其优点是计算较为简单,速度也比较快;缺

点是由于没有先验知识,采用的随机种子点以及随机相似性准则进行分割的分割结果可能会与实际情况不相符合。

3) 分水岭分割方法^{[15][16]},该方法是一种较新的基于区域的分割方法,源自于数学的形态学,思想来源于洼地的积水。具体来说就是:对于一副梯度图像而言,在三维空间中可以看作地形的表面,地势的海拔高度是由表面梯度的大小来决定的,谷底就是海拔低的地方,若向每个谷底注水,那么各个谷底就会形成汇水的盆地,各个汇水的盆地的水位会随着注水量的增加而上升,从而导致不同的盆地的水出现汇合,而在汇合的地方设立堤坝,这样就形成了分水岭,与此同时,该梯度图像也被各个汇水盆地分割成了对应的区域。显然,分水岭是对应于梯度较大的地方,而盆地则是由均匀颜色或者灰度的区域组成,因此,图像的分水岭可以通过该图像的梯度的变换得到,之后图像的各个盆地(区域)可以通过阈值设定分割出来。在分水岭变换之后,一般都要使用形态学操作来减小图像的过分割,因此,最终的图像分割的结果与形态操作算子有关系。

4) 聚类法^[17],该方法是一种无监督分类,是在没有训练样本的情形下,依据数据集的内在结构,依照样本空间中的数据的相似性以及相似性的测量准则来对数据进行自动归类和划分,从而使得同一划分内的样本数据具有相同或者相似的属性,但属于不同划分内的样本数据的差别比较大。 K 均值聚类^[18]是一种具有结构明确,分割快速以及通用性好等的优点的一种基本聚类方法,但其在优化分割的过程中也存在易陷入局部极小值的问题。Dunn 于 1947 年提出了模糊形式 K 均值聚类算法,也称为模糊 C 均值聚类算法^[19]。1981 年 Bezdek 证明了模糊 C 均值聚类算法的收敛性^[20],同时讨论了 K 均值聚类算法与模糊 C 均值聚类算法的关系,并将其进行了进一步的扩展,建立起了模糊聚类的理论。模糊 C 均值聚类引入模糊度量理论,从而使得分割的图像与人眼的视觉分割更加一致,但同时也增加了计算量,也存在易陷入局部最优的问题,基于 K 均值聚类和模糊 C 均值聚类的方法没有考虑到图像像素的空间信息。

还有基于图论的图像分割方法,如规范割(Normalized Cut)准则、最小最大割(Min-max Cut)准则、最小割(Minimum Cut)准则、平均割(Average Cut)准则、比例割(Ratio Cut)准则等。规范割算法可以有效的分割图像中如目标对象边界等的突出特征。由于最小割准则比较容易产生分割碎片,相对于最小割准则而言,规范割准则是一个显著的进步^[21]。规范割算法是一种谱聚类的方法,然而一般的谱聚类方法的计算量都很大。

近年来,对于基于模型的图像分割方法渐渐引起了人们的重视,人们对于混合模型的研究往前最早可以追溯十九世纪末,1894 年彼尔逊采用了具有两个混合成分的高斯混合模型对一组数据进行拟合,并通过矩估计方法对该模型参数进行

估计, 这就是关于混合模型最早的研究。接下来的几十年, 研究人员对混合模型的参数估计一直都是采用矩估计的方法, Wicksell 等人于 1924 年开始讨论二元的高斯混合模型, 在 1928 年 Doetcsch 又把混合成分的个数推广到了多个混合成分的情况。然而矩估计方法的计算太过复杂, 从而导致关于混合模型的研究进展比较缓慢。而计算机的出现使得很多复杂的计算问题得到了解决, 于是在 1972 年 Robertson 与 Tan 等人开始采用极大似然法对混合模型的参数估计进行研究, 并且在此基础上还证明了极大似然法还要优于矩估计方法^[22]。

由于计算机的硬件等设备更新越来越快, 计算机处理的速度也越来越快, 人们渐渐把研究的重点放在了使用极大似估计法进行混合模型参数的估计上。在 1977 年 Dempster 等人采用了期望最大化算法(EM 算法)计算极大似然估计从而使计算困难的问题得到了解决。之后, 关于混合模型的研究进入到了新的阶段, 其研究的领域也扩展到了经济学、医学、计算机视觉等其他的学科。对于混合模型而言, 对数据进行聚类是其较为有代表性作用的。假设一组观测样本值是来自于 k 个总体, 每一个数据可以被认为是来自于 k 个分布总体中的某一个, 各个混合的权值表示的是观测样本值来自于某个分布总体的频数。通过对混合模型进行研究, 可以将来自于不同分布总体的一组数据进行分类。

随着人们对混合模型研究的深入, 基于混合模型的分割方法也越来越受到人们的重视, 国内外许多学者提出了许多种混合模型, 例如高斯混合模型、三角级数混合模型、多项式混合模型以及多样条混合模型等等。在这些混合模型中, 高斯混合模型(Gaussian Mixture Model, GMM)是比较知名也是比较有代表性的一种基于混合模型的分割方法。对于多元数据而言, 该模型是一个较为灵活而且强大的统计建模工具。许多学者都使用高斯混合模型来研究图像分割中的一些关键问题^{[23][24]}。但是, 基于高斯混合模型的图像分割方法是单独对像素进行考虑, 在分割的过程中没有考虑到像素之间的空间信息^[25], 这就导致了其分割的结果对噪声会比较敏感, 对含有噪声的图像其分割结果可能并不准确。

在图像分割方法中, 基于马尔科夫随机场(Markov Random Field, MRF)模型的分割方法也是较为常用的一种分割方法, 其是属于基于统计类的分割方法的范畴。基于马尔科夫随机场(MRF)模型的分割方法是图像处理领域中一个有效的数学工具, 该方法的提出起始于 S.German 等人发表的文献中所提到的关于随机场的论文^[26]。关于马尔科夫随机场, 其有两层含义: 随机场以及马尔科夫性。简单理解就是满足马尔科夫性质的随机场就称之为马尔科夫随机场(MRF)。马尔科夫随机场模型通过定义邻域系统, 确定基团形式, 定义具体的势函数的形式并根据 MRF 与吉布斯(Gibbs)分布的等价性写出其对应的联合吉布斯分布, 根据分布形式定义似然函数从而得出其后验概率的能量函数的表达式, 最后进行参数估计优化求解。马

尔科夫随机场考虑了像素之间的空间信息，能够区分不同的纹理分布，适合于纹理图像分割，对含噪声的图像的分割效果也较好，但 MRF 模型进行图像分割的性能取决于参数估计的精准程度^[27]，若有足够的先验知识就可以较为准确的估计参数，反之就可能面临分割与参数估计互相制约的矛盾。

1.3 论文的主要研究工作

本文对基于高斯混合模型的图像分割方法进行了研究，发现用高斯混合模型进行图像分割比较容易实现，并且需要进行估计的参数也相对较少，可是由于高斯混合模型只是单独的对像素进行考虑，并没有考虑像素之间的空间关系，这就导致了利用高斯混合模型进行图像分割会对噪声比较敏感，对含有噪声的图像用高斯混合模型进行分割可能存在误分的情况。

高斯混合模型(GMM)利用贝叶斯理论进行图像中像素的划分，它是一种通过简单密度分布模拟复杂密度分布的一种有效的数学方法。在图像分割的过程中，每个像素都被赋予了一组先验概率，先验概率代表每个像素属于每个分类的可能性。而在每一组先验概率中的每一个先验概率都只是来自于图像中像素的统计信息，但是没有包含像素之间的空间信息。在对含噪声的图像进行分割的时候，目标区域内的非噪声像素点和与其相邻的噪声像素点的先验分布是不一样的，但是，由于没有将邻域像素之间的空间信息引入到像素的先验分布中，从而导致不能将噪声像素点有效的分割到目标区域中去。在分割的结果中，不能完整、正确地将目标区域分割出来。因此，高斯混合模型对于噪声的抑制能力比较弱。在一副图像中，相邻的像素往往被划分到同一目标区域的可能性比较大，也就是说相近或者相邻的像素更加可能具有相似或者相同的先验概率分布。分析邻域像素之间的空间信息关系，构造出具有更强抗噪性的空间信息关系和更适应于局部空间，并将其引入到图像像素的先验概率分布中去，使得邻域像素具有相似或者相同的先验概率分布。将邻域像素的空间关系信息引入到先验概率分布的高斯混合模型对于噪声的抑制能力会得到加强。

对于如何将邻域像素的空间关系引入到先验概率分布中去，如何进行空间约束的问题，目前主要有两种途径：

其一，第一种途径是采用一种类似于非线性或者线性的空间平滑方法，将邻域像素之间的空间关系引入到它们的先验概率分布中去^{[28][29]}。这种途径的核心思想是相近或者相邻的像素的先验概率分布更可能是趋于相似的或者相同的。这条途径的收敛速度比较快，并且由于空间平滑方法的特点，这条途径也比较容易实现，而且需要估计的参数也相对比较少。

其二, 第二种途径^{[30]-[32]}是认为像素的先验概率分布是服从吉布斯(Gibbs)分布的, 也就是马尔科夫随机场(MRF)。由于似然函数中引入了吉布斯分布函数, 因此, 使用最大后验估计法进行参数的估计而不是用最大似然估计法进行参数估计。这条途径提高了高斯混合模型的抗噪性, 但也会大量增加估计的参数, 使得实现较为复杂, 似然函数的收敛速度也较为缓慢。

本文对这两种途径进行了研究, 其中对第一种途径进行了较为深入的研究, 结合原有方法利用像素在邻域空间内的均值取代像素本身的值的方法构造先验概率分布, 将邻域权重函数的均值作用到先验概率分布中对先验概率进行约束。其次将高斯核函数(高斯径向基函数)作为对像素先验概率分布的空间约束引入到高斯混合模型中对先验概率分布进行约束, 提高模型的抗噪性, 通过梯度下降算法进行模型参数的估计。最后对将马尔科夫随机场引入到高斯混合模型中对像素的先验分布进行约束的第二条途径, 我们进行了一些研究并发现文献[33]中的方法具有一定的缺陷, 在该方法的图像分割的过程中都使用同一个常数 β , 而现实中很多情况却并非如此, 所以对于常数 β 的不同取值以及其取值对于分割结果的不同影响的问题应该是我们下一步将要研究的重点。本文章节的主要安排如下。

1.4 论文的结构安排

本文的结构主要分为五个章节, 具体安排如下:

第一章是本文的引言, 主要介绍了图像分割的研究背景及意义, 图像分割的研究现状, 本文的主要研究内容以及本文的一些结构和安排。

第二章介绍了一些概率知识, 高斯分布, 高斯混合模型, 高斯混合模型和图像分割, 介绍马尔科夫随机场以及将马尔科夫随机场引入到高斯混合模型的先验概率分布。介绍了两种进行参数估计的方法, 期望最大化(Expectation Maximization, EM)算法和梯度下降算法。

第三章首先介绍了将邻域像素的空间关系引入到先验概率分布中去的第一条途径的一种方法, 一种拓展的高斯混合模型, 并提出了一种利用像素在邻域空间内的均值取代像素本身的值的方法构造先验概率分布, 提高模型的抗噪性。

第四章提出了一种基于高斯核函数的先验概率约束的高斯混合模型, 该模型将高斯核函数及拓展的高斯混合模型中每个像素属于每个类别的权重函数一起对像素的先验概率进行约束, 提高模型的抗噪性。进行了人工合成图像的实验, 并对结果进行了分析。

第五章为论文的总结与下一步工作部分, 对论文的主要工作进行总结并对下一步的工作进行展望。

2 基于模型的图像分割

2.1 基础知识

2.1.1 高斯分布

在概率论中，高斯分布是一个非常重要的概率分布，在统计学的许多方面有着重要的影响。高斯分布有着广泛的实际背景，生产以及科学试验中，很多随机变量的概率分布都可以近似的用高斯分布来描述。对于一个随机变量 x ，该随机变量对应的均值为 μ ，方差为 σ 的高斯分布为：

$$P(x|\Theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (2.1)$$

其中， $\Theta = \{\mu, \sigma\}$ 为参数，高斯分布示意图如下图 2.1 所示。

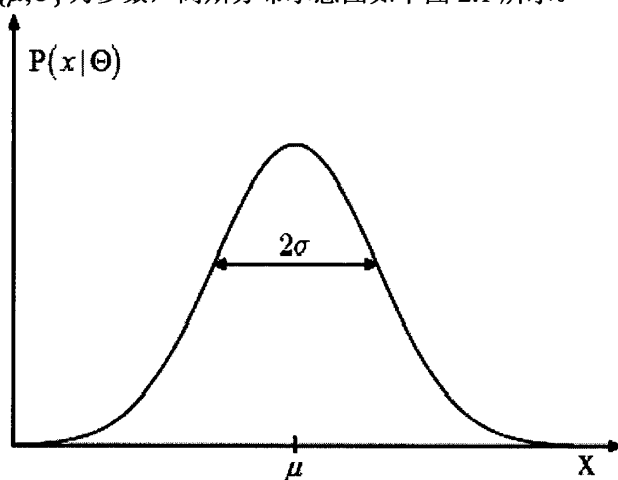


图 2.1 均值为 μ 方差为 σ 的高斯分布示意图

Fig. 2.1 Gaussian distribution diagram of the mean μ and standard deviation σ

高斯分布的概率密度满足：

$$P(x|\Theta) > 0 \quad (2.2)$$

$$\int_{-\infty}^{+\infty} P(x|\Theta) dx = 1 \quad (2.3)$$

随机变量 x 的高斯分布的均值为：

$$\mu = E(x) = \int_{-\infty}^{+\infty} P(x|\Theta) x dx \quad (2.4)$$

方差为:

$$\sigma^2 = E(x^2) - E(x)^2 = \int_{-\infty}^{+\infty} P(x|\Theta) x^2 dx - \mu^2 \quad (2.5)$$

对于 D 维随机变量 x 而言, 其高斯分布 $P(x|\Theta)$ 的形式为:

$$P(x|\Theta) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \quad (2.6)$$

其中, $\Theta = \{\mu, \Sigma\}$ 为参数, μ 为 D 维均值列向量, Σ 是 $D \times D$ 的协方差矩阵。

2.1.2 高斯分布参数估计的极大似然法

由于极大似然具有渐进最优的性质, 已经成为参数估计的一种常用的方法。极大似然估计的准则是使观测值出现的概率最大。

给定一个数据集 $X = \{x_1, x_2, \dots, x_N\}$, 其中 x_i 为观测值且都独立地服从高斯分布, 我们就可以通过极大似然法估计该高斯分布的参数^[34], 其对应的对数似然函数为:

$$L(\Theta|X) = \frac{ND}{2} \log(2\pi) - \frac{N}{2} \log |\Sigma| - \frac{1}{2} \sum_{i=1}^N \frac{(x_i - \mu)^2}{\sigma^2} \quad (2.7)$$

其中 $\Theta = \{\mu, \sigma\}$ 为参数。为了使对数似然函数(2.7)达到最大则我们需要优化估计参数 $\Theta = \{\mu, \sigma\}$, 一般用最大似然法(Maximum Likelihood, ML)来估计, 计算似然函数最大值, 即是求 $\bar{\Theta} = \Theta$ 使得:

$$\log L(\bar{\Theta}|X) = \max_{\Theta \in \Theta} \log L(\Theta|X) \quad (2.8)$$

一般确定极大似然估计值的方法都是首先建立极大似然估计值满足的似然方程, 通过求解似然方程获得极大似然估计值。似然方程的建立为: 对似然函数 $L(\Theta|X)$ 的参数 μ 和 σ 分别求偏导数, 并令其等于 0, 有:

$$\frac{\partial L(\Theta|X)}{\partial \mu} = \sum_{i=1}^N \frac{x_i - \mu}{\sigma^2} = 0 \quad (2.9)$$

$$\frac{\partial L(\Theta|X)}{\partial \sigma} = \sum_{i=1}^N \left(-\frac{1}{\sigma} + \frac{(x_i - \mu)^2}{\sigma^3} \right) = 0 \quad (2.10)$$

上面两个式子称之为似然方程, 我们可以得到参数 μ 和 σ 在下一时刻即 $t+1$ 步

的似然估计值。样本的均值为：

$$\mu^{(t+1)} = \frac{1}{N} \sum_{i=1}^N x_i \quad (2.11)$$

由式(2.9)也可得到(2.11)，通过式(2.10)得

$$\sigma_{(t+1)}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu^{(t+1)})^2 \quad (2.12)$$

对 $\mu^{(t+1)}$ 和 $\sigma_{(t+1)}^2$ 分别求期望得到：

$$E[\mu^{(t+1)}] = \mu \quad (2.13)$$

$$E(\sigma_{(t+1)}^2) = \frac{N-1}{N} \sigma^2 \quad (2.14)$$

这里可以看出，均值 μ 的极大似然估计的期望和真实均值相等，而方差 σ^2 的极大似然估计的方差比真实值稍小。这里可以定义估计^[34]：

$$\hat{\sigma}_{(t+1)}^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu^{(t+1)})^2 \quad (2.15)$$

可以修正误差，且 $E(\hat{\sigma}_{(t+1)}^2) = \sigma^2$ 。

2.2 高斯混合模型

图像的特征向量在其特征空间的分布并不是严格地服从哪一种分布的，但是由概率论中的中心极限定理可以知道，大量相互独立的服从同一分布的随机变量其在总体上是服从高斯分布(也称正态分布)的。由于正态分布既能反映大多数事物的分布规律，且在数学上也能表现出良好的性能从而得到了广泛的应用。

对于一副图像而言，如果能够将从图像中的像素提取出来的特征向量分为若干个标签(类别)，而且标签与标签之间是相互独立的，标签内的向量之间也是相互独立的，那么每个标签内的向量服从同一高斯分布，多个标签的高斯分布进行一定的加权组合，这样就是该图像的特征向量的总体分布，表现了图像的特征。基于高斯混合模型进行图像分割的原理是对图像中存在的每个待分割的区域建一个高斯分布，这个分布的参数由其区域内的特征向量的分布来决定，表征了这个区域的可分割性。一般而言每个待分割区域的概率密度函数形式相同，但是密度函数中的参数不同，则构建的模型是在特定概率密度函数的形式下的一组参数。

待分割图像的各个区域可用某些特征向量来表示一些信息，例如像素值、位置、纹理等等，并能反映图像中区域差异。图像中的区域可用一高斯分布的概率

密度函数来表示,也就是说第 k 个区域可用混合高斯密度函数中第 k 个高斯分量的参数来表示。但不知道第 k 个区域包含哪些像素,也不知道给定像素提取的特征向量属于哪一个区域。这种由基于像素提取特征向量形成的概率密度函数就是高斯混合模型。

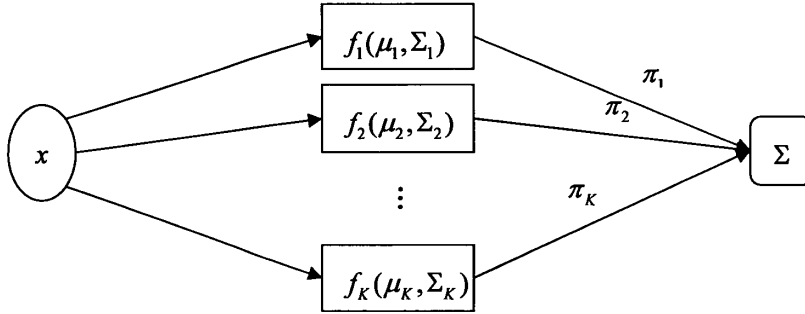


图 2.2 高斯混合模型

Fig 2.2 Gaussian mixture model

图 2.2 中, $x_i, i=1,2,\dots,N$ 表示图像中第 i 个像素的观测值(也叫随机变量), $\Omega_1, \Omega_2, \dots, \Omega_K$ 表示 K 个标记。 $P(x_i | \Theta_j) (j=1,2,\dots,K)$ 是其对应的概率密度函数, 其函数为^{[35][36]}:

$$f(x_i | \Pi, \Theta) = \sum_{j=1}^K \pi_j P(x_i | \Theta_j) \quad (2.16)$$

则像素集服从有限混合分布, 像素集对应的模型就是有限混合模型。其中 $\Pi = \{\pi_1, \pi_2, \dots, \pi_K\}$, π_j 为像素 x_i 属于标记 Ω_j 的成分分布混合的先验概率。 Θ_j 为第 j 个成分分布的参数, $\Theta = \{\Theta_j, j=1,2,\dots,K\}$ 是所有参数的集合。对于先验概率 π_j 而言, 需满足:

$$0 \leq \pi_{ij} \leq 1, \quad \sum_{j=1}^K \pi_{ij} = 1 \quad (2.17)$$

若每个成分分布都是服从高斯分布, 那么其所对应的模型就是高斯混合模型。

$$P(x_i | \Theta_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}\right) \quad (2.18)$$

其中 $\Theta_j = \{\mu_j, \sigma_j\}$ 。像素样本之间具有统计独立性, 则对于像素集 $X = (x_1, x_2, \dots, x_N)$ 的联合条件密度为^[34]:

$$p(X | \Pi, \Theta) = \prod_{i=1}^N f(x_i | \Pi, \Theta) = \prod_{i=1}^N \left(\sum_{j=1}^K \pi_j P(x_i | \Theta_j) \right) \quad (2.19)$$

这是参数 $\{\Pi, \Theta\}$ 关于 X 的似然函数。由于对数函数具有单调性, 对式(2.19)取

对数得到高斯混合模型的对数似然函数:

$$L(\Theta, \Pi | X) = \sum_{i=1}^N \log \left(\sum_{j=1}^K \pi_j P(x_i | \Theta_j) \right) \quad (2.20)$$

其中 $\Theta = \{\Theta_j, j=1, 2, \dots, K\}$ 。接下来可以使用期望最大化算法(EM 算法)对上面的对数似然函数进行参数的估计以实现最大化对数似然函数,从而得到分割结果。由上面的对数似然函数可以看到高斯混合模型的最大优点就是形式较为简单,而且所要求的参数相对也比较少,EM 算法将在下面 2.4 小节进行介绍。但是该模型也有一定的缺陷,那就是把像素作为单独的样本进行考虑,而没有考虑到相邻像素之间的空间关系,没有把邻域像素之间的空间信息关系引入到高斯混合模型的先验概率分布中。因此,对于含有噪声的图像分割较为敏感。

2.3 基于马尔科夫随机场的高斯混合模型简介

2.3.1 马尔科夫随机场基本理论

为了解决高斯混合模型没有考虑图像像素的空间信息,从而导致其对含有噪声的图像分割较为敏感的问题,一些学者相继提出了基于马尔科夫随机场(MRF)的混合模型。在这些模型中,先验概率分布是通过马尔科夫随机场得到,这些混合模型的优点是考虑了像素间空间的信息,从而提高含有噪声的图像的分割质量。

马尔科夫随机场模型可以用于非监督分割,在非监督分割的方法中,标号数不知道,模型的参数需要从观测图像中去估计。因此,非监督分割问题实质上是观测图像与随机场模型的拟合问题。而实际应用中,往往无法获得训练数据,因此非监督分割方法在实际应用中更加广泛。在非监督分割的问题中,主要的困难在于在没有任何的先验知识的情形下要去准确的估计图像中的标号数,该问题也叫做聚类的有效性。马尔科夫随机场模型把不确定性描述和先验知识联系在一起,通过观测图像,估计理论中最优准则和统计决策来确定图像分割问题中的目标,并求出满足条件的最大可能分布,这样就将图像分割的问题转化为最优化的问题。马尔科夫随机场(MRF)是一种概率理论,它为分析一副图像的空间限制提供了一种随机的数学框架。

马尔科夫随机场(MRF)的概念由两个方面组成,其一是随机场,其二是马尔科夫性。换句话说就是满足马尔科夫性质的随机场就称之为马尔科夫随机场,其具体表述如下:

S 表示 N 个点(像素)的有限离散集合, $S = (1, 2, \dots, N)$ 。 $L = (1, 2, \dots, K)$ 表示图

像所包含的标签集(类别集)。点集 S 为像素的位置, L 表示的是分割区域的数目。每个点(像素)对应一个描述算子, 表示这个点的状态, z_i 表示第 i 点的状态, $Q_i = (z_i, z_i \in L)$ 表示 z_i 的相空间。点集的状态值空间为:

$$Q = \prod_{i=1}^N Q_i \quad (2.21)$$

如果对于 $Z \in Q$ 有 $p(Z) > 0$, 则称 $p(Z)$ 为一个随机场。为了表示局部各点间相互作用的空间关系, 可以将邻域 N 看作一个系统, 设 $N = (N_i, i \in S)$ 为定义在 S 上的邻域系统上的集合, 满足以下的特性:

- (1) $N_i \subset S$
- (2) $i \notin N_i$
- (3) $\forall i, j \in S, i \in N_j \Leftrightarrow j \in N_i$

位置 $i \in N_j$ 称为 j 的邻点, N_j 称为 j 的邻点集。当 S 中的子集 c 中每对不同位置总是相邻的时候, 则称 c 为一个势团(Clique), C 为势团的集合。势团为包含了若干个位置的集合, 在某些极端情形下, S 的所有的子集都是势团, 则这时可以认为所有的像素相互影响; 或者每一个位置 S 是一个势团, 则此时可以认为像素间没有相互作用影响。势团具有势能(Potential Energy), 可以用势能表示相邻位置上状态之间的一些局部相互作用。对于势团 c 所对应的势能 U_c 的值是由这个势团像素分量的灰度值来决定的。

随机场 $p(Z)$ 是一个马尔科夫随机场(MRF), 当且仅当“在某个位置上的局部条件概率仅仅随所考虑的位置的邻域状态而变化^[37]”。具体的表示如下:

$$p(z_i | z_{S-\{i\}}) = p(z_i | z_{N_i}) \quad (2.22)$$

即一个点的状态只是取决于这个点的邻域的点的状态, 满足马尔科夫性质。

根据 Hammersley-Clifford 定理^[38]中马尔科夫随机场(MRF)和吉布斯(Gibbs)分布的等价性, 可以将马尔科夫随机场(MRF)的概率分布转化为吉布斯(Gibbs)分布来求解。根据 Hammersley-Clifford 定理可以将式(2.22)表示为:

$$p(Z | \beta) = W^{-1} \exp(-U(Z | \beta)) \quad (2.23)$$

其中, W 是一个归一化常数, 称为 Gibbs 分布的划分函数。 $U(Z | \beta)$ 是一个能量函数, 是所有可能势团的势能的总和, 表述为下:

$$W(\beta) = \sum_{Z \in Q} \exp(-U(Z | \beta)) \quad (2.24)$$

$$U(Z | \beta) = \sum_{c \in C} V_c(Z | \beta) \quad (2.25)$$

其中 V_c 是定义在势团 c 上的势函数, C 为所有势团的集合。 β 为参数, 为了计算(2.24)涉及到 MRF 的状态 Z 的所有可能, 而这个却非常难以实现, 因此 Besag 等人提出了伪似然函数^[39], (2.23)可以写成:

$$p(Z|\beta) = \prod_{j=1}^N p(z_i | z_{N_i}; \beta) \quad (2.26)$$

其中 $p(z_i | z_{N_i}; \beta)$ 为:

$$p(z_i | z_{N_i}; \beta) = \frac{\exp\left(-\sum_{ci} V_c(Z|\beta)\right)}{\sum_{z_i} \exp\left(-\sum_{ci} V_c(Z|\beta)\right)} \quad (2.27)$$

上式的概率分布可以得到马尔科夫随机场的参数估计。

2.3.2 基于马尔科夫随机场的高斯混合模型

为了降低高斯混合模型进行图像分割的结果对噪声的敏感度, 许多学者相继提出了邻域像素的空间信息并将其引入到高斯混合模型的先验概率分布中作为解决该问题的途径, 这些模型认为像素的先验概率分布是服从吉布斯(Gibbs)分布的, 也就是马尔科夫随机场(MRF), 并将其引入到似然函数中。文献[23]提出了一种空间可变的有限混合模型(Spatially Variant Finite Mixture Model, SVFMM)进行图像分割。将 N 个的图像的像素分成 K 类, 像素 x_i 的概率密度函数为:

$$f(x_i | \Pi, \Theta) = \sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \quad (2.28)$$

其中, 高斯分布 $P(x_i | \Theta_j)$ 与式(2.16)相同。 π_{ij} 为像素 x_i 属于标记 Ω_j 的先验分布, 且满足:

$$0 \leq \pi_{ij} \leq 1, \quad \sum_{j=1}^K \pi_{ij} = 1 \quad (2.29)$$

由于标记 Ω_j 中的像素 x_i 相互独立, 则数据集 $X = (x_1, x_2, \dots, x_N)$ 的联合条件密度为:

$$p(X | \Pi, \Theta) = \prod_{i=1}^N f(x_i | \Pi, \Theta) = \prod_{i=1}^N \left(\sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right) \quad (2.30)$$

下面将马尔科夫随机场(MRF)分布^[40]应用于高斯混合模型的先验概率分布, 由 MRF 和 Gibbs 分布的等价性有:

$$p(\Pi) = W^{-1} \exp \left[-\frac{1}{T} U(\Pi) \right] \quad (2.31)$$

其中, W 为归一化常数, T 是一个温度常数。 $U(\Pi)$ 是平滑先验。由贝叶斯公式得后验概率密度函数可写为:

$$p(\Pi, \Theta | X) \propto p(X | \Pi, \Theta) p(\Pi) \quad (2.32)$$

由上式得对数似然函数为:

$$\begin{aligned} L(\Pi, \Theta | X) &= \log(p(\Pi, \Theta | X)) \\ &= \sum_{i=1}^N \log \left[\sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right] + \log p(\Pi) \\ &= \sum_{i=1}^N \log \left[\sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right] - \log W - \frac{1}{T} U(\Pi) \end{aligned} \quad (2.33)$$

由上式可以看到根据 $U(\Pi)$ 的选择不同就会得到不同的模型, 在文献[23]中的空间可变有限混合模型(SVFMM)中 T 的值设为 1, 像素 x_i 的先验概率 $p(\Pi)$ 服从 Gibbs 分布为:

$$p(\Pi) = \frac{1}{Z} \exp(-U(\Pi)) \quad (2.34)$$

这里,

$$U(\Pi) = \beta \sum_{i=1}^N \sum_{j=1}^K \sum_{m \in N_i} (\pi_{ij} - \pi_{mj})^2 \quad (2.35)$$

其中, Π 为参数集合; $\Pi = \{\pi_{ij}\}$, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, K$ 。 β 为先验参数, 控制平滑度。则 SVFMM 模型的对数似然函数为:

$$\begin{aligned} L(\Pi, \Theta | X) &= \sum_{i=1}^N \log \left[\sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right] + \log p(\Pi) \\ &= \sum_{i=1}^N \log \left[\sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right] - \beta \sum_{i=1}^N \sum_{j=1}^K \sum_{m \in N_i} (\pi_{ij} - \pi_{mj})^2 \end{aligned} \quad (2.36)$$

将该式与(2.20)高斯混合模型比较可以看到上式的对数函数非常复杂。可以用迭代的 EM 算法去进行参数 $\{\Theta, \Pi\} = (\mu_j, \Sigma_j, \pi_{ij})$ 的极大似然估计。

在完整数据条件的情况下^[25], 最大化似然函数(2.36)式会导致目标函数 $H(\Pi, \Theta | X)$ 的值增加:

$$H(\Pi, \Theta | X) = \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(0)} \{ \log \pi_{ij} + \log P(x_i | \Theta_j) \} - \beta \sum_{i=1}^N \sum_{j=1}^K \sum_{m \in N_i} (\pi_{ij} - \pi_{mj})^2 \quad (2.37)$$

其中, z_{ij} 为隐含变量, 表示像素 x_i 属于标记 Ω_j 的后验概率, t 表示第 t 步迭代, z_{ij} 的条件期望为:

$$z_{ij}^{(t)} = \frac{\pi_{ij}^{(t)} P(x_i | \Theta_j^{(t)})}{\sum_{k=1}^K \pi_{ik}^{(t)} P(x_i | \Theta_k^{(t)})} \quad (2.38)$$

分别对(2.37)式求参数 μ_j, Σ_j^{-1} 及 π_{ij} 的偏导数, 令其偏导为 0 得到:

$$\begin{aligned} \mu_j^{(t+1)} &= \frac{\sum_{i=1}^N z_{ij}^{(t)} x_i}{\sum_{i=1}^N z_{ij}^{(t)}} \\ \Sigma_j^{(t+1)} &= \frac{\sum_{i=1}^N z_{ij}^{(t)} [(x_i - \mu_j)(x_i - \mu_j)^T]}{\sum_{i=1}^N z_{ij}^{(t)}} \\ \pi_{ij}^{(t+1)} &= \frac{\sum_{m \in N_i} \pi_{mj} + \sqrt{\left(\sum_{m \in N_i} \pi_{mj} \right)^2 + \frac{N_i}{\beta} z_{ij}^{(t)}}}{2N_i} \end{aligned} \quad (2.39)$$

利用上面的式子使用迭代的 EM 算法的 M 步进行参数的更新, 但由于先验分布 π_{ij} 需满足(2.29), 这样, 在计算上就会变得非常复杂。并且比起高斯混合模型来该模型需要非常多的参数, 为了把 N 个像素分成 K 类需要处理 $K \times (2+N)$ 个参数 (K 个 μ_j , K 个 Σ_j 以及 NK 个 π_{ij}), 意味着图像的像素比较多时需要估计的参数也会变得非常多, 优点是该模型提高了图像的抗噪性。

根据式(2.33)中 $U(\Pi)$ 选择的不同得到的模型也会有所不同, 在贝叶斯 auto-logistic 模型中^[40], $U(\Pi)$ 为:

$$U(\Pi) = \sum_{i=1}^N \sum_{j=1}^K \alpha_{ij} \pi_{ij} + \sum_{i=1}^N \sum_{j=1}^K \sum_{m \in N_i} \beta_{ijm} \pi_{ij} \pi_{mj} \quad (2.40)$$

其中, α_{ij} 与 β_{ijm} 为参数集。虽然这个模型涉及到了许多参数, 但是对于含噪声的图像的分割还不是十分具有鲁棒性。

其他的对于先验概率约束的基于马尔科夫随机场的混合模型成功的应用在了图像分割当中, 并且对于 $U(\Pi)$ 的选择都有不同的方式, Blekas K. 等人选择了 $U(\Pi)$ 为^[31]:

$$U(\Pi) = \beta \sum_{i=1}^N \sum_{m \in N_i} \left[1 + \left(\sum_{j=1}^K (\pi_{ij} - \pi_{mj})^2 \right)^{-1} \right]^{-1} \quad (2.41)$$

其中，上面两式中的 β 为常数，Nikou C. 等人选择了 $U(\Pi)$ 为^[32]:

$$U(\Pi) = \sum_{i=1}^N \sum_{j=1}^K \sum_{s=1}^S \left[\frac{1}{2} \log \beta_{js}^2 - \frac{1}{2} \frac{\left(\sum_{m \in \partial_i} (\pi_{ij} - \pi_{mj}) \right)^2}{\beta_{js}^2} \right] \quad (2.42)$$

其中， S 为考虑方向的总数，一般情况下 S 的值为 4(水平方向、垂直方向和两个对角线方向)， β_{js} 为变量参数， ∂_i 为像素 x_i 的邻域且该邻域包括 x_i 在内。

由上面的(2.35)、(2.41)以及(2.42)式所示，增加了空间信息的联系就会使复杂程度增加，由于对数似然函数的复杂性，EM 算法的 M 步不能直接应用于估计先验分布 π_{ij} ，且先验分布需满足限制条件(2.29)，即需要解先验分布 π_{ij} 的有约束的优化问题，则计算的复杂性将会大大增加。

在上面的(2.40)式中，平滑先验的形式比较简单，因此对于最大化对数似然函数的参数集 $\{\Pi, \Theta\}$ 的估计比较容易，但是其对含噪声的图像的分割的鲁棒性不好。对于(2.35)、(2.41)以及(2.42)式而言，其先验平滑项的形式较为复杂，会增加训练的复杂性，且 EM 算法的 M 步不能直接用于估计先验分布，为此文献[33]中提出了一种快速及鲁棒性的空间约束的高斯混合模型。该模型首先定义因子 G_{ij} 为：

$$G_{ij}^{(0)} = \exp \left[\frac{\beta}{2N_i} \sum_{m \in N_i} (z_{mj}^{(0)} + \pi_{mj}^{(0)}) \right] \quad (2.43)$$

其中， z_{mj} 为后验概率， β 为控制平滑先验的温度值。在这里 β 设置为 12，邻域选择 5×5 方格， N_i 为像素 x_i 的邻域像素的数量。仔细观察上式可以发现，因子 G_{ij} 就像一个中值滤波，使某一像素的领域的平均值代替该像素的值更加简便。因子 G_{ij} 的优点是容易实现，并且联系了邻域像素之间的空间关系。

在该方法中的平滑先验 $U(\Pi)$ 为：

$$U(\Pi) = - \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(0)} \log \pi_{ij}^{(t+1)} \quad (2.44)$$

其中， t 表示第 t 步迭代。该模型将马尔科夫随机场(MRF)引入到高斯混合模型的先验概率中，对先验概率进行约束，使用 EM 算法进行参数的优化，且具有分割速度快分割结果的精度高的特点。

但该模型在图像分割的过程中都使用同一个常数 β ，而现实中很多情况却并

如此, 所以对于常数 β 的不同取值问题将是我们下一步将要研究的重点。

2.4 模型的参数估计方法

2.4.1 EM 算法

EM 算法是一种求参数的极大似然估计的迭代算法, 是一种较为简单而且实用的算法, 该算法可以从不完全的数据集中去求解模型的参数。该方法处理不完全数据效果良好, 被广泛的应用于处理截断数据、污染数据及缺失数据等。EM 算法通过最大后验概率(MAP)或者极大似然估计(ML)方法去拟合有限的混合模型时, 极大的简化为 E 步和 M 步两步, 较为方便的估计出参数。

EM 算法的 E 步是将计算完全数据的对数似然函数的期望转化为求 x_i 来自第 j 个高斯分量的贝叶斯后验概率:

$$z_{ij}^{(0)} = \frac{\pi_j^{(0)} P(x_i | \Theta_j^{(0)})}{\sum_{l=1}^K \pi_l^{(0)} P(x_i | \Theta_l^{(0)})} \quad (2.45)$$

上式表示 x_i 来自于 Ω_j 的贝叶斯后验概率。

M 步是求取模型参数的极大似然估计, 并对参数进行更新:

将(2.20)式分别对参数 μ_j, σ_j, π_j 求偏导数, 并令其等于 0 得到:

$$\mu_j^{(t+1)} = \frac{\sum_{i=1}^N z_{ij}^{(t)} x_i}{\sum_{i=1}^N z_{ij}^{(t)}} \quad (2.46)$$

$$[\sigma_j^2]^{(t+1)} = \frac{\sum_{i=1}^N z_{ij}^{(t)} (x_i - \mu_j^{(t+1)})^2}{\sum_{i=1}^N z_{ij}^{(t)}} \quad (2.47)$$

$$\pi_j^{(t+1)} = \frac{1}{N} \sum_{i=1}^N z_{ij}^{(t)} \quad (2.48)$$

其中, t 及 $t+1$ 分别代表迭代的次数, 总结高斯混合模型的 EM 算法如下:

第 1 步 初始化参数 $\{\Theta, \Pi\} = \{\mu_j, \sigma_j, \pi_j\}$ 。

第 2 步(E 步) 通过初始化的参数值计算(2.45)的后验概率的值。

第 3 步(M 步) 通过(2.46)、(2.47)、(2.48)式更新参数 μ_j, σ_j, π_j 。

第4步 计算对数似然函数(2.20)，检测该函数或者参数是否收敛，若没有收敛则返回第2步。

EM 算法执行完之后，使用(2.45)式对图像中的每个像素进行类标记分配。

2.4.2 梯度下降算法

梯度法是计算数学和最优化中求一个函数的极值问题的一个重要的方法，这个小节介绍使用梯度下降算法进行高斯混合模型的参数估计^[41]。

首先，类别 Ω_j 对应的先验概率 π_j 选择：

$$\pi_j = \frac{\exp(\gamma_j)}{\sum_{k=1}^K \exp(\gamma_k)} \quad (2.49)$$

上式叫做 Softmax 函数^[41]，可以确保对于 $-\infty \leq \gamma_j \leq \infty$ ， π_j 满足式(2.17)。根据式(2.16)及(2.49)，需要优化参数 $\Xi = \{\Theta, \Pi\} = \{\mu_j, \sigma_j, \gamma_j\}$ 去最大化对数似然函数(2.20)。由于对数函数是单调递增的函数，因此，可以考虑负对数似然函数作为损失函数^[34]：

$$J(\Theta, \Pi | X) = -L(\Theta, \Pi | X) = -\sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_j P(x_i | \Theta_j) \right\} \quad (2.50)$$

应用于完整的数据，最小化上式的负对数似然函数等价于最小化损失函数 $E(\Xi^{(0)} | \Xi^{(t+1)})$ ：

$$E(\Xi^{(0)} | \Xi^{(t+1)}) = -\sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(0)} \log \left[\pi_j^{(t+1)} P(x_i | \Theta_j^{(t+1)}) \right] \quad (2.51)$$

其中， $z_{ij}^{(0)} = \frac{\pi_j^{(0)} P(x_i | \Theta_j^{(0)})}{\sum_{l=1}^K \pi_l^{(0)} P(x_i | \Theta_l^{(0)})}$ 表示当前步迭代的后验概率，与式(2.45)相同。

这里可以采用梯度下降算法调整参数 $\Xi = \{\Theta, \Pi\} = \{\mu_j, \sigma_j, \gamma_j\}$ 用来最小化损失函数，参数通过下面的规则进行更新：

$$\Xi^{(t+1)} = \Xi^{(t)} - \eta \nabla E(\Xi^{(t)}) \quad (2.52)$$

其中， $\nabla E(\Xi) = \left[\frac{\partial E}{\partial \mu_j}, \frac{\partial E}{\partial \sigma_j}, \frac{\partial E}{\partial \gamma_j} \right]$ ， η 是学习率，也叫步长，一般都是一个非常小的值。损失函数对各个参数的偏导数为：

$$\frac{\partial E}{\partial \mu_j} = -\sum_{i=1}^N z_{ij}^{(0)} \frac{x_i - \mu_j}{\sigma_j^2} \quad (2.53)$$

$$\frac{\partial E}{\partial \sigma_j} = -\sum_{i=1}^N z_{ij}^{(0)} \left(-\frac{1}{\sigma_j} + \frac{(x_i - \mu_j)^2}{\sigma_j^3} \right) \quad (2.54)$$

$$\frac{\partial E}{\partial \gamma_j} = -\sum_{i=1}^N \left(z_{ij}^{(0)} - \pi_j^{(0)} \right) \quad (2.55)$$

对于高斯混合模型的梯度下降算法总结如下：

第 1 步 初始化参数 $\Xi = \{\Theta, \Pi\} = \{\mu_j, \sigma_j, \gamma_j\}$ 。

第 2 步 使用当前参数计算式(2.45)的值。

第 3 步 通过式(2.52)重新估计参数 $\mu_j, \sigma_j, \gamma_j$ 。

第 4 步 计算对数似然函数(2.20)，检测该对数似然函数或者其参数是否收敛，若没满足收敛条件，则返回第 2 步。

在参数学习阶段完成以后，通过后验概率 z_{ij} 给每个像素进行类标记分配。对于每个像素 x_i ，通过 z_{ij} 给出其对应的每个类标记的后验概率。通过最大后验概率准则将 N 个像素的图像分割成 K 个标记(类别)，具体表示为：

$$x_i \in \Omega_j : \text{If } z_{ij} \geq z_{ik}, j, k = 1, 2, \dots, K \quad (2.56)$$

2.5 本章小结

这一章首先较为详细的介绍了高斯混合模型及其图像分割，并探讨了高斯混合模型的一些优点和缺点，高斯混合模型比起其他模型来具有形式较为简单，对参数的需求比较少，比较容易实现；缺点是高斯混合模型是对每个像素单独进行考虑，然而一张图像的某些区域是相似的不能仅仅基于特征属性去给像素进行分类，因此高斯混合模型对含噪声的图像分割较为敏感。接着介绍了马尔科夫随机场，将马尔科夫随机场分布引入到混合模型的先验概率中，比起高斯混合模型，引入马尔科夫随机场到混合模型的先验概率的方法考虑了邻域像素之间的空间关系，而将马尔科夫随机场引入到高斯混合模型的方法及其对现有方法的一些改进将是我们下一步研究的重点。对高斯混合模型的先验概率进行约束，其每个像素的先验分布 π_{ij} 是不同的，是依据像素的邻域空间信息和对应的参数决定的，该方法不像高斯混合模型一样对噪声很敏感，其对含噪声的图像分割的敏感度降低了，但是这个模型方法比起高斯混合模型来形式非常复杂，计算量很大，并且要求的

参数也比较多。接着介绍了两种模型参数估计的两种方法，EM 算法和梯度下降算法，介绍这两种方法在高斯混合模型中的应用。

3 高斯混合模型的图像分割及其先验概率约束

图像分割的其目的是根据图像的区域满足一定的灰度、纹理等特征的相似性准则，划分图像的各个特征区域并提取出感兴趣的目标。之前许多学者提出了很多图像分割的方法，例如基于阈值的方法，基于 K-means 聚类方法等。然而这些方法对噪声都比较敏感，这是由于这些方法没有考虑到像素的空间信息。在这些分割方法中，贝叶斯方法被广泛的应用于不确定的概率模型中。

高斯混合模型(GMM)是一种基于贝叶斯的方法，被许多学者用于研究图像分割领域的许多关键问题。该模型易于实现、参数较少，并且其对数似然函数参数估计也比较简单，但是该模型没有考虑空间因素导致对噪声比较敏感。为了解决该问题而提出的考虑空间关系的马尔科夫随机场 (MRF) 模型在图像分析、图像恢复和图像分割中被广泛使用。

这一章我们通过对像素的邻域信息进行考虑，将像素的邻域信息引入到高斯混合模型的先验概率分布中，对先验概率进行约束，像素的先验概率分布由邻域像素的空间信息来决定。将邻域像素之间的空间信息引入到模型的先验概率分布中的主要目的是，使得分割图像中在空间上连续的区域被分割到同一类别中的可能性更大，这样可以抑制由于噪声干扰导致的空间上像素的先验概率分布的不连续和不平滑，提升高斯混合模型的抗噪性。本章通过对于高斯混合模型的空间约束进行建模的第一种途径的一个方法——扩展的高斯混合模型^[42]进行研究，进而提出一种利用像素在邻域空间内的均值取代像素本身的方法构造先验概率分布，将邻域权重函数的均值作用到先验概率分布中，提高模型的抗噪性，通过梯度下降算法进行模型参数的估计，提高模型的抗噪性。

3.1 扩展的高斯混合模型

在高斯混合模型对于先验概率分布约束的两条途径的其中第一条途径中的其中一个方法就是扩展的高斯混合模型^[42]，在这个模型中，首先定义一个表示像素 x_i 属于 Ω_j 类的权重的函数：

$$\xi_j(x_i) = \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \quad (3.1)$$

这里 c_j 和 b_j ， $j = (1, 2, \dots, K)$ 是参数，可以用梯度下降算法最大化该模型的似然函数获取该参数的最优值。对于像素 x_i 的邻域 N_i ， Ω_j 类的权重函数定义为：

$$g_j(x_i) = \left[\sum_{m \in N_i} \exp \left(-\frac{(x_m - c_j)^2}{2b_j^2} \right) \right]^\alpha \quad (3.2)$$

这里 α 为参数, N_i 为像素 x_i 的邻域 (该模型的 N_i 为 5×5 邻域)。将邻居像素之间的空间关系结合到先验概率分布 π_{ij} 为:

$$\pi_{ij} = \frac{g_j(x_i)}{\sum_{k=1}^K g_k(x_i)} \quad (3.3)$$

由上式可以看出对于图像中每个像素对应的每个标记的先验分布有不同的值, 上式的先验分布 π_{ij} 满足限制条件:

$$0 \leq \pi_{ij} \leq 1, \quad \sum_{j=1}^K \pi_{ij} = 1$$

x_i 的密度函数为:

$$f(x_i | \Pi, \Theta) = \sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \quad (3.4)$$

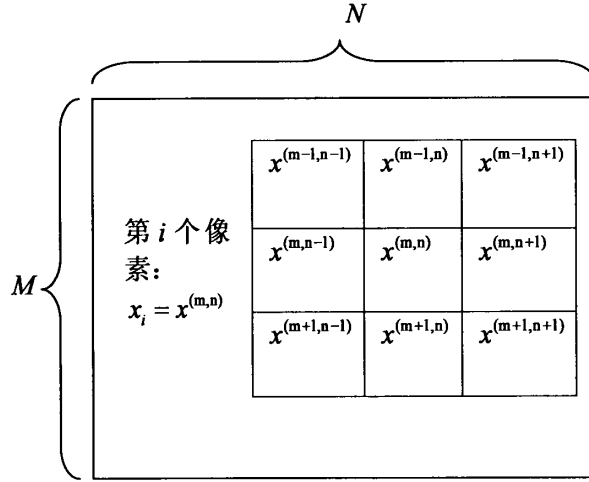
其对数似然函数为:

$$L(\Theta, \Pi | X) = \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right\} \quad (3.5)$$

当 $\pi_{ij} = \pi_{cj}$ 时, $\forall i, c = (1, 2, \dots, N)$, (3.5) 式的对数似然函数与 (2.20) 式是一样的。可以说高斯混合模型是拓展的高斯混合模型的一种特殊情形。结合 (3.2) 与 (3.3) 式, 我们可以看到先验分布 π_{ij} 近似于一个均值滤波器, 因此, 使用拓展的高斯混合模型的图像分割的结果对于噪声应该是具有鲁棒性的、具有抑制作用的。之后, 该模型的似然函数采用了梯度下降算法代替期望最大化算法 (EM) 进行计算并且估计该模型的参数, 从而进一步地提高了模型的效率。

3.2 基于空间邻域关系的高斯混合模型

根据扩展的高斯混合模型我们提出的这个基于空间邻域关系的高斯混合模型是利用像素在邻域空间内的均值取代像素本身的值从而构造先验概率分布, 对于一副 $M \times N$ 的灰度图像, 为了便于理解, 用 $x^{(m,n)}$ 表示图像的第 i 个像素 x_i , 其中 $i = M(m-1) + n$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$ 。选择 3×3 邻域, 如图 3.1 所示。

图 3.1 第 i 个像素的 3×3 邻域Fig 3.1 3×3 window within the neighborhood of the i -th pixel

对于第 i 个邻域，用 $W^{(m,n)}$ 表示该邻域像素的集合为：

$$W^{(m,n)} = (x^{(m-1,n-1)}, x^{(m-1,n)}, x^{(m-1,n+1)}, x^{(m,n-1)}, x^{(m,n)}, x^{(m,n+1)}, x^{(m+1,n-1)}, x^{(m+1,n)}, x^{(m+1,n+1)})$$

令 $\overline{x^{(m,n)}}$ 表示邻域内所有像素的均值：

$$\overline{x_i} = \overline{x^{(m,n)}} = \frac{1}{9} \sum_{v=-1}^1 \sum_{u=-1}^1 x^{(m+v,n+u)} \quad (3.6)$$

定义 $\xi_j(\overline{x_i})$ ，表示每个邻域是属于每个类别 Ω_j 的权重：

$$\xi_j(\overline{x_i}) = \xi_j(\overline{x^{(m,n)}}) = \exp \left(- \frac{(\overline{x^{(m,n)}} - c_j)^2}{2b_j^2} \right) \quad (3.7)$$

其中 b_j 、 c_j ($j=1,2,\dots,K$) 是两个参数，我们在后面去获取该参数的最优值，对于第 i 个邻域而言，邻域权重 $\xi_j(\overline{x_i})$ 的均值可以表示如下：

$$\overline{\xi_j(\overline{x_i})} = \overline{\xi_j(\overline{x^{(m,n)}})} = \frac{1}{9} \sum_{v=-1}^1 \sum_{u=-1}^1 \exp \left(- \frac{(\overline{x^{(m+v,n+u)}} - c_j)^2}{2b_j^2} \right) \quad (3.8)$$

将邻域像素之间的空间关系结合到先验概率分布 π_{ij} ，对于每个像素对应于每个类别 Ω_j ，先验概率分布具有不同的值，表示如下：

$$\pi_{ij} = \frac{\overline{\xi_j(\overline{x_i})}}{\sum_{k=1}^K \overline{\xi_k(\overline{x_i})}} \quad (3.9)$$

这种将邻居像素之间的空间关系结合到先验概率分布的方法的核心思想是：相邻的像素的先验概率分布更可能是趋于相同或者相似的。基于这种思想，在该方法中每个像素本身被像素在邻域中的均值所取代来构造权重函数，像素权重函数被邻域权重函数的均值所取代构成先验概率分布的主要内容。

先验概率分布满足：

$$0 \leq \pi_{ij} \leq 1, \quad \sum_{j=1}^K \pi_{ij} = 1 \quad (3.10)$$

x_i 的密度函数为：

$$f(x_i | \Pi, \Theta) = \sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \quad (3.11)$$

高斯分布 $P(x_i | \Theta_j)$ 为：

$$P(x_i | \Theta_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}\right) \quad (3.12)$$

对于给定的先验概率分布(3.9)，对数似然函数为：

$$L(\Theta, \Pi | X) = \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right\} \quad (3.13)$$

接下来为了最大化对数似然函数(3.13)需要优化参数 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$ ，由于对数函数是单调递增的函数，我们可以考虑求似然函数的负对数^[34]：

$$J(\Theta, \Pi | X) = -L(\Theta, \Pi | X) = -\sum_{i=1}^N \log \left(\sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \right) \quad (3.14)$$

应用于完整数据^[43]，得到新旧参数所产生的损失函数为：

$$J(\Theta^{(t+1)}, \Pi^{(t+1)} | X) - J(\Theta^{(t)}, \Pi^{(t)} | X) = -\sum_{i=1}^N \log \left(\frac{\sum_{j=1}^K \pi_{ij}^{(t+1)} P(x_i | \Theta_j^{(t+1)})}{\sum_{k=1}^K \pi_{ik}^{(t)} P(x_i | \Theta_k^{(t)})} \times \frac{z_{ij}^{(t)}}{z_{ij}^{(t+1)}} \right) \quad (3.15)$$

上式的 $z_{ij}^{(t)}$ 与(2.38)一样，总是满足条件：

$$0 \leq z_{ij}^{(t)} \leq 1, \quad \sum_{j=1}^K z_{ij}^{(t)} = 1$$

在此处介绍琴生不等式^[44](Jensen's inequality)，琴生不等式表述为：一个数据集满足 $\lambda_j \geq 0$ 且 $\sum_{j=1}^K \lambda_j = 1$ ，则有：

$$\log \left(\sum_{j=1}^K \lambda_j y_j \right) \geq \sum_{j=1}^K \lambda_j \log(y_j) \quad (3.16)$$

由琴生不等式得到:

$$J(\Theta^{(t+1)}, \Pi^{(t+1)} | X) - J(\Theta^{(t)}, \Pi^{(t)} | X) \leq - \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \log \left(\frac{\pi_{ij}^{(t+1)} P(x_i | \Theta_j^{(t+1)})}{z_{ij}^{(t)} \sum_{k=1}^K \pi_{ik}^{(t)} P(x_i | \Theta_k^{(t)})} \right) \quad (3.17)$$

丢掉在 t 步迭代处的旧的数的项, 损失函数可以表示如下:

$$E(\Theta^{(t)}, \Pi^{(t)} | \Theta^{(t+1)}, \Pi^{(t+1)}) = - \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \log(\pi_{ij}^{(t+1)} P(x_i | \Theta_j^{(t+1)})) \quad (3.18)$$

因此最大化似然函数(3.13)等价于最小化损失函数(3.18)。下面讨论最小化损失函数及其参数的调整 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$, $j = 1, 2, \dots, K$, 最后使用后验概率 z_{ij} 给每个像素分配标记:

$$z_{ij} = \frac{\pi_{ij} P(x_i | \Theta_j)}{\sum_{k=1}^K \pi_{ik} P(x_i | \Theta_k)} \quad (3.19)$$

3.3 参数学习

对于每个像素 x_i 应该分配给哪个标记 Ω_j 的问题是一个后验概率 z_{ij} 的问题, 为了得到后验概率 z_{ij} 我们需要优化参数 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$, $j = 1, 2, \dots, K$, 以实现最小化损失函数(3.18)或者最大化似然函数(3.13)。

这里使用梯度下降算法^[34]调整参数以实现损失函数的最小化, 梯度下降算法总结如下:

第 1 步 使用 K-means 算法初始化参数均值 μ_j 和方差 σ_j , $c_j = \mu_j$, $b_j = \sigma_j$, 这时获得初始化参数 $\Theta^{(0)}$ 。

第 2 步 利用式(2.18)与(3.9)分别计算高斯分布 $P(x_i | \Theta_j^{(0)})$ 和先验概率分布 $\pi_{ij}^{(0)}$, 接着计算后验概率 $z_{ij}^{(0)}$:

$$z_{ij}^{(0)} = \frac{\pi_{ij}^{(0)} P(x_i | \Theta_j^{(0)})}{\sum_{k=1}^K \pi_{ik}^{(0)} P(x_i | \Theta_k^{(0)})}$$

第 3 步 根据梯度下降方法, 更新参数 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$, 通过下面的规则获得新的参数 $\Theta^{(t+1)}$:

$$\Theta^{(t+1)} = \Theta^{(t)} - \eta \nabla E(\Theta^{(t)})$$

其中, η 为学习率, 也叫步长, 控制每次下降的大小, 一般学习率是一个比较小的值, 在该方法中学习率设置为 $\eta = 10^{-7}$ 。 $\nabla E(\Theta^{(t)})$ 表示当前参数的梯度, 且:

$$\nabla E(\Theta^{(t)}) = \left[\frac{\partial E}{\partial u_j}, \frac{\partial E}{\partial \sigma_j}, \frac{\partial E}{\partial c_j}, \frac{\partial E}{\partial b_j} \right] \quad (3.20)$$

第 4 步 检测负对数似然函数或者参数值是否收敛, 若没收敛则将参数设置为当前 $t+1$ 步的参数 ($\Theta^{(t)} = \Theta^{(t+1)}$), 返回第 2 步。

第 3 步中损失函数关于参数的梯度的具体表达式参见附录。本章节的伪代码实现如表 3.1 所示:

表 3.1 伪代码实现

Table 3.1 Pseudo code implementation

伪代码实现:
读入 N 个像素的图像; 输入 K 个类别; FOR(遍历 N 个像素点){ FOR(遍历 K 个类别){ 使用 K-means 算法初始化参数 μ_j 、 σ_j ; $c_j = \mu_j$, $b_j = \sigma_j$; do{ 利用式(2.18)与(3.9)分别计算 $P(x_i \Theta_j^{(t)})$ 和 $\pi_{ij}^{(t)}$; 通过(3.19)计算后验概率 $z_{ij}^{(t)}$; 更新参数 $\Theta^{(t+1)} = \Theta^{(t)} - \eta \nabla E(\Theta^{(t)})$; }while(参数不收敛) 计算后验概率 z_{ij} ; } Max(z_{ij}); 将第 i 个像素分配给第 j 个类别; } 分割并输出图像; END;

3.4 本章小结

这一章主要介绍了高斯混合模型中对于先验概率分布约束的两条途径中的其中第一条途径，利用空间平滑方法，将邻域像素之间的空间关系引入到高斯混合模型的先验概率分布中去，对先验概率进行约束。首先介绍了一种拓展的高斯混合模型，该模型采用了一种新的方法将邻域像素之间的空间关系引入到高斯混合模型的先验概率分布中。之后提出了一种利用像素在邻域空间内的均值取代像素本身的值的方法构造先验概率分布——基于空间邻域关系的高斯混合模型，提高模型的抗噪性，该模型最后使用梯度下降算法对模型的似然函数的参数进行估计。并给出该模型方法的伪代码实现。

4 基于高斯核函数的先验概率约束的高斯混合模型

4.1 高斯混合模型及图像分割

有限混合模型(FMM)作为一种无监督的软划分的聚类方法被提出,在最近的几十年里,有限混合模型(FMM)已经被广泛的研究,并且应用在各种类型的图像分割当中,包括自然图像分割、医学图像分割以及视频监控中移动景物分割等。

有限混合模型^[46](FMM)由两部分组成:第一部分利用贝叶斯定理计算每个像素的类标记的概率,也就是像素的后验概率,以此对像素进行类标记的分配。其中,先验分布是按照经验或者人为产生的一组概率去预先定义每个像素归属于某个类别的可能性,我们一般是利用 K 均值算法产生初始的先验概率分布。对于不同的数据分布,在有限混合模型中可以采用不同的混合条件概率密度函数,譬如,高斯混合模型(GMM)采用高斯函数作为混合条件密度函数,而非高斯混合模型则采用非高斯分布的条件概率密度函数。后验概率则由贝叶斯公式将先验概率分布和条件概率密度函数进行计算得出,后验概率则是直接指定每个像素点到某个标记类别的决定因素。第二部分则是利用期望最大化算法最大化似然函数,或是使用梯度下降算法求解负对数似然函数。

在这个章节我们提出了一种基于高斯核函数的先验概率约束的高斯混合模型,图像中的像素,若它们的距离越是接近那么它们的先验概率分布就可能越相似甚至相同。基于此我们将高斯核函数^{[47][48][49]} (高斯径向基函数)作为对像素先验概率分布的空间约束引入到高斯混合模型中对先验概率分布进行约束,并且根据扩展的高斯混合模型引入每个像素属于每个类别的权重函数一起对先验概率分布进行约束。最后通过梯度下降算法对参数进行优化。

在一副图像中, $x_i, i=1,2,\dots,N$ 表示图像中第 i 个像素的观测值(或随机变量), $\Omega_1, \Omega_2, \dots, \Omega_K$ 表示 K 个标记。高斯混合模型中像素 x_i 的密度函数为:

$$f(x_i | \Pi, \Theta) = \sum_{j=1}^K \pi_{ij} P(x_i | \Theta_j) \quad (4.1)$$

其中 π_{ij} 为像素 x_i 属于标记 Ω_j 的成分分布混合的先验概率, Θ_j 为第 j 个成分分布的参数, $\Theta = \{\Theta_j, j=1,2,\dots,K\}$ 是所有参数的集合, $P(x_i | \Theta_j)$ ($j=1,2,\dots,K$) 是第 j 类对应的概率密度函数。

高斯分布 $P(x_i | \Theta_j)$ 为:

$$P(\mathbf{x}_i | \Theta_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(\mathbf{x}_i - \mu_j)^2}{2\sigma_j^2}\right) \quad (4.2)$$

先验概率分布满足条件:

$$0 \leq \pi_{ij} \leq 1, \quad \sum_{j=1}^K \pi_{ij} = 1 \quad (4.3)$$

对于先验概率分布 π_{ij} , 其对数似然函数为:

$$L(\Theta, \Pi | X) = \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} P(\mathbf{x}_i | \Theta_j) \right\} \quad (4.4)$$

根据琴生不等式(Jensen's inequality), 对数似然函数(4.4)可以得到:

$$\begin{aligned} L(\Theta, \Pi | X) &= \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} P(\mathbf{x}_i | \Theta_j) \right\} \\ &\geq \sum_{i=1}^N \sum_{j=1}^K \pi_{ij} \log P(\mathbf{x}_i | \Theta_j) \end{aligned} \quad (4.5)$$

令 $L^*(\Theta, \Pi | X)$ 为:

$$L^*(\Theta, \Pi | X) = \sum_{i=1}^N \sum_{j=1}^K \pi_{ij} \log P(\mathbf{x}_i | \Theta_j) \quad (4.6)$$

从式(4.5)可以看到, 最大化对数似然函数 $L(\Theta, \Pi | X)$ (4.4) 等价于最大化对数似然函数 $L^*(\Theta, \Pi | X)$ (4.6)。式(4.6)是一个关于参数 π_{ij} 和 Θ_j 的对数似然函数。

4.2 基于高斯核函数的先验概率约束的高斯混合模型

在这个模型中, 首先定义一个表示像素 \mathbf{x}_i 属于 Ω_j 类的权重的函数^[42]:

$$\xi_j(\mathbf{x}_i) = \exp\left(-\frac{(\mathbf{x}_i - \mathbf{c}_j)^2}{2b_j^2}\right) \quad (4.7)$$

其中 \mathbf{c}_j 和 b_j , $j = (1, 2, \dots, K)$ 是第 j 类的参数。将高斯核函数^{[46][47]}(高斯径向基函数)作为对像素先验概率分布的空间约束引入到高斯混合模型中对先验概率分布进行约束, 先验概率分布为:

$$\pi_{ij} = \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \xi_j(\mathbf{x}_q) \quad (4.8)$$

其中:

$$Z_i = \sum_{q \in N_i} h(i, q) \quad (4.9)$$

是归一化因子, $h(i, q)$ 表示像素 i 和像素 q 之间的几何距离^{[48][49]}(欧式距离)。几何距离 h 是像素 i 到像素 q 的相对位置向量的幅度的高斯函数, $\|u_i - u_q\|$ 。高斯核函数 $h(i, q)$ 表示如下:

$$h(i, q) = \exp\left(\frac{-\|u_i - u_q\|^2}{2\sigma_g^2}\right) \quad (4.10)$$

其中, σ_g 是函数的宽度参数, 控制了相邻像素之间的径向作用范围, u_i 和 u_q 分别是像素 i 和像素 q 的位置。在这个模型里 σ_g 取值为 10, 从式(4.10)可以看出当距离 $\|u_i - u_q\|$ 增加时, 函数 h 会减小。如果两个像素之间的几何距离越近, 那么这两个像素属于同一类别的概率就越大; 而两个像素之间的几何距离越远, 那么这两个像素属于同一类别的概率就越小。基于这种思想我们引入了高斯核函数, 同时对于(4.18)式也能保证先验概率 π_{ij} 满足条件(4.3)。对于给定的先验概率分布(4.8), 对数似然函数(4.6)可以写成:

$$\begin{aligned} L^*(\Theta, \Pi | X) &= \sum_{i=1}^N \sum_{j=1}^K \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \xi_j(x_q) \log P(x_i | \Theta_j) \\ &= \sum_{i=1}^N \sum_{j=1}^K \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \log P(x_i | \Theta_j) \end{aligned} \quad (4.11)$$

参数 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$, $j=1, 2, \dots, K$ 可以通过最大化对数似然函数(4.11)来估计, 由于对数函数是单调递增的函数, 我们可以考虑求似然函数的负对数为:

$$E(\Theta, \Pi | X) = - \sum_{i=1}^N \sum_{j=1}^K \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \log P(x_i | \Theta_j) \quad (4.12)$$

现在最大化对数似然函数(4.11)等价于最小化负对数似然函数(4.12), 下面我们使用梯度下降算法调整参数 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$, $j=1, 2, \dots, K$, 使得负对数似然函数(4.12)最小化。梯度下降算法总结如下:

第 1 步 使用 K-means 算法初始化参数均值 μ_j 和方差 σ_j , $c_j = \mu_j$, $b_j = \sigma_j$, 这时获得初始化参数 $\Theta^{(0)}$ 。

第 2 步 利用式(2.18)与(4.8)分别计算高斯分布 $P(x_i | \Theta_j^{(0)})$ 和先验概率分布 $\pi_{ij}^{(0)}$, 接着计算后验概率 $z_{ij}^{(0)}$:

$$z_{ij}^{(0)} = \frac{\pi_{ij}^{(0)} P(x_i | \Theta_j^{(0)})}{\sum_{k=1}^K \pi_{ik}^{(0)} P(x_i | \Theta_k^{(0)})}$$

第 3 步 根据梯度下降方法,更新参数 $\Theta = (\mu_j, \sigma_j, c_j, b_j)$, 通过下面的规则获得新的参数:

$$\Theta^{(t+1)} = \Theta^{(t)} - \eta \nabla E(\Theta^{(t)}) \quad (4.13)$$

其中, η 为学习率, 也叫步长, 控制每次下降的大小, 一般学习率是一个比较小的值, 本文中学习率设置为 $\eta = 10^{-7}$ 。 $\nabla E(\Theta^{(t)})$ 表示当前参数的梯度, 且:

$$\nabla E(\Theta^{(t)}) = \left[\frac{\partial E}{\partial \mu_j}, \frac{\partial E}{\partial \sigma_j}, \frac{\partial E}{\partial c_j}, \frac{\partial E}{\partial b_j} \right] \quad (4.14)$$

第 4 步 检测负对数似然函数或者参数值是否收敛, 若没收敛则将参数设置为当前 $t+1$ 的参数 ($\Theta^{(t+1)} = \Theta^{(t+1)}$), 返回第 2 步。

关于第 3 步中负对数似然函数关于参数的梯度的具体表达式参见附录。

在估计完参数之后, 通过计算得到的后验概率 z_{ij} 给每个像素进行类标记分配。对于每个像素 x_i , 通过 z_{ij} 给出其对应的每个类标记的后验概率。通过最大后验概率(MAP)准则将 N 个像素的图像分割成 K 个标记(类别), 具体表示为(2.56)式。本章节的伪代码的实现于第三章类似, 只是最优化的目标函数不一样, 所以在这里不再列出伪代码。

4.3 实验结果及分析

传统的高斯混合模型采用 EM 算法对似然函数进行最大化, EM 算法表述如下:

表 4.1 期望最大化算法描述

Table 4.1 The description of Expectation-Maximization algorithm

期望最大化算法:
输入图像;
初始化参数 $\{\Theta, \Pi\} = \{\mu_j, \sigma_j, \pi_j\}$;
do {
通过(2.45)计算后验概率的值;
通过(2.46)、(2.47)、(2.48)式更新参数 μ_j, σ_j, π_j ;
计算对数似然函数;
}while(对数似然函数不收敛)
通过(2.45)利用最大后验概率准则对像素进行分配;
输出分割结果图;
END;

为了验证具有空间信息约束的高斯混合模型对含有噪声的图像的分割的有效性，采用了人工合成图像进行实验，实验对比了传统的高斯混合模型(GMM)以及高斯核函数约束的高斯混合模型(RBFGMM)。梯度下降算法伪代码如第三章表 3.1 所示，实验结果如图 4.1、4.2 所示。

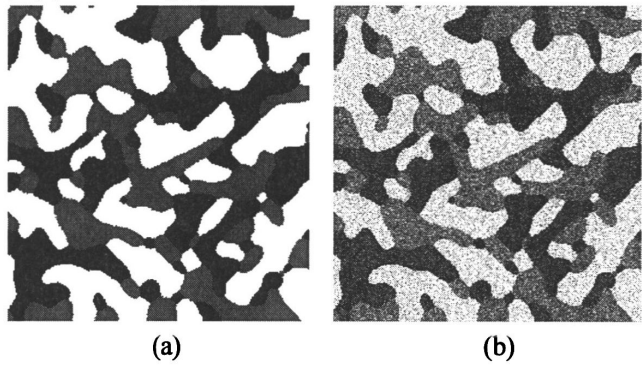


图 4.1 人工合成图像的源图像和其受高斯噪声 Gaussian(0,0.02)干扰的图像

Fig 4.1 Synthetic image and it corrupted original image with Gaussian noise (0 mean, 0.02 variance)

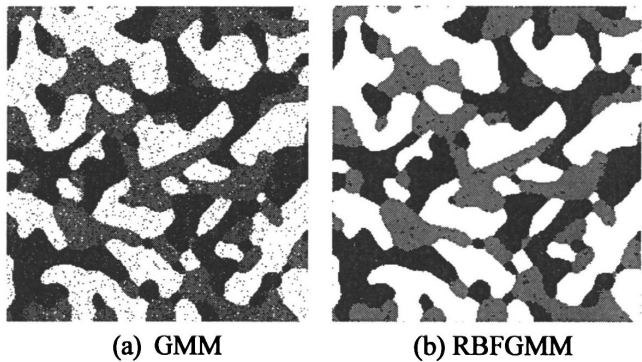


图 4.2 分割结果图

Fig 4.2 Segmentation results

该实验是在 Core i5 2.8GH 处理器、8G 内存的机器 Windows 环境下用 MATLAB 实现，人工合成源图像具有灰色、白色和黑色三个类别 ($K=3$)。图 4.1 中(a)为源图像，(b)为加高斯噪声 Gaussian(0,0.02)的图像。图 4.2 中(a)为用传统的高斯混合模型进行分割所得的图像，(b)为用高斯核函数进行约束的高斯混合模型进行分割所得的图像。传统的高斯混合模型与 RBFGMM 模型进行实验结果的比较，从实验结果可以看出，没有利用高斯混合模型进行含噪声图像的分割会有许多错误分割的像素，例如许多灰色的像素被分割到了白色的区域；而利用具有空间约束的 RBFGMM 进行分割的结果可以看到这种类型的错误分割像素会减少许多。这正是由于传统的高斯混合模型只是单独的对像素点进行考虑而没有进行一些空间信息方面的约束导致的。在本文中采用误分率(MCR Misclassification Ratio)^[45]表示分割

的精度，误分率越低表示错误分割的像素越少，分割的精度就越高，误分率(MCR)表示为：

$$MCR = \frac{\text{错误分割的像素数目}}{\text{总像素数目}} \times 100\%$$

(4.14)

传统的高斯混合模型与 RBFMM 模型对于含不同噪声水平, 的图像分割的误分率见表 4.2。

表 4.2 高斯混合模型(GMM)，RBFMM，对不同噪声水平的误分率 MCR(%)

Table 4.2 MCR(%) for different noise levels of GMM, RBFMM

Gaussian noise (0 mean, var)	GMM	RBFMM
var=0.01	11.41	1.08
var=0.02	19.97	2.12
var=0.03	25.33	4.29
var=0.04	30.45	7.42
var=0.05	32.98	9.65

从表 4.2 可以看到，随着噪声强度的增加，GMM 与 RBFMM 对图像的误分率也随之增加，但对于同一噪声水平而言，带有空间信息约束的 RBFMM 要比传统的 GMM 的误分率要低，抗噪性比 GMM 强。

4.4 本章小结

在这个章节里面我们提出了一种基于高斯核函数的先验概率约束的高斯混合模型，图像中的像素，若它们的距离越是接近那么它们的先验概率分布就可能越相似甚至相同。基于此我们将高斯核函数(高斯径向基函数)作为对像素先验概率分布的空间约束引入到高斯混合模型中对先验概率分布进行约束，并且根据扩展的高斯混合模型引入每个像素属于每个类别的权重函数一起对先验概率分布进行约束。推导出该模型的对数似然函数并对似然函数取负对数，最后通过梯度下降算法对参数进行优化。对传统的高斯混合模型以及 RBFMM 进行人工合成图像的实验，实验结果表明高斯混合模型中对先验概率进行空间方面的约束确实能取得比较好的效果。

5 总结与下一步工作

5.1 总结

随着对图像分割方法的研究的不断深入, 基于混合模型, 特别是高斯混合模型的聚类分析的图像分割方法已经渐渐成为解决图像分割问题的重要组成部分, 这种方法的主要特点是根据图像像素之间特征的相似性进行图像分割。基于高斯混合模型的聚类分析方法可以解决存在不确定性分类以及具有复杂内容的图像分割的问题, 从而得到更加令人满意的图像分割结果。可是基于高斯混合模型的图像分割方法只单独考虑了像素的统计信息, 但是忽略了像素之间的空间信息, 从而导致了使用该方法进行图像分割得到的结果对于噪声的干扰异常敏感。

本文首先对当前的图像分割的各种方法进行介绍, 并对基于高斯混合模型的图像分割方法进行了研究, 介绍了一些概率论的知识, 高斯分布, 高斯混合模型, 高斯混合模型和图像分割, 介绍马尔科夫随机场以及将马尔科夫随机场引入到高斯混合模型的先验概率分布。介绍了两种进行参数估计的方法, 期望最大化(EM)算法和梯度下降算法。并对高斯混合模型进行图像分割的优点以及缺点进行阐述。

其次介绍了进行先验概率约束的第一条途径的一种方法, 一种拓展的高斯混合模型。基于相邻的像素的先验概率分布更可能是趋于相似或者相同的思想, 提出了一种利用像素在邻域空间内的均值取代像素本身的方法构造先验概率分布, 将邻域权重函数的均值作用到先验概率分布中, 提高模型的抗噪性, 通过梯度下降算法进行模型参数的估计, 并给出了算法的伪代码实现。

最后将高斯核函数及拓展的高斯混合模型中每个像素属于每个类别的权重函数一起对像素的先验概率进行约束, 提出了一种基于高斯核函数的先验概率约束的高斯混合模型。最后分别通过人工合成图像进行实验分析, 结果表明高斯混合模型中对先验概率进行空间方面的约束确实能取得比较好的效果。

5.2 下一步工作

由于基于空间邻域的高斯混合模型在本文中只有理论上的说明与推导, 但还没有进行实验, 所以本文下一步首先需要对第三章的基于空间邻域的高斯混合模型进行实验, 并与其它模型进行比较, 以验证模型方法的真实有效性。

在本文的模型以及参数估计的方法中还有一些问题可以进行研究:

首先在梯度下降算法中,对于步长的选择可以进行适当的变动。在算法开始执行的时候让步长相对较大,之后每迭代一次步长就相对减小,可以通过最优化方法以及数学的一些知识可以在每次迭代之后计算出适合的步长,使得梯度下降算法的迭代次数更少就可以达到最优点。这应该是一个可以值得研究的问题。

其次在文献[33]中,该文献讲述了将马尔科夫随机场引入到高斯混合模型中对像素的先验分布进行约束的第二条途径。该文中作者在图像分割的过程中整幅图像都使用同一个常数 β 。常数 β 在分割过程中对于同一个区域来说对噪声有抑制作用,但是对不同区域的边界常数 β 将会产生不好的影响。对噪声水平比较高的图像,可能分割结果中不同区域的边界会变得模糊不清甚至被消除,因此在图像分割的过程中常数 β 应该具有不同的值,这样效果可能会更好。因此对于将马尔科夫随机场引入到高斯混合模型中对像素的先验分布进行约束的第二条途径的方法进行研究,以及文献[33]中常数 β 的取值问题也将是我们下一步将要研究的内容。

参考文献

- [1] 章毓晋.图像理解与计算机视觉[M].北京: 清华大学出版社, 2000:4
- [2] Gonzalez R C, Wood R E. Digital Image Processing.3rd.ed, Addisonwesley, 1992:40-48
- [3] Ahmed M. N., Yamany S. M., Mohamed N., Farag A. A., and Moriarty T., "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data," IEEE Transactions on Medical Imaging, vol.21, no. 3, pp. 193-199, 2002.
- [4] Ashburner J., and Friston K. J., "Unified segmentation," NeuroImage, vol. 26, no. 3, pp. 839-851, 2005.
- [5] Thanh N. M., and Wu Q. M. J., "Robust Student's-t Mixture Model with Spatial Constraints and Its Application in Medical Image Segmentation," IEEE Transactions on Medical Imaging, Accepted, 2011.
- [6] Fionn M., Adrian E. R., and Starck J. L., "Bayesian inference for multiband image segmentation via model-based cluster trees," Image and Vision Computing, vol. 23, no. 6, pp. 587-596, 2005.
- [7] 章毓晋.图像分割[M].北京: 科学出版社, 2001.
- [8] D. Man', E. Hildreth. Theory of edge detection[J]. Proc. Roy. Soc. London B, 1980, 207(1167):187-217.
- [9] I. Sobel. Camera models and machine perception [D]. Stanford University, Starldford CA, 1970.
- [10] J. Canny. Finding edges and lines in images [R]. MIT Artificial Intelligence Lab, Cambridge MA, Technology Report, 1983.
- [11] J. Olszewska, C. De Vleeschouwer, B. Macq. Speeded up gradient vector flow b-spline active contours for robust and real-time tracking[C]. IEEE Proc. Int. Conf. Acoustics, Speech and Signal Processing, 2007, 1:905-908
- [12] D. Cremers, M. Rousson, R. Dedche. A review of statistical approaches to level set Segmentation: integrating color, texture, motion and shape[J]. Int. J. Comput. Vis., 2007, 72(2): 195-215.
- [13] J. Prewitt, M. Mencilsohn. The analysis of cell images[J]. Annals of the New York Academy of Sciences, 1966, 128: 1035-1053.
- [14] J. Chassory, C. Garbay, G. Brugal. An iterative segmentation method based on a contextual color and shape criterion[J]. IEEE PAMI, 1984, 6(6): 794-800.
- [15] S. Beucher, F. Meyer. The morphological approach to segmentation: The watershed transformation[J]. Math. Morphol. Image Process., 1993, 34: 433-481.
- [16] 强柯南. 形态学分水岭结合谱聚类的图像分割算法. 长安大学. 2014.
- [17] R. Duda, P. Hart, D. Stork. Pattern classification[M]. Hoboken, NJ: Wiley, 2001.
- [18] J. MacQueen. Some methods for classification and analysis of multivariate observations[C]. Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, 1967, 1: 281-296.
- [19] J. Bezdek. Pattern recognition with fuzzy objective function algorithms[M]. Plenum, New York, 1981.

- [20] J. Bezdek. A convergence theorem for the fuzzy ISODATA clustering algorithm[J]. IEEE PAMI, 1980, 1(2): 1-8.
- [21] Z. Wu, R. Leahy. An optimal graph theoretic approach to data clustering: theory and its application to image segmentation[J]. IEEE Trans. Pattern Anal. Mach.Intell., 1993,15(11): 1101-1113.
- [22] Andrew R. Webb. 统计模式识别[M]. 电子工业出版社 2003.
- [23] 杨红喆, 赵立辉. “融合高斯混合模型的测地线脑肿瘤分割方法”, 计算机工程, 第 40 卷, 第 2 期, 2014 年 2 月.
- [24] Wenlong Fu, Mark Johnston and Mengjie Zhang. “Gaussian Mixture Models and Information for Image Segmentation using Particle Swarm Optimisation”, 28th International Conference on Image and Vision Computing New Zealand. pp. 328-333, 2013.
- [25] Sanjay G. S., and Hebert T. J., “Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm,” IEEE Transactions on Image Processing, vol. 7, no. 7, pp. 1014-1028, 1998.
- [26] LEIT, SEWCHAND W. Statistical approach to X-ray CT imaging and its applications in image analysis-Part 2: A new stochastic model-based image segmentation technique for X-ray CT images [J]. IEEE Transactions on Medical Image, 1992,11(1):62-69.
- [27] Fabien Salzenstein, and Wojciech Pieczynski, Parameter Estimation in Hidden Fuzzy Markov Random Fields and Image Segmentation, Graphical Models and Image Processing 1997, 59(4): 205-220.
- [28] Christophoros Nikou, Nikolaos P. Galatsanos, and Arirtidis C. Likas, “A Class Adaptive Spatially Variant Mixture Model for Image Segmentation”, IEEE Transactions on Image Processing, vol. 16, no. 4, pp. 1121-1130, April 2007.
- [29] Hui Tang, Jean-Louis Dillenseger, Xu Dong Bao, and Li Min Luo, “A vectorial image soft Segmentation method based on neighborhood weighted Gaussian mixture model”, Computerized Medical Imaging and Graphics, vol. 33, no. 8, pp. 644-650, December 2009.
- [30] Kyungsuk (Peter) pyun, Johan Lim, Chee Sun Won, and Robert M. Gray, “Image Segmentation Using Hidden Markov Gauss Mixture Models”, IEEE Transactions on Image Processing, vol. 16, no. 7, pp. 1902-1911, July 2007.
- [31] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, “A Spatially Constrained Mixture Model for Image Segmentation”, IEEE Transactions on Neural Networks, vol. 16, no. 2, pp. 494-498, March 2005.
- [32] Thanh Minh Nguyen, Q. M. Jonathan Wu, “Dirichlet Gaussian Mixture Model: Application to Image Segmentation”, Image and Vision Computing, vol. 29, no. 12, pp. 818-828, November 2011.
- [33] Thanh Minh Nguyen. Q. M. Jonathan Wu., “Fast and Robust Spatially Constrained Gaussian Mixture Model for Image Segmentation,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no.4, pp. 621-634, April 2013 .
- [34] Bishop C. M., “Pattern Recognition and Machine Learning”, Springer, 2006.
- [35] Tittertington D. M., Smith A. F. M., and Makov U. E., “Statistical Analysis of Finite Finite Mixture Distributions”, Hoboken, NJ: Wiley, 1985.
- [36] McLachlan G., and Peel D., “Finite Mixture Models”, New York, Wiley, 2000.

- [37] 张磊. 基于 TS-MRF 模型的图像分割. 西安电子科技大学. 2012.
- [38] Clifford P., "Markov random fields in statistics," *Disorder in Physical Systems, A Volume in Honour of John M. Hammersley on the Occasion of his 70th Birthday*, G. Grimmett and D. Welsh, Oxford, UK, Clarendon, 1990.
- [39] Besag J., "Statistical analysis of non-lattice data," *Journal of the Royal Statistical Society, Series D*, vol. 24, no. 3, pp. 179-195, 1975.
- [40] Li S. Z., "Markov Random Field Modeling in Image Analysis", Springer Verlag, 2009.
- [41] Bishop C. M., "Neural Networks for Pattern Recognition", Oxford University Press, Walton Street, Oxford, 1995.
- [42] Thanh Minh Nguyen, Q. M. Jonathan Wu, and Siddhant Ahuja, "An Extension of the Standard Mixture Model for Image Segmentation", *IEEE Transactions on Neural Networks*, vol. 21, no. 8, pp. 1326-1338, August 2010.
- [43] Titterton D. M., Smith A. F. M., and Makov U. E., "Statistical Analysis of Finite Mixture Distributions", Hoboken, NJ: Wiley, 1985.
- [44] Rudin W., "Real and Complex Analysis", Third Edition, McGraw-Hill, pp. 62-65, 1987.
- [45] Zhang Y., Brady M., and Smith S., "Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45-57, 2001.
- [46] G. J. McLachlan and D. Peel, "Finite Mixture Models", New York: Wiley, 2000.
- [47] Deng Cai, Xiaofei He, Jiawei Han, and Thomas S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 33, no. 8, pp. 1548-1560, 2011.
- [48] W. C. K. Wong and A. C. S. Chung, "Bayesian image segmentation using local iso-intensity structural orientation," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1512-1523, 2005.
- [49] 邹立颖, 郝冰, 沙丽娟. "基于快速高斯核函数模糊聚类算法的图像分割", *化工自动化及仪表*, 第 37 卷, 第 11 期, 2010 年 11 月.

附录

梯度表达式

基于空间邻域关系的高斯混合模型梯度下降算法中损失函数对各个参数的偏导数如下：

$$\frac{\partial E}{\partial u_j} = -\sum_{i=1}^N z_{ij}^{(0)} \frac{x_i - u_j}{\sigma_j^2} \quad (\text{A.1})$$

$$\frac{\partial E}{\partial \sigma_j} = -\sum_{i=1}^N z_{ij}^{(0)} \left(-\frac{1}{\sigma_j} + \frac{(x_i - u_j)^2}{\sigma_j^3} \right) \quad (\text{A.2})$$

$$\frac{\partial E}{\partial c_j} = -\sum_{i=1}^N \left(z_{ij}^{(0)} \frac{\delta_{ij}}{v_{ij}} \right) + \sum_{i=1}^N \sum_{k=1}^K \left(z_{ij}^{(0)} \frac{\delta_{ik}}{\sum_{l=1}^K v_{il}} \right) \quad (\text{A.3})$$

$$\frac{\partial E}{\partial b_j} = -\sum_{i=1}^N \left(z_{ij}^{(0)} \frac{\tau_{ij}}{v_{ij}} \right) + \sum_{i=1}^N \sum_{k=1}^K \left(z_{ij}^{(0)} \frac{\tau_{ik}}{\sum_{l=1}^K v_{il}} \right) \quad (\text{A.4})$$

其中 δ_{ij} , v_{ij} 和 τ_{ij} 分别为：

$$\begin{aligned} \delta_{ij} &= \sum_{v=-1}^1 \sum_{u=-1}^1 \left[\frac{(x^{(\overline{m+v, n+u})} - c_j)}{2b_j^2} \times \exp \left(-\frac{(x^{(\overline{m+v, n+u})} - c_j)^2}{2b_j^2} \right) \right] \\ v_{ij} &= \sum_{v=-1}^1 \sum_{u=-1}^1 \exp \left(-\frac{(x^{(\overline{m+v, n+u})} - c_j)^2}{2b_j^2} \right) \\ \tau_{ij} &= \sum_{v=-1}^1 \sum_{u=-1}^1 \left[\frac{(x^{(\overline{m+v, n+u})} - c_j)^2}{2b_j^3} \times \exp \left(-\frac{(x^{(\overline{m+v, n+u})} - c_j)^2}{2b_j^2} \right) \right] \end{aligned}$$

基于高斯核函数的先验概率约束的高斯混合模型梯度下降算法中负对数似然函数对各个参数的偏导数表达式如下：

$$\frac{\partial E}{\partial u_j} = -\sum_{i=1}^N \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \frac{x_i - u_j}{\sigma_j^2} \quad (\text{A.5})$$

$$\frac{\partial E}{\partial \sigma_j} = -\sum_{i=1}^N \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \left(-\frac{1}{\sigma_j} + \frac{(x_i - u_j)^2}{\sigma_j^3}\right) \quad (\text{A.6})$$

$$\frac{\partial E}{\partial c_j} = -\sum_{i=1}^N \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \left(\frac{x_i - c_j}{b_j^2}\right) \log P(x_i | \Theta_j) \quad (\text{A.7})$$

$$\frac{\partial E}{\partial b_j} = -\sum_{i=1}^N \sum_{q \in N_i} \frac{h(i, q)}{Z_i} \exp\left(-\frac{(x_i - c_j)^2}{2b_j^2}\right) \left(-\frac{1}{b_j} + \frac{(x_i - c_j)^2}{b_j^3}\right) \log P(x_i | \Theta_j) \quad (\text{A.8})$$

作者简历

一、作者简历

欧垚江，男，2012 级硕士研究生，就读于北京交通大学计算机与信息技术学院，计算机科学与技术专业。

教育经历：

2007.9-2011.7.中央民族大学，理学院，信息与计算科学专业，理学学士；

2012.9-至今.北京交通大学，计算机与信息技术学院，计算机科学与技术专业，工学硕士。

二、发表论文

[1]欧垚江，于双元.基于高斯混合模型的图像分割方法的研究[EB/OL].北京：中国科技论文在线[2015-04-24].

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，也不包含为获得北京交通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：欧燕江 签字日期：2015 年 6 月 11 日

学位论文数据集

表 1.1： 数据集页

关键词*	密级*	中图分类号	UDC	论文资助
图像分割；高斯混合模型；马尔科夫随机场；梯度下降算法	公开			
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*
北京交通大学		10004	工学	硕士
论文题名*		并列题名		论文语种*
基于高斯混合模型的图像分割的研究				中文
作者姓名*	欧垚江		学号*	12120441
培养单位名称*		培养单位代码*	培养单位地址	邮编
北京交通大学		10004	北京市海淀区西直门外上园村 3 号	100044
学科专业*		研究方向*	学制*	学位授予年*
计算机科学与技术		并行与分布式计算	2.5 年	2015 年
论文提交日期*	2014.6.10			
导师姓名*	于二元		职称*	副教授
评阅人	答辩委员会主席*		答辩委员会成员	
	赵耀		王移芝 田丽霞	
电子版论文提交格式 文本（√） 图像（） 视频（） 音频（） 多媒体（） 其他（） 推荐格式：application/msword； application/pdf				
电子版论文出版（发布）者		电子版论文出版（发布）地		权限声明
论文总页数*	54			
共 33 项，其中带*为必填数据，为 21 项。				