

摘 要

乳腺癌是中老年妇女中发病率和死亡率最高的恶性肿瘤之一，其早发现及早治疗是提高治愈率和降低死亡率的关键。乳腺钼靶摄影是普查和诊断乳腺癌的最有效方法。但纯人工阅读钼靶片存在易疲劳、耗时多、误诊和漏诊率较大等问题。利用计算机图像处理技术辅助检测，可以提高乳腺癌早期发现率和诊断准确率，使得诊断结果更具有客观性及精确性。

本文以乳房肿块检测展开，依照乳腺钼靶 X 射线图像去噪、乳腺钼靶 X 射线图像去噪增强、可疑肿块的分割的次序，进行了乳腺钼靶 X 射线图像乳房肿块的辅助检测研究。

在乳腺钼靶 X 射线图像去噪方面，针对乳腺钼靶 X 射线图像的肿块区域和背景区域都有较大噪声，影响阅片的问题，本文提出了基于小波变换的软阈值化去噪方法，在保持诊断影像特征特别是给定轮廓线内的毛刺和小叶征象的前提下，消除了噪声和伪影带来的影响。在乳腺钼靶 X 射线图像增强方面，本文采取了简单实用的直方图均衡化方法使乳腺钼靶 X 射线图像对比度低的问题得到了很好的解决。对于可疑肿块的分割，本文采取了基于水平集理论的窄带算法，实现了全自动和半自动的分割处理，通过几十组良性和恶性肿瘤图像数据的测试，程序可以使肿块区域准确地从乳腺区域分割出来。

关键词：乳房肿块；钼靶 X 射线；小波变换；直方图均衡化；水平集分割

ABSTRACT

Breast cancer is one of the most common malignancies with high incidence and mortality rates among older women. To increase the cure rate and reduce the mortality from this disease, early detection and treatment of breast cancer is the important premise. Mammography is considered as the most reliable and cost-effective method for early screening and diagnosis of breast cancer. But the problems, such as fatigue, time-consuming, higher rates of misdiagnosis and missed diagnosis, will occur if the detection of breast cancer is solely performed by a single radiologist. Computer-aided detection, which is based on computer image processing techniques, has the potential to improve the rate of early detection and make diagnosis results more objective and accurate.

This dissertation, rooting from the detection of breast lumps, with the sequence of breast X-ray images denoising, X-ray images enhancement, segmentation of suspicious mass, make a series study about aided detection of breast lumps.

For mammography X-ray image denoising, to reduce the noise in the mammography of breast mass regions and background regions, this dissertation presents a wavelet-based soft-thresholding denoising method, while maintaining diagnostic imaging features, especially within a given contour and the burr under the premise of lobular signs, eliminating the noise and the impact of artifacts. For mammography X-ray in image enhancement, this paper adopted a simple and practical method of histogram equalization and found a good solution to the problem of mammography X-ray image of low contrast. For the segmentation of suspicious mass, this theory has taken the narrowband level set based algorithm to achieve a fully automatic and semi-automatic segmentation, benign and malignant tumors through the dozens of groups of test image data, the program can make the tumor region from the breast region accurately split out.

Key words: breast mass; mammogram; wavelet transform; histogram

equalization; level set segmentation

目 录

第 1 章 绪论	1
1.1 研究的背景、目的及意义	1
1.2 国内外研究现状	3
1.3 本文主要的研究工作	6
1.4 论文结构	6
第 2 章 乳腺癌和钼靶 X 射线图像	8
2.1 乳腺癌的概述	8
2.2 乳腺癌的诊断方法	8
2.3 乳腺癌钼靶 X 射线影像特点分析	11
2.3.1 乳腺影像对应病理特征分析	11
2.3.2 钼靶 X 射线影像的诊断难点	14
2.4 本章小结	16
3.1 小波去噪	17
3.1.1 小波去噪的优势	17
3.1.2 基于小波变换的阈值化去噪算法	20
3.1.3 实验结果	23
3.2 直方图均衡化	24
3.2.1 直方图均衡化原理概述和算法流程	24
3.2.2 实验结果	25
3.3 本章小结	27
第 4 章 乳腺可疑肿块区域分割	28
4.1 医学图像分割的概念、分类和研究现状	28
4.2 水平集方法理论基础	29
4.3 水平集窄带算法流程	32
4.4 实验结果	34

4.5 本章小结.....	37
结论.....	38
参考文献.....	39
致谢.....	43

第 1 章 绪论

随着经济的快速发展和国民生活水平的不断提高，人们生活方式发生了巨大变化，我国已成为乳腺癌发病率增长最快的国家^[1]。目前乳腺癌治疗还未研究出有效的方法，早期发现、诊断和治疗可以降低乳腺癌的发病率和死亡率，改善患者愈后生活质量，提高医疗资源的社会利用率，降低病人个人和家庭的痛苦及社会的负担。

1.1 研究的背景、目的及意义

乳腺癌是女性最常见的恶性肿瘤之一，具有发病隐匿，愈后差等特点。据统计，全球每年约有 120 余万妇女患乳腺癌，50 万妇女死于乳腺癌^[2-3]。近 30 年来随着世界范围内人们生活水平的提高，乳腺癌的发病率表现出持续的高增长。在美国，专业机构和研究人员从 1973 年开始统计，女性乳腺癌发病率每年都有上升，其中 1982 年到 1987 年间每年增加 4%，到了 1998 年以后，这个数字又有提高^[4]。北美、北欧历来是乳腺癌的高发地区，东欧、南欧以及南美次之，亚洲、非洲的发病率最低。但从 20 世纪 70 年代起，原先为低发区的亚洲乳腺癌发病率出现上升趋势，尤其是日本、新加坡和我国沿海地区^[5]。随着物质经济的迅猛发展和生活水平的快速提高，这些地区的乳腺癌发病率正在以同样快的速度持续增长，而且越来越趋于年轻化。近几年来，在所有的女性恶性肿瘤中，乳腺癌的发病率已跃居到了首位，超过了原先的肺癌，其死亡率高达 40% 以上，对人类健康的严重危害已引起了世界卫生组织和医疗界人士的高度重视。

我国妇女乳腺癌的发病率目前正以每年 3%~4% 的增长率急剧上升，每年约有 18 万妇女罹患乳腺癌，1.3 万多妇女死于乳腺癌。20 世纪的最后 10 年，我国城市妇女乳腺癌的死亡率增长了 38.9%，农村妇女乳腺癌的死亡率增长了 39.7%，乳腺癌已经成为我国改革开放以后死亡率增长最快的癌症^[6]。沪、京、津是我国乳腺癌高发地区，其中上海最高，在所有女性恶性肿瘤中居首

位。现在我国乳腺癌的发病率虽然还不及欧美等发达国家，但死亡率却已经节节攀升，增长速度惊人，而欧美等发达国家在 20 世纪 90 年代之后，其乳腺癌死亡率已经开始明显下降。这主要是因为跟欧美等发达国家相比，我国妇女预防和检查乳腺癌的意识还比较落后，同时医疗卫生水平发展也不均衡，因此大力宣传预防乳腺癌的知识、加强乳腺癌诊断与治疗工作，将成为我国接下去较长时期内癌症防治的重要任务，对乳腺癌进行系统深入的研究刻不容缓。

目前，在乳腺癌早期诊断的多种方法中，乳腺钼靶 X 射线摄影术被认为是检测早期征象与病灶的最可靠、最有效的方法，其最大优点是可以对各种乳腺疾病做出明确诊断，在临床上摸不到肿块的早期乳腺癌，甚至在没有肿块的情况下，基于乳腺组织内微钙化灶的形态做出早期乳腺癌的诊断，其检出率达 80%~90%，诊断符合率 88%，且定位准确，可降低 30%以上的死亡率。

如果 40 岁以上女性每年进行一到两次的乳腺 X 射线摄影检查，每年会产生大量的乳腺 X 射线图像，由于普查环境中的恶性病例毕竟很少，从复杂的正常解剖背景发现异常的钙化和肿块，是绝大多数医生的一项艰巨和费时的的工作。研究发现，10%至 30%的乳腺癌不是被直接检测出的；此外，由于乳腺 X 射线图像分辨率较低，利用乳腺 X 射线摄影术进行乳腺癌检查，假阳性诊断率会很高，推荐进行活组织穿刺病理检查的病例只有 15%~30%被证实为恶性病灶，多达 65%~85%的乳腺癌活检结果为良性病变^[7]。不必要的活检不仅增加了患者的经济负担，而且导致患者生理和心理创伤，给社会医疗保健制度、病人和医师增加了不必要的负担；即使以两名医生判读图像，虽然可能会增加诊断的准确性，但由于乳腺 X 射线图像的复杂性和早期乳腺癌症状的隐匿性，使遗漏和假阳性成为乳腺钼靶 X 射线检查中不可避免的问题，由此，为了减轻放射科医师的工作强度，同时提高乳腺癌检查的准确率，利用计算机进行乳腺癌辅助诊断（Computer Aided Diagnosis, CAD）成为一种临床辅助诊断的有效手段。乳腺 X 射线图像计算机辅助检测方法是应用数理统

计理论、计算机图像处理技术及人工智能技术建立的计算机辅助诊断方法，它可以给放射科医生提供一个宝贵的“第二意见”^[7-8]，帮助医生提高早期乳腺癌诊断和检测的准确性和检测效率，减少观察者之间阅片的差异。

近年来，各地研究机构和研究人员在乳腺 X 射线图像的计算机辅助检测与诊断方面已经进行了很多研究工作，并取得了大量研究成果，西方国家研究已证实，计算机辅助诊断和检测系统可以很大程度减少漏诊率和误诊率，在早期发现与诊断乳腺癌中具有重要的价值。

乳腺 X 射线图像中的微钙化点和肿块是计算机辅助检测方法需要识别的主要对象^[9]。微钙化点通常表现为一些面积较小、灰度和对比度较低的不规则形状小块，肉眼不易发现，即使经验丰富的放射科医生也难在早期发现乳腺 X 射线图像的微小钙化点，延误病人的治疗时机。因此，及时发现乳腺 X 射线图像的微钙化点，并确定是否有恶性病变，是乳腺癌早期诊断的关键技术。在肿块检测方面，目前乳腺 X 射线图像计算机辅助检测方法灵敏度还比较低，大多在 60%至 85%，假阳性的检出率也比较高。对于计算机检测出的假阳性区域，放射科医生需要花费很多时间来重新评估测试提示的结果，为减少患者的痛苦，避免不必要的穿刺活检，有时放射科医生会对计算机检测的结果产生怀疑，把真阳性病灶区域当作假阳性来处理，造成对病人的误诊^[10]。这些都导致了放射科检测的准确性和检测效率的下降，使他们对计算机辅助检测方法失去信心。因此，如何提高乳腺癌计算机辅助检测方法的敏感性，降低假阳性的检出率，提高检测方法的整体性能，提高放射科医生使用计算机辅助检测的信心，是一项紧迫的研究课题。

1.2 国内外研究现状

计算机辅助乳腺癌检测和诊断技术的研究始于上世纪 60 年代末^[11]。随着医学影像数字化技术的发展和计算机性能的不断提高，从 90 年代初开始，学术界和企业界掀起了计算机辅助乳腺癌检测与诊断技术的研究高潮^[12-14]。纵观现有乳腺癌辅助检测如诊断肿块系统流程概况，一般分为乳腺图像增强、

可疑病灶分割、特征提取与选择、分类筛选等步骤^[12]。

1、 乳腺图像增强：图像增强是肿块检测中非常关键的一步，其目标是增大肿块与周围背景组织的对比度。肿块与周围组织对比度低，特别是在致密型乳腺组织中的对比度更低，往往造成肿块漏检。研究表明，如果一个区域的灰度与背景的灰度差小于 2%，人眼就很难区分两者。临床实践也证明简单采用人工调整窗宽窗位的方法就可以提高诊断准确性^[15]。

与其他图像增强技术类似，乳腺图像增强方法可以分成基于时域和基于频域的方法。基于时域的方法又分为基于全局信息和基于局部信息的方法。全局方法有灰度线性拉伸、直方图均衡化等，它们对整幅对比度低的图像较有效。局部方法将图像局部特征，包括梯度、灰度均值、标准差、灰阶熵等融入到增强函数中，如基于树型结构的非线性滤波增强、灰度加权的对比度增强(density-weighted contrast enhancement, DWCE)、基于多分辨率的灰度非线性调整后再加权综合的增强等，可有效增强局部纹理。该类方法虽然不能很好的增强整幅图像，但是应用到检测出的感兴趣区域上可以达到较好的效果。频域增强方法的研究集中在基于小波变换的多分辨率分析上。

2、 可疑肿块分割：肿块分割中广泛使用的方法有阈值分割、基于区域的分割、基于边缘的分割、聚类分割和基于模型的分割^[16]。

(1) 阈值分割：最早期最简单的方法是直接采用灰度统计信息进行阈值分割，但由于肿块不一定比腺体或者脂肪区域的灰度值高，该方法识别感兴趣区域的效果并不理想。

(2) 基于区域的分割：基于区域的分割方法将图像分成一些空间上连通且同质的区域。该类方法可分为区域生长法、分裂合并法。该类方法对于边缘模糊的区域进行分割效果良好，但关键是种子点的选取，这往往需要专业人士的经验知识。

(3) 基于边缘的分割：基于边缘的方法是图像处理中的常用方法。总体上有两大类，一是先找到一些边缘点，然后将边缘点连接成封闭的曲线；二是直接先给出一条封闭的曲线，然后优化其位置。由于后一种方法具备高级视

觉特性，可以将从图像中获得的约束信息和领域的先验知识有机的结合在一起，因而得到了更广泛的研究和应用。

(4) 聚类分割和基于模型的分割：基于聚类的分割方法将具有相似特征的像素分成一类，典型的方法有 K-means、模糊 K-means、基于 Markov 随机场的聚类方法等。该类方法较好的解决了边缘检测中边不连续的问题，其缺点是对初始轮廓位置要求较高。

在应用一种分割方法时，不同的参数设置对分割结果有很大影响，参数调优是一项非常重要的工作。令人遗憾的是，目前少有比较不同分割算法和不同参数对分割性能影响的系统研究文献。

3、可疑肿块特征提取和选择

特征提取对于可疑肿块的鉴别和分类是非常重要的。常用特征包括灰度统计特征、纹理特征以及形态学特征等。这些特征可用于鉴别可疑区域有无异常、肿块的良恶性、去除假阳性区域等。特征提取时可以计算出很多特征，这些特征中可能存在具有强相关性或不可识别性的冗余特征，将它们全部用于分类并不可取。为了降低分类器的复杂性和提高信噪比，需要进行特征选择和优化。目前，广泛采用的特征选择方法有逐步判别法(Stepwise Feature Selection)^[17]和基于遗传算法(Genetic Algorithm)的选择法^[18]。

4、可疑肿块分类

在特征提取和选择之后，需要一个分类器依据选出的特征将可疑区域分成正常组织、良性肿块和恶性肿块。分类器的构造方法很多，线性判别分类器(Linear Discriminate Analysis, LDA)、人工神经网络(Artificial Neural Networks, ANN)、支持向量机、Bayesian 网络、决策树、K-近邻分类器等都得到了广泛应用，其中线性分类器和神经网络得到了广泛的认可。为了提高分类的准确性，可以将多个分类器串联、并联和级联在一起。

对于一个 CAD 而言，评价其性能是非常重要的工作。通常用 ROC(Receiver Operating Characteristic)曲线和 FROC(Free-response Receiver Operating Characteristic)曲线来衡量系统的性能^[19]。但是由于各个研究团队使

用的图像数据库不同，加上分类方法的参数可以调整，给客观比较不同系统的性能带来了困难。在临床应用中，即使使用同一商用 CAD 系统，由于操作者不同给出的评价也不相同。如能构造标准的测试集将为系统而公正的比较 CAD 性能奠定基础。

1.3 本文主要的研究工作

针对乳腺钼靶射线图像的以下特点：

- 1、肿块和背景的对比度比较低。
- 2、肿块的边缘比较模糊，而肿块的边缘信息是判断肿瘤良性还是恶性的重要依据之一。
- 3、在肿块区域和背景区域通常都有较大的噪声点

本文的主要研究为：

- 1、通过小波变换进行乳房肿块的去噪。
- 2、通过直方图均衡化的方法提高图像对比度，使乳房肿块更清晰地脱离背景。
- 3、通过水平集技术，对乳房肿块图像进行分割。

1.4 论文结构

本文主要针对小波阈值化去噪、直方图均衡化，水平集分割算法等问题进行研究。全文安排如下：

第 1 章，介绍课题的研究背景。结合乳腺 X 射线 CAD 的背景、意义，详细介绍了乳腺 X 射线辅助检测技术的现状，并对其若干研究方向作了总结。

第 2 章，系统介绍了有关乳腺癌的医学知识、诊断知识；并对诊断工具钼靶 X 射线作了相应介绍。重点分析了乳腺癌病理信息在钼靶影像上的显示特点和钼靶影像本身的不足及其导致的诊断难点。系统了课题研究方向的实际背景及方向。

第 3 章，围绕钼靶 X 射线图像的预处理这一主题，详尽阐述了小波阈值

化去噪和直方图均衡化的原理知识及算法实现的流程，并给出了实验结果。

第 4 章，就图像分割的概念，分类，意义作了简要介绍，接下来给出了本设计采取的基于水平集理论的分割方法的理论知识，算法流程，最后给出了现阶段完成的程序运行结果，虽然理论上是可行的，但实际结果还是有些不理想，需要进一步的深入研究。

最后，对全文工作进行总结，指出研究工作的成绩和有待于改进和进一步研究的地方。

第2章 乳腺癌和钼靶 X 射线图像

2.1 乳腺癌的概述

目前乳腺癌的真正病因还未完全清楚，但是，许多因素能影响乳腺癌的发生与发展，却已经被国内外科研人员所发现。它们被称作致病因素，主要有：家族易感性、激素紊乱、月经、生育和哺乳、脂肪摄入过多、乳腺良性疾病、电离辐射、病毒感染等。

乳腺癌的分类方法纷繁复杂，尚未有统一的体系，根据 1990 年中国肿瘤防治办公室、中国抗癌协会合编的《中国常见恶性肿瘤诊治规范》第 8 册中介绍的形态学分类方法[中国肿瘤防治办公室 1990]，大体分为：

1、非浸润性癌（原位癌）：癌瘤最早阶段，病变局限于乳腺导管或腺泡内，未突破基底膜。

2、早期浸润癌：原位癌突破管壁基底膜，开始向间质浸润的阶段。

(1)早期浸润小叶癌：小叶原位癌穿过基底膜，向小叶内间质浸润，但尚未浸润至小叶范围之外。

(2)早期浸润导管癌：导管内癌少量癌细胞突破导管基底膜，向间质浸润，但浸润范围小。

3、浸润性癌：癌组织向间质内广泛浸润，形成各种形态癌组织与间质相混杂的图像。浸润型癌又分为浸润性非特殊型癌和浸润性特殊型癌。

此外还有其他一些罕见的乳腺癌种类，包括分泌性乳腺癌（幼年性癌）、富脂质癌(脂质分泌癌)、印戒细胞癌、乳头状瘤病癌变、富糖原透明细胞癌、腺纤维瘤癌变、伴化生的癌等。

2.2 乳腺癌的诊断方法

乳腺癌作为一种进行性以临床和病理为特征的疾病，具有很多病因病理方面差异，随着病变组织来源不同和病情进展速度不同、病变位置不同等常

常出现临床变异的特点。到目前为止生物学标志物还没有被发现，所以，乳腺癌常用的诊断方法需要具有较高的诊断准确性、良好的特异性和敏感性，使乳腺癌及早被发现、及时被治疗。目前广泛应用于乳腺癌诊断的方法有四大类：细胞学诊断、组织学诊断、临床医生的触摸式诊断和影像学诊断^[20]。

细胞学诊断主要包括穿刺检查和涂片检查。其中，针穿刺细胞学检查与活检类似，它具有轻微的损伤，且精度高的特点，是外科医生手术前病理检查的重要手段。但肿块的大小、癌组织的类型、组织的坏死、针吸技术的操作熟练程度和涂片质量，以及其它因素将直接影响到针穿刺检查的准确性；另外针穿刺检查有可能在组织内形成肿胀、造成癌细胞的扩散及体内出血，对医生的进一步检查造成影响；涂片检查一般针对出现乳头溢液患者进行。

组织学诊断是从人体器官内部提取部分病理组织，用于活检来进行病变组织性质的确定，为医生确定诊断方案提供数据。组织学诊断准确性最高，但是提取病理组织会对病人造成创伤，对患者造成身体的伤害。通常用于中晚期乳腺癌诊断，医生已基本确定病人患有乳腺癌，需要进行进一步确认或者病灶切除手术后复查时使用。常用方法有：穿刺活检、乳腺 X 线立体定位穿刺和冰冻切片检查等。

触摸式诊断是医生通过触摸疑似病变部位来判断诊断乳腺癌的方法。它对医生的水平要求很高，而且准确率只有 70%~80%，具有多年临床专业经验的老医生、老专家会先进行触摸检查，再做进一步的影像学诊断，并且极易出现漏诊或误诊。

与病理检查相比较，临床检查有一定的误差，即使有丰富临床经验的医师对原发性病灶检查的正确率约为 70%~80%。临床检查腋窝淋巴结约有 30% 假阴性和 30~40% 假阳性，故尚需其他辅助诊断方法，以提高诊断的正确率。影像学诊断是利用医学影像学的光透视技术对乳腺疾病的症状和病变进行直接检测的方法，使用方便而且对身体无创伤，应用范围最广，而且对早期隐匿性乳腺癌具有较好的诊断能力，非常适合用于适龄女性的乳腺癌普查^[21]。

乳腺癌的医学影像学诊断技术主要包括：

1、乳腺 X 射线摄影：是一种低剂量乳腺 X 光拍摄乳房的技术，它能清晰显示乳腺各层组织，可以发现乳腺增生、各种良恶性肿瘤以及乳腺组织结构紊乱。乳腺 X 射线摄影技术诊断乳腺病变，正确率可达 90%以上，是公认的筛查和探测乳腺病变的最有效、最可靠、最经济的方法^[22]。

2、乳腺导管造影：正常乳腺导管造影是导管的各分支显影，充盈良好，从粗到细，走向柔软、自然，未见狭窄或充盈缺损区。导管癌表现为导管扩张后狭窄、中断，呈刀切样或鼠尾状，导管内亦可见不规则充盈缺损，导管僵硬，局部见致密团块影，还伴有精盐样或泥沙样钙化。

3、红外光谱检查：主要利用人体不同结构的软组织对红外线的吸收率不同而进行诊断。具有一定的准确性，但很多情况下无法作出准确的判断，其优点是价格低廉，对人体毫无损伤。

4、超声检查：是利用人体对超声波的反射进行检查，用弱超声波照射到身体上，将组织的反射波进行图像化处理。随着实时灰阶超声的不断发展，在乳腺癌检查中开始使用超声造影，诊断的准确性可达到 80%~85%，超声检查可以显示乳腺的各层结构、肿块的形态及其密度。相较于乳腺 X 射线摄影，乳腺超声不会含有放射性物质，可重复应用，且对肿块的囊性/实性的检测准确，可对腋淋巴结的状况进行检查，但其对直径小于 1.0 厘米的小癌变、微钙化点和毛刺样结构不易显示，对早期乳腺癌诊断价值不大。

5、其余影像学技术：远红外线乳腺检查是通过吸收人体辐射热的远红外乳腺癌诊断技术成像的方法，可以用来区分良性/恶性病变，但它的灵敏度低，设备非常昂贵，无法广泛使用。CT 和 MRI 等现代最常用的医学影像学技术，其对乳腺癌的检出率非常高，成像清晰，但检查费用非常高，时间较长，且对细小微钙化的敏感性不如钼靶 X 线摄影^[23]。

总之，目前乳腺癌特点各异，诊断方法种类繁多。触摸式诊断由于容易造成误诊与漏诊，现在不作为主流的诊断方法；对于中晚期乳腺癌、病灶已经形成之后的检测，组织学和细胞学诊断方法有着最高的准确性；影像学诊断，因为它方便非侵入性的特点，可用于早期诊断和普查。在影像学诊断方

法中，远红外乳腺扫描、液晶热成像和超声波扫描没有放射性，不伤害人体，将成为乳腺癌筛查的第一选择，但这些技术存在着图像分辨率低、检测精度不高和对微钙化点不敏感的缺点，容易错过早期的隐匿性乳腺癌，CT，MRI 检查效果好，但检查时间长、成本高，可以用于乳腺癌的检测中，但不适合用于普查，且对微钙化的检测效果不好；钼靶 X 线摄影，虽然会带来一些辐射，但它对微钙化簇最敏感，而且图像比较清晰，多年的临床使用，大量经验丰富的放射学专家可以较准确读取乳腺 X 射线图像信息，做出判断^[24]。因此在众多的检查方法中，钼靶 X 线摄影技术是最常用的乳腺癌临床诊断工具，也是乳腺癌普查的首选检查手段^[25]。

2.3 乳腺癌钼靶 X 射线影像特点分析

2.3.1 乳腺影像对应病理特征分析

胡永升主编的《现代乳腺影像诊断学》报告，钼靶 X 线度乳腺良、恶性肿瘤的诊断的正确率分别为 92.1%和 84.3%；另有报告分别为 89.6%和 84.0%。虽然钼靶摄片诊断有较高的正确率，但仍然存在着误诊问题。为了总结经验，提高乳腺癌的诊断符合率，下面就五种病样影像进行讨论。

1、结节肿块型



图 2.1 结节肿块性

本型是乳腺癌最直接的 X 线征象，主要表现为大小不等的结节或肿块阴影，部分呈分页状。多数肿瘤的边缘有细短毛刺及周围结构紊乱，其病理基础是肿瘤细胞向周围组织浸润生长致肿瘤周围结缔组织增生性反应。肿瘤密度多为高密度，原因是由于癌细胞浸润使单位体积的细胞数增加及周围组织水肿所致。肿瘤可浸润皮肤使皮肤增厚，阻塞导管使导管扩张，压迫血管使其增粗，牵拉血管使其迂曲。这些间接征象对乳腺癌的诊断均有非常重要的意义。

2、星型肿块型

此征象为浸润型乳腺癌的典型 X 线征象。病变以不规则纤维增生为主，在肿瘤边缘形成粗长毛刺，毛刺的长度多超过肿瘤的直径，并可呈放射状排列，此征象主要反映了癌细胞浸润性生长的特征。浸润性癌伴结缔组织增生性反应，并能使邻近 Cooper's 韧带增厚或缩短，导致皮肤回缩，局部凹陷形成“酒窝征”。浸润性导管癌易朝导管方向浸润，当癌细胞侵犯乳晕区时致纤维组织收缩导致乳头回缩，形成“漏斗征”。

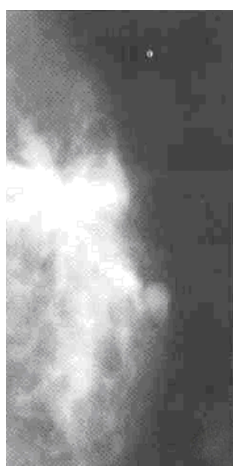


图 2.2 星性肿块型

3、增生样肿块型

此症状产生原因可能是因为癌细胞浸润性生长致大量的纤维组织形成，

同时与周围增生的乳腺融合的缘故。X 线主要表现为大片状密度增高阴影，高于周围腺体组织的密度，无清楚肿块轮廓，中心密度较高，边缘毛糙模糊，瘤周结构紊乱，类似较重的乳腺增生，易误诊。不易做出诊断者，可双侧乳腺对照研究，一般两侧乳腺的外形及腺体密度基本对称，一旦不对称或明显不对称，应考虑乳腺癌。伴有泥沙样钙化和腋淋巴结肿大者可确诊乳腺癌。

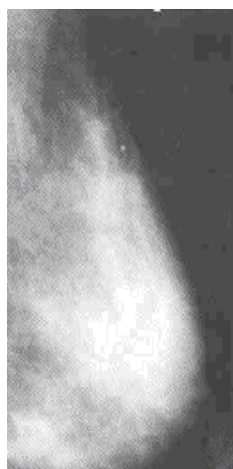


图 2.3 增生肿块型

4、钙化无肿块型

钙化是乳腺癌的重要 X 线征象。X 线上不显示肿块的原因，可能与投照条件有关系，也可能是乳腺组织较密实或伴有乳腺炎、乳腺增生时，与肿瘤不能形成密度对比，显示不出肿瘤的典型征象，此待进一步研究。钙化的机制可能是瘤灶区域局部营养不良或坏死，坏死区内细胞的溶解，核酸分解出大量磷酸根，同时局部钙离子增加，碱性磷酸酶增加导致磷酸钙盐沉着的缘故。乳腺癌的钙化多位于瘤体坏死区内，也可位于瘤周围结缔组织内。细小泥沙样钙化是乳腺癌的特征性钙化，无论有无肿块，见类似钙化者应引起重视。

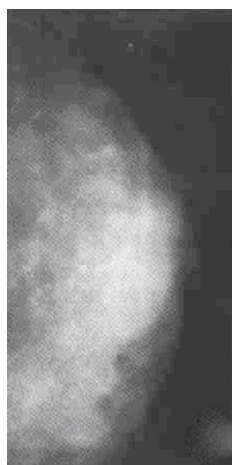


图 2.4 钙化无肿块型

5、不典型型

临床触摸均有肿块或结节，拟诊乳腺癌，在钼靶乳腺 X 片上未发现明显肿块或结节，仅在其相应部位显示斑片状和不规则小片状阴影。其原因可能是乳腺癌的发生多起源于导管或者乳腺小叶组织内，可沿乳腺导管壁和小叶组织呈片状浸润生长，不形成具体肿块，所以在钼靶片上不显示肿块。也可能是因为肿瘤生长在乳腺组织较深或者组织结构较厚的部位，同时并伴有乳腺增生与肿瘤融合，因此在钼靶 X 线片上显示不出具体肿块轮廓，只显示密度较高的不规则大片状阴影，该型乳腺癌易误诊和漏诊。

2.3.2 钼靶 X 射线影像的诊断难点

钼靶摄影产生的乳腺 X 射线图像中，空间分辨率和灰度分辨率较高，图像较为清晰，肿块和微钙化两种病灶较为明显。即便如此，乳腺癌诊断的误诊率和漏诊率仍然无法获得更大的降低。据了解，在乳腺 X 射线图像中一次性检测出乳腺癌的概率只有 70%~85%，其余情况下，约 2/3 可在第二次检查发现，对隐匿性乳腺癌的临床症状检出率只有 5%~59%。此外，乳腺 X 射线图像也有约 4%~10%的误诊率，出现这些问题的原因是由于恶性病变和良性病变出现的迹象过小，乳腺癌的一些迹象可能会模糊，类似的肿块密度和正

常腺体相仿或肿瘤形态不是那么明显，所以很难用肉眼察觉，医师在阅读了大量放射学影像，视觉的疲劳和疏忽也是一方面原因。因此，有必要用计算机对乳腺 X 射线图像进行预检测。

目前，乳腺 X 射线图像的 CAD 技术在检测方面还有许多困难，通过钼靶 X 射线摄影和计算机技术结合，常用的数字乳腺 X 射线图像相对于其他医疗成像技术虽然已经具有很高的清晰度和分辨率，但毕竟不像 X 射线可通过放大镜对图像的精细结构观察，而图像有大量的冗余信息，图像质量仍然是有限的，计算机自动检测的视觉识别能力，与人眼的能力还有一定的差距^[26]。

对于微钙化点，在 0.1 毫米/像素的数字图像中，只有直径大于 100 微米的微钙化点有可能出现在图像中，更小的微钙化点将无法显示，即使利用放大镜观察微钙化点，它也似乎过于微小，很难察觉，有些微钙化点和正常组织的结构具有相似性，在该区域的一些特征非常类似于乳腺的其它组织结构，这些特点都造成了微钙化点自动检测难度的增加。除此之外，不像人的视觉，计算机视觉很难根据目标区域背景的变化自动适应和调整，因此，不同区域的不同背景特征及其微钙化点，难以用统一的检测标准，这样使微钙化点检测更加困难。

对于肿块，计算机辅助诊断的困难亦不少，主要原因是：病变组织和乳腺区域内广泛存在的腺体组织特征接近，很难区分，相对于其背景，肿块呈现为白色区域，对比度比较低，形态不一致，且大小差别很大，因此感兴趣区域自动提取难度较大；同时，肿块边缘与正常组织相互交叠，边缘比较模糊，从坏死中心到过渡灰色背景周边区域的灰度往往难以准确地分别，与微钙化点检测类似的难点是位于不同区域的肿块特点是不相同的，这些很难使用统一标准来诊断。

从上述乳腺病变检测的困难可以看出，理想的乳腺 X 射线图像 CAD 技术需克服各种困难实现微钙化簇和肿块的检测，以确定微钙化簇和肿块与其他腺体组织之间的本质区别，使用大小不固定的检测窗口，自动调节适应参

数, 实现对微钙化簇和肿块做出精确的定位和分割, 准确提取感兴趣区域实现最终诊断识别^[27-28]。

2.4 本章小结

本章对乳腺癌进行了病理学研究, 介绍了乳腺癌的病理、乳腺癌的分类以及乳腺癌的临床医学诊断方法, 分析了乳腺钼靶 X 射线的主要影像病灶特点与诊断难点, 阐述了当前乳腺 X 射线图像计算机辅助诊断方法的主要问题。

第3章 乳腺钼靶X射线图像的预处理

3.1 小波去噪

3.1.1 小波去噪的优势

长期以来,在信号处理方面,人们所使用的最基本的数学工具就是傅立叶变换。但是,傅立叶变换是一种全局的变换,无法表述信号的时频双域局部特性,只能反映信号的整体特征。而在不少实际问题中,人们所关心的却是信号在局域范围中的特征,例如音乐和语音信号中,人们关心的是在什么时刻演奏什么音符,发出什么音节;对人上地震波的记录,人们关心的是在什么位置出现了什么反射波;图像处理中边沿识别,关心的是信号的突变位置等。小波变换正是弥补这一不足强有力的工具。

小波变换理论是近二十年中才发展起来的一个数学分支。目前它仍然是国际上极为活跃的研究领域。小波变换的分析窗口大小(即窗口面积)固定但其形状可改变,小波变换是时间窗和频率窗都可改变的时频局部化分析方法,对不同的频率在时域上的取样步长是可调节的,即在低频时小波变换的时间分辨率较差,而频率分辨率较高;在高频时小波变换的时间分辨率较高,而频率分辨率较低。这种自适应性正符合低频信号变化缓慢而高频信号变化迅速的特点,所以,被誉为数学显微镜^[29]。

小波变换在时域和频域同时具有良好地局部化性质,不仅可将图像的结构和纹理可以分别表现在不同分辨率层次上,而且具有检测边沿(局域突变)的能力,因此,利用小波变换在去除噪声时,可提取并保存对视觉起主要作用的边沿信息。而传统的傅立叶变换去噪方法在去除噪声和边沿保持上存在着矛盾,原因是傅立叶变换方法在时域不能局部化,难以检测到局域突变信号,在去除噪声的同时,也损失了图像边沿信息。由此可见,与傅立叶变换去噪方法相比,小波变换去噪方法具有明显的优越性。它的成果已被广泛应用于图像处理、语声人工合成、地震勘探、大气湍流、天体识别以及机器视

觉等众多领域。

表 3-1 傅里叶变换、短时傅里叶变换和小波变换的对比

变换类型	分解种类	分析函数	变量	适应场合
傅里叶变换	频率	正弦函数，余弦函数	频率	平稳信号
短时傅里叶变换	时间-频率	由三角震荡函数复而成的时间有限的合波	频率 窗口的位 置	次稳定信号
小波变换	时间-尺度 或时间-频率	具有固定震荡次数的时间有限的波	尺度 小波位置	非稳定信号

3.1.2 小波变换理论

小波是指函数空间 $L^2(R)$ 中满足下述条件的一个函数或者信号 $\Psi(x)$,

$$C_{\psi} = \int_{R^3} \frac{|\psi(x)|}{|\omega|} d\omega < \infty \quad (3.1)$$

这里, $R^3 = R - \{0\}$ 表示非零实数全体。

对于任意的函数或者信号 $f(x)$, 其小波变换定义为:

$$w_f(a, b) = \int_R f(x) \overline{\varphi_{(a,b)}(x)} dx = \frac{1}{\sqrt{|a|}} \int_R f(x) \overline{\varphi\left[\frac{x-b}{a}\right]} dx \quad (3.2)$$

因此, 对任意的函数 $f(x)$, 它的小波变换是一个二元函数。另所谓多分辨率分析是指设 $\{V_j; j \in Z\}$ 是 $L^2(R)$ 上的一列闭子空间, 其中的一个函数, 如果它们满足如下五个条件, 即

(1) 单调性: $V_j < V_{j+1}$, $j \in Z$;

(2) 惟一性: $\bigcap_{j \in Z} V_j = \{0\}$;

(3) 稠密性: $\overline{V_j} = L^2(R)$;

(4) 伸缩性: $f(x) \in V_j, f(2x) \in V_{j+1}$, $p_j \in Z$;

(5) Riesz基存在性: 存在 $\phi(t) \in V_0$, 使得 $\{\phi(2^j x - n); n \in Z\}$ 构成 V_j 的 Riesz 基, 称 $\phi(t)$ 为尺度函数。那么, 称 $\{\{V_j; j \in Z\}, \phi(x)\}$ 是 $L^2(R)$ 上的一个多分辨分析。

若定义函数 $\phi_{j,n}(x) = 2^{\frac{j}{2}} \phi(2^j x - n)$, $p_j, n \in Z$; 则由多分辨分析的定义, 容易得到一个重要结果, 即函数族

$$\{\phi_{j,n}(x) = 2^{\frac{j}{2}} \phi(2^j x - n); n \in Z\} \quad (3.3)$$

是空间 V_j 的标准正交基。关于多分辨分析, 在这里以一个三层的分解进行说明, 多分辨分析只是对低频部分进行进一步分解, 而高频部分则不予考虑。分解具有关系

$$S = A_3 + D_3 + D_2 + D_1 \quad (3.4)$$

另外强调一点, 这里只是以一个层分解进行说明, 如果要进行进一步分解, 则可以把低频部分分解成低频部分和高频部分, 以下再分解, 依次类推。在理解多分辨分析时, 必须牢牢把握一点, 即分解的最终目的是力求构造一个在频率上高度逼近空间的正交小波基, 这些频率分辨率不同的正交小波基相当于带宽各异的带通滤波器。多分辨分析只对低频空间进行进一步的分解, 使频率的分辨率变得越来越高^[30]。

而关于Mallat算法是将 $L^2(R)$ 上的多分辨分析记为 $\{\{V_j; j \in Z\}; \phi(x)\}$, 尺度方程和小波方程为

$$\phi(x) = \sqrt{2} \sum_{n \in Z} h_n \phi(2x - n) \text{ 和 } \psi(x) = \sqrt{2} \sum_{n \in Z} g_n \phi(2x - n) \quad (3.5, 3.6)$$

其中，系数关系是 $g_k = (-1)^{1-k} \overline{h_{1-k}}$, $k \in Z$ ，对任意的整数 j 和 k ，沿用记号

$\bar{\phi}_{j,n}(x) = 2^{\frac{j}{2}} \phi(2^j x - n)$ ， $\bar{\psi}_{j,n}(x) = 2^{\frac{j}{2}} \psi(2^j x - k)$ 和

$$\left\{ \begin{array}{l} V_j = \overline{\{\phi_{j,n}(x); n \in Z\}} \\ W_j = \overline{\{\psi_{j,n}(x); n \in Z\}} \\ L^2(R) = \overline{\bigcup_{j \in Z} W_j} \end{array} \right\}$$

对于任意信号 $f(x) \in L^2(R)$ 引入记号

$$C_{j,k} = \int_R f(x) \bar{\phi}_{j,k}(x) dx, d_{j,k} = \int_R f(x) \bar{\psi}_{j,k}(x) dx \quad (3.7)$$

称为 $f(x)$ 的尺度系数和小波系数，同时，将 $f(x)$ 在闭子空间 V_j 和 W_j 上的正交投影记为 $f_j(x)$ 和 $g_j(x)$ ，这样

$$f_j(x) = \sum_{k \in Z} C_{j,k} \phi_{j,k}(x), g_j(x) = \sum_{k \in Z} d_{j,k} \psi_{j,k}(x) \quad (3.8)$$

根据空间正交值和分解关系 $V_{i+1} = V_i \oplus W_i$ 可得 $f_{j+1} = f_j(x) + g_j(x)$ 因此，信号的尺度变换系数和小波变换系数之间的关系现在可以写成

$$\sum_{k \in Z} C_{j+1,k} \phi_{j+1,k}(x) = \sum_{k \in Z} C_{j,k} \phi_{j,k}(x) + \sum_{k \in Z} d_{j,k} \psi_{j,k}(x) \quad (3.9)$$

3.1.2 基于小波变换的阈值化去噪算法

传统的去噪方法是将被白噪声干扰的信号通过一个滤波器，滤掉噪声频率成分，但对于脉冲信号、白噪声、非平稳过程信号等，传统方法存在一定局限性，对这类信号，在低信噪比情况下，经过滤波处理，不仅信噪比得不到较大改善，而且信号的位置信息也被模糊掉了。基于小波变换的去噪方法利用小波变换中的变尺度特性对确定信号具有一种“集中”的能力。如果一个信号的能量集中于小波变换域少数系数上，那么对这些系数的取值必然大于在小波变换域内能量分散于大量小波系数上的信号或噪声的小波系数值^[31]。

在这里,如何对小波变换域内的系数进行筛选是小波阈值化去噪的关键步骤。小波系数筛选主要有阈值化与门限的选取两个问题,下面就这两个问题进行详细讨论。

1、阈值化

在小波系数进行取舍之前,实际上按着一定准则,将小波系数换划分为两类:一类是重要的、规则的小波系数;另一类被看作是非重要的或者受噪声干扰较大的小波系数^[32]。通常以小波系数的绝对值作为小波系数的分类单元。小波系数绝对值趋向零,意味着小波系数所包含的信息量少并且强烈地受噪声干扰。Donobo^[33]等人证明这种判断方法具有良好的统计优化特性。小波系数的绝对值是一个局部测度,每个小波系数被看成是独立变量。给定一个阈值 δ ,所有绝对值小于 δ 的小波系数被划为“噪声”,它们的数值用零代替;而超过阈值的小波系数的数值用阈值 δ 缩减后再重新取值。这种方法意味着阈值化将在小波域中移去小幅度的噪声或者非期望的信号,在小波的逆变换中,将得到所需要的信号。“软阈值化”(soft-thresholding)和“硬阈值化”(hard-thresholding)是对超过阈值之上小波系数进行缩减的两种主要方法。

软阈值化:

$$W_{\delta} = \begin{cases} \text{sgn}(W)(|W| - \delta), & |W| \geq \delta \\ 0, & |W| < \delta \end{cases} \quad (3.10)$$

硬阈值化:

$$W_{\delta} = \begin{cases} W, & |W| \geq \delta \\ 0, & |W| < \delta \end{cases} \quad (3.11)$$

两种阈值化方法各有差异,前者具有连续性,从数学上容易处理;后者更接近实际应用。

阈值化处理的关键问题是选择合适的阈值 δ 。如果阈值太小,去噪后的信号仍然有噪声存在;相反地,阈值太大,重要的图像特征将被滤掉,引起偏差。从直观上,对于给定小波系数,噪声越大,阈值就越大。大多数阈值

选择过程是针对一组小波系数，即根据本组小波系数的统计特性，计算出一个阈值 δ 。如果噪声是加性的、随即平稳的，则可证明：在小波变换域中，每个子带或者每层分辨率上的噪声仍然保持加性与随机平稳。在第 j 层子带上，当无干扰噪声的小波系数与去噪后的小波系数之间均方误差最小时，存在一个最优的阈值 $\delta_{j,opt}$ 。由此，阈值的选择过程可以通过一个风险函数来定义：

$$R_j(\delta) = \frac{1}{N_j} \|W_{j,\delta} - V_j\|^2 \quad (3.12)$$

式中， N_j 代表在第 j 层子带上的小波系数个数； $W_{j,\delta}$ 表示阈值化得小波系数矢量； V_j 表示无噪声干扰时小波系数矢量。Donobo^[33]等已提出了一种典型阈值选取方法，从理论上给出并证明阈值与噪声的方差成正比，其大小为

$$\delta_j = \sqrt{2 \log N_j \sigma_j} \quad (3.13)$$

事实上，对于有限长度 N_j 的信号，上式仅仅是阈值优化的上界，阈值优化是随信号长度渐近变化的，当信号为无限长时，才能符合上式的阈值优化条件，因此，信号足够长，去噪效果才明显。

有人提出另一类阈值选择过程，其风险估计函数定义为

$$R(\delta) = \frac{\frac{1}{N} \|W_j - W_{j,\delta}\|^2}{\left[\frac{N_{j0}}{N_j}\right]^2} \quad (3.14)$$

在这里， N_{j0} 是小波系数被置为零的个数。上式不需要估计每一层噪声的方差 σ_j^2 ，因此，它的计算复杂性是线性的。

综上所述，基于小波变换的阈值化去噪方法步骤概括如下：

1) 计算小波变换系数 W 。

2) 根据(3.13)或(3.14)，计算每一层小波系数的阈值 δ_j 。

3) 利用“软阈值化”式，或者“硬阈值化”式，对小波系数进行阈值化，得到新的小波系数 W_δ 。

4) 计算小波逆变换，完成信号的去噪处理。

3.1.3 实验结果

本程序采用的是软阈值法。

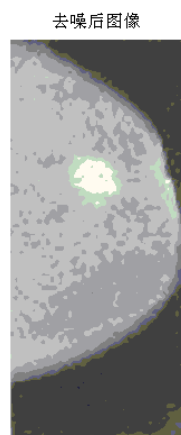


图 3.1 原噪声图像 a

图 3.2 图 a 软阈值化去噪后图像

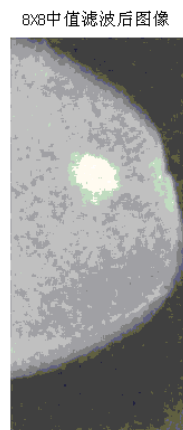
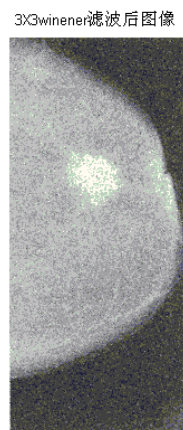


图 3.3 图 a Winener 滤波去噪后图像

图 3.4 图 a 中值滤波去噪后图像

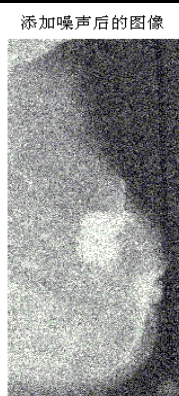


图 3.5 原噪声图像 b

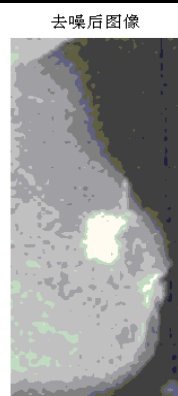


图 3.6 图 b 软阈值化去噪后图像

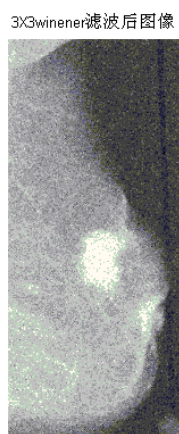


图 3.7 图 b Winener 滤波去噪后图像

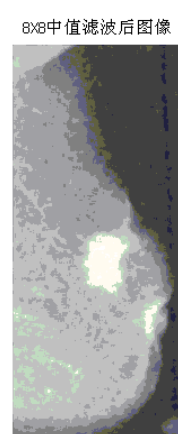


图 3.8 图 b 中值滤波去噪后图像

3.2 直方图均衡化

3.2.1 直方图均衡化原理概述和算法流程

直方图均衡化是一种简单而常用的图像对比度增强算法，也是增强乳腺钼靶 X 线图像效果较好的方法之一^[34]。

直方图均衡化是将原图像的直方图改变为在整个灰度范围内基本均匀地分布的形式，由此扩大了像素灰度的动态范围，从而增强了图像的对比度。

直方图均衡化的实施可以分为以下几个步骤进行：

第一步，根据原图像计算灰度直方图

$$p_s(s_k) = \frac{n_k}{n} \quad (3.14)$$

第二步，计算

$$r_k = \sum_{i=0}^k p_s(s_i) = \sum_{i=0}^k \frac{n_i}{n} \quad (3.15)$$

称为原图像的累积灰度直方图，这样求得的 r_k 就是直方图均衡化后的灰度的均匀分布；

第三步，按公式

$$t_k = \text{int}[(N-1)r_k + 0.5] \quad (3.16)$$

确定经直方图均衡化后的灰度级别，其中，运算符号 int 表示取其整数部分， N 为原图像中灰度级别数；

第四步，确定原图像灰度级别 s_k 转换到 t_k 的映射关系，然后按此映射关系将原图像每一像素的灰度逐一转换。

需要指出的是：对一副图像进行直方图均衡化处理可以增加原图像的对比度，从而使该图像的细节显示得更清楚。但与此同时，也增加了原图像的可视粒度，从而使其显得更加粗糙^[35]。

3.2.2 实验结果

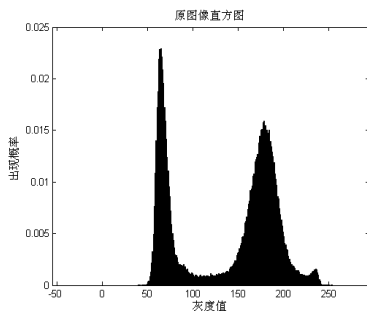


图 3.9 原图像 a 的直方图

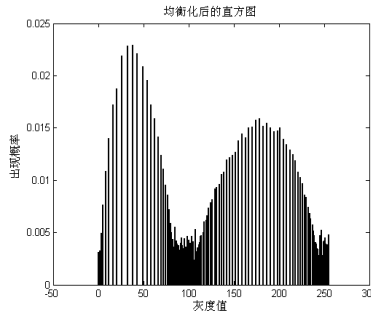


图 3.10 图 a 均衡化后的直方图

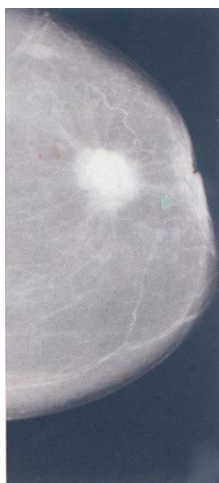


图 3.11 原图 a

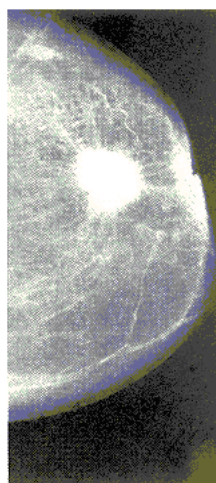


图 3.12 图 a 均衡化后的图像

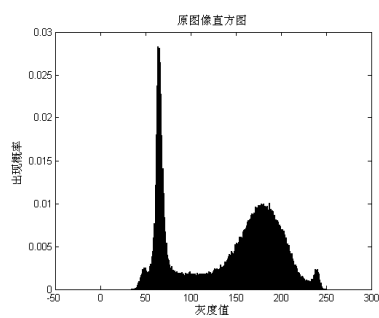


图 3.13 原图像 b 的直方图

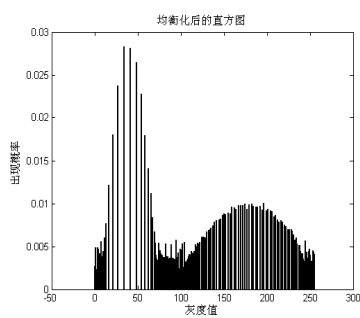


图 3.14 图 b 均衡化后的直方图

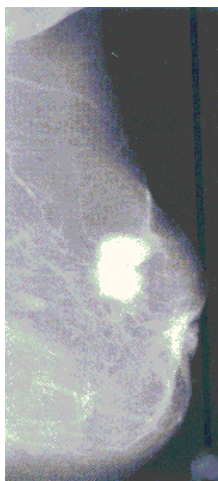


图 3.15 原图 b

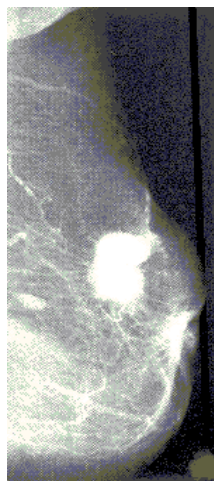


图 3.16 图 b 均衡化后的图像

3.3 本章小结

本章主要内容是对乳腺钼靶射线图像进行预处理中去噪和增强两步处理，去噪采取了基于小波变换的软阈值化去噪方法，增强采取了效果较好且简单易行的直方图均衡化均取得了不错的效果。这为后续的分割打下了基础。

第4章 乳腺可疑肿块区域分割

4.1 医学图像分割的概念、分类和研究现状

所谓医学图像分割，就是根据医学图像的某种相似性特征（如亮度、颜色、纹理、面积、形状、位置、局部统计特征或频谱特性等）将医学图像划分为若干互不相交的联通的区域的过程，相关特征在同一区域内表现为一致性或相似性，而在不同区域间表现出明显的不同，也就是说在区域边界上的像素存在某种不连续性。一半说来，有意义的图像分割结果中至少存在一个包含感兴趣目标的区域。

通常，医学图像分割方法可以划归为三大类：基于阈值的分割方法、基于边缘的分割方法和基于区域的分割方法^[35-36]。在理想的情况下，医学图像中的每一个区域都是由相应的封闭轮廓线包围着。原则上，使用边界跟踪算法可以得到区域的边缘（或封闭的轮廓线）；反过来，使用简单的区域填充算法也可以得到边缘所包围的区域。但在实际的医学图像中，很少能够从区域中得到连续、封闭的边缘，反之亦然。由于受人体内外环境中种种确定性、不确定性因素的干扰和成像噪声的影响，实际所获得的医学图像不可避免地具有模糊、不均匀等缺陷；另外，人体的解剖结构比较复杂而且因个体的病理或生理的差异有很大的不确定性、这在医学图像中引入了新的复杂性，同时也给医学图像分割带来了很大的困难；目前，现有医学图像分割的基本方法多数是针对二维图像进行的，当推广到三维乃至四维医学图像分割应用场合时，在数据结构和算法处理上不可避免地导致更大的复杂性，使得医学图像的分割更为困难。近年来大量学者致力于将新概念、新思想和新方法应用于复杂二维医学图像和高维医学图像或者图像序列的分割，其中包括数学形态学、模糊理论、神经网络、遗传算法、蚁群算法、粗糙集理论、水平集理论、支持向量机、马尔科夫随即场理论、核函数方法、小波分析和小波变换等，其间有很大一部分属于基于模式识别原理的医学图像分割方法。各种分

割方法或数学工具的有效应用，极大地改善了医学图像的分割效果。

纵观医学图像分割技术的发展历程，实际上是一个从人工分割到半自动分割和全自动分割逐步发展的过程。人工分割是指由经验丰富的临床医生在原始胶片图像上直接勾画出组织的边界，或者通过图像编辑器用鼠标在计算机显示器上勾画出组织的边界或者感兴趣的区域。半自动分割技术是随着计算机科学的发展而产生的，它把计算机强大的数据处理、算法分析和智能计算能力以及自动存储和记忆功能与医学专家的知识 and 经验有机地结合起来，通过人机交互的方式完成图像分割。全自动分割则彻底脱离了人为干预，完全由计算机自动完成图像分割的全过程。由于全自动分割方法不存在人为因素的影响，为图像中感兴趣区域的自动精确测量奠定了基础。但是绝大多数自动分割算法实现复杂，运算量大，在很多情况下，分割结果尚不理想，分割速度和性能也有待提高。

从目前图像分割技术在临床上的应用情况来看，人工分割的精度在所有分割方法中是最高的，被视为金标准，但该方法费时、费力，其分割结果的优劣完全取决于操作者的经验和知识，且分割结果难以重视。半自动方法与人工分割相比，分割速度明显提高，但其分割结果在很大程度上仍然依赖于操作者得主观经验和知识，这在一定程度上影响了半自动分割技术在临床上的广泛应用。研究高效、实用的半自动分割方法并最终取代繁琐的人工分割和主观性依然很强的半自动分割一直是人们追求的目标，也是近年来图像分割方法的研究重点。由于医学图像经常存在对比度低、组织特征的可变性大以及不同软组织之间或软组织与病灶之间边界模糊，微细血管或神经结构形状分布复杂等特点，加上成像中种种客观因素的制约，使得医学图像分割成为医学图像处理过程中的一个经典难题，到目前为止既不存在一种通用的分割方法，也不存在一种准确评价的分割成功与否的客观标准。

4.2 水平集方法理论基础

水平集(Level Set)方法是 1988 年由 Osher 和 Sethian 首次提出^[37]，其思想

是将曲线或曲面看成一个函数的零水平集，通过函数的水平集演化来确定曲线和曲面的演化。水平集方法自提出以来，已经在图像处理和计算机视觉等领域得到了广泛的应用，除应用于图像分割之外，还可应用于图像去噪与增强，图像修复，运动目标跟踪等。

Osher 和 Sethian 提出可以将平面闭合曲线隐含地表达为三维连续曲面 $\varphi(x, y)$ 的具有相同函数值的同值曲线，通常选 $\{\varphi=0\}$ ，称为零水平集， $\varphi(x, y)$ 则称为水平集函数。



图 4.1 曲线

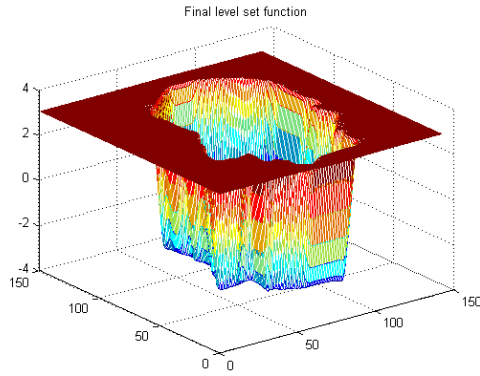


图 4.2 曲线的水平集函数

下面简要说明水平集方法的数学表达及其与曲线演化的关系。

设连续函数 $\varphi(x, y, t) : R^2 \times R^+ \rightarrow R$ 是闭合曲线 $c(p, t) : 0 \leq p \leq 1$ 在 t 时刻的隐含表达，即 t 时刻 $c(p, t)$ 对应于 $\varphi(x, y, t)$ 的零水平集

$$\begin{cases} c(p, t) = \{(x, y) | \varphi(x, y, t) = 0\} \\ c(p, 0) = \{(x, y) | \varphi(x, y, 0) = 0\} \end{cases} \quad (4.1)$$

水平集方法的目的就是要使水平集函数 $\varphi(x, y, t)$ 在演化过程中， φ 的零水平集的平面闭合曲线

$$\varphi(c(t), t) = 0 \quad (4.2)$$

始终满足曲线演化的偏微分方程 $\frac{\partial c}{\partial t} = v(k)N$ ，其中 $v(k)$ 是速度函数， N

是曲线的内向单位法向量，对方程(3.2)求全微分得

$$\nabla \varphi \frac{\partial c}{\partial t} + \frac{\partial \varphi}{\partial t} = 0 \quad (4.3)$$

其中 $\nabla \varphi$ 是 φ 的梯度。

设 S 是闭合曲线 c 的弧长参数, 根据水平集函数的定义, φ 沿曲线 c 的方向的变化量为 0, 即 $\varphi_s = 0 = \varphi_x \chi_s + \varphi_y y_s = \langle \nabla \varphi, c_s \rangle$, 这样 $\nabla \varphi$ 就垂直于闭合曲线 c 的切线, 因此, $\nabla \varphi$ 和 c 的法线同向。假设函数 φ 位于 c 内部的部分为负, 外部部分为正, 则水平集 c 的内向单位法向量就是

$$N = -\frac{\nabla \varphi}{|\nabla \varphi|} \quad (4.4)$$

将(1.1)式和(3.4)式带入(3.3)式, 整理后, 可得

$$\frac{\partial \varphi}{\partial t} = v(k) |\nabla \varphi| \quad (4.5)$$

这样任意的水平集函数 $\varphi_0(x, y): \varphi_0(c_0) = 0$, 方程(3.5)可以保证水平集函数 $\varphi(x, y, t)$ 随时间的演化满足 $\{\varphi(c[(p, t), t]) = 0\}$ 的条件, 即 φ 的零水平集始终是活动轮廓线 $c(p, t)$, 上述方程是曲线演化方程的欧氏表达。曲线的曲率可又水平集函数表示为

$$k = \nabla \frac{\nabla \varphi}{|\nabla \varphi|} = \frac{\varphi_{xx}\varphi_y^2 - 2\varphi_{xy}\varphi_x\varphi_y + \varphi_{yy}\varphi_x^2}{(\varphi_x^2 + \varphi_y^2)^{\frac{3}{2}}} \quad (4.6)$$

在具体实现水平集方法时, 通常将水平集函数选取为由初始曲线生成的符号距离函数(Signed Distance Function), 记为 SDF。

设 $\varphi(x, y, t=0) = d(x, y)$ 是 SDF, 则

$$\varphi(x, y, t=0) = d(x, y) \quad (4.7)$$

其中 $d(x, y)$ 是点 (x, y) 到初始曲线 c 的距离, 其符号根据点 (x, y) 在 c_0 的内外部而定, 如果 (x, y) 位于 c_0 的内部, 可取正号(或负号), 反之取负号(或正号)。在曲线演化过程中要求水平集函数始终保持为符号距离函数。因此, SDF 的快速稳定计算, 对于提高水平集方法的效率和稳定性来说是十分重要

的。因为演化方程仅从零水平集推导而来，一般速度函数只定义在零水平集，其它水平集没有定义，不过，单位法矢和曲率在所有的水平集都定义。因此，演化 ϕ 时需要一种方法将速度函数 $v(k)$ 推广到所有的水平集，变为延拓的速度场 $v_{ext}(k)$ 。然而许多速度场的延拓的方法常导致水平集函数不能保持符号距离函数，从而引起单位法矢和曲率计算上的误差。为减小误差，在经过若干步更新迭代后要重新初始化。

4.3 水平集窄带算法流程

水平集方法的基本思想是将平面闭合曲线隐含地表达为二维曲面函数的水平集，即具有相同函数值的点集，通过水平集函数曲面的进化隐含地求解曲线的运动。尽管这种转化使得问题在形式上变得复杂，但在问题的求解上带来很多优点，其最大的优点在于曲线的拓扑变化能够得到很自然的处理，而且可以获得唯一的满足熵条件的弱解；另外，界面的几何属性（法向量、曲率等）可以很容易得到；最后，这种思想可以不做任何改变推广到高维空间。

水平集函数的演化满足如下的基本方程：

$$\Phi_t + F|\nabla\Phi| = 0 \quad (4.8)$$

其中， Φ 为水平集函数，其零水平集表示目标轮廓曲线，即：

$$\Gamma(t) = \{x | \Phi(x, t) = 0\} \quad (4.9)$$

$|\nabla\Phi|$ 表示水平集函数的梯度范数； F 为曲面法线方向上的速度函数，控制曲线的运动，一般 F 包括 3 项：与图像有关的项(如梯度信息等)，与曲线的几何形状有关的项(如曲线的曲率等)以及附加的演化项(additional propagation terms)。

$$\frac{\partial\Phi}{\partial t} \approx \frac{\Phi_{i,j}^{t+1} - \Phi_{i,j}^t}{\Delta t} \quad (4.10)$$

$$\begin{aligned}
 (|\nabla\Phi|)^2 &\approx \left(\frac{\partial\Phi}{\partial x}\right)^2 + \left(\frac{\partial\Phi}{\partial y}\right)^2 \\
 &\approx \left(\frac{\Phi_{i+1,j}^t - \Phi_{i-1,j}^t}{2\Delta x}\right)^2 + \left(\frac{\Phi_{i,j+1}^t - \Phi_{i,j-1}^t}{2\Delta y}\right)^2
 \end{aligned} \tag{4.11}$$

我们可以得到下面的离散方程：

$$\Phi_{i,j}^{t+1} = \Phi_{i,j}^t + F(K) \left| \left(\frac{\Phi_{i+1,j}^t - \Phi_{i-1,j}^t}{2\Delta x} \right)^2 + \left(\frac{\Phi_{i,j+1}^t - \Phi_{i,j-1}^t}{2\Delta y} \right)^2 \right| \Delta t \tag{4.12}$$

水平集方法虽然在处理拓扑变化方面具有参数形变模型所无可比拟的优势，但是由于它是将问题转化到高一维空间来处理，必然会带来计算量方面的大幅增加。窄带(narrow band)算法是水平集方法实现中常用的算法，其主要思想是只更新零水平集附近点的水平集函数值，而无需计算整个图像空间中每一点的水平集函数值，以此来提高水平集方法的计算效率。如何选取窄带宽度成为提高窄带算法效率的关键，过宽必然加大计算量，过窄必然影响水平集函数值的差分迭代。

图 4.3 给出了水平集算法的主框架流程。该算法的输入数据为距离图，距离图是这样定义的，给定一条轮廓线，图像上每个像素的值表示该点到轮廓线上最近一点的距离，如果该点在轮廓线外部，其值为正，如果在内部，其值为负。输出数据仍为一副距离图，只是零水平线已经被更新过。该算法主要分为三个模块：

(1)初始化部分

一般在这个部分主要对输入图像进行处理，找出零水平线和窄带，初始化的方法有很多种，可以通过不同的子类来定义。

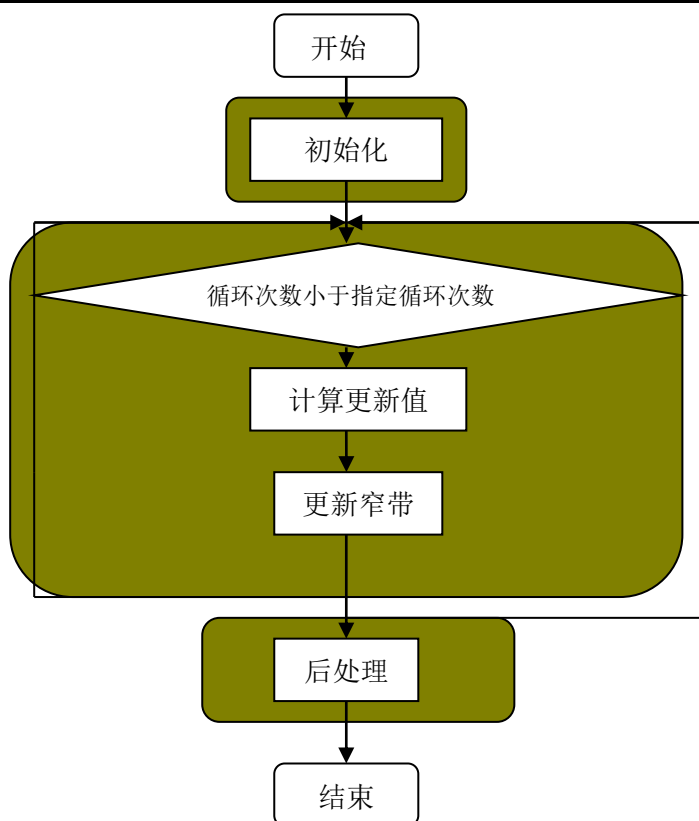


图 4.3 水平集算法主框架

(2)循环部分

这部分包括三块：

- 循环终止条件：一般用循环次数来定义。当循环次数小于指定循环次数时继续循环，大于时终止循环，这个条件由子类具体定义；
- 计算窄带内的点的更新值：应用数学公式，更新窄带内每一点的值；
- 更新窄带：根据上一步计算出来的更新值，重新计算零水平线和窄带。

(3)后处理部分。

4.4 实验结果

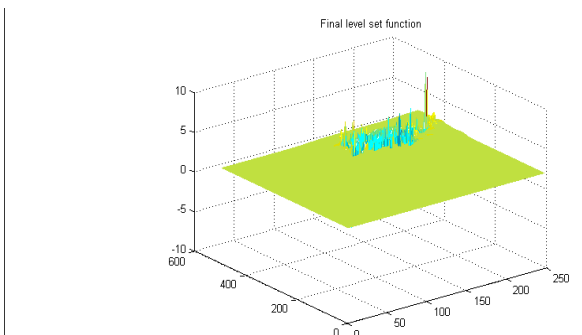
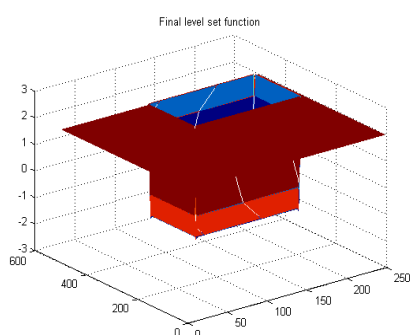
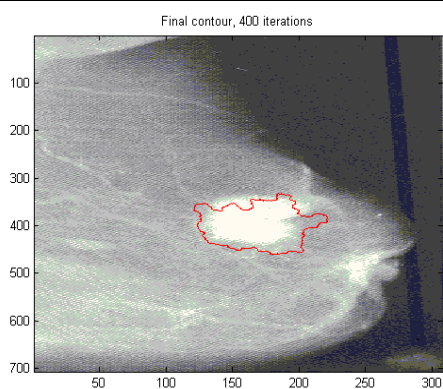
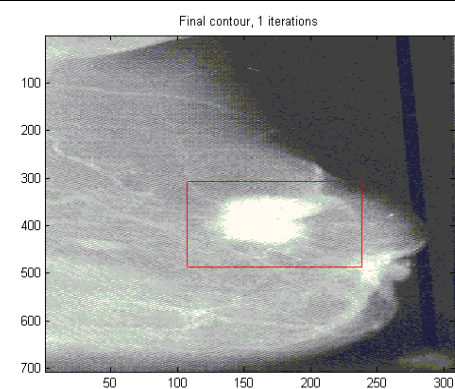


图 4.4 开始设定区域及其水平集函数

图 4.5 经过 400 次迭代后的分割区域及其水平集函数

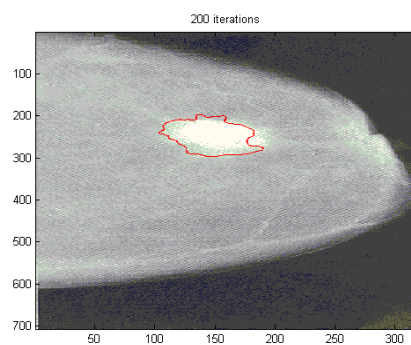
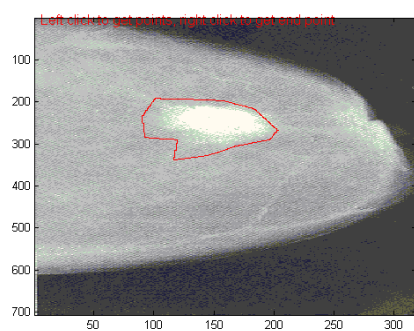


图 4.6 人工设定初始区域

图 4.7 200 次迭代后分割区域

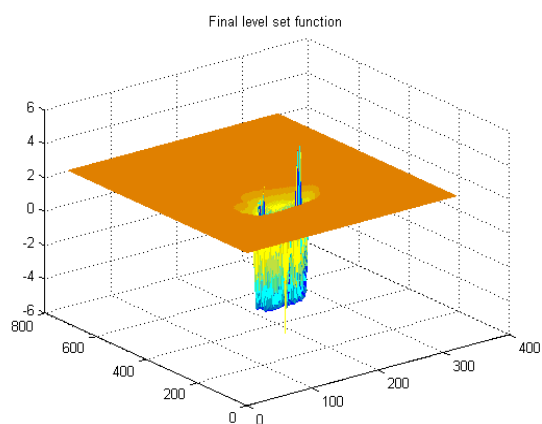


图 4.8 分割后区域曲线水平集函数

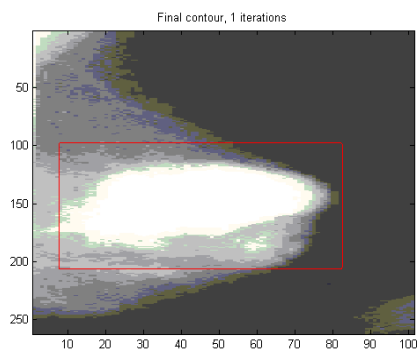


图 4.9 原始设定区域

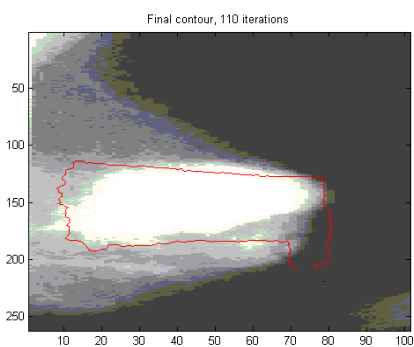


图 4.10 经 110 次迭代自动分割结果

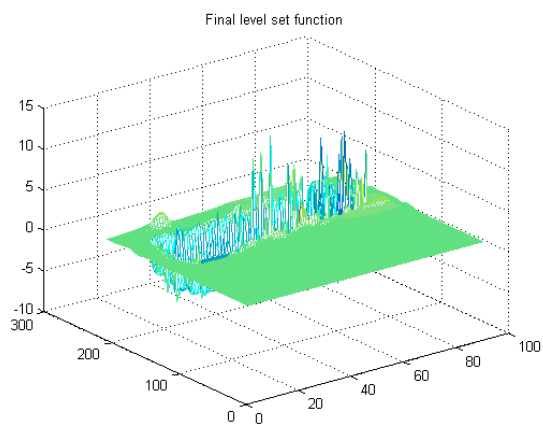


图 4.11 最后结果区域的水平集函数

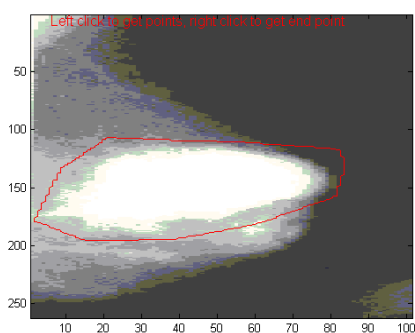


图 4.12 人工设定区域

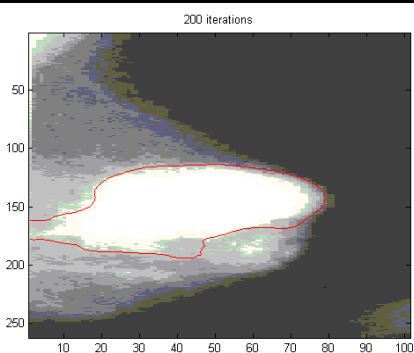


图 4.13 经过 200 次迭代后的结果

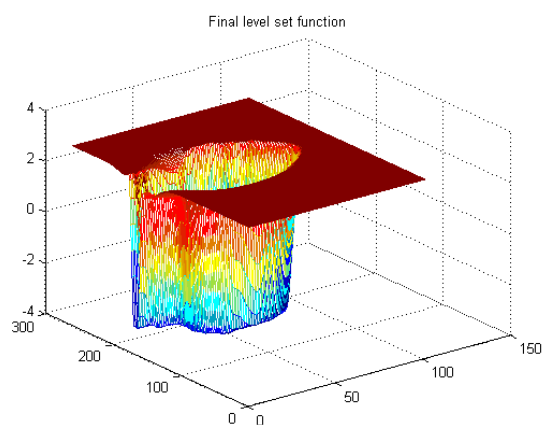


图 4.14 最终分割结果的水平集函数

4.5 本章小结

本章一开始介绍了医学图像分割的概念、分类和研究现状。医学图像分割是医学图像处理和分析的关键步骤，也是其他高级医学图像分析和解释系统的核心组成部分。本章主要以水平集分割对乳腺癌毛刺图片进行了肿块区域的分割，给出了全自动和半自动分割算法各一个，由于病理信息不确定因素过多等因素导致实际分割效果并不是完全理想，可以进行更深入的研究。

结 论

本文的主要研究内容是对乳腺癌肿块区域的辅助检测，首先利用基于小波变换的阈值化方法对乳腺钼靶 X 射线图像进行了去噪处理，通过与中值滤波、Winener 滤波的结果进行比较，显示这种方法不仅较好地消除了噪声和伪影，实现了乳腺区域的去噪处理，而且更好地保留了肿块区域的小叶与毛刺等细节特征，为后继肿块区域的检测等乳腺 X 线图像的处理提供了便利；然后本文采取了直方图均衡化这种简单实用的增强方法，增加了肿块区域，乳腺区域，背景区域的对比度，使图像的 ROI 在计算机视觉和人的视觉上都变得更加“醒目”。最后本文尝试利用基于水平集理论的分割方法对肿块区域进行分割，分别实现了全自动和半自动的区域分割，通过几十幅良性和恶性肿瘤图像的测试，证明本文采取的方法具有比较好的灵活性，通用性，更适应医疗中人机交互理念的发展。

可是本文的研究还存在着一些值得完善的地方，首先水平集分割算法还有待完善；其次还未能实现乳腺癌的辅助诊断，如果能继续进行研究，下一步就会在现有工作的继续完善的基础上，开始对乳腺肿块进行标定，模板匹配的研究；最后钙化点是乳腺癌钼靶 X 射线影像辅助检测中的另一个主要征象，往后的研究中会把这一病灶信息的辅助检测考虑在内。总之，辅助检测也许只是乳腺 CAD 系统研究的第一步，但诚如著名放射学家 Etta D. Pisano 所言：任何可以提高乳腺图像组织病变部位醒目性的显示技术都是非常有意义的进展。

相信会有更多的研究人员涉足乳腺 CAD 领域的研究，为女性带来更多福音。

参考文献

- [1] Marcano-Cedeño A, Quintanilla-Domínguez J, Andina, D. WBCD breast cancer database classification applying artificial metaplasticity neural network[J]. Expert Systems with Applications 38 (8), 2011: 9573-9579.
- [2] Muraca L, Leung D, Clark A, Beduz M.A, Goodwin P. Breast cancer survivors: Taking charge of lifestyle choices after treatment[J]. European Journal of Oncology Nursing 15 (3), 2011: 250-253.
- [3] <http://www.doh.gov.tw/static/data/卫生动向/一.pdf>, 2011-05-27/2011-06-4.
- [4] <http://www.doh.gov.tw/static/data/死因摘要/7.女性主要癌症死亡率.xls>, 2011-05-10/2011-06-4.
- [5] American Cancer Society, Breast cancer facts&figures 2007-2008 [OL]. <http://www.cancer.org/downloads/STTBCEFF-Final.pdf>, 2011-05-23/2011-06-4.
- [6] 罗拥志, 欧国军.乳腺癌的 X 线常见表现及分析[J].中国医药指南, 2008, 6(3): 96-97.
- [7] Carlos AP. Principles and practice of radiation technology[M]. Philadelphia: Lippincott Williams and Wilkins, 2004: 131-215.
- [8] FJ Gilbert, SM Astley, MG Gillan, et al. Single reading with computeraided detection and double reading of screening mammograms in the United Kingdom National Breast Screening Program [J]. Radiology, 2006, 241: 47-53.
- [9] N Karssemeijer, JD Often, AL Verbeek, et al. Computer aided detection versus independent double reading of masses on mammograms [J]. Radiology, 2003, 227 192-200.
- [10] MIRZAALIAN H, AHMADZADEH MR, SADRI S. Pre-processing Algorithms on Digital Mammograms[C], MVA2007 IAPR Conference on Machine Vision Applications, Tokyo, Japan, 2007, 118-121.
- [11] Kunio D. Computer-aided diagnosis in medical imaging: historical review,

- status and future potential. *Computerized Medical Imaging and Graphics*, 2007, 31(4): 198-211.
- [12] Cheng H D, classification Shi X J, Min R, et al of masses in mammograms. *Approaches for automated detection and Pattern Recognition*, 2006, 39(4): 646-668.
- [13] Rangaraj M R, Fabio J A, Desautels JEL. A review of computer-aided diagnosis of breast cancer: toward the detection of subtle signs[J]. *Journal of the Franklin Institute*, 2007, 344(3): 312-348.
- [14] Arnau O, Jordi F, Joan M, et al. A review of automatic mass detection and segmentation in mammographic images[J]. *Medical Image Analysis*, 2010, 14(2): 87-110.
- [15] 廖嘉煦, 李林. 调整窗宽窗位对 CR 乳腺铝靶影像诊断的应用价值[J]. *现代临床医学*, 2008, 34(6): 439-440.
- [16] R F, Hoffmeister J W, Zisman GS et al. A computer-aided detection system for the evaluation of breast cancer by mammographic appearance and lesion size[J]. *American Journal of Roentgenology*, 2005, 184(3): 893-896.
- [17] Huo Z, Giger M L, Wolverton D E, et al. Computerized analysis of mammographic parenchymal patterns for breast cancer risk assessment: Feature selection[J]. *Medical Physics*, 2000, 27(1): 4-12.
- [18] Sahiner B, Chan H P, Wei D, et al. Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue[J]. *Medical Physics*, 1996, 23(10): 1671-1684.
- [19] 邹莉玲, 沈其君, 陈峰, 郑凯尔. ROC 曲线下面积的 ML 估计与假设检验[J]. *中国公共卫生*. 2003, 19(1): 17-23.
- [20] 方志沂, 刘君. 乳腺癌诊断进展[J]. *中国病灶临床*. 2002, 29(11): 815-817.
- [21] 于金明, 孙晓蓉. 乳腺癌射线学诊断技术新进展[J]. *中华医学全科杂志*, 2003, 2(2): 44-47.

- [22] 冯敢生, 韩萍.重视乳腺射线学[J].临床放射学杂志, 2001, 20(1): 5.
- [23] 徐伟栋.乳腺 X 线图像的计算机辅助诊断技术研究[D].杭州.浙江大学, 2005.
- [24] 马鸣远.人工智能与专家系统导论[M].北京:清华大学出版社 2006, 11: 2-3.
- [25] T Uematsu, S Yuen, S Yukisawa, et al. Comparison of FDG PET and SPECT for detection of bone metastases in breast cancer [J]. AJR, 2005, 184(4): 1266-1273.
- [26] 张雪林.医学影像学[M].北京: 高等教育出版社, 2007: 288-300.
- [27] 刘佩芳.乳腺癌影像学诊断现状[J].中国医刊, 2003, 38(9): 16-18.
- [28] WT Yang. Sonographic, mammographic, and histopathologic corelation of symptomatic ductal carcinoma in situ [J]. AJR, 2004, 182: 101-110.
- [29] 成礼智, 王红霞, 罗永.小波的理论与应用[M].北京: 科学出版社.2006: 312-331.
- [30] 倪林.小波变换与图像处理[M].合肥: 中国科学技术大学出版社, 2010: 2-24, 140-155.
- [31] Walid Ahmed Abd Elhamid Dabour. Research on Multiscale Denoising and Mass Segmentation Methods for Digital Mammograms[D].Wuhan.Huazhong University of Science and Technology.2010.
- [32] 陈武凡.小波分析及其在图像处理中的应用[M].北京: 科学出版社, 2002: 152-166.
- [33] David L. Donobo, De-noising by soft-thresholding[J].IEEETrans. On Information Theory. 1995, 41(3): 617-627.
- [34] Rafael C.Gonzalez, Richard E.Woods. Digital Image Processing[M]. Dalian: Dalian University of Technology, 2010: 34-75.
- [35] 聂生东, 丘建峰, 郑建立.医学图像处理[M].上海: 复旦大学出版社, 2010: 58-64.

- [36] 田捷, 薛健, 戴亚康. 医学影像算法平台设计与平台构建[M].北京: 清华大学出版社, 2007: 208-216
- [37] J.A.Sethian. Level Set Method and Fast Marching Method [M]. London: Cambridge University Press, 1996: 54-103.

致 谢

四年的大学生活即将结束，首先我要向培育我的母校——哈尔滨工程大学致以诚挚的谢意，在此我度过了人生中一段美好而难忘的时光。

本论文是在王磊老师和李金老师的悉心指导和帮助下完成的，从论文选题，定稿到具体的研究工作都得到了王磊老师的悉心指导。在整个论文成稿过程中，王磊老师经常和作者就论文中涉及的一些学术问题进行讨论，指出作者研究中的不足，使作者受益匪浅。初稿完成后，老师们又数次认真仔细的审阅，每次都提出了一些有学术价值的建议与改进意见。值此论文成文之际，在此对王磊老师和李金老师致以深深的祝福和衷心的感谢。

这里我还要感谢我的家人、朋友，在我求学期间，他们在我生活和学习上给予了充分的关怀、支持与鼓励。

谨向在百忙之中抽出时间审阅论文和参加答辩的各位老师表示衷心的感谢。