

新疆地方性肝包虫 CT 图像的灰度直方图特征提取与分析

木拉提·哈米提¹,周晶晶¹,严传波¹,李莉¹,陈建军¹,胡彦婷¹,孔德伟²

1. 新疆医科大学医学工程技术学院,乌鲁木齐 830011
2. 新疆医科大学第一附属医院影像中心,乌鲁木齐 830054

摘要 图像特征提取是图像识别、图像数据挖掘、基于内容的图像检索等工作的基础,是模式识别和分类中的关键问题。本文运用灰度直方图法提取新疆地方性肝包虫 CT 图像特征,对图像进行尺寸归一、去噪和增强的预处理,并对灰度直方图特征进行统计分析,用最大类间距法获取图像分类的主要特征,同时使用判别分析法对特征的分类能力进行评价。结果表明,灰度直方图法提取的特征在统计分析中存在差异,且提高图像分类的准确率,一定程度上有助于对肝包虫病 CT 图像进行分类和检索。

关键词 灰度直方图;新疆地方性肝包虫;特征提取;最大类间距

中图分类号 R532.32, TP751.1

文献标识码 A

doi 10.3981/j.issn.1000-7857.2012.06.011

Feature Extraction and Analysis on CT Image of Xinjiang Local Liver Hydatid by Using Gray-scale Histograms

HAMIT Murat¹, ZHOU Jingjing¹, YAN Chuanbo¹, LI Li¹, CHEN Jianjun¹, HU Yanting¹, KONG Dewei²

1. College of Medical Engineer Technology, Xinjiang Medical University, Urumqi 830011, China
2. Department of Radiology, The First Affiliated Hospital, Xinjiang Medical University, Urumqi 830054, China

Abstract The feature extraction of images is a foundational work for image recognition, image data mining, and content-based image retrieval, and it is also the key issues of pattern recognition and classification. Feature extraction based on gray-scale histograms is a typical algorithm for the medical image feature extraction. For features of liver hydatid CT images that is extracted by using different gray-scale histograms are normalizing scale by uniform quantization, the noise is removed by using a median filter, the contrast is enhanced by limited adaptive histogram equalization; and then the gray-scale histograms is used to get the features of the image. The main features of the image classification are obtained by using statistical and maximum classification distance analysis on the histogram features, and then the classification ability of features is evaluated by discriminant analysis. The result shows that there is a certain discrepancy of statistical analysis for the features extracted by gray-scale histograms; features selected by maximum classification distance enhance the accuracy of image classification. This study would lay a solid foundation for the content-based medical image retrieval and the computer-aided diagnosis system to a certain extent.

Keywords gray-scale histogram; Xinjiang local liver hydatid; feature extraction; maximum classification distance

0 引言

肝包虫病是畜牧地区的常见病,是棘球绦虫的囊状幼虫(棘球蚴)寄生在肝脏内的寄生虫病。肝包虫病在新疆有较高的发病率,属于新疆地方性特色病。肝包虫病早期临床表现不明显且不易发觉,需要借助医学影像技术以确立诊断,其中 CT 检查是该病早期诊断和筛查的有效方法。

特征提取是对研究对象固有的、本质的及重要的特征或属性进行量测并将结果数值化,或将对象分解并符号化,形成特征矢量或符号串、关系图的过程^[1],是模式识别和分类中关键的一步^[2]。基于灰度直方图的特征提取是医学图像颜色特征提取中的一个典型算法。任何一幅图像的灰度直方图都包含了丰富的信息。Hu^[3]于 1961 年首次提出了矩不变量的概

收稿日期:2011-12-05;修回日期:2012-02-10

基金项目:国家自然科学基金项目(30960097,81160182)

作者简介:木拉提·哈米提,教授,研究方向为医学图像处理及信号分析,电子邮箱:murat.h@163.com

念,但是Hu矩不变量不具有正交性,所以包含大量冗余信息。Zemike矩^[4]不变量是一种正交的矩不变量,正交矩在信息冗余度、图像表达以及在识别效果方面比其他类型的矩要好。随着小波分析理论的发展,Shen等^[5]提出了基于小波变换的小波矩。由于小波变换具有时频局部化特征,所以小波矩不仅可以描述图像的全局特征,还可以描述图像的局部特征,不易受噪声干扰,因而在识别相似物体时有更高的识别率。图像的灰度直方图提供了该图像外观的一个全局描述,所提取的特征具有RST(旋转、比例和位移)不变性^[6]。

图像的灰度直方图(简称直方图)是表示一幅图像灰度

分布情况的统计特性,是图像处理中一种十分重要的分析工具。医学图像直方图特征是医学图像的基本而重要的特征数据,在肝脏疾病中使用,如脂肪肝识别。目前,对肝包虫CT图像进行特征提取的研究较少。肝包虫CT图像具有特有的病理特征,分为不同类型的包虫囊肿,且灰度分布不同,如图1所示。本文根据不同类型肝包虫的灰度分布,首先对肝包虫CT图像进行尺寸归一化、去噪和自适应增强的预处理;其次,进行灰度直方图特征提取;最后,对特征向量进行描述性统计分析和最大类间距分析,获得分类能力最好的特征,并用判别分析验证特征的分类能力。

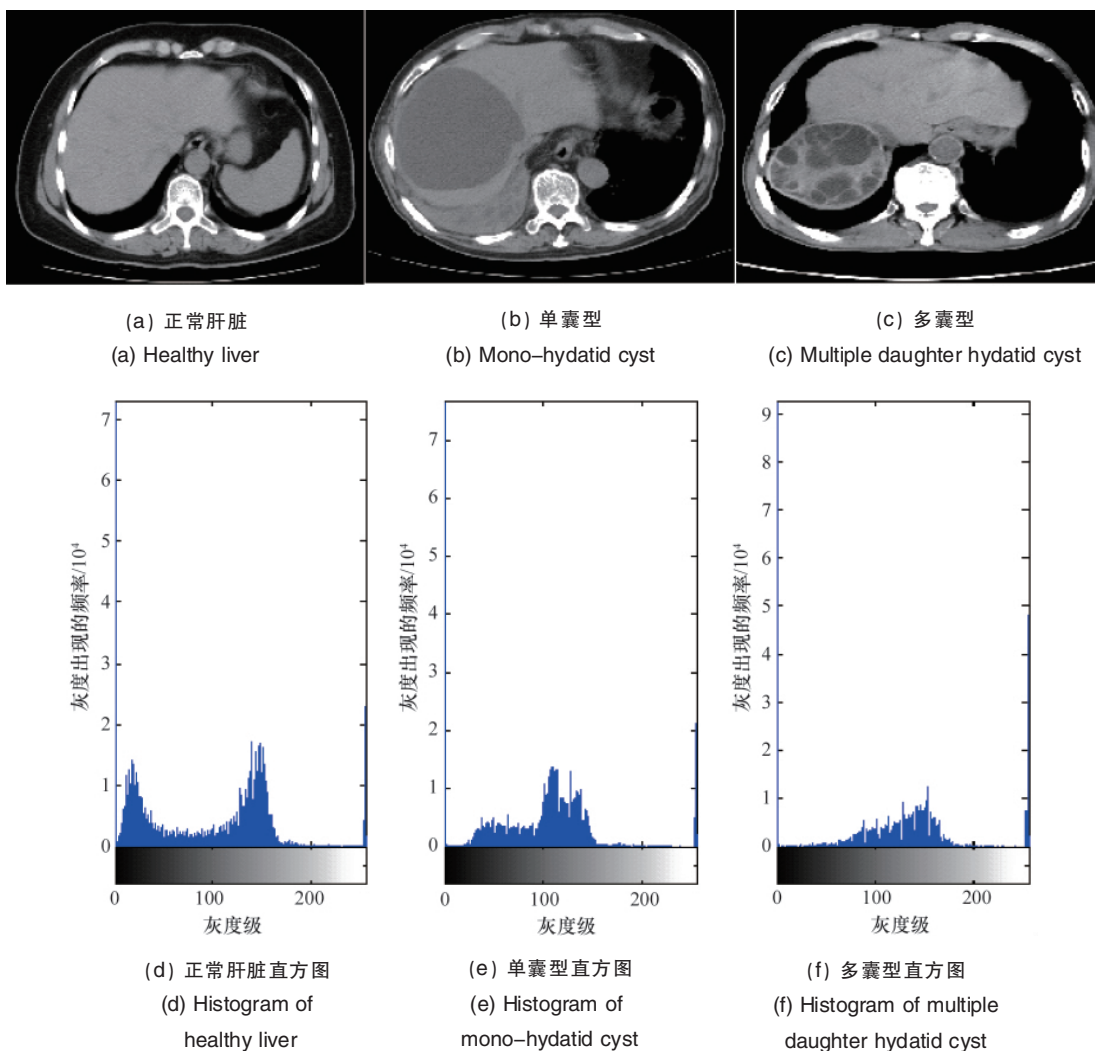


图1 不同类型肝包虫CT图像及其灰度直方图

Fig. 1 CT images and their gray-scale histograms for the different kinds of liver hydatid diseases

1 灰度直方图法的特征提取

1.1 灰度直方图的定义

一幅数字图像的灰度直方图就是一个灰度级的离散函数,可以用式(1)表示图像灰度直方图的定义。

$$H(i) = \frac{n_i}{N}, i=0, 1, \dots, L-1 \quad (1)$$

其中, i 表示灰度级, L 表示灰度级种类数, n_i 表示图像中具有

灰度级 i 的像素的个数, N 表示图像总的像素数。式(1)描述的是图像中具有该灰度级的像素的个数占图像总像素的百分比,即图像中具有灰度级 i 的像素出现的频率。

1.2 图像预处理

图像预处理的目的是对从医学设备获得的图像进行处理从而去除由于外界干扰出现的噪声得到更好的感兴趣区域的可视性和对比度,从而对医生的临床诊断和后续的图像

处理提供更好的支持。本文对肝包虫 CT 图像采取均匀量化的方法对图像进行尺寸归一化^[7]、中值滤波去噪和对比度自适应直方图均衡的预处理。

图 2、图 3 和图 4 分别表示正常肝脏、单囊型肝包虫病和多囊型肝包虫病的 CT 图像预处理结果, (a)为原始图像, (b)

为尺寸归一化后图像, (c)为去噪后图像, (d)为对比度自适应直方图均衡后图像。从图中可以看出, 图像预处理后, 图像的质量明显改善, 病理组织与正常组织之间的对比度明显增强, 为后续的图像特征提取工作打下良好基础, 进而提高图像分类和检索的准确率。

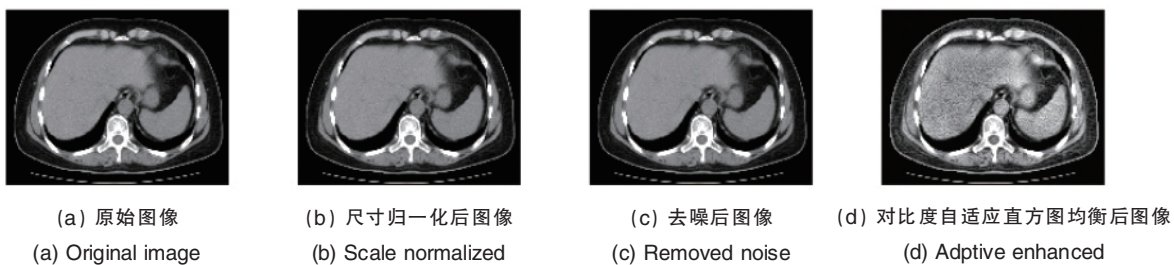


图 2 正常肝脏 CT 图像预处理结果
Fig. 2 Preprocessing results of healthy liver CT

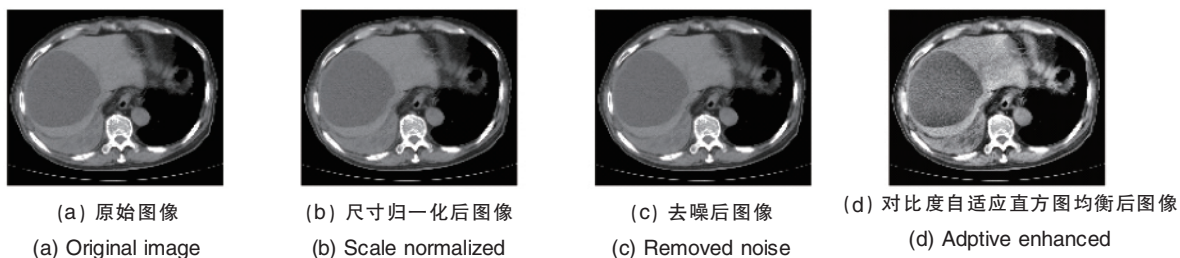


图 3 单囊型肝包虫病 CT 图像预处理结果
Fig. 3 Preprocessing results of mono-hydatid cyst CT

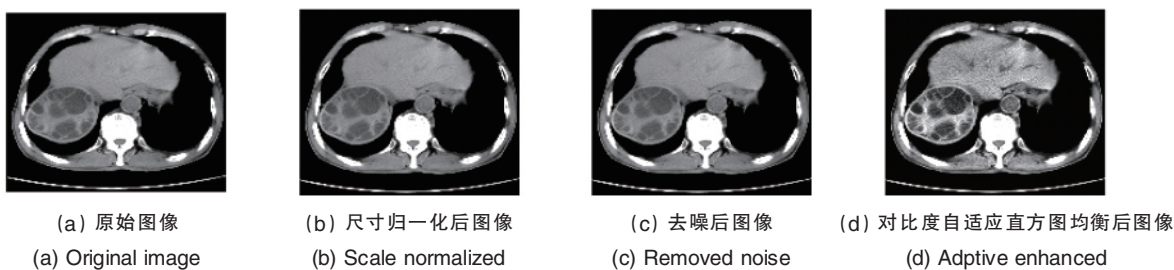


图 4 多囊型肝包虫病 CT 图像预处理结果
Fig. 4 Preprocessing results of multiple daughter hydatid cysts CT

1.3 基于直方图的统计特征提取

一般不是直接将医学图像的直方图作为特征, 而是通过用一些统计量来反映图像的直方图特征, 这些统计量通常称为直方图的统计特征^[8-9]。常用下列几种统计量反映图像的直方图特征。

均值: 反映一幅图像的平均灰度值。

$$\mu = \sum_{i=0}^{L-1} iH(i) \quad (2)$$

方差: 反映一幅图像的灰度在数值上的离散分布情况。

$$\sigma^2 = \sum_{i=0}^{L-1} (i-\mu)^2 H(i) \quad (3)$$

歪斜度: 反映图像直方图分布的不对称程度, 歪斜度越大表示直方图分布越不对称, 反之越对称。

$$\mu_3 = \frac{1}{\sigma^3} \sum_{i=0}^{L-1} (i-\mu)^3 H(i) \quad (4)$$

峰态: 反映图像的灰度分布在接近均值时的大致状态, 用以判断图像的灰度分布是否非常集中于平均灰度附近。峰态越小, 表示越集中; 反之, 表示越分散。

$$\mu_4 = \frac{1}{\sigma^4} \sum_{i=0}^{L-1} (i-\mu)^4 H(i) - 3 \quad (5)$$

能量: 反映灰度分布的均匀程度, 灰度分布较均匀时能量较大, 反之, 较小。

$$\mu_n = \sum_{i=0}^{L-1} [H(i)]^2 \quad (6)$$

熵: 反映了直方图灰度分布的均匀性。

$$\mu_e = \sum_{i=0}^{L-1} H(i) \log_2 [H(i)] \quad (7)$$

2 最大类间距法的特征筛选及判别分析

2.1 特征提取结果

实验对象为正常肝脏、单囊型肝包虫病和多囊型肝包虫病 3 类 CT 图像, 图像均由新疆医科大学第一附属医院放射

科提供。从影像数据库中抽取正常肝脏、单囊型和多囊型 3 种类型图像各 50 幅, 进行灰度直方图法的特征提取, 分别提取均值(P1)、方差(P2)、歪斜度(P3)、峰态(P4)、能量(P5)和熵(P6), 共计 6 个特征, 组成图像的特征向量。特征提取结果

表 1 灰度直方图法提取的 3 种图像特征
Table 1 Three kinds of images features extracted by gray-scale histograms method

图像类型	P1	P2	P3	P4	P5	P6
正常肝脏	99.34989	4025.94848	0.18586	-0.65086	0.00937	7.18320
	102.36463	3865.74939	0.07280	-0.65104	0.00974	7.13471
	103.65386	3814.20571	0.05690	-0.55980	0.01059	7.02389
			⋮			
单囊型	103.61481	3665.38157	-0.02546	-0.62284	0.00995	7.09428
	108.30674	2572.84977	0.89258	1.66969	0.01091	7.03385
	108.39285	2552.62787	0.85574	1.56133	0.01062	7.04146
			⋮			
多囊型	107.61914	2421.68464	0.77990	1.62148	0.01007	7.06857
	130.61447	4021.58532	0.32444	-0.05638	0.01898	6.92223
	128.67866	3189.53907	0.17943	0.54266	0.01256	7.00553
			⋮			
多囊型	130.08860	3073.30114	0.05771	0.64386	0.01245	6.98351
	129.55317	2986.67549	-0.01878	0.71697	0.01141	7.04164
	127.03650	3101.35274	0.08096	0.69476	0.01192	7.01374

如表 1 所示。

2.2 特征分析

本文采用描述性统计分析和类间距分析, 对表 1 中的实验结果进行分析。

2.2.1 描述性统计分析

通过描述性统计分析, 描述不同类型图像灰度直方图均值(P1)、方差(P2)、歪斜度(P3)、峰态(P4)、能量(P5)和熵(P6)的集中趋势和离散趋势, 计算图像特征向量的均值和方差。

2.2.2 最大类间距法的分类能力分析

适当地选择特征很重要, 因为它是区分对象的唯一依据。方差归一化间距是衡量用一个特征区分两类能力的一种指标^[10-11], 在对乳腺灰度和纹理特征筛选中取得较好效果。本

文采用归一化间距来评价和选取特征。对于某一个特征 x 而言, 第 i 类和第 j 类类间距离 D 为:

$$D = \frac{|u_{xi} - u_{xj}|}{\sqrt{\sigma_{xi}^2 + \sigma_{xj}^2}} \quad (8)$$

其中, u_{xi}, u_{xj} 和 $\sigma_{xi}^2, \sigma_{xj}^2$ 分别为第 i 类和第 j 类 x 特征的均值和方差。 D 越大, 说明该特征的分类能力越强; 反之, 该特征的分类能力越弱。表 2 为各特征的均值、方差及 3 种类型图像两两分类的类间距。

由表 2 中的统计分析结果可以看出, 不同类别图像之间的类间距值分布不同。在正常肝脏与单囊型类间距中, 歪斜度(P3)和熵(P6)的类间距大于 1, 说明在正常肝脏与单囊型分类情况下, 这两个特征具有较好的分类能力; 在正常肝脏与

表 2 特征统计分析
Table 2 Statistical analysis on gray-scale histogram features

特征	正常肝脏		单囊型		多囊型		类间距		
	均值	方差	均值	方差	均值	方差	正常肝脏与单囊型	正常肝脏与多囊型	单囊型与多囊型
P1	120.05990	1.12901	110.03210	69.59245	126.01659	131.68308	0.14407	0.04523	0.10732
P2	2611.74182	10002.90058	2734.59231	312175.43087	3290.63053	215365.49258	0.00039	0.06787	0.00178
P3	0.31508	0.02620	0.67472	0.03515	0.33412	0.02502	8.20404	0.00000	0.00000
P4	1.35008	0.00398	1.17280	0.31538	0.58743	0.14851	0.56207	5.13348	1.67922
P5	0.01132	0.00000	0.01132	0.00002	0.01383	0.00001	0.00618	461.13509	150.81088
P6	6.94920	0.00055	7.06130	0.00935	6.96350	0.00294	11.97237	4.77661	9.97948

多囊型类及单囊型与多囊型间距中,均为峰态(P4)、能量(P5)和熵(P6)的类间距大于1,说明在正常肝脏与多囊型及单囊型与多囊型分类情况下,这3个特征具有较好的分类能力。

2.3 Bayes 判别分析

根据表2中正常肝脏、单囊型和多囊型两两类间距计算所得结论,采用 Bayes 判别分析法,对上述结论进行验证。判别分析使用正常肝脏、单囊型和多囊型图像各50幅,通过本实验方法对图像进行处理,提取灰度直方图特征。

2.3.1 正常肝脏与单囊型包虫病

正常肝脏与单囊型类间距计算结果中,歪斜度(P3)和熵(P6)为类间距值大于1的特征。在正常肝脏与单囊型分类中,使用歪斜度和熵两个特征进行判别分析,由表3可看出分类正确率达92%和88%。

表3 正常肝脏与单囊型的 Bayes 判别分析
 Table 3 Bayes discriminant analysis on healthy liver and mono-hydatid cyst

原分类	判归分类	
	正常肝脏	单囊型
正常肝脏	46	4
单囊型	6	44

2.3.2 正常肝脏与多囊型包虫病

正常肝脏与多囊型类间距计算结果中,峰态(P4)、能量(P5)和熵(P6)为类间距值大于1的特征。在正常肝脏与多囊型分类中,使用峰态、能量和熵3个特征进行判别分析,由表4可看出分类正确率达90%和86%。

表4 正常肝脏与多囊型的 Bayes 判别分析
 Table 4 Bayes discriminant analysis on healthy liver and multiple daughter hydatid cysts

原分类	判归分类	
	正常肝脏	多囊型
正常肝脏	45	5
多囊型	7	43

2.3.3 单囊型包虫病与多囊型包虫病

单囊型与多囊型类间距计算结果中,峰态(P4)、能量(P5)和熵(P6)为类间距值大于1的特征。在单囊型与多囊型分类中,使用峰态、能量和熵3个特征进行判别分析,由表5可以看出分类正确率达82%和74%。

表5 单囊型与多囊型的 Bayes 判别分析
 Table 5 Bayes discriminant analysis on mono-hydatid cyst and multiple daughter cysts

原分类	判归分类	
	单囊型	多囊型
单囊型	41	9
多囊型	13	37

综合上述实验结果,通过类间距对特征进行筛选,可以提高图像分类的准确率,具有统计学意义,为进一步肝包虫图像分类提供依据。

3 结论

根据新疆地方性包虫病不同类型在灰度直方图分布中存在差异,结合肝包虫病的特点,提出基于灰度直方图的特征提取方法,并且通过类间距对各个特征进行评价。由实验结果可以看出,将灰度直方图与类间距结合的方法,可以在一定程度上提高分类能力,满足分类要求,对肝包虫类型判断提供一定的依据。这为后续基于内容的医学图像检索系统研究提供了一个方向。

参考文献(References)

- [1] 刘益新, 郭依正. 灰度直方图特征提取的 Matlab 实现 [J]. 电脑知识与技术, 2009, 5(32): 9032-9034.
Li Yixin, Guo Yizheng. *Computer Knowledge and Technology*, 2009, 5(32): 9032-9034.
- [2] 翟俊海, 赵文秀, 王熙照. 图像特征提取研究 [J]. 河北大学学报: 自然科学版, 2009, 29(1): 106-112.
Zhai Junhai, Zhao Wenxiu, Wang Xizhao. *Journal of Hebei University: Natural Science Edition*, 2009, 29(1): 106-112.
- [3] Hu M K. Visual pattern recognition by moment invariant [J]. *IRE Transactions on Information Theory*, 1962, 2(8): 179-187.
- [4] Khotanzad A, Hong Y H. Invariant image recognition by zernike moments [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, 12(5): 489-497.
- [5] Shen D G, Horace H S. Ip. Discriminative wavelet shape descriptors for recognition of 2-D patterns[J]. *Pattern Recognition*, 1999, 32(2): 151-165.
- [6] 李顺山, 庄天戈, 陈辉. 基于灰度直方图和互相关方法的医学图像检索[J]. 上海交通大学学报, 2001, 35(5): 694-698.
Li Shunshan, Zhuang Tiange, Chen Hui. *Journal of Shanghai Jiao Tong University*, 2001, 35(5): 694-698.
- [7] 李莉, 木拉提·哈密提, 艾克热木·阿西木, 等. 基于灰度共生矩阵的新疆地方性肝包虫 CT 图像特征提取方法 [J]. 科技导报, 2010, 28(16): 31-35.
Li Li, Hamit M, Aximu A, et al. *Science and Technology Review*, 2010, 28(16): 31-35.
- [8] 王姝勤. 肝脏 CT 辅助诊断系统中特征选择和提取研究 [D]. 上海: 上海交通大学, 2010.
Wang Shuqin. Feature extraction and selection in CT image based hepatic lesion diagnosis system [D]. Shanghai: Shanghai Jiao Tong University, 2010.
- [9] 黄亚丽, 李芬华, 赵真. B 超肝脏图像纹理特征提取方法的实验研究[J]. 中国医学影像技术, 2004, 20(12): 1937-1939.
Huang Yali, Li Fenhua, Zhao Zhen. *Chinese Journal of Medical Imaging Technology*, 2004, 20(12): 1937-1939.
- [10] 尹剑仑, 卫武迪, 梁永贵. 肿瘤超声图像灰度和纹理特征提取方法研究[J]. 通信技术, 2009, 42(9): 168-170.
Yin Jianlun, Wu Wudi, Liang Yonggui. *Communications Technology*, 2009, 42(9): 168-170.
- [11] 王庆福, 徐晓男, 李新天, 等. 基于 EMD 的水声目标信号的特征提取[J]. 声学及电子工程, 2009, 24(2): 1-3.
Wang Qingfu, Xu Xiaonan, Li Xintian, et al. *Acoustics and Electronics Engineering*, 2009, 24(2): 1-3.

(责任编辑 吴晓丽)