

多尺度脊波字典的构造及其在图像编码中的应用

邓承志¹⁾ 曹汉强^{1,2)}

¹⁾(华中科技大学电子与信息工程系,武汉 430074) ²⁾(国家防伪工程技术研究中心,武汉 430074)

摘要 在变换域图像编码技术中,图像的稀疏表示是编码的关键。该文在分析了脊波函数的缺陷的基础上,首先提出了一种多尺度脊波字典的构造方法,并通过树形结构对原子进行组织,加快了图像稀疏分解中最匹配原子的搜索速度;然后提出了一种基于树形多尺度脊波字典的匹配追踪静态图像编码方法;最后通过对量化失真与编码速率的分析,结合稀疏分解系数的分布,提出了系数的自适应量化和编码方案。实验结果表明,多尺度脊波字典能够对图像进行有效的稀疏表示;与 JPEG2000 相比,新的编码算法具有更好的编码性能,尤其在低比特率条件下。

关键词 图像编码 多尺度脊波字典 稀疏表示 匹配追踪

中图法分类号:TN919.81 文献标识码:A 文章编号:1006-8961(2009)07-1273-06

Construction of Multiscale Ridgelet Dictionary and Its Application for Image Coding

DENG Cheng-zhi¹⁾, CAO Han-qiang^{1,2)}

¹⁾(Department of Electronic & Information, Huazhong University of Science and Technology, Wuhan 430074)

²⁾(National Anti-counterfeit Engineering Research, Wuhan 430074)

Abstract The sparse representation of image is the key to the image coding technique in transform domain. In this paper, based on the analysis of the limitation of ridgelet function, a new construction method of multiscale ridgelet dictionary is firstly proposed. The speed of search of the most matching atoms is improved based on the tree structured organization of atoms. And then a tree structure multiscale ridgelet dictionary based matching pursuit image coding scheme is proposed. In the last the decomposed coefficients are adaptively quantized and encoded based on the distribution of coefficients and the rate-distortion analysis. Experiment results show that the new multiscale ridgelet dictionary can represent the images sparsely. On the other hand, the performances of the new method are shown to compare favorably against those of JPEG2000 scheme, especially at low bit rate.

Keywords image coding, multiscale ridgelet dictionary, sparse representation, matching pursuit

1 引言

变换域图像编码技术中,图像内容的稀疏表示是编码的关键。然而,由 1 维小波变换张成的 2 维可分离小波变换仅有有限的方向,用它来对图像进行表示时,并不能稀疏地表示图像中的曲线状的奇

异特征^[1]。为了更稀疏地表示图像等高维数据,人们提出了多尺度几何分析方法,其中常用的有脊波 (ridgelet) 变换^[2], 曲线 (curvelet) 变换^[3] 和 Contourlet 变换^[4] 等。

脊波变换是一种非自适应的高维数据表示方法,它能有效地处理 2 维空间中具有线状奇异特性的信号^[5-6]。遗憾的是,对于含曲线奇异特征的图

基金项目:国家自然科学基金项目(60772091,60462003);江西省教育厅科技项目(GJJ09366)

收稿日期:2007-12-10;改回日期:2008-04-08

第一作者简介:邓承志(1980~),男。现为华中科技大学电子与信息工程博士研究生。研究方向为多尺度几何分析、图像处理等。

E-mail: dengchengzhi@126.com

像,脊波变换的稀疏表示性能仅相当于小波变换。于是,Candes 又提出了单尺度脊波变换的概念,并给出了其构建的方法,它能较好地解决含曲线奇异特性的图像的稀疏表示问题^[7]。单尺度脊波变换的基本思想是利用剖分的方法,用直线来逼近曲线。当曲率足够小时,在图像的局部区域,曲线就可以近似地当作直线处理。由于单尺度脊波变换是采用分块方式处理图像,因此图像编码时会带来严重的块效应。

本文主要研究多尺度脊波字典的构造方法和基于树结构多尺度脊波字典的匹配追踪图像编码算法。论文首先构造出了多尺度脊波字典,从而为图像稀疏表示提供了基础;然后采用树结构对多尺度脊波字典进行组织来加快图像稀疏分解的速度;最后通过分析量化失真与编码速率模型,实现了稀疏分解系数的自适应量化和编码。

2 多尺度脊波字典的构造

脊波变换最早由 Candes 提出,它采用一系列脊波函数叠加的形式来表示函数^[5-7]。由于脊波变换可对线状奇异特征进行稀疏表示,因此被广泛应用于图像处理领域。然而,脊波函数尺度的“单一性”,导致它对曲线状奇异特征表示的不稀疏。

脊波函数尺度的“单一性”与脊波本身的定义有关。如果容许神经激活函数 $\psi(t)$ 满足 $\int \psi(t) dt = 0$ 及 $\int |\hat{\psi}(\xi)|^2 / |\xi|^2 d\xi < \infty$, 则称 $\psi_{a,b,\theta}(x,y) = a^{-1/2} \psi[(x \cos \theta + y \sin \theta - b)/a]$ 为脊波。由定义可知,由于尺度参数 a 的变化,仅仅改变的是垂直于脊线方向的尺度,并没有改变沿着脊线方向的尺度,因此脊波函数的长度布满整个空间,不利于图像中曲线特性的表示,正会引入“裂痕”噪声。图 1 给出了脊波函数随尺度参数 a 的变化情况。图 1 准确地反映了脊波函数尺度的“单一性”。

本文通过构造多尺度脊波字典来实现对图像中的曲线状奇异特征进行稀疏表示。与 2 维可分离张量积小波函数的构造类似,本文通过引入脊波尺度函数来构造多尺度脊波原子。称 $\varphi_{a,b,\theta}(x,y) = a^{-1/2} \varphi[(y \cos \theta - x \sin \theta - b)/a]$ 为脊波尺度函数,如果容许神经激活函数 $\varphi(t)$ 满足 $\int \varphi(t) dt = 0$ 且 $\int |\hat{\varphi}(\xi)|^2 / |\xi|^2 < \infty$, 则多尺度母脊波原子和多尺度父脊波原子分别定义为

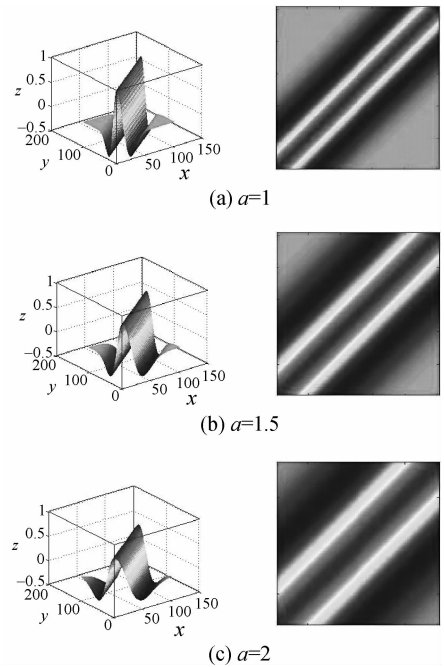


图 1 脊波函数随尺度参数 a 的变化

Fig. 1 Variance of ridgelet function with scale parameter a

$$g_{a,b,\theta}^M(x,y) = \psi_{a_1,b_1,\theta}(x,y) \varphi_{a_2,b_2,\theta}(x,y) \quad (1)$$

$$g_{a,b,\theta}^F = \varphi_{a_1,b_1,\theta}(x,y) \varphi_{a_2,b_2,\theta+\pi/2}(x,y) \quad (2)$$

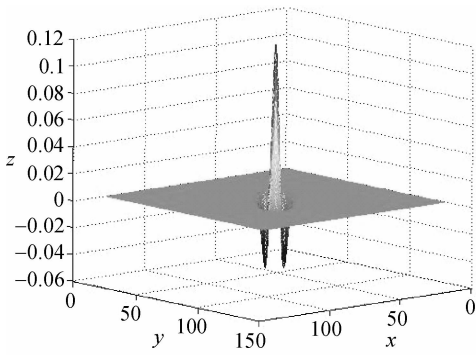
多尺度母脊波原子用来表示图像的高频分量,而多尺度父脊波原子则用来表示图像低频分量。由于图像的低频分量是各向同性的,并且不具有方向性,因此可以假定 $\theta = 0, a_1 = a_2$ 。此时多尺度父脊波原子为

$$g_{a_1,b}^F = a_1^{-1} \varphi((-x - b_1)/a_1) \varphi((y - b_2)/a_1) \quad (3)$$

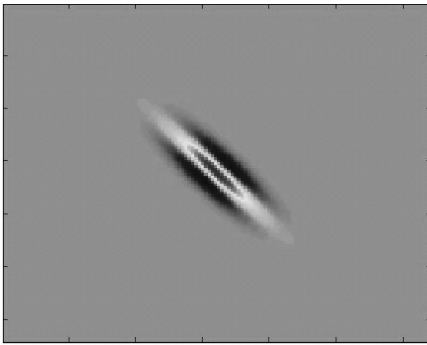
为了更有效地对图像中的曲线等结构特征进行稀疏表示,则要求多尺度母脊波原子具有多方向、多尺度和各向异性等特性,且必须满足:(1)沿边缘轮廓的方向上为平滑的低分辨率函数;(2)垂直与边缘的方向上函数能近似边缘的跃变。基于以上两点考虑,本文将容许神经激活函数 $\psi(t)$ 和 $\varphi(t)$ 分别选取为 Marr 小波和 Gaussian 小波。图 2 为 $a_1 = 1, a_2 = 4, \theta = 3\pi/4$ 的多尺度母脊波原子示意图。

通过对多尺度母脊波原子和父脊波原子中的尺度参数向量 $\mathbf{a} = (a_1, a_2)$, 平移参数向量 $\mathbf{b} = (b_1, b_2)$ 和方向参数 θ 进行“稠密”的过采样,就可以生成以下稀疏分解的多尺度脊波字典:

$$\mathbf{D} = \{g_{a,b,\theta}^M(x,y)\} \cup \{g_{a,b,\theta}^F(x,y)\} \quad (4)$$



(a) 3 维视图



(b) 平面图

图 2 多尺度母脊波原子

Fig. 2 Multiscale mother ridgelet atom

3 图像编码算法

为了测试多尺度脊波字典对图像表示时的稀疏性,本文提出了一种基于树形脊波字典的匹配追踪静态图像的编码算法,该算法流程如图 3 所示。算法首先利用匹配追踪对图像进行稀疏分解;然后对投影系数进行自适应量化;最后将最匹配原子的索引和量化后的投影系数进行自适应算术编码。

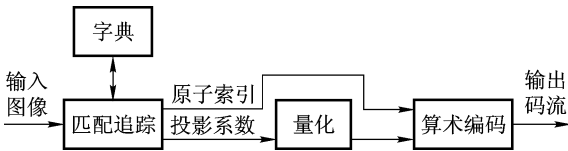


图 3 编码算法流程

Fig. 3 Flow chart of encoding scheme

3.1 字典的树形聚类

由于匹配追踪^[8]图像稀疏分解的运算量非常大,因而不利于图像编码的应用。本文采用对字典原子进行树形聚类^[9]的方式来减少匹配原子的搜索空间,从而降低了匹配追踪算法的运算量。

由于字典的冗余性使得原子间存在很强的相干性,因此可以用其对相干性强的原子进行聚类。聚类时,首先对相干性强的原子进行聚类,并形成分子(或节点);然后以此对相干性强的分子进行逐级聚类;最后形成一个树结构的字典。为了算法计算方便,每次聚类时,仅将 M 个相干的原子或分子进行聚类,即每个树节点仅包含 M 个子节点。

聚类过程中,可采用原子或分子的相干性距离 $d(g_i, g_j)$ 作为聚类的标准,其定义如下。

$$d(g_i, g_j) = 1 - |\langle g_i, g_j \rangle|^2 / (\|g_i\|_2 \|g_j\|_2) \quad (5)$$

$0 < d(g_i, g_j) \leq 1$, 该值越小,原子间的相干性越强,反之,相干性越弱。

3.2 系数的量化与编码

图像经匹配追踪稀疏分解后,将得到一组最佳的匹配原子位置索引和投影系数集。原子位置索引表征了原子的形状,必须对其进行无损压缩,这里采用自适应算术编码对原子索引集进行编码。投影系数是最佳匹配原子与图像间的内积,应先进行自适应量化,然后再对量化后的系数进行自适应算术编码。

稀疏分解迭代过程中,由于随着迭代的进行,投影系数逐渐变小,因此不同迭代阶段的投影系数对图像失真的影响不同。由于最初迭代系数的量化误差会引起较大的失真,因此本文采用可变量化阶的方法对不同迭代阶段的投影系数进行自适应量化。

通常,投影系数主要集中在某一特定的系数均值范围内;偏离均值越大,投影系数出现的概率越小^[10]。基于此,本文通过设计系数均值 $(m_i - \alpha\sigma_i, m_i + \alpha\sigma_i)$ 变化范围内的量化器的方式来实现投影系数的量化,其中 m_i 和 σ_i 为第 i 次迭代后投影系数的均值与标准差。设 $(m_i - \alpha\sigma_i, m_i + \alpha\sigma_i)$ 范围内的量化等级为 q_i , 则 k 次迭代的量化失真为

$$D(q_i) = \sum_{i=1}^k \frac{\alpha^2 \sigma_i^2}{3q_i^2} \quad (6)$$

假定字典大小为 M , 则 k 次迭代的原子索引值和量化投影值编码后的码率模型为

$$R = \sum_{i=1}^k \text{lb}q_i + k \text{lb}M \quad (7)$$

从式(6)和式(7)可以看出,量化等级 q_i 越大,失真越小,编码所需的码流越多。基于以上量化失真与码率模型,最优量化等级 q_i 可通过 Lagrange 乘子法得到,即

$$q_i = \arg \min_{q_i} \left\{ \sum_{i=1}^k \frac{\alpha^2 \sigma_i^2}{3q_i^2} + \lambda \left(\sum_{i=1}^k \lg q_i + k \lg M \right) \right\} \quad (8)$$

#对式(8)求导可得

$$q_i = \sqrt{\frac{2\alpha^2 \sigma_i^2 \ln 2}{(3\lambda)}} \quad (9)$$

由此可以得到

$$q_{i+1}/q_i = \sigma_{i+1}/\sigma_i \quad (10)$$

从上式可知,一旦 q_1 确定,则整个迭代过程中的量化等级 q_i 就可确定,而 q_1 可根据实际的比特率来确定。

4 实验结果与分析

为了验证多尺度脊波字典和本文图像编码算法的有效性,以 $256 \times 256 \times 8$ bits 的 Lena, Cameraman 和 Barbara 图像为例进行了实验。在编解码实验中,分别从客观保真度的衡量指标——峰值信噪比 (PSNR) 和主观视觉质量两方面,与 JPEG2000 和文献[10]算法的解码结果进行比较。

4.1 多尺度脊波字典稀疏性能测试

为了测试多尺度脊波字典对图像表示时的稀疏性,以 Lena 图为例进行了非线性近似实验,并通过实验对比了小波变换和有限脊波变换的非线性逼近性能。实验结果如图4所示,图4中,WT表示小波非线性逼近,FRIT为有限脊波变换^[11]非线性逼近,BFRIT为块有限脊波变换非线性逼近,MRD为本文算法的非线性逼近。

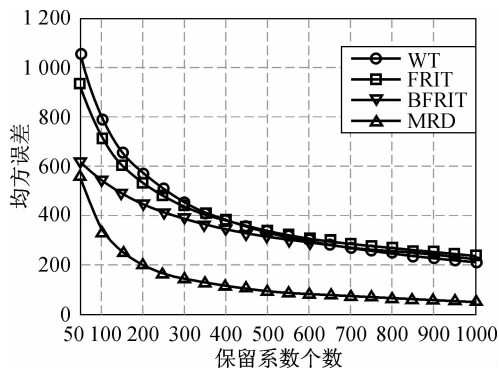


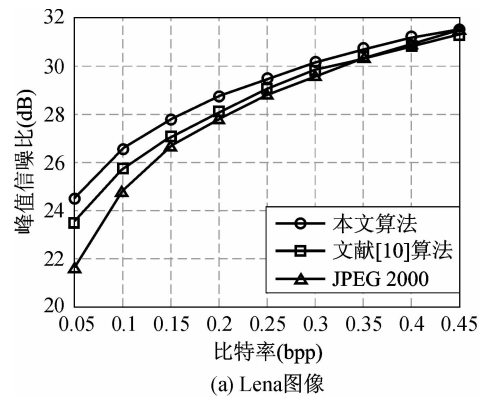
图4 非线性逼近性能比较

Fig. 4 Comparisons of nonlinear approximation performance

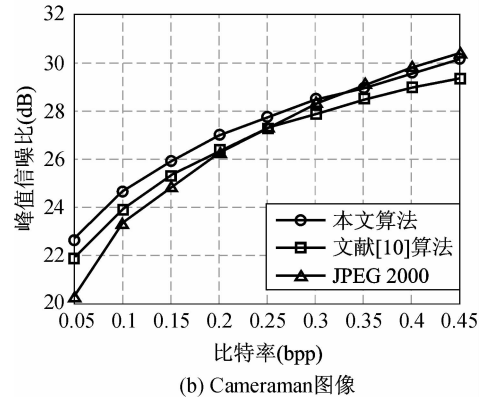
从图4可以看出,基于多尺度脊波字典的非线性逼近的性能最好,块有限脊波变换次之,有限脊波变换与小波变换两者的性能相当。

4.2 编码率失真性能比较

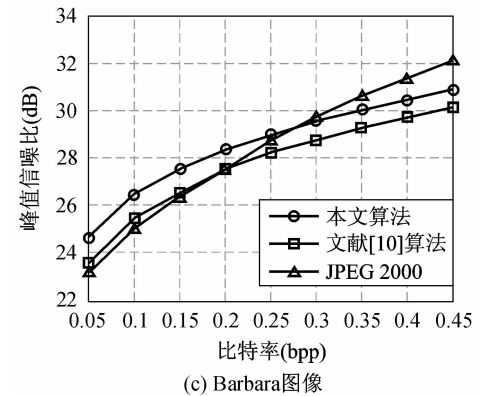
图5给出了不同编码比特率(比特/像素)下,3种算法的PSNR比较;其中图5(a)、图5(b)和图5(c)分别对应Lena, Cameraman 和 Barbara 图像的率失真性能曲线。从图5可知,在低比特率下,本文算法与文献[10]算法的编码性能都要优于JPEG2000。随着比特率的增加, JPEG2000 的性能将超过文献[10]算法,而与本文算法的性能相当。另外,在任意比特率下,本文算法都比文献[10]算法具有更好的编码性能。



(a) Lena图像



(b) Cameraman图像



(c) Barbara图像

图5 各算法编码性能比较

Fig. 5 The coding performance of various methods

由于多尺度脊波字典中的母脊波原子能很好地捕获图像中的边缘等结构特征,因此本文算法能用很少的比特率对这些信息进行编码。对于结构特征丰富的 Lena 图像,在较宽的比特率范围内,本文算法的编码性能优于 JPEG2000。对于 Cameraman 图像,由于图像中整个地面都是纹理特征,因此本文算法仅在比特率小于 0.35 bpp 时的编码性能才优于 JPEG2000。对于纹理和结构更为复杂的 Barbara 图像,本文算法在比特率小于 0.3 bpp 时的编码性能优于 JPEG2000。

对于 Lena 图像,当比特率大于 0.45 bpp 时, JPEG2000 的编码性能超过本文算法的编码性能。对于 Cameraman 和 Barbara 图像,仅当比特率大于 0.35 bpp 和 0.3 bpp 时, JPEG2000 的编码性能才优于本文算法。由此可见,本文算法非常适合低比特率或者甚低比特率图像或视频的编码。

4.3 视觉质量比较

图 6 为 Lena 图像在比特率为 0.35 bpp 时,各算法解码后的重建图像;图 7 为解码重建的 Lena 图像的局部细节。图 6(a)~图 6(d)分别为原图、JPEG2000、文献[10]算法和本文算法的解码重建图像。从这些图可以看出,本文算法与文献[10]算法可获得更为平滑的重建图像,且重建图像不存在

“振铃”噪声;而 JPEG2000 的重建图像则在边缘区域存在严重的“振铃”噪声。造成这一结果的主要原因是由于 JPEG2000 中的小波变换不能很好地对图像中的方向边缘等结构特征进行稀疏表示所致,且有损编码的数据丢失会导致误差的扩散。而本文构造的多尺度脊波字典则具有各向异性和方向性,其能用很少的数据对边缘特征进行稀疏表示。另外,从图 7 可以看出,对于图像边缘等结构特征,本文算法可获得最好的图像重建效果。



图 7 比特率为 0.35 bpp 时,解码后 Lena 图像的局部细节

Fig. 7 Details of reconstructed Lena images at 0.35 bpp



图 6 比特率为 0.35 bpp 时,各算法解码重建的 Lena 图像

Fig. 6 Reconstructed Lena image of various methods at 0.35 bpp

另外,国内外已对脊波变换域静态图像编码进行了一定的研究^[12-14]。由于这些方法都是基于传统的脊波变换,因此无法避免有限脊波变换存在的问题。这给实际的编码应用带来不良效果,易产生“裂痕”噪声。研究表明,对于曲线状的奇异特征,传统的脊波变换的非线性逼近性能仅相当于小波变换。传统的脊波变换域图像编码算法仅对含线状奇异特征的图像有效,而对含曲线状奇异特征的图像,其编码性能并不理想。本文避开了传统的脊波变换,构造出新的脊波函数。本文构造的脊波函数具有真正意义上的多尺度和多方向特征,且它对曲线状奇异特征的非线性逼近性能与曲线波变换相当。对于自然图像,本文的编码算法可获得更好的编码性能。算法更适用于实际的图像压缩编码系统。

5 结 论

多尺度几何分析由于能更加有效地对 2 维图像进行稀疏表示,从而利于编码的应用。脊波变换作为多尺度几何分析中的一种,更加适于处理具有线状和超平面状奇异性的数据。本文首先构造出多尺度脊波字典,这种新构造的字典,除了具有线状和超平面状奇异性外,还能对曲线状奇异特征进行稀疏表示;然后对字典原子进行树结构组织,从而加快了稀疏分解的速度;最后结合匹配追踪稀疏分解方法,提出了静态图像编码方案,即通过对稀疏分解数据的自适应量化和编码,实现了静态图像的编码。实验结果表明,多尺度脊波字典能对图像边缘和轮廓等特征进行稀疏表示。在低比特率编码环境中,本文算法编码的图像可取得更高的峰值信噪比;同时解码后的重建图像不存在“振铃”噪声,具有更好的视觉质量。如何提高本文算法在高比特率条件下的编码性能,将是下一步研究的主要目标。

参考文献 (References)

- 1 Candes E J, Donoho D L. Curvelets- a surprisingly effective nonadaptive representation for objects with edges [A]. In: Cohen A, Rabut C, Schumaker L, Editors; Curves and Surface Fitting [C], Nashville, Tennessee, USA: Vanderbilt University Press, 1999.
- 2 Candes E J. Ridgelet: Theory and Applications [D]. Department of Statistics, Stanford University, Stanford, CA, USA, 1998.
- 3 Candes E J, Demanet L, Donoho D L, *et al.* Fast discrete curvelet transforms [J]. Multiscale Modeling Simulation, 2005, 5(3): 861-899.
- 4 Do M N, Vetterli M. The contourlet transform: an efficient directional multiresolution image representation [J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091-2106.
- 5 Candes E J, Donoho D L. Ridgelets: A key to higher-dimensional intermittency [J]. Philosophical Transactions of the Royal Society of London Series A, 1999, 357(1760): 2495-2509.
- 6 Candes E J. Harmonic analysis of neural networks [J]. Applied and Computational Harmonic Analysis, 1999, 6(2): 197-218.
- 7 Candes E J. Monoscale ridgelet for the representation of images with edges [R]. Technology Report, 1999-27, Department Statistic, Stanford University, Stanford, CA, USA, 1999.
- 8 Mallat S, Zhang Z. Matching pursuit with time-frequency dictionaries [J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3397-3415.
- 9 Philippe J, Vandergheynst P, Frossard P. Tree-based pursuit: algorithm and properties [J]. IEEE Transactions on Signal Processing, 2006, 54(12): 4685-4697.
- 10 Frossard P, Vandergheynst P, Ventura R M, *et al.* A posteriori quantization of progressive matching pursuit streams [J]. IEEE Transactions on Signal Processing, 2004, 52(2): 525-535.
- 11 Do M N, Vetterli M. The finite ridgelet transform for image representation [J]. IEEE Transactions on Image Processing, 2003, 12(1): 16-28.
- 12 Si Jing-jing, Wang Cheng-ru, Cheng Yin-bo. Multi-layered image coding based on modified orthonormal FRIT [J]. Chinese Journal of Scientific Instrument, 2006, 27(10): 1283-1287. [司菁菁, 王成儒, 程银波. 基于改进的正交 FRIT 的分层图像编码算法 [J]. 仪器仪表学报, 2006, 27(10): 1283-1287.]
- 13 Shu Zheng-hua, Zou Dao-wen, Wang Sheng-qian, *et al.* A hybrid coder scheme based on the ridgelet and wavelet transforms [J]. Journal of Jiangxi Normal University (Natural Science), 2007, 31(2): 111-114. [舒正华, 邹道文, 汪胜前等. 一种基于 ridgelet 和 wavelet 的混合编码压缩算法 [J]. 江西师范大学学报 (自然科学版), 2007, 31(2): 111-114.]
- 14 Malmurugan N, Shanmugam A, Jayaraman, *et al.* A Novel Image Compression Algorithm using Ridgelet Transformation with Modified SPIHT [EB/OL]. <http://www.acadjournal.com/2004/V13/Part6/P2.html>. 2004-11-6.