

基于分段式拟合的低比特率图像压缩编码方法

李恒建^{①②} 陈怀新^② 张家树^①

^①(西南交通大学信号与信息处理四川省重点实验室 成都 610031)

^②(中国电子科技集团公司第10研究所 成都 610036)

摘要: 为了提高图像在低比特率条件下的解码质量和视觉效果, 该文提出一种基于稀疏分解的低比特率图像压缩编码新方法。利用二维不可分离、具有各向异性尺度的墨西哥草帽小波作为生成函数, 这种函数构建的冗余字典能够有效捕获图像边缘轮廓特征。为降低原子投影系数的冗余度和减小其编码量, 对图像稀疏分解投影的系数采用分段式拟合。获得的图像压缩码流具有渐进特性的, 满足现代无线通信对可伸缩码流的要求。实验结果表明, 在低比特率下, 该文方法与 JPEG2000 方法和通常的稀疏分解编码法相比, 有更高的峰值信噪比; 解压图像无振铃效应, 主观效果好。

关键词: 图像压缩; 稀疏分解; 分段式匹配追踪; 各向异性原子

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2008)09-2211-05

A Low-bit Rate Image Compression Algorithm Based on Piecewise Matching Pursuits

Li Heng-jian^{①②} Chen Huai-xin^② Zhang Jia-shu^①

^①(Sichuan Province Key Laboratory of Signal and Information Processing,
Southwest Jiaotong University, Chengdu 610031, China)

^②(China Electronics Technology Group Corporation No.10th Research Institute, Chengdu 610036, China)

Abstract: A novel image coding scheme based on sparse decomposition is proposed to improve decoded image quality and vision impact at low bit-rates. The over-complete dictionary, which is built on anisotropic refinement and rotation of 2D nonseparable Mexican hat wavelet, can efficiently capture edge characteristics present in natural images. And to improve the coding efficiency, an atom modulus which is piecewise approximated by the least squares line is proposed. Simultaneity, the stream can simply be truncated at any point to provide a SNR-adaptive bit-stream, which is very suitable for transmission in wireless communication. Experimental results show that the objective and subjective quality of decoded image is, at low bit rates, comparable to the state-of-the-art image coder, represented here by JPEG2000 and other based on Matching Pursuit methods. Also, the coding artifacts in the MP algorithm are less annoying than the ringing introduced by wavelets.

Key words: Image compression; Sparse decomposition; Piece-wise matching pursuit; Anisotropic refinement atom

1 引言

传统上, 图像压缩采用的是正交变换如离散余弦变换或小波变换来处理, 然而, 基于这两种变换的图像压缩方法在低比特率压缩下的解码图像质量差。对于高斯信源, DCT变换能够充分去除其相关性, 而图像是一种典型的非稳定信源, 呈超高斯分布, 因此二维正交DCT变换不是最优的。另外, JPEG在低比特率压缩下(如0.2bpp)会出现方块效应。二维张量小波变换不能有效捕获图像的正则性以及边缘轮廓特征, 使以小波变换为基础的JPEG2000在图像的低比特率压缩下会

出现蚊式噪声^[1]。因此, 在低比特率下, 图像压缩编码十分依赖于图像分解的选择以及对分解系数的有效编码。

稀疏分解以其良好的视觉特性和客观质量为低比特率图像和视频压缩编码提供了潜在的解决方案。在稀疏分解中, 冗余字典的构建方式和结构至关重要。研究表明, 具有各向异性结构和多方向的基原子更能稀疏表达图像^[2-4]。目前获取冗余字典主要有两种途径, 一种是通过神经网络学习的方法得到冗余字典, 然而该方法只适合对特殊的图像进行压缩编码^[5-7]。另外一种是通过基本函数生成冗余字典^[8-11], 由二维Gabor和正交小波构建的冗余字典, 在视频的运动补偿中取得了比较好的效果, 提高视频在低比特率下的解码质量^[8, 9]。近年来, 由不可分离的二维墨西哥草帽函数生成的冗余字典由于具有良好的视觉特性, 在图像稀疏分解的低比特率压缩编码中得到很好的应用^[10, 11]。

2007-02-07 收到, 2007-11-08 改回

四川省青年基金项目(03ZQ026-033), 四川省应用基础研究项目(2006 J13-10), 教育部新世纪优秀人才支持计划项目(NECT-05-0794)和西南交通大学博士生创新基金(2007)资助课题

匹配追踪算法是稀疏分解常用的算法,由此算法得到的图像分解系数收敛曲线的上界是指数递减的($2^{-\lambda}$)^[12]。根据分解系数这一变化规律,在图像稀疏分解上提出自适应指数量化方案,在低比特率下取得了良好的效果^[10]。进一步的研究表明,由匹配追踪得到的码流很容易产生渐进传输的特性^[11]。然而,自适应量化计算复杂度比较高,同时引入量化误差,使截断码流解码图像质量下降。为此,本文采用最小二乘法分段拟合图像稀疏分解系数,对原子索引参数按照原子对解码重建图像的重要性优先编解码,因此获得的码流具有可伸缩性。冗余字典由不可分离的二维墨西哥草帽函数生成,库中的原子的尺度具有各向异性特征,对边缘轮廓特征有很好的捕获能力。文章最后针对不同的图像类型,给出了本文提出方法与基于稀疏分解的压缩方法^[10, 11]及 JPEG2000^[13]方法在低比特率下图像压缩的对比结果。

2 图像稀疏逼近及其基于分段式匹配追踪的低比特率图像编解码

设研究图像为 f , 设 $D = \{g_\gamma, \|g_\gamma\| = 1\}_{\gamma \in \Gamma}$ 为冗余字典, Γ 表示所有索引构成的集合。该集合包含的向量可构成整个图像空间的一组基,能够张成整个图像空间。如果图像的维数为 d , 则 $|\Gamma| \geq d$ 。在冗余字典中,找到图像最稀疏的表达,这等同于解决下述问题:

$$\min \|c\|_0 \quad \text{s.t.} \quad f_N \approx \sum_{n=0}^{N-1} c_n g_{\gamma_n} \quad (1)$$

其中 $\|c\|_0$ 是序列 $\{c_n\}$ 中非零项的个数, $d \geq N$ 。选择最优的逼近方式已经证明是 NP 难题^[12]。Mallat 给出解决此问题的次优解,即匹配跟踪分解算法。该算法采用基于迭代的贪婪方式,通过每一步的最优代替整个过程的最优,进而寻找图像的最优逼近。

在稀疏分解中,虽然只要集合满足完备性条件,任意大小和任何波形的函数集合都可作为字典的生成函数,但是冗余字典的构建方式和结构非常重要。一般来说,如果字典结构符合图像的特征,那么得到的图像表示就越稀疏,对图像压缩编码就越有利。基于匹配追踪图像稀疏分解“贪婪”特性能够很快捕获到图像大部分有用信息,因此,稀疏分解适合图像低比特率压缩编码。图像稀疏分解的编解码系统包括 3 个方面:(1)冗余字典构建方式和字典生成函数选取;(2)稀疏分解算法收敛性和分解系数以及字典索引编码方法;(3)图像编码的失真和码流所具有特征。

2.1 冗余字典生成函数的选择及其构建

图像的边缘轮廓是图像的主要成分和特征。由于冗余字典没有正交的限制,选取的生成函数应在沿着图像轮廓的方向上为光滑的低分辨率函数,而与图像边缘垂直的方向上是具有与小波相似特性的函数,这样才能够有效逼近图像的非奇异特征。考虑到高斯函数的时频特性最好,符合人眼视觉特性,在表示图像有好的视觉效果,本文选择由高斯核函数和与其相垂直方向上一维高斯函数的二阶导数的乘积得到

函数作为生成函数,显而易见,该函数是不可分离的:

$$g(x, y) = \frac{2}{\sqrt{3\pi}} (4x^2 - 2) \exp(-x^2 - y^2) \quad (2)$$

图像是由多尺度和多方向的特征组成,因此,如果表示图像的基或者冗余字典具有这些特性就能够提高图像的非线性逼近速率^[2-4]。类似于图像的多尺度几何表示采用的基或者冗余字典的特点,本文采用的冗余字典 $D = \{U_\gamma : \gamma \in \Gamma\}$ 包括以下算子 U_γ : 平移、旋转和各向异性尺度伸缩。由于平移算子保证原子函数能够遍历整个图像空间,多方向使原子适应图像空域的灰度结构变化,多尺度适合图像渐进逼近的特点。如果两个正交方向上的尺度满足 $s_y = s_x$, 则称各向异性,各向异性符合图像边缘的支集特性,如图 1 所示。初步研究表明,用各向异性尺度原子表示具有段光滑边缘曲线的图像的系统能够显著提高编码增益,其率失真为 $D(R) \propto \log R \times R^{-2}$ ^[14]。

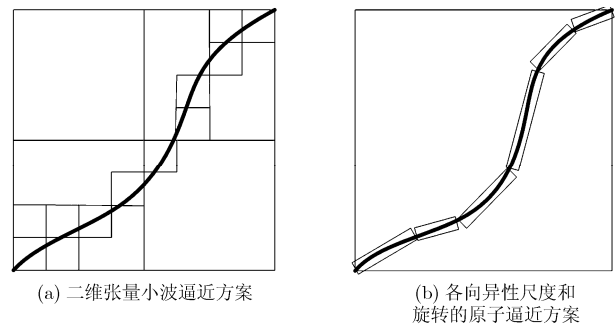


图 1 图像边缘的不同逼近方案

2.2 分段式匹配追踪的特点及其分解参数编码

设待分解的图像为 f , $R_0 = f$, $f_0 = 0$ 。在匹配追踪第 $m(m=1)$ 次迭代中,通过解决以下优化问题找到最佳匹配原子索引:

$$\gamma_m \in \arg \max_{\gamma \in \Gamma} \left| \langle R_{m-1}, g_\gamma \rangle \right| \quad (3)$$

更新图像残差:

$$R_m = R_{m-1} - \langle R_{m-1}, g_{\gamma_k} \rangle g_{\gamma_k} \quad (4)$$

匹配追踪就是通过不断在残差图像上迭代更新实现图像的稀疏表示。深入分析发现图像分解具有层次性,在小波分解上表现多分辨率特性,而在匹配追踪迭代呈现出分阶段的特征。在图像稀疏分解某一个局部分解阶段,认为原子系数的衰减速度相同,记为 λ , 则

$$\|R_{m+l}\| \approx 2^{-\lambda l} \|R_m\| \quad (5)$$

$l=1, 2, \dots, L$, L 为当前阶段总的迭代次数。匹配追踪的收敛特性 λ 可以表示为 $2^{-\lambda} = (1 - \alpha^2 \beta^2)^{1/2}$, 其依赖于字典的冗余度和优化搜索策略的选择。参数 $\alpha \in (0, 1]$ 表示优化因子,如果 MP 是全局最优,那么 $\alpha = 1$ 。参数 β 是字典的冗余因子,取决于原子库的结构特性,它反映原子库捕获输入信号特征的能力。设 c_{γ_m} 表示第 m 次匹配追踪迭代时的原子

投影系数, 即 $c_{\gamma_m} = \langle R_m, g_m \rangle$, 则由式(5)可得:

$$|c_{\gamma_{m+1}}| \approx 2^{-M} |c_{\gamma_m}| \quad (6)$$

式(6)表明, 匹配追踪得到的原子模值(modulus, 即系数的绝对值)曲线局部趋近于指数规律递减, 因此可以用分段的指数曲线进行拟合。为了尽可能减少逼近误差, 从而获得好的拟合效果, 本文选用一阶最小二乘法进行拟合, 得到的拟合参数即为模值取对数后的分段线段斜率和截距。每一个原子模值的拟合精度, 采用下式衡量:

$$\eta = \left| \frac{|c_\gamma| - \tilde{c}_\gamma}{|c_\gamma|} \right| \quad (7)$$

\tilde{c}_γ 为拟合得到的系数, c_γ 为原子的投影系数。在拟合过程中, 如果 $\eta > T$, 则把 η 对应的原子放入下一阶段进行拟合。对原子系数的模值具体拟合如下:

步骤 1 原子数据的预处理和算法初始化。输入阈值 T 和原子模值 $|c_\gamma|$ 以及索引参数 γ 。由于匹配追踪算法得到的原子模值并不是严格递减的, 因此首先需要对原子模值 $|c_\gamma|$ 做降序处理, 并记录排序后的原子序号。同时重新排列原子索引参数值, 使之与原子系数相对应, 并保存原子系数的符号。最后对原子模数 $|c_\gamma|$ 取以 2 为底的对数, 记为 $|\tilde{c}_\gamma|$, 即 $|\tilde{c}_\gamma| = \log_2 |c_\gamma|$ 。记分段拟合的原子序号起始位置为 i , 结束位置为 j , 当前阶段为 s 。初始化 $i=1, j=i+2, s=1$;

步骤 2 采用一阶最小二乘法对原子序号从 i 到 j $|\tilde{c}_\gamma|$ 进行拟合。计算拟合值 c_p 与对应原始的原子模值 $|c_\gamma|$ 差值的比率 $\eta(l), l=1 \cdots j-i+1$, 计算其最大值为 η_m , 比较 η_m 与预先设定的阈值 T 的大小;

步骤 3 若 η_m 大于阈值 T 时, 当前原子和前 $(j-i)$ 个原子模值在不同的匹配追踪阶段, 而序号从 i 到 $j-1$ 的原子处于同一线段, 计算其拟合系数和当前阶段的原子总数 $L(s)=j-i$ 。同时设 $i=j+1, j=j+3, s=s+1$, 转向步骤 2, 进行下一阶段的原子模值拟合;

步骤 4 若 η_m 小于阈值 T 时, 如果 j 达到原子序号 N (即匹配追踪的最大的迭代次数), 则算法结束; 否则 $i=j+1, j=j+3$ 转向步骤 2。

整个拟合过程中, 有 $\sum_s L(s) = N$ 成立。虽然拟合整个原子系数所得的斜率参数呈现递减的特点, 然而所得的整个参数的个数非常少, 可直接转化为二进制数据编码存储。对原子参数中的尺度参数 s , 首先用 $[s_x, s_y - s_x]$ 代替原始尺度参数 s , 然后用更新概率的算术编码算法进行编码。原子索引的其它参数满足均匀分布, 直接对其进行转化为二进制存储。

2.3 编码器的失真分析和码流可伸缩性

在解码端, 首先解码出原子系数, 其次依次解码出图像原子的索引值, 最后由原子系数和索引值生成原子图像, 重建图像为各个解码原子图像的线性叠加和:

$$\tilde{f}_N = \sum_{n=0}^{N-1} \text{sign}(c_{\gamma_n}) c_{p_n} g_{\gamma} \quad (8)$$

记单个原子系数的拟合误差为 $\xi_n = c_{\gamma_n} - \tilde{c}_{\gamma_n}$, 由 2.2 节可得, 系数的最大差值为 $\xi_n = \eta_m |c_{\gamma_n}|$, 拟合得到的系数与原始的系数之间满足

$$(1 - \eta_m) |c_{\gamma_n}| \leq c_{p_n} \leq (1 + \eta_m) |c_{\gamma_n}| \quad (9)$$

结合式(9)可得, 拟合编码和直接重建图像满足 $(1 - \eta_m)f \leq \tilde{f} \leq (1 + \eta_m)f$ 。如果利用图像的峰值信噪比衡量图像失真(本文考虑的量化等级为 8 的灰度图像, 即 $f_m = 255$), 那么:

$$\begin{aligned} \text{PSNR} &= 10 \lg \left(\frac{255^2}{\text{Mean}(\tilde{f} - f)^2} \right) \geq 20 \times \lg \frac{255}{\eta_m f_m} \\ &= -20 \times \lg \eta_m \end{aligned} \quad (10)$$

f 为原始图像, \tilde{f} 为重建图像。一般情况下, T 的取值很小, 如 $T=0.01$, 则 $\eta_m \leq T=0.01$, 即使系数在最差拟合的情况下, 拟合图像和直接重建图像的失真仍然远远大于 40dB。对于这样的拟合结果, 拟合误差在人眼的接受之外。在基于稀疏分解的图像压缩中, 图像的失真由两部分组成, 即拟合误差和匹配追踪截尾引起的失真 $\|R_N\|$ 。在本文提出的低比特率图像压缩编码算法中, 失真主要是由匹配追踪截尾引起的, 即

$$D = \sum_{j=0}^{N-1} |\xi_j|^2 + \|R_N\|^2 \approx \|R_N\|^2 \quad (11)$$

在解码端, 重要原子系数首先被解码, 并且索引参数随之被解码, 并不需要上下文的其它辅助信息支撑, 这样获得的码流以最小的代价给图像最好的逼近。同时产生的图像压缩编码具有可伸缩性, 符合渐进传输码流的特点。此外, 原子之间的独立性能够保证图像码流传输的稳健性和提高在有错误发生时解码图像的可视效果。设编码原子索引所需要的比特数为 R_γ , 另加一个符号位, 压缩后的图像的数据量 R 约为 $R \approx N \times (R_\gamma + 1)$, N 为编码的原子个数, 原子索引 $R_\gamma = R_p + R_s + R_\theta$ 。通过控制 N 的数目就可改变压缩图像的码率。

3 试验结果及分析

目前基于匹配追踪算法的稀疏分解图像压缩都是基于低比特的, 因此在实验时只考虑解码图像的比特率小于 0.25bpp 的情况。测试图像为 Resolutionchart 和标准 Lena, Peppers, Barbara 和 Goldhill, 大小为 256×256 , 原始图像如图 2 所示。图 3 和图 4 给出与基于稀疏分解的其它图像压缩方案^[8, 9]和 JPEG2000(采用默认的 CDF97 小波, 分解层数是五层)的比较结果^[14]。

在图 3 中, 在压缩倍数为 80 时, 本文压缩方法得到的图像 3(a)可以分辨出阿拉伯数字 3, 4, 5, 而 JPEG200 不能分辨出, 说明本文压缩方法有较好的图像空间分辨率。这是由于在冗余字典中引入旋转算子, 原子比张量小波基能有效地捕获数字的多方向特征。图 4 是 Lena 和 Peppers 在相同的比特率下本文和 JPEG2000 压缩方法下的主观效果比较。图 3(d)和 3(f)采用的是不可分离的原子逼近图像的特征,

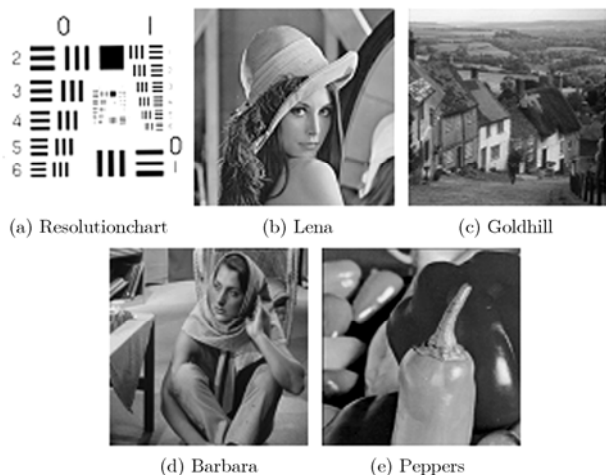
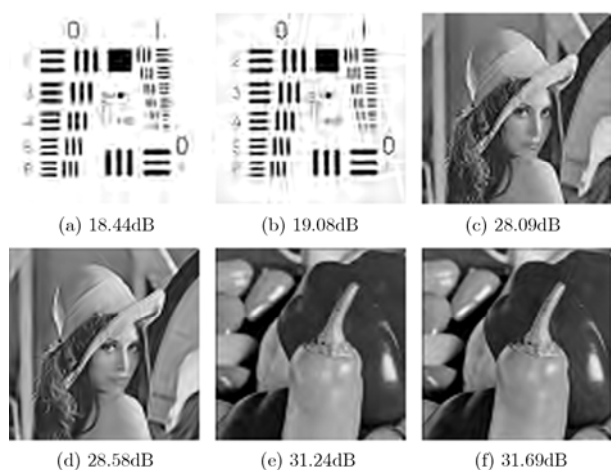
图2 标准的测试图像, 图像大小为 256×256 

图3(a)、3(c)和3(e)是 JPEG2000 压缩得到的图像。图3(b)、3(d)和3(f)是本文方法得到的压缩图像。图3(c)、3(d)、3(e)和3(f)在 0.19bpp 下的解码图像。图3 JPEG2000 和本文方法的对比结果

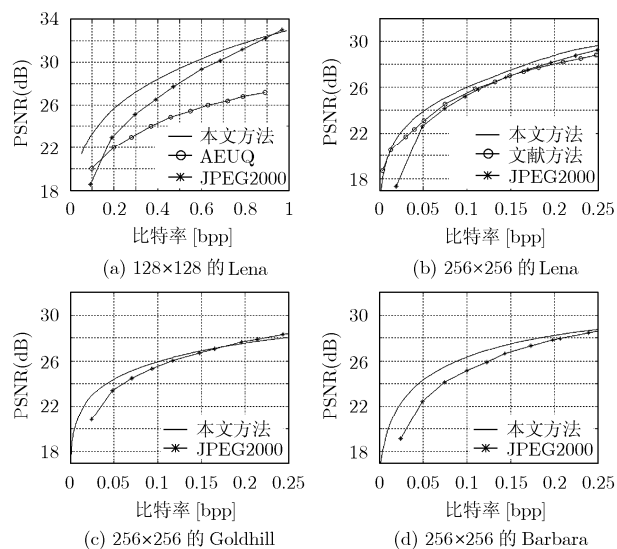


图4 不同压缩方法的峰值信噪比(PSNR)的比较

在图像边缘没有图3(c)和3(e)出现的毛刺现象, 同时原子具有好的时频特性, 解码图像有很好的视觉效果。

图4(a)采用 $128 \times 128 \times 8$ 的 Lena 图像, 在比特率小于 0.9bpp 时, 本文算法要优于 JPEG2000 方法和文献[11]方法。图4(b)采用 $256 \times 256 \times 8$ 的 Lena 图像在低比特率 (≤ 0.25 bpp) 下, 本文算法的 PSNR 高于 JPEG2000 算法和其它基于稀疏分解的压缩方法^[8, 9]。图4(c)(Goldhill, 大小 256×256) 和图4(d)(Barbara, 大小 256×256) 是本文方法与 JPEG2000 在低比特率下比较结果, 由于边缘轮廓特征没有 Lena 丰富, PSNR 的增益没有 Lena 图像高。由于匹配追踪在冗余字典中各向异性原子首先捕获图像边缘轮廓等重要信息, 因此在甚低比特率下, 稀疏分解的压缩方法解码图像的峰值信噪比要远远高于 JPEG2000 压缩方法。随着图像压缩的比特率增加, 需要大量比特对原子索引进行编码, 相对而言原子对解码图像的贡献比较小, 在 $\text{bpp} > 0.25$ 时, 本文方法解码图像的峰值信噪比小于 JPEG2000。

4 结束语

现有图像压缩标准采取正交变换的图像分解方式, 而这种方式难以对图像进行有效的稀疏表示, 因此在低比特率图像压缩编码时几乎得不到一个好的解码图像质量。稀疏冗余扩展方式一方面可以实现数据压缩的高效性, 更重要的是可以利用字典的冗余特性捕捉原始信号的自然特征, 在低比特图像压缩上能够获得好的客观质量和主观效果。自然图像往往含有多种成分, 单一生成函数构建的冗余字典还无法捕获整个图像的特征。下一阶段我们研究目标是采用多个生成函数构建的冗余字典来获得低比特率图像压缩的更好效果。

参考文献

- [1] Taubman D and Marcellin M. JPEG2000: Image Compression Fundamentals, Standards and Practice [M]. Boston, MA: Kluwer, 2001.
- [2] Peyre G and Mallat S. Surface compression with geometric bandelets[J]. *ACM Trans on Graphics*, 2005, 24(3): 601-608.
- [3] Velisavljevic V, Beferull-Lozano B, Vetterli M, and Dragotti P L. Directionlets: Anisotropic multi-directional representation with separable filtering[J]. *IEEE Trans. on Image Proc.*, 2006, 15(7): 1916-1933.
- [4] Do M N and Vetterli M. The contourlet transform: An efficient directional multiresolution image representation[J]. *IEEE Trans. on Image Processing*, 2005, 14(6): 760-769.
- [5] Murray J F and Kreutz-Delgado K. Learning sparse overcomplete codes for images [J]. *Journal of VLSI Signal Processing*, 2006, 45(1): 97-110.
- [6] Aharon M, Elad M, and Bruckstein A M. The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation [J]. *IEEE Trans. on Signal Processing*, 2006, 54(11): 4311-4322.

- [7] Elad M, Goldenberg R, and Kimmel R. Low bit-rate compression of facial images. *IEEE Trans. on Image Processing*, 2007, 16(9): 2376-2383.
- [8] Neff R and Zakhor A. Very low bit-rate video coding based on matching pursuits [J]. *IEEE Trans. on Circuits and Systems for Video Technology*, 1997, 7(1): 158-171.
- [9] 廖斌, 许刚, 王裕国. 基于非抽样小波冗余字典的低速率视频编码[J]. *软件学报*, 2004, 15(2): 221-228.
Liao Bin, Xu Gang, and Wang Yu-guo. Low Bit-Rate video coding based on undecimated wavelet dictionary[J]. *Journal of Software*, 2004, 15(2): 221-228.
- [10] Frossard P, Vandergheynst P, Ventura R, and Kunt M. A posteriori quantization of progressive matching pursuit streams [J]. *IEEE Trans. on Signal Processing*, 2004, 52(2): 525-535.
- [11] Figueras i Ventura R M, Vandergheynst P, and Frossard P. Low-rate and flexible image coding with redundant representations [J]. *IEEE Trans. on Image Processing*, 2006, 3(15): 726-739.
- [12] Mallat S and Zhang Z. Matching pursuit with time-frequency dictionaries. *IEEE Trans. on Signal Processing*, 1993, 41(12): 3397-3415.
- [13] <http://jj2000.epfl.ch>.
- [14] Figueras i Ventura R M, Granai L, and Vandergheynst P. R-D analysis of adaptive edge representations[C]. in Proc. MMSP, St Thomas, US Virgin Islands, 2002: 130-133.
- 李恒建: 男, 1981年生, 博士生, 研究方向为图像编码和安全传输.
- 陈怀新: 男, 1963年生, 研究员, 博士生导师, 研究方向为图像压缩编码和智能图像处理.
- 张家树: 男, 1965年生, 教授, 博士生导师, 研究方向为视频图像处理及其安全应用等.