

基于多原子快速匹配追踪的图像编码算法

邓承志^① 汪胜前^① 曹汉强^②

^①(南昌工程学院计算机科学与技术系 南昌 330099)

^②(华中科技大学电子与信息工程系 武汉 430074)

摘要: 该文提出一种多原子快速匹配追踪信号稀疏分解算法,并将其应用于静态图像编码。多原子匹配追踪通过每次迭代选取多个原子的形式,实现信号的快速稀疏分解。在此基础上,通过构造多尺度脊波字典实现图像的稀疏分解,并对稀疏分解的数据进行自适应量化和编码。实验结果表明,多原子匹配追踪获得了与匹配追踪相当的逼近性能,同时极大地提高了稀疏分解的速度。新的编码算法在低比特率情况下,获得了比 JPEG2000 更理想的编码性能。

关键词: 图像编码; 稀疏表示; 冗余字典; 匹配追踪; 多尺度脊波字典

中图分类号: TP911.73

文献标识码: A

文章编号: 1009-5896(2009)08-1807-05

Multi-atoms Rapid Matching Pursuit-Based Image Coding Algorithm

Deng Cheng-zhi^① Wang Sheng-qian^① Cao Han-qiang^②

^①(Department of Computer Science & Technology, Nanchang Institute of Technology, Nanchang 330099, China)

^②(Department of Electronics & Information Engineering, Huazhong Univ. of Science & Technology, Wuhan 430074, China)

Abstract: In this paper, a Multi-Atoms rapid Matching Pursuit (MAMP) algorithm for signal sparse decomposition and its application to image coding are proposed. The MAMP algorithm decompose signal sparsely by selection several atoms at each iteration. A multiscale ridgelet dictionary is constructed and used to represent image based MAMP. The sparsely decomposed data are adaptively quantized and encoded. Experimental results show that the approximation performances of the MAMP algorithm are comparable with those of the matching pursuit. Meanwhile, the computation speed is greatly improved. On the other hand, the performances of the new coding scheme are shown to compare favorably against those of the state of the art JPEG2000 scheme at low bit rate.

Key words: Image coding; Sparse representation; Redundant dictionary; Matching pursuit; Multiscale ridgelet dictionary

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。小波变换也因其良好的空间/频率局部化特性而成为静止图像编码新标准 JPEG2000^[1]的基础。然而,由一维小波变换张成的二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奇异特征;因此不利于低比特率环境下的图像编码。

超完备信号稀疏表示通过引入超完备冗余字典来实现对信号的最优非线性稀疏展开。自 1993 年 Mallat 和 Zhang 首次提出基于超完备冗余字典的信号稀疏表示思想和匹配追踪(Matching Pursuit, MP)算法以来,超完备信号稀疏表示^[2-5]已成为调和分

析和信号处理领域的一个研究热点。近年来,超完备稀疏表示理论^[6-8]被广泛应用于图像处理各领域,并取得了较好的成果。然而,由于匹配追踪算法的运算复杂度,给图像编码的实用化带来相当的困难。为此,人们提出了一些匹配追踪的改进方法。这些方法主要通过对字典进行修正的方式,加快匹配追踪中最优原子的搜索速度。已提出的字典修正方法有, Gribonval 和 Vleeschouwer 等人^[9,10]构造多尺度字典和子带字典, Schmid-Saugeon 和 Chou 等人^[11,12]提出的矢量量化法,以及 Jost 等人^[13]提出的树结构字典等。

以上方法虽说一定程度上降低了匹配追踪算法的运算复杂度,但稀疏分解的速度仍然非常缓慢。造成这一问题的主要原因是:在每次迭代过程中,匹配追踪算法仅挑选出一个与残差最匹配的原子。本文提出了多原子快速匹配追踪算法;算法通过每

次迭代选取多个匹配原子的形式,实现图像的快速稀疏分解。在此基础上,通过构造多尺度脊波字典实现图像的稀疏分解。最后通过分析量化失真与编码速率之间的关系和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。

2 多原子匹配追踪(MAMP)

2.1 字典的非相干分解

每次迭代过程中,多原子匹配追踪算法选取若干个匹配原子对信号进行稀疏分解;为了保证分解的稀疏性,要求选取的原子间的相干性足够小。为此,算法首先对字典进行非相干分解,即将字典分解成若干非相干的子字典;然后进行多原子稀疏分解。给定冗余字典 D ,将其分解成 p 个子字典 D_k ,即 $D = \bigcup_{k=1}^p D_k$,且当 $i \neq j$ 时满足 $D_i \cap D_j = \emptyset$ 。

子字典相干系数定义如下:

$$\mu_p \triangleq \max_{i \neq j, 1 \leq i, j \leq p} \max_{k, l} \left| \langle g_k^i, g_l^j \rangle \right| \quad (1)$$

其中 $\langle \cdot, \cdot \rangle$ 为内积, g_k^i 和 g_l^j 分别表示子字典 D_i 和 D_j 的第 k 个和第 l 个原子。

子字典相干系数 μ_p 的大小反映了子字典间的相干性强弱; μ_p 越大,子字典间的相干性越强,反之,越弱。给定字典 $D = \bigcup_{k=1}^p D_k$,如果子字典相干系数 μ_p 足够小,则定义字典 D 为子字典非相干字典。由定义可知, $\mu_p \leq \mu_1$,其中 μ_1 为字典 D 的相干系数,即 $\mu_1 = \max_{i, j \in D, i \neq j} \left| \langle g_i, g_j \rangle \right|$ 。

本文通过K均值聚类的方式将字典分解成 M 个非相干的子字典。聚类过程中,采用相干性距离 $d(g_i, g_j)$ 作为原子相似性的度量。

$$d(g_i, g_j) = 1 - \frac{\left| \langle g_i, g_j \rangle \right|^2}{\left(\|g_i\|^2 \|g_j\|^2 \right)} \quad (2)$$

其中 $0 < d(g_i, g_j) \leq 1$ 。字典分解过程中,首先将字典聚类成 M 个子字典;然后根据设定的子字典相干系数 μ_p 对 M 个子字典进行聚类修正。聚类修正过程中,如果任意两个子字典中原子的相干系数大于 μ_p ,则将这两个原子放在同一个子字典中,直到子字典相干系数等于 μ_p 。

2.2 多原子稀疏分解

对字典 D 进行非相干分解后,可以得到 M 个非相干的子字典;此时,可以对信号进行多原子匹配追踪稀疏分解。假定 f 为待分解信号,多原子匹配追踪算法的具体步骤如下:

(1)初始化参数。令迭代次数 $i = 0$,初始逼近信号 $f^{(0)} = 0$,初始残差信号 $R^{(0)} = f$,迭代终止误差 ϵ_{stop} ,阈值参数 γ ,相干系数 μ_{stop} ,多原子集 $g = \emptyset$ 。

(2)找出字典 D 中与残差 $R^{(i)}$ 最匹配的原子 g_{opt} ,

即 $g_{\text{opt}} = \arg \sup_{g \in D} \left| \langle R^{(i)}, g \rangle \right|$ 。

(3)分别找出各子字典 D_k 中与残差 $R^{(i)}$ 最匹配的原子 g_{opt_k} ,即 $g_{\text{opt}_k} = \arg \sup_{g \in D_k} \left| \langle R^{(i)}, g \rangle \right|$;如果 $\left| \langle R^{(i)}, g_{\text{opt}_k} \rangle \right| \geq \gamma \left| \langle R^{(i)}, g_{\text{opt}} \rangle \right|$,则将原子 g_{opt_k} 加入到多原子集 g ,即 $g \leftarrow g \cup \{g_{\text{opt}_k}\}$ 。

(4)根据原子与残差间相干性的强弱,对 g 中的 m 个原子进行排序,如果 m 个原子的累积相干系数 $\mu_{\text{cum}}(m) > m \cdot \mu_{\text{stop}}$,则舍去相干性最差的原子,直到 $\mu_{\text{cum}}(m) \leq m \cdot \mu_{\text{stop}}$ 。其中累积相干系数定义为

$$\mu_{\text{cum}}(m) \triangleq \max_{|g|=m} \max_{\omega \in \Lambda} \sum_{\lambda \in \Lambda} \left| \langle g_\omega, g_\lambda \rangle \right| \quad (3)$$

(5)计算残差 $R^{(i)}$ 在最后选定的若干原子上的正交投影 $P_V R^{(i)}$,更新稀疏信号逼近 $f^{(i+1)} = f^{(i)} + P_V R^{(i)}$ 和逼近残差 $R^{(i+1)} = f - f^{(i+1)}$ 。

(6)判定终止条件。如果 $\|R^{(i+1)}\|^2 < \epsilon_{\text{stop}}$,终止迭代;否则, $g = \emptyset$, $i = i + 1$,并转到第(2)步继续迭代。

2.3 算法复杂度分析

多原子匹配追踪算法主要包含两个步骤,即多个原子的选取和正交投影。假定字典 D 包含 N 个原子,每次内积运算需要 p 次操作,则匹配追踪每次迭代选取一个匹配原子的运算复杂度为 $O(Np)$ 。由于多原子匹配追踪算法中的第(2)和第(3)步可同时进行,因此每次迭代选取 n 个原子的运算复杂度仍为 $O(Np)$ 。 $P_V R^{(i)}$ 是残差 $R^{(i)}$ 在选定的若干原子所张成的子空间上的正交投影。多原子匹配追踪选定 n 个匹配原子时, $P_V R^{(i)}$ 的运算复杂度为 $O(nq)$,其中 q 为常数。

因此,对于信号的 n 个原子的稀疏逼近,多原子匹配追踪算法的运算复杂度为 $O(Np + nq)$,而匹配追踪算法的运算复杂度为 $O(Nnp)$ 。由于字典的超完备性,即 $N \gg n$;则 $Np \gg nq$, nq 相对于 Np 可以忽略。因此对于 n 项稀疏逼近,多原子匹配追踪算法的运算速度约为匹配追踪算法的 n 倍。

3 基于多原子匹配追踪的图像编码算法

多原子匹配追踪极大地提高了信号稀疏分解的运算速度,因此有利于图像编码应用的开展。本文通过构造多尺度脊波字典,实现基于多原子匹配追踪的静态图像编码算法,算法流程如图1所示。算法首先利用多原子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;最后将最佳匹配原子索引和量化后的投影系数进行自适应算法术编码,得到输出码流。

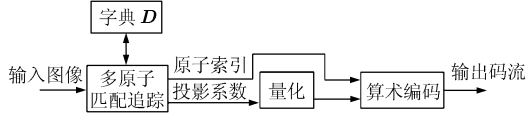


图 1 图像编码算法流程

3.1 多尺度脊波字典的构造

脊波变换^[14]最早由 Candes 于 1996 年提出, 它采用一系列脊波函数叠加的形式来表示函数。脊波变换由于对线状奇异特征的稀疏表示, 而被广泛应用于图像处理领域。然而, 由于脊波函数尺度的“单一性”, 导致它对曲线状奇异特征的表示的不稀疏性。脊波函数尺度的“单一性”与脊波自身的定义有关。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, 尺度参数 a 的变化, 仅仅改变的是垂直于脊线 $x \cos \theta + y \sin \theta = \text{const}$ 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, 不利于曲线特性的表示, 并引入“裂痕”噪声。本文通过构造多尺度母脊波原子和父脊波原子来实现图像的稀疏表示。

与二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 $\int (|\hat{\varphi}(\xi)|^2 / |\xi|^2) d\xi < \infty$ 。则多尺度母脊波原子和多尺度父脊波原子分别定义为

$$g_{a,b,\theta}^M(x,y) = \psi_{a_1,b_1,\theta}(x,y) \varphi_{a_2,b_2,\theta}(x,y) \quad (4)$$

$$g_{a,b,\theta}^F = \varphi_{a_1,b_1,\theta}(x,y) \varphi_{a_2,b_2,\theta+\pi/2}(x,y) \quad (5)$$

多尺度母脊波原子用来表示图像高频分量, 而多尺度父脊波原子用来表示图像低频分量。通常图像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

$$g_{a,b}^F = a_1^{-1} \varphi((-x - b_1)/a_1) \varphi((y - b_2)/a_1) \quad (6)$$

为了更有效地对图像中的曲线等结构特征进行稀疏表示, 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)垂直于边缘的方向上函数能近似边缘的跃变。基于以上两点考虑, 容许神经激活函数 $\psi(t)$ 和 $\varphi(t)$ 分别选取为 Marr 小波和 Gauss 小波。

最后, 通过对多尺度母脊波原子和父脊波原子

中的尺度参数 a , 平移参数 b 和方向参数 θ 进行“稠密”的过采样, 可以生成用于稀疏分解的多尺度脊波字典 D 。

$$D = \{g_{a,b,\theta}^M(x,y)\} \cup \{g_{a,b,\theta}^F(x,y)\} \quad (7)$$

3.2 系数的自适应量化与编码

投影系数是最佳匹配原子与图像间的内积, 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。量化时, 最初迭代投影系数的量化误差对引起的失真较大。基于此, 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^[15]。因此, 本文通过设计 $(m_i - \alpha\sigma_i, m_i + \alpha\sigma_i)$ 范围内的量化器来实现对投影系数的量化, 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

$$D(q_i) \triangleq \sum_{i=1}^k \alpha^2 \sigma_i^2 / (3q_i^2) \quad (8)$$

假定字典大小为 M , 则对 k 次迭代的原子索引值和量化投影值编码后的码率可描述为

$$R = \sum_{i=1}^k \log_2 q_i + k \log_2 M \quad (9)$$

从式(8)和式(9)可以看出, 量化等级 q_i 越大, 失真越小; 此时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 乘法得到。

$$q_i = \arg \min_{q_i} \left\{ \sum_{i=1}^k \alpha^2 \sigma_i^2 / (3q_i^2) + \lambda \left(\sum_{i=1}^k \log_2 q_i + k \log_2 M \right) \right\} \quad (10)$$

对式(10)求导可得, $q_i = \sqrt{2\alpha^2 \sigma_i^2 \ln 2 / (3\lambda)}$ 。由此可以得到

$$q_{i+1} / q_i = \sigma_{i+1} / \sigma_i \quad (11)$$

从式(11)可知, 一旦 q_1 确定, 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

为了验证算法的性能, 以 $256 \times 256 \times 8$ bit 的 Lena 和 Barbara 图像为例进行以下实验。所有实验均在主频为 2.0 GHz, 内存为 512 M 的 PC 机 Matlab 环境下进行。采用 K 均值聚类法将多尺度脊波字典分解成 64 个非相干子字典。

4.1 多原子匹配追踪算法性能的测试与分析

本实验以 Lena 为例, 分别对 MAMP 与 MP 算法稀疏逼近的速度, 参数 γ 和 μ_{stop} 对 MAMP 稀疏

逼近性能的影响进行比较与分析。图2为给定稀疏逼近性能(峰值信噪比, PSNR)条件下, MAMP和MP两种算法时间花费的比较, 其中 $\gamma = 0.7$, $\mu_{stop} = 0.01$ 。由图可知, 随着PSNR的增加, MP花费时间成指数级增长; 当PSNR=30 dB时, MP耗时约为MAMP的35倍。

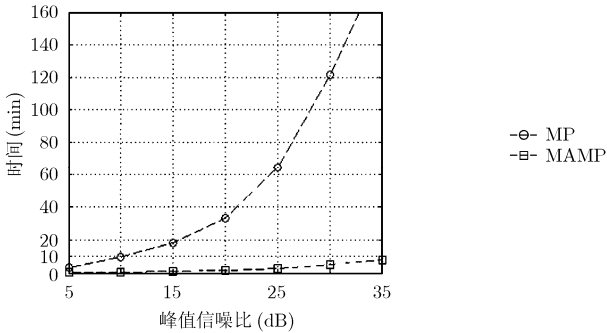


图2 MAMP算法与MP算法稀疏分解耗时比较

图3为 $\mu_{stop} = 0.01$ 时, MAMP算法中参数 γ 对稀疏逼近性能的影响, 其中实线为匹配追踪算法。随着 γ 的增大, MAMP的逼近性能接近于MP。特殊地, 当 $\gamma = 1$ 时, MAMP算法每次迭代仅选取一个最匹配原子, 此时MAMP脱化成MP。图4为 $\gamma = 0.7$ 时, 参数 μ_{stop} 对稀疏逼近性能的影响, 其中实线为匹配追踪算法。与图3相似, 随着 μ_{stop} 的逐渐减小, MAMP与MP性能越接近; 当 μ_{stop} 足够小时, MAMP等同于MP。由图3和图4可知, MAMP算法在提高速度的同时, 并没有牺牲算法的太多性能。

4.2 编码性能测试与分析

为了测试基于多原子匹配追踪图像编码算法的性能, 进行如下实验。实验分别从客观保真度(峰值信噪比, PSNR), 与JPEG2000和文献[8]进行比较。图5为3种编码算法率失真性能的比较, 其中图5(a)

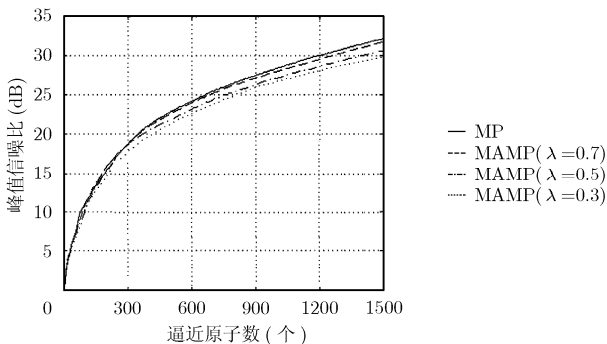


图3 不同 γ 时, MAMP算法的性能随逼近原子数的变化

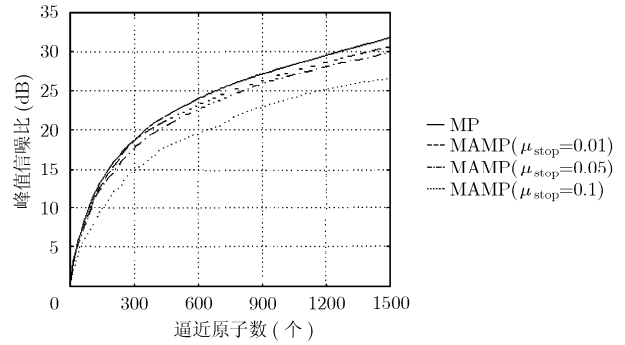


图4 不同 μ_{stop} 时, MAMP算法的性能随逼近原子数的变化

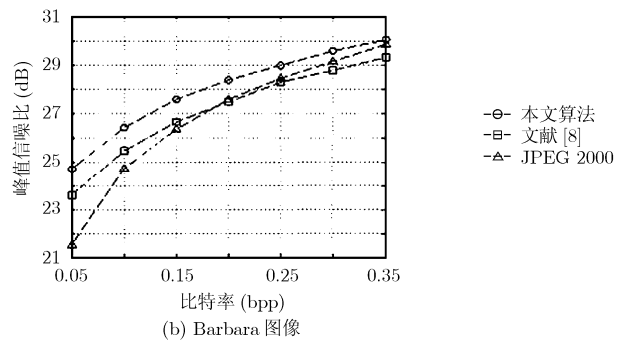
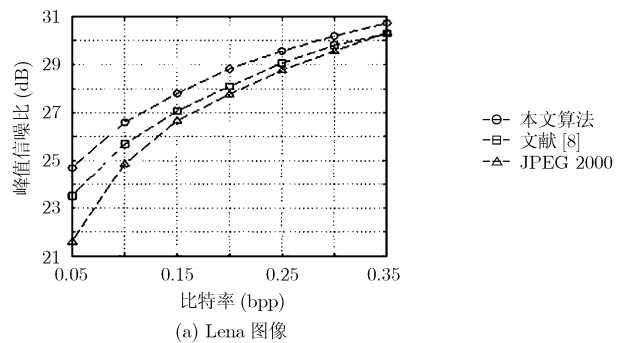


图5 率失真性能比较

和图5(b)分别对应Lena和Barbara图像。从图可以看出, 无论是对结构特征丰富的Lena图像还是纹理特征丰富的Barbara图像, 在任意低比特率下, 本文算法都比文献[8]和JPEG2000获得更好的编码性能。通常, 人眼最先捕获到的是图像目标的轮廓, 即边缘等结构特征。因此, 对于低比特率或者甚低比特率编码, 最先编码的信息应该是边缘等特征。而本文构造的多尺度母脊波原子能很好地捕获图像中的边缘等特征, 并能用很少的数据对这些特征进行稀疏表示。由此可知, 本文算法非常适合低比特率或者甚低比特率的图像或视频编码。

5 结论

本文提出一种新的信号稀疏分解方法—多原子匹配追踪算法。与传统的匹配追踪算法不同, 本算法通过每次迭代选取若干个最匹配原子来构造信号

的稀疏逼近。实验结果表明, 多原子匹配追踪算法在保证稀疏逼近性能的同时, 大大降低了运算复杂度。在此基础上, 构造了多尺度脊波字典; 它能够稀疏地对图像进行表示。通过对稀疏分解数据的自适应量化和编码, 实现基于多原子匹配追踪的静态图像编码。对测试图像的实验结果表明, 在低比特率下本文编码算法比 JPEG2000 获得更好的客观保真度和主观视觉质量。如何实现高比特率图像或视频编码, 是我们下一步研究的主要内容。

参 考 文 献

- [1] Taubman D and Marcellin M. JPEG-2000: Image Compression Fundamentals, Standards and Practice [M]. Massachusetts: Kluwer Academic Publishers, 2001: 212-379.
- [2] Mallat S and Zhang Z. Matching pursuit with time-frequency dictionaries [J]. *IEEE Transactions on Signal Processing*, 1993, 41(12): 3397-3415.
- [3] Gribonval R and Vandergheynst P. On the exponential convergence of matching pursuits in quasi-incoherent dictionaries [J]. *IEEE Transactions on Information Theory*, 2006, 52(1): 255-261.
- [4] Peotta L and Vandergheynst P. Matching pursuit with block incoherent dictionaries [J]. *IEEE Transactions on Signal Processing*, 2007, 55(9): 4549-4557.
- [5] Aharon M and Elad M. Sparse and redundant modeling of image content using an image-signature-dictionary [J]. *SIAM Journal on Imaging Sciences*, 2008, 1(3): 228-247.
- [6] Mairal J, Sapiro G, and Elad M. Learning multiscale sparse representations for image and video restoration [J]. *SIAM Multiscale Modeling and Simulation*, 2008, 7(1): 214-241.
- [7] Mairal J, Elad M, and Sapiro G. Sparse representation for color image restoration [J]. *IEEE Transactions on Image Processing*, 2008, 17(1): 53-69.
- [8] Ventura R M Figueras I, Vandergheynst P, and Frossard P. Low-rate and flexible image coding with redundant representation [J]. *IEEE Transactions on Image Processing*, 2006, 15(3): 726-739.
- [9] Gribonval R. Fast matching pursuit with a multiscale dictionary of gaussian chirps [J]. *IEEE Transactions on Signal Processing*, 2001, 49(5): 994-1001.
- [10] De Vleeschouwer C and Macq B. Subband dictionaries for low-cost matching pursuit of video residues [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1999, 9(7): 984-993.
- [11] Schmid-Saugeon P and Zakhor A. Dictionary design for matching pursuit and application to motion-compensated video coding [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, 14(6): 880-886.
- [12] Chou Y T, Hwang W L, and Huang C L. Gain-shape optimized dictionary for matching pursuit video coding [J]. *Signal Processing*, 2003, 83(9): 1937-1943.
- [13] Jost P, Vandergheynst P, and Frossard P. Tree-based pursuit: algorithm and properties [J]. *IEEE Transactions on Signal Processing*, 2006, 54(12): 4685-4697.
- [14] Candes E J. Harmonic analysis of neural networks [J]. *Applied and Computational Harmonic Analysis*, 1999, 6(2): 197-218.
- [15] Shoa A and Shirani S. Adaptive quantization for matching pursuit [C]. *IEEE International Workshop on Multimedia Signal Processing*, Victoria, Canada, Oct. 3-6, 2006: 71-74.

邓承志: 男, 1980年生, 博士, 研究方向为信号与图像的稀疏表示理论及其应用等。

汪胜前: 男, 1965年生, 博士, 教授, 硕士生导师, 研究方向为图像处理、小波分析理论等。

曹汉强: 男, 1953年生, 博士, 教授, 博士生导师, 研究方向为图像处理、信息安全等。