

# 基于互相关估计的快速图像逼近算法

甘 涛<sup>1</sup>, 何艳敏<sup>2</sup>, 朱维乐<sup>1</sup>

(1. 电子科技大学电子工程学院, 四川成都 610054; 2. 电子科技大学自动化工程学院, 四川成都 610054)

**摘 要:** 目前图像稀疏分解的应用还受到过大运算量的阻碍. 基于对库原子间互相关信息的估计, 提出一种改进的匹配搜索算法. 通过自适应预测, 有效地降低了单次迭代的内积运算. 并行一次选择多个原子, 显著地减少了迭代次数, 从而大幅度降低了总运算复杂度. 实验结果表明, 与原算法相比, 改进算法在微小精度损失的情况下, 表现出明显的速度优势. 如进行 800 个原子的分解, 速度提高近 43 倍. 将该算法应用于图像编码中, 在低码率下获得了与 JPEG2000 相当的编码性能.

**关键词:** 稀疏分解; 匹配搜索; 互相关

中图分类号: TN919.81 文献标识码: A 文章编号: 0372-2112(2008)05-1019-05

## Fast Algorithm for Image Approximation Based on Cross-Correlation Estimation

GAN Tao<sup>1</sup>, HE Yamin<sup>2</sup>, ZHU Weile<sup>1</sup>

(1. School of Electronic Engineering, University of Electronic Science and Technology, Chengdu, Sichuan 610054, China;

2. School of Automation Engineering, University of Electronic Science and Technology, Chengdu, Sichuan 610054, China)

**Abstract:** The main obstacle to the application of image sparse representation nowadays is the enormous computational complexity. Based on the estimation of the dictionary cross correlation information, an improved matching pursuit algorithm is proposed for image approximation. At each iteration, the adaptive prediction is introduced to effectively reduce the inner product computation load. Meanwhile, the number of iterations is decreased significantly by picking a group of atoms at a time. As a result, the total computational complexity is greatly reduced. Experimental results show that the proposed algorithm yields a significant speed improvement over the original one, while maintaining the approximation quality. For instance, it achieves a speed up gain of near 43 times when performing 800 atoms decomposition. The algorithm is applied to image coding and shown to provide results comparable to JPEG2000 at low bitrate.

**Key words:** sparse representation; matching pursuit; cross correlation

## 1 引言

近年来, 稀疏逼近 (sparse approximation) 的新信号表示方法成为人们关注和研究的热点<sup>[1]</sup>. 其基本思想是用冗余函数系统 (或原子库) 取代传统的正交基函数, 并从中找到具有最佳线性组合的  $M$  项原子来表示一个信号. 目前稀疏分解的主要障碍在于运算量大. 国内有研究人员指出, 当信号长度为 1024 采样点时, 信号的稀疏分解的难度将十分巨大<sup>[2]</sup>. 在众多的稀疏分解方法中, 匹配追踪 (Matching Pursuit, MP) 算法<sup>[3]</sup> 原理简单, 运算量相对较低, 已在多个领域得到了应用. 但对于大数据量图像信号, MP 分解的运算量仍然大得惊人. 这也是长期以来图像稀疏分解发展缓慢的主要原因之一. 在降低运算复杂度方面, 学者们已提出了众多改进方法. 快速傅立叶变换 (Fast Fourier Transform, FFT) 的引入<sup>[4]</sup>, 简化了

匹配运算, 极大地提高了分解速度.  $M$  项追踪 (M-Term Pursuit, MTP) 算法<sup>[5]</sup> 一次并行选择多个原子, 成倍地加快了搜索过程. 另外, 对原子库的结构加以约束, 如选取二维可分离 Gabor 函数<sup>[6]</sup> 作为原子库, 也成为减小运算量的有效途径. 但这样往往会造成逼近精度的下降, 削弱了稀疏分解的优势. 作为图像稀疏分解新的突破, Vandergheynst 等提出了基于几何原子库的高伸缩性的图像编码方案<sup>[7,8]</sup>, 获得了与 JPEG2000 国际标准相当的性能, 表现出非常好的发展前景. 目前该方案由于采用二维不可分离原子库, 计算复杂度较高, 离实际应用还有距离. 本文就针对这个问题展开工作, 对其分解算法进行改进.

## 2 全局匹配搜索算法 (FSMP)

为有效地表征图像平滑区域和二维奇异性 (如边

缘轮廓等)特征,文献[8]算法采用二维高斯函数和其二阶导数作为母函数,通过平移、拉伸、旋转变换形成归一化的原子库,并运用基于 FFT 的全局匹配追踪(Full Search Matching Pursuit, FSMP)算法对图像进行稀疏分解.

FSMP 采用迭代的方式从原子库  $D$  中逐一选出与图像匹配的原子. 初始时, 余量信号  $R_0$  为输入图像. 在第  $n$  次迭代, 从  $D$  中选出最匹配的原子  $g_s^n$  (上标  $n$  代表第  $n$  次的最佳) 使得:

$$|\langle R_n, g_s^n \rangle| = \sup_{g_s \in D} |\langle R_n, g_s \rangle| \quad (1)$$

其中  $R_n$  为当前残余图像. 然后对  $R_n$  作如下分解, 并进入下一次迭代.

$$R_{n+1} = R_n - \langle R_n, g_s^n \rangle g_s^n \quad (2)$$

为降低运算量, 在计算内积时引入 FFT<sup>[4]</sup>. 对中心原子  $g_{y,0}$  (即不作平移或平移为 0 的原子) 进行一次 FFT, 就可计算出其所有位移原子  $g_{y,p}$  ( $p = (px, py)$  代表平移位置) 的内积. 这样参与运算的原子集合由原来的  $D$  变为较小的中心原子集合  $D_0$ , 运算量得到极大地减少.

### 3 快速 M 项匹配搜索算法 (FMTP)

从算法流程可知, FSMP 的复杂度主要来自于原子匹配和原子选择两个过程, 即每一次迭代, 都需要进行大量的匹配(内积)运算以决定当前最优的原子; 而且一次迭代只有一个原子被选中. 对于较大的图像, 即使是通过 FFT, 其运算量还是相当可观的. 为此, 我们对 FSMP 进行了改进, 提出快速 M 项匹配搜索算法 (Fast M-Term matching Pursuit, FMTP). 为保证逼近精度, 我们采用与文献[8]相同的库原子结构, 在分析 MP 的特性和挖掘库相关信息特点的基础上, 从原子匹配和原子选择两方面显著地降低了运算量, 取得了逼近精度和运算复杂度之间的较好折中.

#### 3.1 MP 的特性

对图像进行 MP 分解可以看作是逐步提取图像中与库原子相匹配的二维结构的过程. 因为库原子是能量归一化的, 在第  $n$  次迭代, 残余图像  $R_{n+1}$  将与选出的匹配原子  $g_s^n$  正交, 如式(2)所示. 在  $n+1$  次迭代, 设我们将库原子  $g_y$  沿  $g_s^n$  作如下正交分解:

$$g_y = \langle g_y, g_s^n \rangle g_s^n + K_y \quad (3)$$

则其当前投影系数  $\langle R_{n+1}, g_y \rangle$  为  $\langle R_{n+1}, K_y \rangle$ . 由于我们选择的原子库是过冗余的且库中绝大多数原子是空域高度局域化的, 若相关系数  $\langle g_y, g_s^n \rangle$  越大, 即  $g_y$  与前次选中的原子  $g_s^n$  越相关,  $K_y$  的能量就越小, 该原子被选中的机率将越小. 可见, 在稍后的相邻迭代中, 被 MP 提取出来的用以逼近残余图像的原子将是几乎无关的. 为了验证该想法, 我们选取一系列标准图像, 对 MP 算法

做了大量统计实验, 实验结果是令人鼓舞的. 比如对 Lena 图像, 发现除了前几次迭代, 相邻迭代中选中原子的相关系数接近于 0, 投影系数幅值的相对变化小于 10%. 对不同测试图像, 其结果基本一致. 为此我们可得出如下结论: MP 在相邻迭代中, 以很高的概率选择互不相关的原子, 且相应投影系数幅值的变化较小. 该结论也在文献[4, 5]中得到证实.

#### 3.2 库互相关信息的估计

原子间的相关信息对匹配运算的简化很重要. 正如文献[3]提到, 如果原子间互相关提前计算或用解析式快速求出, 匹配搜索将变得非常简单. 然而对于当前过冗余的原子库, 这两种方法都不现实. 由原子的产生方式可知, 原子能量大部分都集中在中心附近, 远离中心呈指数衰减, 且众多原子都是通过中心原子位移形成的, 对于中心距离稍大的原子, 其互相关很小, 甚至为零, 故我们可用原子中心距离来估计其是否相关. 具体做法是, 对中心原子集合的每一对原子, 我们先确定其相干的方形区域, 而后记录该区域的顶点坐标. 在分解过程中, 对于给定的两个原子, 直接取出相应的相干区域, 并与其实际中心距离做比较, 即可判断它们的相关性. 我们下面提出的两个加速方法正是借助于库原子相关信息的估计, 在保证精度的前提下, 大大简化了匹配运算.

#### 3.3 快速原子匹配

原子的匹配运算是 MP 最耗时的部分. 在本小节, 我们集中讨论其加速方法. 从式(3)知, 如果某原子与  $n-1$  次选中的原子不相关, 其对应的投影系数将保持不变, 即若  $\langle g_y, g_s^{n-1} \rangle = 0$ , 则  $\langle R_n, g_y \rangle = \langle R_{n-1}, g_y \rangle$ . 结合前述的 MP 的特性, 我们提出如下的快速匹配方案.

在  $n$  次迭代, 我们并不像 FSMP 那样对中心集中的每一个原子  $g_{y,0}$  运用 FFT 搜索其最优的位移形式  $g_{y,p}^n$ . 在运算之前, 先判断原子的前次最优位移形式  $g_{y,p}^{n-1}$  与前次选中原子  $g_s^{n-1}$  的相关性, 如果它们的相关程度足够小, 即

$$|\langle g_{y,p}^{n-1}, g_s^{n-1} \rangle| < \delta, \delta \in [0, 1] \quad (4)$$

我们将跳过对当前原子的匹配运算, 预测其最优位移  $g_{y,p}^n$  为前次迭代的值  $g_{y,p}^{n-1}$ .

#### 3.4 M 项并行原子选择

现在我们来讨论原子的选择过程及其加速方法. 前面已经分析过, MP 在相邻迭代中趋向于选择互不相关的原子. 直观地, 如果每次选择两个或者更多的不相关原子, 将不会对逼近质量造成较大的影响. MTP 算法正是基于这个想法而提出的, 它在一次迭代中并行选择 M 个原子来加快运算. 但是由于缺乏对原子间相关信息的有效获取机制, MTP 通过原子库划分, 候选原子集合相

干性检测和正交投影等步骤来确保选出一组相干性小的原子. 由于这些代价, 其速度提升没有达到  $M$  倍的理想效果.

有了 3.2 节的库互相关信息估计, 我们可避免这些步骤而采取非常简单的方法达到相同的目的. 即在 FSMP 的构架下, 一次迭代直接选取一组满足相关性要求的原子, 精度同样可以得到保证. 为此, 我们提出了如下  $M$  项并行搜索方案.

(1) 在第  $n$  次迭代, 当完成匹配运算后, 一次选择  $m_n$  个原子  $g_{sk}^n (k \in \Lambda_n)$ , 它们满足以下两条件:

$$|\langle R_n, g_{sk}^n \rangle| \geq \gamma \sup_{g \in \mathcal{B}} |\langle R_n, g \rangle|, \forall k \in \Lambda_n, \forall \gamma \in [0, 1] \quad (5)$$

$$|\langle g_{si}^n, g_{sj}^n \rangle| < \delta, \forall i, j \in \Lambda_n, i \neq j, \delta \in [0, 1] \quad (6)$$

其中  $\Lambda_n = \{1, \dots, m_n\}$  是选中原子的指标集,  $\gamma$  与  $\delta$  为的门限参数.

(2) 对余量信号的更新改为如下方式:

$$R_{n+1} = R_n - \sum_{k=1}^{m_n} \langle R_n, g_{sk}^n \rangle g_{sk}^n \quad (7)$$

## 4 算法分析

### 4.1 复杂度分析

这里我们将 FSMP, MTP 和本文提出的 FMTP 算法的复杂度作分析比较. 设图像数据大小为  $N$ , 中心原子库的大小为  $P$ , 在第  $n$  次迭代, 并行选中的原子数目为  $m_n$ , 则 FSMP 的复杂度为  $o(P \cdot m_n \cdot N \log_2 N)^{[4]}$ ; MTP 算法的复杂度为  $o(PN \log_2 N + m_n q)^{[5]}$ , 其中第二项对应于集合划分, 正交投影等附加运算,  $q$  为某常量.

FMTP 的两个加速方案都是建立在对原子库相关信息的估计基础上的. 运用 3.2 节的方法, 式(4)和式(6)的判断可很方便地得出, 其代价可以忽略. 这样算法总体复杂度的降低将直接受益于快速原子匹配和并行原子选择两方面的贡献. 一方面, 在快速原子匹配中, 由于库原子具有高度的空间局域性, 式(4)的预测条件在很多时候是成立的, 这样算法将在每次迭代自适应地选择参与运算的原子, 相对于全局搜索, 大量的 FFT 运算可被跳过, 算法速度将会有大幅度的提高. 设参与运算的中心原子数目为  $P_1 (P_1 < P)$ , 速度增益为  $P/P_1$  倍. 另一方面, 在并行原子选择中,  $m_n$  个原子将被一次选中, 并且没有显著的附加操作, 速度增益为  $m_n$  倍. 一般地, 后者的速度增益将大于前者. 然而 FMTP 总体的速度提升并没有达到两者叠加的理想效果. 原因是, 由于应考虑多个原子的相关性, 式(4)的预测条件得以加强, 使得满足条件的原子减少, 快速原子匹配的加速效果受到抑制. 设实际参与匹配运算的中心原子数目为  $P_2 (P_1 < P_2 < P)$ , FMTP 算法的复杂度为  $o(P_2 N \log_2 N)$ , 其中  $P_2$  与库原子的局域化程度、图像大小等因素有关. 图像越大,

库原子的局域化程度越高,  $P_2$  就越小. 总之, 相对于 FSMP, FMTP 的速度增益为  $P \cdot m_n / P_2$ . 当  $P$ ,  $m_n$  和  $N$  较大时, FMTP 的复杂度将明显低于 FSMP, 这在下面的实验中得到了证实.

### 4.2 逼近质量分析

FMTP 采用局部和并行搜索来降低运算复杂度, 相对于全局和串行搜索的 FSMP, 逼近精度一般会有所下降. 局部匹配搜索可能引起所选原子的变化, 但它没有改变 MP 的本质, 3.1 节的分析和结论仍然成立, 故不会影响并行搜索的效果. 而并行原子选择对局部匹配的影响是双方面的. 一方面, 多重原子的选择会造成预测误差的累积, 降低精度. 另一方面, 如前讨论, 由于满足预测条件的原子减少, 局部搜索范围扩大, 质量损失得到了控制.

总体上, 我们通过调节式(4)~(6)中的参数  $\delta$  和  $\gamma$ , 来对算法的质量与复杂度进行权衡.  $\delta$  变大,  $\gamma$  变小时, 参与匹配运算的原子减少, 并行选择的原子数目变多, 速度得到提升, 但精度下降. 反之, 随着  $\delta$  的变小,  $\gamma$  的变大, FMTP 算法逼近质量提高, 速度下降, 性能逐渐接近 FSMP. 当  $\delta$  降至 0,  $\gamma$  增至 1 时, FMTP 就退化为原始的 FSMP 了.

### 4.3 适用性分析

FMTP 利用库原子的空域局域化特点来加快 MP 过程, 原子的局域化程度越高, 加速效果越好. 算法对原子库的结构并没有限制, 如本文所用的原子库为表征图像边缘和平滑区域的两种原子组成, 它还可以级联用以逼近纹理的 Gabor 原子<sup>[4]</sup>; 和 MTP 一样, FMTP 也适用于表征视频的时空 (spatio-temporal) 原子库<sup>[5]</sup>.

## 5 实验结果

为验证 FMTP 算法的有效性, 我们对其进行了仿真. 测试图像为标准  $512 \times 512$  灰度图像: Lena, Barbara, Goldhill, Peppers 及  $256 \times 256$  灰度图像: Lena 和 Cameraman.

### 5.1 FMTP 算法的图像逼近性能

我们将 FMTP 与 FSMP<sup>[4]</sup> 的实验结果进行了比较. FMTP 的参数由实验选定为:  $\gamma = 0.6$ ,  $\delta = 0.01$ . 图 1 显示了对 Peppers ( $512 \times 512$ ) 图像逼近过程的比较. 从参与匹配运算的原子数目  $p$  来看, FSMP 恒为库中心原子数目  $P$ ; 而 FMTP 只有部分原子参与 FFT 运算, 平均数目  $\bar{P}_2$  占  $P$  的 60%. 从单次迭代选择的原子数目  $m$  来看, FSMP 每次只选择一个原子, 而 FMTP 适应了 MP 系数开始下降陡, 后来趋于平缓的变化特点. 在初始迭代阶段仅选择少量的原子, 以保证精度. 随着迭代次数的增加, 投影系数幅度减小, 越来越多原子被选中, 在精度损失较小的情况大大加快迭代过程.

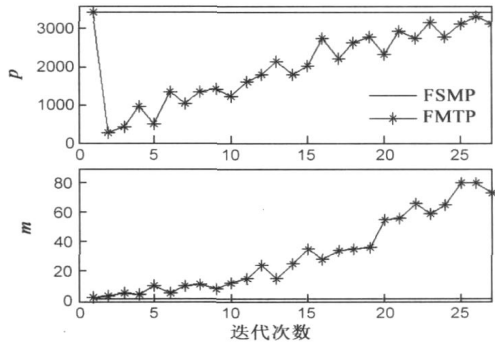


图 1 对 Peppers(512×512)图像,参与 FFT 运算的原子数  $p$  和选中的原子数  $m$  随迭代变化的比较

表 1 列出了两算法的运行时间和逼近质量的测试结果. 正如理论分析一样,与 FSMP 相比,FMTP 在微小的精度损失的情况下,具有明显的速度优势,且该优势随着逼近原子数目的增多而迅速增强. 如用 800 个原子的图像逼近,平均 FMTP 的速度是 FSMP 的 43.98 倍,而 PSNR 平均下降 0.21dB. 图 2 给出了两算法对 Peppers (512×512) 作 800 个原子的分解合成的恢复图像. 可以

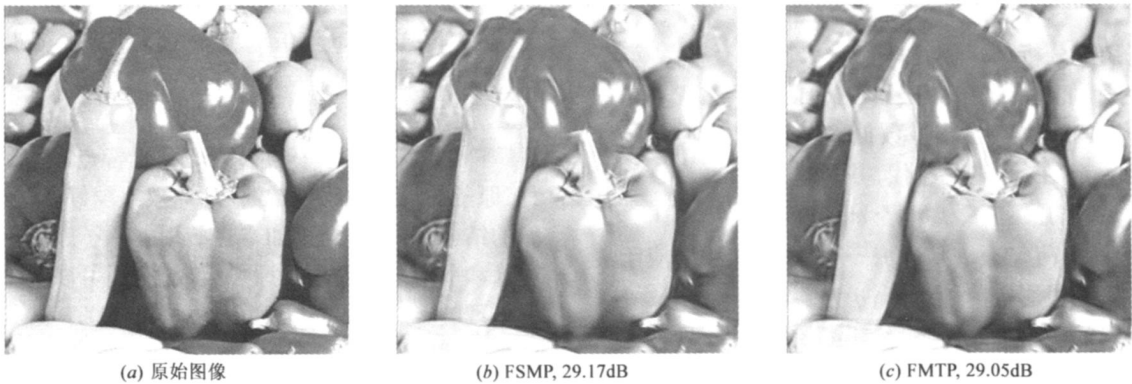


图 2 Peppers (512×512) 800 个原子的合成恢复图像

## 5.2 FMTP 算法的图像编码性能

本实验我们将 FMTP 算法运用到图像压缩编码中来进一步验证其有效性. 我们仿真重现了文献[8]的工作,并将 FMTP 应用于相同的编码框架中. 具体地,用 FMTP 替代 FSMP 进行图像分解,而后对分解系数进行指数型区间的均匀量化 (exponentially bounded uniform

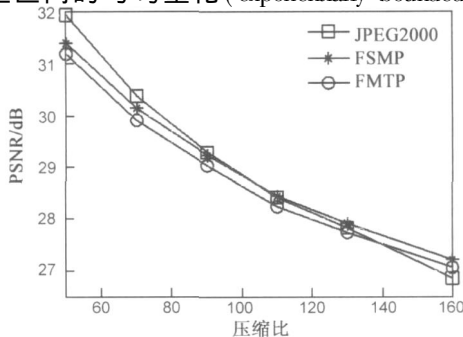


图 3 JPEG2000,基于 FSMP 和 FMTP 的编码器对 Lena (512×512) 图像的率-失真性能比较

看到,一方面,两算法都能很快捕捉到图像的重要结构信息,获得稀疏的图像表征;另一方面,相对于 FSMP,FMTP 算法在保持主观质量的前提下,大大提高了图像分解的速度.

表 1 对测试图像 FSMP 和 FMTP 的运行时间和逼近质量的结果比较

测试图像	算法	100 个原子		200 个原子		400 个原子		800 个原子	
		时间 单元	PSNR (dB)	时间 单元	PSNR (dB)	时间 单元	PSNR (dB)	时间 单元	PSNR (dB)
Lana 512×512	FSMP	100	23.17	200	25.29	400	27.58	800	30.14
	FMTP	6.12	23.13	8.36	25.13	12.55	27.42	17.21	29.91
Barbara 512×512	FSMP	100	20.44	200	21.78	400	23.12	800	24.84
	FMTP	5.10	20.37	7.69	21.74	10.97	23.02	16.36	24.61
Goldhill 512×512	FSMP	100	23.10	200	24.50	400	25.95	800	27.58
	FMTP	6.09	22.84	7.95	24.25	10.81	25.77	15.38	27.40
Peppers 512×512	FSMP	100	21.27	200	23.74	400	26.36	800	29.17
	FMTP	5.52	21.15	8.28	23.57	12.25	26.24	17.76	29.05
Lena 256×256	FSMP	100	23.54	200	25.74	400	28.40	800	31.53
	FMTP	6.99	23.35	10.58	25.60	14.72	28.16	21.29	31.25
Cameraman 256×256	FSMP	100	22.06	200	24.16	400	26.45	800	29.04
	FMTP	7.58	21.89	11.76	24.00	16.42	26.36	23.61	28.85

quantization), 最后对原子参数和量化系数进行自适应算术编码. 我们将实验结果与文献[8]算法, JPEG2000<sup>[9]</sup> 进行比较.

图 3 给出了在不同压缩比情况下三种编码器的率-失真性能. 可见,基于本文 FMTP 的编码器与基于 FSMP 的编码器客观 PSNR 差别不大,且在压缩较大时,二者都超过了 JPEG2000. 图 4 显示了压缩比为 64 倍时,各编码器对 Lena (512×512) 的编码重建图像. 可以看到,虽然 JPEG2000 的客观 PSNR 较高,但由于出现小波分解的振铃 (ringing) 效应,其主观效果不如本文编码器,特别是在图像的轮廓边缘等地方,这体现了基于冗余几何原子库的分解能更好地逼近图像二维特征,获得较小波分解更稀疏的表征.

## 6 结论

信号和图像的稀疏分解近年来在国内正在兴起,已表现出极大的发展潜力,对其快速算法的研究意义

重大. 本文就此方面做出工作, 提出  $M$  项并行搜索算法. 在保持逼近精度的情况下, 同时有效地降低了原子匹配和原子选择的复杂度, 取得了明显的速度提升效果. 将该算法运用于图像编码中, 在低码率下获得了与 JPEG2000 相当的编码性能. 另外, 由于采用基于几何原

子库的分解方法, 本文编码器保持 MP 编码器的独特优势<sup>[8]</sup>, 如编码流具有灵活的伸缩性, 可支持任意分辨率解码且解码速度快等, 可在网络或移动终端的图像业务中得到很好的应用.



图 4 压缩比为 64 时 Lena (512×512) 编码重建图像

#### 参考文献:

- [1] 王建英, 尹忠科, 等. 信号与图像的稀疏分解及初步应用 [M]. 成都: 西南交通大学出版社, 2006.
- [2] 邹红星, 周小波, 等. 时频分析回溯与前瞻 [J]. 电子学报, 2000, 28(9): 78–84.  
Zou Hongxing, Zhou Xiaobo, et al. Which time frequency analysis a survey [J]. Acta Electronica Sinica, 2000, 28(9): 78–84. (in Chinese)
- [3] Mallat S, Zhang Z. Matching pursuits with time frequency dictionaries [J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3397–3415.
- [4] Figueras i Ventura R M. Sparse image approximation with application to flexible image coding [D]. Switzerland: EPFL, 2005.
- [5] Rahmoune A. Image and video representation and scalable cod-

- ing using redundant dictionaries [D]. Switzerland: EPFL, 2005.
- [6] 廖斌, 许刚, 等. 二维匹配跟踪自适应图像编码 [J]. 计算机辅助设计与图形学学报, 2003, 15(9): 1084–1090.  
Liao Bin, Xu Gang, et al. Adaptive image coding based on matching pursuit [J]. Journal of Computer aided Design & Computer Graphics, 2003, 15(9): 1084–1090. (in Chinese)
- [7] Vandergheynst P, Frossard P. Efficient image representation by anisotropic refinement in matching pursuit [A]. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Salt Lake City: IEEE, 2001. 1757–1760.
- [8] Figueras i Ventura R M, Vandergheynst P, et al. Low rate and flexible image coding with redundant representations [J]. IEEE Transaction on Image Process, 2006, 15(3): 726–739.
- [9] Taubman D, Marcellin M. JPEG2000: Image Compression Fundamentals, Standards and Practice [M]. Boston: Kluwer Academic Publishers, 2001.

#### 作者简介:



甘涛男, 1975 年生于四川成都, 电子科技大学电子工程学院讲师, 博士研究生. 主要研究领域为图像压缩、数字音视频技术和多媒体通信. E-mail: garretc@infomedia.net.cn



何艳敏女, 1977 年生于四川南充, 电子科技大学自动化学院讲师, 博士研究生. 主要研究领域为图像处理、模式识别.