

基于稀疏分解的可伸缩图像编码

甘 涛¹,何艳敏²

(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: 图像编码技术的新的突破可寄希望于信号表示的深刻变革. 采用基于冗余原子库的快速匹配追踪算法对图像进行稀疏分解, 在分析和总结原子空间位置分布规律的基础上, 提出与之相适应的块划分编码方法, 节约了用于表示原子参数和投影系数的比特数. 实验结果表明, 本文编码器在计算复杂度、编码效率和伸缩性能等方面都优于当前同类型编码器, 特别是在前两方面, 其优势十分明显. 比如对 512×512 测试图像, 编码率为 0.5bpp 时本文编码器的平均 PSNR 增益达 1.73dB. 特别地, 凭借原子库的几何特性, 该编码器提供了较传统方法更灵活的伸缩性, 允许通过简单的参数变换来获得任意分辨率大小的重建图像.

关键词: 冗余原子库; 稀疏分解; 伸缩性

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 0372-2112 (2010) 01-0156-05

Scalable Image Coding Based on Sparse Decomposition

GAN Tao¹, HE Yan-min²

(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: New breakthroughs in image coding may rely on deep changes in the signal representation. The fast matching pursuit (MP) algorithm is first employed to get the sparse image decomposition over a redundant dictionary. The distribution of the spatial position of selected atoms is studied and a novel block partitioning coding method is then proposed. The bit savings in representations of atom parameters and projection coefficients are obtained. Experimental results show that the new coding scheme has striking advantages over the latest MP coder in computational complexity, coding efficiency as well as scalability. For instance, for 512×512 test images an average PSNR gain of 1.73dB is achieved at 0.5bpp. Notably, thanks to the geometrical structure of the dictionary, the new coder provides attractive adaptability features which allow the codestream to be easily and efficiently decoded at any spatial resolution.

Key words: redundant dictionary; sparse decomposition; scalability

1 引言

目前, 在图像压缩编码中, 基于 DCT、小波等正交线性变换的主流编码技术已发展到了成熟阶段. 近来的研究表明, 传统的二维小波变换并不能有效地捕获图像边缘、轮廓上的光滑性, 故不能达到精简或稀疏的图像表征. 可见, 编码技术的新的突破可寄希望于在信号表示上的深刻变革. 近年来, 稀疏逼近 (sparse approximation) 或高度非线性逼近成为信号表示的一个新的研究热点^[1]. 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. 在该类稀疏分解方法中, 匹配追踪 (Matching Pursuit, MP) 算法^[2]原理简单, 在信号处理的多个领域得到了应用. 特别地, 基于冗余原子库的 MP

分解方法, 在图像和视频编码中已表现出较传统方法更优的性能^[3,4].

从综合性能上看, 目前处于领先地位的 MP 图像编码方法为文献[3]提出的基于冗余表征的可伸缩低码率编码, 简称为 MP_FLC. 首先, 该方法运用基于几何原子库的全局匹配追踪 (Full Search Matching Pursuit, FSMP) 对图像进行分解, 达到了较高的逼近精度; 采用指数型衰减区间均匀量化 (Exponentially bounded Uniform Quantization, EUQ) 方法^[5]对 MP 系数进行量化, 在低码率下获得与 JPEG2000 国际标准相当的客观率失真性能和更好的主观质量. 其次, MP_FLC 还具有灵活的质量和分辨率可伸缩性. 这些特点使得 MP_FLC 成为很有发展潜力的图像编码方法. 然而, 目前该方法还存在如下问题: (1) 分解参数的编码效率不高, 在中码率段, 如 0.25bpp 到

0.5bpp,其率失真性能远低于传统编码方法;(2)FSMP复杂度高,使编码器的应用受到很大限制;(3)码流是非嵌入式的,虽然可支持质量渐进传输,但在截断点性能有所下降。

在先前的工作中,我们针对FSMP复杂度高的问题,提出了快速图像逼近算法,并将其运用到图像编码中,取得了较好的效果^[6]。之后我们提出快速M项追踪(Fast M-Term Pursuit, FMTP)算法,在保持逼近精度的情况下,进一步提高了稀疏分解的速度^[7]。本文将在此基础上提出一个基于块划分的快速MP图像编码(Fast Matching Pursuit based on Block Partitioning Coding, FMP_BPC)。与MP_FLC相比,FMP_BPC具有如下特点:(1)运用类似小波编码的集合划分方法,将原子的位置参数和MP系数进行有效的组织和编码,编码效率较MP_FLC有了显著的提高,在中码率段获得了与JPEG2000和SPIHT相当的率失真性能;(2)采用FMTP分解方法,在保证编码性能的前提下大大降低了编码复杂度;(3)生成码流为嵌入式的,可以在任意位置处截断。值得一提的是,FMP_BPC在达到以上性能的同时,保持了MP_FLC的优点,如高的图像主观质量,灵活的伸缩特性等,在更大程度上体现了稀疏分解较传统方法的优势。

2 基于块划分的快速MP编码

基于MP的图像编码是运用MP对图像进行分解,然后对分解原子的参数和相应投影系数两类信息进行编码。已有的研究表明,在总编码流中原子位置参数所占的比例最大^[4],其次是MP系数,故对这两个部分的有效编码,能带来整体性能的提高。立足于使MP码流具有高的质量可伸缩性,我们放弃了用差分^[8]或游长编码^[9]来提高编码效率的做法,而按原子投影系数幅度从大到小的顺序进行编码。在对MP原子空间位置分布规律的分析 and 总结基础上,提出与之相适应的块划分编码(Block Partitioning Coding, BPC)方法;将BPC与快速MP分解算法FMTP相结合,最终形成FMP_BPC编码器。相对于已有的MP编码器,FMP_BPC在计算复杂度和率失真性能两方面都有了显著的改善。

2.1 MP原子位置和能量分布

为了找到MP原子分布的特点,我们用FSMP和FMTP两种MP算法对测试图像进行分解实验。实验得出了一致的结论:总体上,原子在图像的边缘、轮廓等地方呈聚集分布,且能量较高,这与小波系数分布有着相似之处;同时,在局部范围内,原子分布并不像小波系数那么密集。

事实上,上述实验结果是可以得到合理解释的。由于我们的几何原子库主要包含用于捕获图像边缘的原子,故算法所选的匹配原子应出现在图像的相应位置

处,原子分布表现出聚集特性。另外,由于在每次迭代中,所选的原子都将从当前图像中被剔除,导致了在局部范围内原子的分布是稀疏的,且原子能量差别较大。

2.2 一般位平面的块划分

根据MP原子位置和能量的分布特点,我们提出与之相适应的编码算法。首先作一些定义。我们将原子按其位置放置于平面上,形成一张原子图。图中每个位置 (i, j) 对应着一个原子集合 $A_{i,j} = \{a_k\}_{k \in \Lambda}$, $A_{i,j}$ 包含位置参数为 (i, j) 的所有原子, Λ 为 $A_{i,j}$ 的指标集。定义原子集合 $A_{i,j}$ 的能量系数 $\epsilon_{i,j}$ 为

$$\epsilon_{i,j} = \begin{cases} \max_{a_k \in A_{i,j}} \{ |p_k| \}, & A_{i,j} \neq \phi \\ 0, & A_{i,j} = \phi \end{cases} \quad (1)$$

其中 p_k 为原子 a_k 对应的投影系数。由能量系数,我们定义集合 S 的重要性。如果

$$\max_{(i,j) \in S} \{ \epsilon_{i,j} \} \geq \alpha^n \quad (2)$$

则称 S 相对于当前 n 是重要的,反之 S 是不重要的。其中, α 为常数。为简化表达,我们将集合 S 的重要性写为:

$$\Gamma_n(S) = \begin{cases} 1, & \alpha^n \leq \max_{(i,j) \in S} \{ \epsilon_{i,j} \} < \alpha^{n+1} \\ 0, & \text{其它} \end{cases} \quad (3)$$

为实现嵌入式码流,我们需要将原子按投影系数幅度(或能量系数)递减的顺序进行编码。这里的排序过程是通过将能量系数与一系列递减的阈值 $T_0, T_1, T_2, \dots, T_{N-1}$ 相比较来实现的。阈值满足 $T_n = T_{n-1}/\alpha$ 。初始阈值 T_0 选择为让所有的能量系数 $\epsilon_{i,j}$ 满足 $\epsilon_{i,j} < 2T_0$ 的关系。对每个阈值 T_n ,我们需要判断并传输原子图中每个原子集合 S 的重要性 $\Gamma_n(S)$ 。借鉴成功的小波编码算法如SPECK^[10]的思想,我们可以利用原子分布的特点,减少传输重要性图(significant map)的代价。

前面我们已看到,原子分布总体表现出聚集特性,因此在每次比较,原子重要性图中都有大量的零值存在,把这些零值成组的表示将有利于提高编码效率。为此我们用一组大小不同的方块 S_i 来组织这些零点。初始时,我们将整个原子图看成一个方块 S_0 ,在每次阈值比较时,判断每个 S_i 的重要性,若不重要,只需传输一个符号0;若重要,则将 S_i 均分为四个小块,然后对每个小块继续作重要性检测。由于编解码双方具有相同的集合划分规则,能保持完全相同的状态,不需要传输任何附加信息。

这里存在一个问题,即当集合 S_i 被划分得很小时,如 8×8 大小,由于原子图分布是局部稀疏的,该块很有可能已不包含原子,或只含一个原子,对 S_i 的继续划分会产生很多小的零块,很大程度上降低了编码效率。我们的做法是设置一最小的块大小 Z_{\min} ,当块尺寸达到

Z_{\min} 后就停止划分,转到对块内原子参数和投影系数的编码。 Z_{\min} 的设置与原子图(或图像)大小有关。

可以看到,以上提出的 BPC 与 SPECK 算法在集合划分上存在相似之处,它们最大不同点在于: SPECK 是按位平面方式对小波系数进行排序的,其阈值选为 2^n 的形式, BPC 则是对原子投影系数进行排序,其阈值设定为更一般的 α^n 形式. 当 $\alpha = 2$ 时,即为位平面方式, $\alpha < 2$ 时,可以提供更细的划分,故我们称 BPC 的排序是基于一般位平面的。

2.3 原子参数和 MP 系数的编码

我们对原子的位置、拉伸和旋转等各类参数按照其变化范围设置相应的符号表,并采用与文献[3]类似的方式对其进行编码. 特别地,由于前面的块划分过程已将编码位置定位到了最小块,所以对原子位置参数,我们只需考虑块内的相对位置. 这样,位置参数的动态范围由原来的图像大小 N 缩小到了 Z_{\min} , 编码效率大为提高。

就 MP 系数的量化上,目前的方法基本上是利用系数的指数分布特性来设计量化器. 然而我们知道,该分布是在理论极限情况下推导出来的,与实际存在差距. 另外,即使采用均匀量化器,从率失真意义上选择最优的量化步长也非易事,一般涉及到较复杂的优化问题. 借助位平面块划分方法,我们很好地避免了上述问题. 由于原子是按一般位平面进行排序的,投影系数的最高位平面在排序中已确定,系数的非确定区域大为缩小,简化了量化器的设计并使编码效率得到提高. 在设计 EUQ 算法的理论推导中,文献[5]得出了最优量化阶数应该使得各系数的平均量化误差相同的结论. 基于此,我们提出如下的系数量化方案:对当前检测为重要的原子系数,将以位平面的方式逐比特的对其进行细化,直到某设定的阈值 Δ . 这样我们可将各系数的量化误差控制在 $\Delta/2$ 内。

在编码的最后环节,我们使用基于上下文的自适应算术编码器对输出的各类符号进行熵编码. 由于块划分、原子参数及系数量化的符号分布随图像的变化而变化,我们对各类符号设置独立的上下文并将初始符号概率都设置成均匀分布. 算术编码器的概率随编码的进行而自动更新. 编码器按位平面的扫描顺序将原子的编码信息逐一送出,生成嵌入式码流。

3 码流的高伸缩性

除了具备较高的编码效率外, FMP_BPC 的另一大特点是码流具有灵活的伸缩性. 与传统的方法不同, FMP_BPC 的伸缩性是信号分解所固有的,不需要借助于对变换系数的精心组织. 而且, FMP_BPC 提供的质量

和分辨率可伸缩性是完全分离的,从而避免了传统小波编码中两种可伸缩性相互抑制和干扰的问题。

3.1 质量可伸缩性

FMP_BPC 利用 MP 分解固有的逐步逼近特点,生成嵌入式码流,提供了良好的质量可伸缩功能. 虽然 MP_FLC 也具有质量可伸缩性,但由于它的码流是在特定的目标码率下生成的,不具备嵌入特性,其中间截断会带来恢复质量的下降。

3.2 分辨率可伸缩性

当前优秀的小波编码器如 JPEG2000 具备了分辨率可伸缩性能,但同时其伸缩性也受到了信号分解的限制,即通常只能提供倍频带对应的分辨率,其它大小的分辨率需要通过一定的转码来获得. 而 FMP_BPC 提供了一种更灵活的分辨率可伸缩功能。

原子库固有的结构赋予了 MP 对平移、各向同性拉伸和旋转等几何操作的不变性. 因此对逼近(或重建)图像施加的几何变换可以通过变换原子的参数来获得^[3]. 例如,现有尺寸为 $W \times H$ 的编码图像,我们想将其解码重建为 $\beta W \times \beta H$ 的图像,则我们只需在合成前将所选原子的位置和尺度参数以及相应的投影系数扩大 β 倍即可. 凭借这个特性, FMP_BPC 码流可在任意分辨率下被解码,且操作十分方便有效. 由于该操作并没有改变编码流中原子的相对大小和顺序,故不会对质量伸缩性产生影响。

4 实验结果

本文提出的 BPC 是适合于 MP 图像编码的通用方法. 在图像分解环节,我们采用 FSMP 和 FMTP 两种 MP 算法,进而形成两种编码器,即 MP_BPC(FSMP + BPC) 和 FMP_BPC(FMTP + BPC). 这里我们对其进行了仿真。

4.1 BPC 算法的有效性

对 BPC 算法各环节的参数,我们由实验设定如下: $\alpha = 1.25$; $\Delta = 16$; 对 512×512 大小的图像, $Z_{\min} = 16$, 对 256×256 的图像, $Z_{\min} = 8$. 为验证 BPC 的效率,我们将 MP_FLC、MP_BPC 和 FMP_BPC 三种编码器的输出码流进行分析比较. 表 1 列出了在不同逼近原子数目时,编码流中各组成成分的比特分配情况。

可以看到,对 MP_FLC, 原子位置(平移)参数和投影系数是码流的主体,而在 MP_BPC 和 FMP_BPC 中,这两部分码率得到了显著的降低,特别是位置参数部分. 两算法也增加了额外的块划分表示代价,对 800 个原子的分解,其总体码率平均分别下降 25.53% 和 26.35%. 这个优势随原子数目(或编码率)的增加而越发显著. 注意到,对类型、拉伸、旋转等原子参数,由于采用类似方法,各算法的差别较小。

表 1 对 Lena 图像(256 × 256)各编码器输出码流的比特分配

原子个数	编码器	位平面块划分	原子参数				投影系数	码率 (bpp)
			类型	平移	拉伸	旋转		
400	MP_FLC	0	97	6453	2941	1566	3130	0.2165
	MP_BPC	2862	97	2532	2944	1569	1376	0.1736
	FMP_BPC	2707	78	2533	2954	1567	1366	0.1710
800	MP_FLC	0	135	12892	5916	3193	5795	0.4262
	MP_BPC	4461	130	5136	5928	3194	1955	0.3174
	FMP_BPC	4276	113	5110	5949	3175	1946	0.3139

4.2 编码器的复杂度

我们通过测试编码时间来对编码器的计算复杂度进行评估.表 2 列出了 MP_FLC 和 FMP_BPC 两种编码器对不同码率的编码时间比较,实验 CPU 为 Pentium M 处理器(1.5GHz).可以看到,由于采用了快速分解方法,FMP_BPC 较 MP_FLC 具有明显的速度优势.

表 2 各编码器编码时间(分)的比较

图像(256 × 256)	码率(bpp)	MP_FLC(A)	FMP_BPC(B)	B/A(%)
Lena	0.0625	1013.78	16.02	1.58
	0.125	1013.80	16.05	1.58
	0.25	1013.85	16.12	1.59
Cameraman	0.0625	1014.83	16.40	1.62
	0.125	1014.85	16.43	1.62
	0.25	1014.90	16.50	1.63

4.3 编码器率失真性能

这里我们将提出的编码器与 MP_FLC 及标准编码器 JPEG2000, SPIHT 进行率失真性能比较.表 3 给出了各编码器实验结果.首先我们看到,同样是基于 FSMP 分解的方法,MP_BPC 性能比 MP_FLC 有大幅度的提高,且随着编码率的增加,其优势越发明显,这主要归功于 BPC 算法的贡献.由于采用快速搜索算法,FMP_BPC 的图像质量较之于 MP_BPC 有微小的下降,但仍然保持着相对于 MP_FLC 的优势,如 0.5bpp 时其平均 PSNR 增益达 1.73dB.其次,提出的两编码器获得了与 JPEG2000, SPIHT 相当的客观质量,且在低码率表现出优势.从主观效果上看,由于采用了适合表示图像边缘的基函数,MP 编码器的主观质量在低码率优于 JPEG2000 和 SPIHT,即使在 PSNR 相对较低的时候.图 1 给出了在 0.125bpp 下 Lena(512 × 512)恢复图像的比较.我们可以较清楚看到,FMP_BPC 的质量好于 JPEG2000



图 1 码率为 0.125 bpp 时 Lena(512 × 512) 的恢复图像

和 SPIHT,特别是在图像的边缘轮廓,如在帽檐、镜框等几何特征上.FMP_BPC 引入的误差只是对图像描述的粗略化,而没有出现类似于小波的恼人的振铃(ringing)现象.

表 3 各编码器的 PSNR 比较

图像	编码器	PSNR(dB)			
		0.0625bpp	0.125bpp	0.25bpp	0.5bpp
Lena	SPIHT	28.38	31.10	34.11	37.21
	JP2K	27.94	30.79	33.97	37.17
	MP_FLC	27.70	30.40	32.39	34.68
	MP_BPC	28.46	31.45	34.31	36.90
	FMP_BPC	28.33	31.27	34.18	36.69
Barbara	SPIHT	23.35	24.86	27.58	31.40
	JP2K	23.21	25.28	28.30	32.14
	MP_FLC	23.26	25.06	27.45	30.21
	MP_BPC	23.64	25.89	28.89	32.31
	FMP_BPC	23.49	25.55	28.37	31.68
Goldhill	SPIHT	26.73	28.48	30.56	33.13
	JP2K	26.55	28.36	30.50	33.15
	MP_FLC	26.07	27.79	28.96	30.73
	MP_BPC	26.58	28.48	30.51	32.74
	FMP_BPC	26.54	28.35	30.29	32.47
Mandrill	SPIHT	20.74	21.72	23.27	25.65
	JP2K	20.60	21.57	23.14	25.52
	MP_FLC	20.63	21.45	21.93	23.29
	MP_BPC	20.82	21.88	23.35	25.38
	FMP_BPC	20.71	21.72	23.05	25.00

4.4 分辨率可伸缩性能

图 2 显示了将图像进行缩放解码的一个例子.这里的编码率为 0.5bpp,放缩因子 β 取 $\sqrt{1/2}$ 和 $\sqrt{2}$.可见,FMP_BPC 用一种非常简单和有效的方式获得了高度灵活的伸缩性能,这是它相对于传统编码方法的一个显著优点.



图 2 FMP_BPC 对 Cameraman(256 × 256) 在不同分辨率下的恢复图像

5 结论

本文提出了一种基于 MP 分解的图像编码方法 - FMP_BPC.相对于同类方法,FMP_BPC 在计算复杂度、编码效率和伸缩性能等方面都表现出优势.FMP_BPC 在低到中码率段获得了与先进小波算法相当的客观质

量和更好的主观效果,且提供了更灵活的质量和分辨率伸缩性能,可在网络或移动终端的图像业务中得到很好的应用.

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,2006.49-61.
- [2] MALLAT S,ZHANG Z. Matching pursuits with time-frequency dictionaries[J]. IEEE Transactions on Signal Processing, 1993, 41(12):3397-3415.
- [3] FIGUERAS I VENTURA R M, VANDERGHEYNST P, et al. Low-rate and flexible image coding with redundant representations[J]. IEEE Trans Image Process, 2006, 15(3):726-739.
- [4] RAHMOUNE A, VANDERGHEYNST P, et al. Flexible motion-adaptive video coding with redundant expansions[J]. IEEE Trans Circuits Syst for Video Technol, 2006, 16(2):178-190.
- [5] FROSSARD P, VANDERGHEYNST P, et al. A posteriori quantization of progressive matching pursuit streams[J]. IEEE Trans. Signal Process, 2004, 52(4):525-535.
- [6] 甘涛,何艳敏,等.基于互相关估计的快速图像逼近算法[J].电子学报,2008,36(5):1019-1023.
GAN Tao, HE Yan-min, et al. Fast algorithm for image approximation based on cross-correlation estimation[J]. Acta Electronica Sinica, 2008, 36(5):1019-1023. (in Chinese)
- [7] GAN T, HE Y M, et al. Fast M-term pursuit for sparse image representation[J]. IEEE Signal Processing Letters, 2008, 15: 116-119.

- [8] NEFF R, ZAKHOR A. Modulus quantization for matching-pursuit video coding[J]. IEEE Trans Circuits Syst Video Technol, 2000, 10(6):895-912.
- [9] NEFF R, ZAKHO A. Very low bit-rate video coding based on matching pursuits[J]. IEEE Trans Circuits Syst. Video Technol, 1997, 17(1):158-171.
- [10] PEARLMAN W, ISLAM A, et al. Efficient, low-complexity image coding with a set-partitioning embedded block coder[J]. IEEE Trans Circuits and Syst for Video Technol, 2004, 14(11):1219-1235.

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.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: garretegt@gmail.com



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.

E-mail: yanmin_he@163.com