

一种改进的最佳时频原子搜索策略

刘利雄¹⁾ 贾云得¹⁾ 廖斌²⁾ 张敏¹⁾

¹⁾(北京理工大学计算机科学与工程系, 北京 100081) ²⁾(中国科学院软件研究所, 北京 100080)

摘要 在极低编码速率条件下, Neff 和 Zahor 提出的基于匹配跟踪信号分解的视频编码器不仅具有比 H. 263 编码器更高的编码性能, 而且能够避免产生人眼敏感的方块效应, 但由于该算法需要在一个冗余字典里搜索最佳匹配误差结构的原子函数, 其实现所需要的运算量比传统的编码器要高很多, 因而影响了该编码器的效率。为了提高编码效率, 在对能量优先原子搜索策略进行分析的基础上, 提出了一种改进的全搜索策略和加权能量优先搜索策略, 从而改进了最佳时频原子搜索策略。最后还对搜索策略的编码性能和运算效率进行了评价和实验。

关键词 视频编码 原子搜索 匹配跟踪

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2004)07-0873-05

An Improved Searching Scheme Using Optimal Time-frequency Atoms

LIU Li-xiong¹⁾, JIA Yun-de¹⁾, LIAO Bin²⁾, ZHANG Min¹⁾

¹⁾(Department of Computer Science and Engineering, Beijing Institute of Technology, Beijing 100081)

²⁾(Institute of Software, Chinese Academic Science, Beijing 100080)

Abstract At very low bit rates, the video coder based on matching pursuits proposed by Neff and Zahor can avoid producing visual-sensible block artifacts and has been proved to have better coding performance, compared with H. 263 standards. But, the computation complexity of its realization is much too higher than traditional coders, because it needs to search optimal matching atoms in a redundant dictionary, and the searching burden strongly affects the efficiency of the coder. In order to advance the efficiency of traditional coders, The author first analyzes the atoms searching scheme based on energy priority, then develops a improved full searching scheme and a weighted searching scheme based on energy priority. Both novel time-frequency atoms searching algorithms improve the optimal time-frequency atoms searching scheme. At the same time, we compare and experiment the coding performance and computation efficiency of our algorithms.

Keywords video coding, atom searching, matching pursuit

1 引言

在低速率编码应用当中, 如果采用基于冗余字典的匹配跟踪信号分解算法来取代基于正交分块 DCT 变换对运动残余编码, 则能够获得更好的视觉效果^[1~4]。在 Neff 和 Zahor 给出的低速率匹配跟踪视频编码算法^[3]中, 就是先选择有限可分离二维 Gabor 窗函数集合来建立冗余字典, 然后通过不断地在该字典中搜索最佳匹配原始图像的时频原子, 再利用这些最佳时频原子的线性组合来逼近原始图

像。虽然依据匹配跟踪算法的要求, 可以先在图像中的每个像素位置计算字典中所有时频原子与图像的内积, 然后将内积最大的时频原子作为此次匹配的结果, 但是, 如果每次迭代过程都在整幅图像范围内进行搜索, 则所需要的计算量是相当大的。为了减少计算量, 可在原文给出的算法中, 采用一种基于能量优先的搜索策略, 其基本思想是先将图像划分为固定尺寸的宏块, 然后选择其中能量最大的块作为搜索起点, 然而由于采用这种方式搜索得到的最佳时频原子不一定准确, 从而将降低算法收敛的速度, 并将进一步影响最终的编码效率。本文提出了一种改

进的全搜索以及加权能量搜索策略,并对它们之间的性能进行了比较。为了讨论方便,首先给出相干比和能量衰减程度的定义。

对于一个冗余的字典集合,若采用匹配跟踪算法来对信号 f 进行逼近扩展,则经过 $n+1$ 次迭代后的能量衰减程度^[4] $\Delta R^{(n+1)}$ 定义为

$$\Delta R^{(n+1)} = \|R^{(n)}f\|^2 - \|R^{(n+1)}f\|^2 \geq w[m] \|R^{(n)}f\|^2 \quad (1)$$

其中, $\|R^{(n)}f\|^2$ 和 $\|R^{(n+1)}f\|^2$ 分别是信号 f 的 n 阶和 $n+1$ 阶残余能量, $w[m] = \lambda^2(R^{(n)}f)$ (m 表示宏块中所包含的最佳原子数目)表示信号衰减的速率,它依赖于信号残余 $R^{(n)}f$ 与基函数 g_γ ($\gamma \in \Gamma = R^+ \times R^2$) 之间的相关性;相干比(coherency ratio)^[4] 定义为

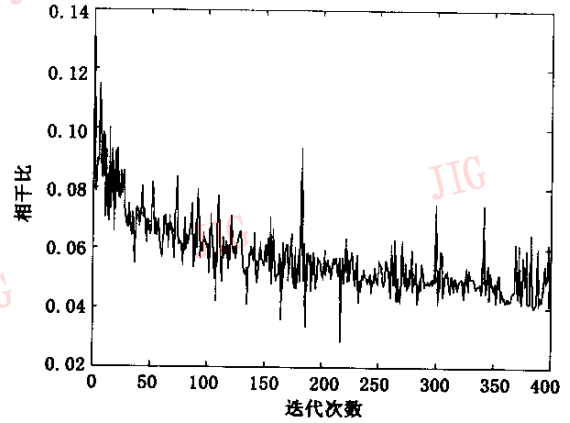
$$\lambda(R^{(n)}f) = \sup_{\gamma \in \Gamma} \frac{|\langle R^{(n)}f, g_\gamma \rangle|}{\|R^{(n)}f\|} \quad (2)$$

2 能量最大优先搜索策略

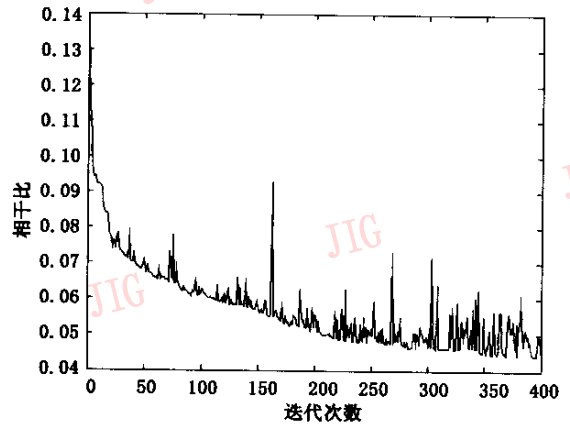
Neff 和 Zahor 认为,在预测误差帧(Missa)残余能量分布不均匀、稀疏的前提下,首先将图像划分为 12×12 大小的重叠块,并作出如下假设,即,若在具有能量最大的块范围内搜索到最佳时频原子的可能性最大,则就可以避免在整幅图像内搜索,而只局限于一个局部的搜索窗口内搜索,这样就可使得整个搜索过程所需要的运算量大大降低。

然而,通过对计算的相干比的分析,如图 1(a)所示,采用这种方式搜索到的时频原子并不一定是最佳时频原子,由于相干比曲线并不呈现应有的递减趋势,且波动很大,而且在迭代过程中,往往会突然出现一个相干比值很大的原子,这说明这样的选择过程没有能够将最能够反映图像特征的原子优先挑选出来,从而也就直接影响到匹配跟踪算法的收敛速度。

为了进一步考察该算法,可尝试将图像块按照能量值排序,然后选择在能量值最大的前 k 个块中搜索最佳时频原子,这样做实际上是针对原有的假设作出修改,即认为假使最佳时频原子没有出现在最大能量块范围内,那么至少也应该出现在前 k 个最大能量块中(当然 k 的选择不可能太大,一般而言 $k < 10$, 当 $k=1$ 时,相当于原有的搜索策略;而当 k 取比较大的值时,则将会使得搜索运算量大幅上升,并将逼近在整幅图像内搜索的情况)。图 1(b)给出的相干比曲线表明,在初始的迭代过程中,已基本上能够挑选出近似最佳的时频原子,其相干比曲线变化比较平滑,然而随着迭代次数的增加,而在迭代过程后期,相关曲



(a) 搜索范围内的最大能量块数 $k=1$,
均值 $\mu=0.057$, 方差 $\delta=0.0126$



(b) 搜索范围内的最大能量块数 $k=10$,
均值 $\mu=0.0575$, 方差 $\delta=0.0119$

图 1 预测误差帧(Missa)的相干比曲线

线则出现明显的波动,这反映出该搜索策略的前提假设不总是成立,即并不符合实际情况。

3 改进的全搜索策略

大家知道,在整个图像区域内搜索最佳时频原子所需的运算量非常高。考虑到因时频原子本身的局部特性而造成其空间支撑范围的有限性,如,在实验采用的有限离散 Gabor 字典当中,由于最长的时频原子长度为 $L_{\max} = 35$,因此并不需要在每次迭代过程中,都在整个图像范围内搜索,而可以利用上一次的搜索结果。具体算法流程如下:

(1) 初始化 将尺寸为 $W \times H$ 的图像依然划分为 $B \times B$ ($B=16$) 的宏块,并建立一个用于存储与

每个宏块对应的最佳时频原子的数组,初始化为空, $t=0$ (t 为算法循环迭代次数);

(2) 检查每个宏块的最佳时频原子是否为空,如果为空,则在围绕着该块中心位置的 $B \times B$ 的窗口区域内搜索出一个最佳时频原子,并记录其参数信息,包括系数幅值、位置 (x, y) 、水平和垂直原子索引,同时通过比较每个宏块的最佳时频原子的系数幅值,然后将内积最大的时频原子作为整个图像块的最佳时频原子;

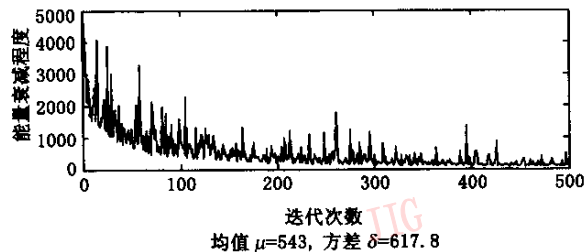
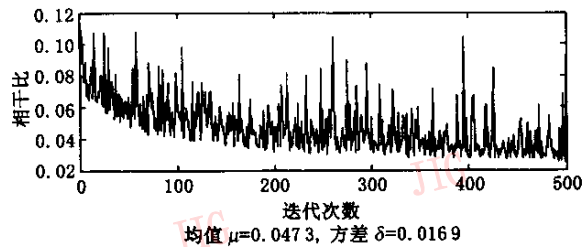
(3) 更新误差图像,并将处于该时频原子空间支撑范围内的所有宏块的最佳时频原子标记为空,因为这些图像块的最佳时频原子需要重新搜索。这样在以后的搜索过程中,如果任何一个块中的元素值没有发生改变,则该块内的最佳时频原子也不可能发生变化,这样就不需要在该块中重新搜索原子;

(4) $t=t+1$;如果没有达到预定迭代次数,则返回步骤 2;

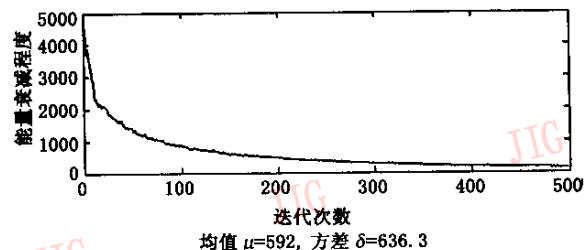
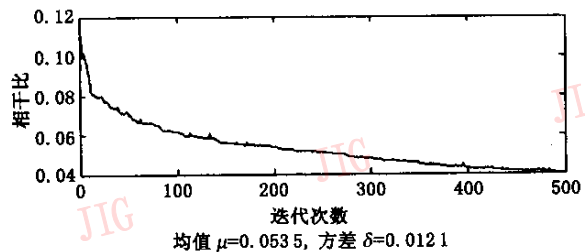
(5) 结束。

采用这种方式,就可以有效利用前次计算的内积结果,除了迭代过程的第 1 步,因需要在全局范围内搜索而计算量较大以外,在后续的每次迭代过程中,则只需要重新搜索 $(L_h \cdot L_v)/B^2$ 个宏块内的最佳时频原子(L_h 为水平滤波器长度, L_v 为垂直滤波器长度)。从图 1 中给出的相干比以及能量衰减程度分析来看,随着迭代次数的增加,相干比的变化基本上呈现出单调递减的趋势,这表明与预测误差图像残余结构的特征最匹配的原子可被优先挑选出来。图 2 中的误差能量衰减曲线也证明改进后的搜索策略可使得匹配跟踪算法实现的收敛速度大大加快。对于 QCIF 格式的 missa 图像序列中某预测误差图像帧,用该方法进行了实验,经过 500 次迭代过程,能量最大优先搜索算法重构图像的峰值信噪比和运算时间为: $PSNR=41.15$ dB, $T=1.6$ s;而用改进全搜索策略重构图像的峰值信噪比和运算时间为: $PSNR=42.29$ dB, $T=3.7$ s。由此可见,后者重构图像的质量虽提高了大约 1dB 左右,但运算时间也相应增加了不少,这主要是由于为了提高搜索的精度,所需要搜索的宏块数目比以前增加较多所致,这种搜索策略平均每次迭代过程需要搜索 2.6 个宏块。

为了进一步降低需要重新搜索的宏块的数目,笔者又通过进一步分析发现,如果宏块内当前的最佳时频原子所支撑的空间区域与整个图像最佳时频原子支撑的空间区域并没有发生重叠,那么即使该



(a) 能量优先搜索策略



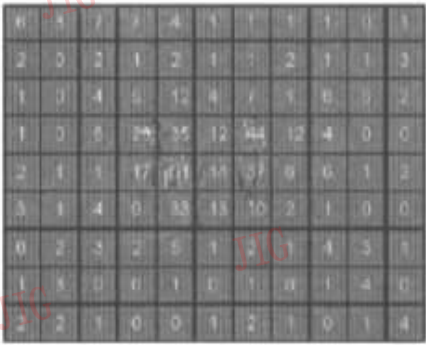
(b) 改进的全搜索策略

图 2 预测误差帧(Missa)的相干比 $\lambda(R^{(n)}f)$ 曲线及能量衰减程度 $\Delta R^{(n)}$

宏块内的某些数据发生变化,由于仍然可以认为该宏块内的最佳时频原子不会发生变化,因此不需要重新搜索。经过这样的处理以后重构图像的峰值信噪比和运算时间为: $PSNR=42.16$ dB, $T=2.4$ s。据统计,除了第 1 次迭代过程以外,其余迭代过程平均每次需要搜索的宏块数目降低了 1.8 个,而原有的搜索策略每次只选择一个能量最大的宏块进行搜索,这也基本上与所增加的运算时间相符。图 3 为经过 500 次迭代过程以后,两幅误差图像各宏块内的原子数目分配图。



(a) Missa(176×144)



(b) Hallmonitor(176×144)

图 3 各宏块内原子数目分布图

4 加权能量优先搜索策略

该搜索策略的主要思想在于选择初始搜索块时,除了应考虑宏块本身包含的能量因素以外,还应考虑宏块所包含的原子数目。笔者认为,每个宏块所期望的能量衰减程度是随着该宏块原子数目的增加而减少的,换句话说,出于能量衰减的考虑,那些包含较少数目原子的宏块成为初始搜索块的可能性要大于具有较多数目原子的能量更大的宏块。上述思想可以应用于预测那些能够引起最大能量衰减的宏块中,例如

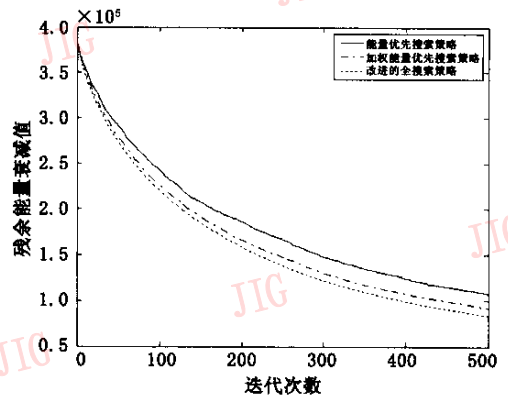
$$\Delta R_i^{(n+1)} \geq w_i[m] \|R_i^{(n)} f\|^2 \quad (4)$$

其中, $R_i^{(n)}$ 表示宏块 i 的第 n 阶残余; m 表示宏块 i 中所包含的最佳原子数目,且

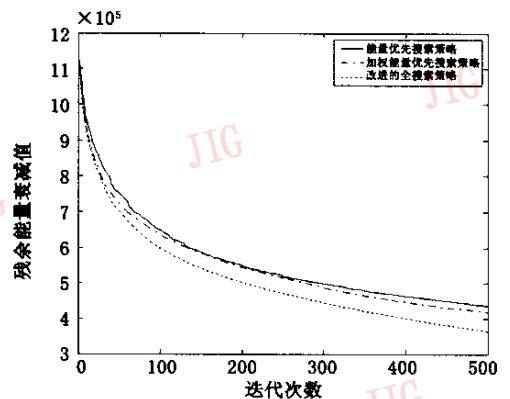
$$w_i[m] = \lambda^2 (R_i^{(n)} f)$$

因此,如果能够得到关于加权值 w_i 的一个有效估计,则利用公式(4)就可以确定各宏块能量衰减的下界。如果直接利用公式(2)计算权值,则需要对每个宏块都进行匹配跟踪的迭代过程,这样所需要的运算量是不可以接受的。为了减少计算复杂度,可以考

虑假设所有宏块的加权值只与宏块所包含的原子数目有关,而与宏块的具体位置无关,即令 $w_i[m] = w[p]$ (p 代表了宏块中包含的原子数目)。其计算方法是先采用改进的全搜索算法对预测误差残余作匹配跟踪分解,然后计算每个宏块相应的归一化能量衰减因子 $\Delta[p]$ ($\Delta[p] = \frac{\|R^{(n+1)} f\|^2 - \|R^{(n)} f\|^2}{\|R^{(n)} f\|^2}$)。为了获得统计意义上有效性,可以通过选择几幅测试序列来计算平均 $\Delta[p]$,并进行归一化处理,以使得最大加权值为 1,这样就可获得关于能量加权值的一个估计 $\hat{w}[p] = \Delta[p]$,而利用加权能量优先搜索策略,在相同条件下,就可对 Missa 图像序列中某预测误差帧进行匹配跟踪分解,其重构图像的峰值信噪比和运算时间为: $PSNR=41.87\text{dB}$, $T=1.7\text{s}$ 。从图 4 所示的残余能量衰减曲线分析结果可见,加权能量搜索策略的衰减速率要比原有的能量优先策略提高很多,虽与改进的全搜索策略相比仍有一些差距,但其优势在于运算复杂度较低,这主要是由于在每次迭代过程中,只需要对一个宏块进行最佳原子的搜索即可。



(a) Missa 图像



(b) Hallmonitor 图像

图 4 残余能量 $\|R^{(n)} f\|^2$ 随迭代次数衰减曲线

5 结 论

本文对基于冗余字典的匹配跟踪所涉及到的时频原子的搜索问题进行了分析。由于在匹配跟踪视频编码算法中,需要在原子字典中搜索最佳匹配原始图像的时频原子,而按照原始匹配跟踪信号分解算法的要求,则将会使得整个视频编码算法实现的运算量比较高,因此,为了减少搜索实现的运算量, Neff 和 Zahor 给出了能量最大优先搜索的思想,它虽然降低了搜索的运算量,但是同时也影响了算法收敛的速度。为了提高搜索策略的运算效率,本文提出了一种改进的全搜索以及加权能量搜索策略,并对它们之间的性能进行了比较和分析。实验结果表明,采用以上两种搜索方式,不仅使得整个算法实现的计算量显著下降,同时也保持了原有的视频编码性能,这使得匹配跟踪算法在视频编码方面得到进一步推广。

参 考 文 献

- 1 Vetterli M, Kalker T. Matching pursuit for compression and application to motion compensated video coding[A]. In: IEEE International Conference on Image Processing[C], Austin, TX, USA, 1994:725~729.
- 2 Neff R, Zahor A. Very low bit rate video coding based on matching pursuits [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997,7(1):158~171.
- 3 Banham M, Braillean J. A selective update approach to matching pursuits video coding [J]. IEEE Transactions on Circuits and Systems for video Technology, 1997,7(1):119~129.
- 4 Gharavi-Alkhansari M. A model for entropy coding in matching pursuit [A]. In: Proceedings of IEEE International Conference Image Processing'98[C], Chicago, Illinois USA, 1998:778~782.

- 5 Rijse, Karel. H.263: Video coding for low bit rate communication [J]. IEEE Communications Magazine, 1996, 34(12):42~45.
- 6 Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries [J]. IEEE Transactions on Signal Proceedings, 1993,41(12):3397~3415.



刘利雄 1974 年生,2001 年获武汉大学计算机应用技术专业硕士学位,现为北京理工大学计算机科学与工程系讲师,在职博士。主要研究方向为视频图像编码、计算机视觉、人工智能。

E-mail:lxlin@bit.edu.cn



贾云得 1962 年生,工学博士,北京理工大学计算机科学与工程系教授,博士生导师。主要研究方向为计算机视觉、媒体计算和智能系统。



廖 斌 1974 年生,2003 年获中国科学院软件研究所计算机应用技术专业博士学位。主要研究方向为图像编码、多媒体技术。



张 敏 1975 年生,2001 年获郑州工业大学计算机应用技术专业硕士学位,现为北京理工大学计算机科学与工程系博士研究生。主要研究方向为计算机视觉、机器学习、人工智能。