

# 一种基于贝叶斯决策的自适应视频编码方法

柳葆芳 平西建 李弼程 邵美珍

(郑州信息工程大学信息技术学院信息科学系, 郑州 450002)

**摘要** 针对图象质量和传输码率在不同的视频传输系统中的不同要求, 给出了一种用于视频编码宏块分类的贝叶斯代价函数, 并依此构成了基于贝叶斯决策的自适应编码算法。试验表明, 该算法可以实现图象质量和传输码率的主观折衷, 并能提高重建图象的质量和平均峰值信噪比。

**关键词** 视频编码 图象质量 传输码率 贝叶斯代价 折衷参数

中图法分类号: TP919.8 文献标识码: A 文章编号: 1006-8961(2001)01-0065-05

## Adaptive Video Coding Based on Bayes-Cost Function

LIU Bao-fang, PING Xi-jian, LI Bi-cheng, SHAO Mei-zhen

(Zhengzhou Information Engineering University Information Technology  
Institute Information Science Department, Zhengzhou 450002)

**Abstract** Both picture quality and coder's outputting bit-rate are important in the video coding process. An adaptive video coding algorithm is proposed in this paper. The algorithm takes into account both picture quality and coder's outputting bit-rate. It gives out a reasonable Bayes-cost function and introduces Bayes decision method into the video coding process. The algorithm uses Bayes decision method to choose coding strategy in order to agree with the statistical sort of image macroblock. The algorithm chooses the coding strategy whose Bayes-cost is the minimum and the subjective assessment of the coding strategy is the best. At the same time, the algorithm gives out the trade-off parameter  $\lambda$  between picture quality and coder's outputting bit-rate. The trade-off parameter  $\lambda$  is updated according the state of the buffer and the coder's outputting bit-rate is smoothed in order to make the best use of the channel. This method can get better quality of reconstructed picture and higher average PSNR.

**Keywords** Video coding, Picture quality, Translating bit-rate, Bayes-cost, Trade-off parameter

## 0 引言

近年来, 视频数据压缩方法有了令人瞩目的发展, 且一系列活动图象编码标准有力地促进了数据压缩技术的进步和商品化进程; 同时视频通信传输码率也从 H. 261 的中等码率(64k~1 920kbit/s), 发展至 MPEG-2 的高码率(达到 80M bit/s), 如今低码率端的视频通信(H. 263)又成了新的研究热点之一, 而目前 MPEG-4 中对可移动性的视频编码速率仍仅仅为 5~64kbit/s。

随着国际标准的发展, 视频编码算法的研究也

很活跃。由于块匹配算法定义简单且有效, 因此其在运动估计和补偿中得到广泛的应用, 并且人们已提出了很多的改进算法<sup>[1,2]</sup>。H. 263 中提出了用两步的块匹配法来进行运动估计和补偿<sup>[3]</sup>, 通过半像素精度的运动估计, 不仅提高了估计精度, 还改善了预测效果。而且在这些运动估计和补偿算法中, 运动矢量的搜索仅从图象质量出发, 并以均方误差(函数)(MSE)或绝对平均误差(MAD)作为块匹配过程中的匹配准则。

一般, 实际的图象信源是非平稳的, 即不同时刻、帧内不同位置的图象信号可能表现出不同的统计特性。为了提高压缩编码的性能, 应当针对图象数

据局部的统计特性,采用自适应编码技术来进行压缩编码.一般的自适应编码方法是先将子图象信源进行分类,然后根据图象数据局部的统计关系来选取适当的编码策略,但在这样的自适应处理中,难以结合不同的应用条件,给出符合某种主观要求的图象分类结果.因此,在编码系统设计中,编码策略应当与实际要求相联系,还需要使用具有一定程度主观折衷能力的自适应编码技术,且这种编码策略不应当局限于严格统一的控制参数,而是要能够按照主观需要进行修正,以便进行与主观评价对应的最佳决策.由于贝叶斯决策是一种主观评价的分类,因此本文给出一种基于贝叶斯决策的自适应视频编码方法,并从视频信源的非平稳性和信道码率恒定之间的矛盾分析出发,给出了一种贝叶斯代价函数,然后依据贝叶斯决策来进行视频数据分类和选择适当的编码策略,由于实现了重建图象质量与输出码率之间的折衷,从而平滑了编码器的输出比特率,和实现了视频数据的自适应编码.

## 1 贝叶斯代价函数的定义

由于在视频编码过程中,人们不仅希望得到高质量的重建图象,而且希望编码器的输出比特率较小,因此编码器的输出比特率在视频编码过程中也是很重要的指标,但要想得到较好的质量,那么输出比特率必然增大;要想使输出比特率减小,那么图象质量必然降低,这两者是彼此矛盾的,特别是在低码率视频编码中,显得更为突出,所以,在视频编码的自适应处理中,合理的编码方法应当能够尽量同时满足图象质量和编码器输出比特率的要求,但是在通常的运动估计和补偿算法中,匹配准则主要从图象质量的角度出发,没有考虑视频编码过程中的另一个重要因素——编码器的输出比特率.为了能满足这两方面的要求,人们已提出了一些改进的算法<sup>[4,5]</sup>.本文在综合考虑图象质量和编码器的输出比特率的前提下,提出了基于贝叶斯决策的自适应编码算法,并给出了视频编码的贝叶斯代价函数 $C(Q,R)$ ,其中 $Q$ 表示图象质量, $R$ 表示编码输出的比特数.一般情况下,对图象质量的评价用失真 $D$ 来表示,这样,贝叶斯代价函数又可表示为 $C(D,R)$ ,同时,对贝叶斯代价函数提出以下要求:

$$\textcircled{1} C(D,R) \geq 0;$$

$$\textcircled{2} C(D,R) \text{ 随图象失真的增大而增大;}$$

$$\textcircled{3} C(D,R) \text{ 随输出比特率的增大而增大.}$$

由于图象信源的非平稳性,因而致使编码器的输出比特率随时间变化而不同,然而信道的传输码率是恒定的,显然这两者之间存在着矛盾,而且这种矛盾必然导致信道资源不能充分利用.为了缓解这种矛盾,可在贝叶斯代价函数中,给出折衷参数 $\lambda$ ,以便在图象质量与传输码率之间进行折衷, $\lambda$ 可根据输出缓存器的存储量大小取不同的值,并根据当前传输码率状况不断调整图象质量和传输码率在贝叶斯代价中的比重,从而使编码器的输出尽量平稳,以保证充分利用信道资源.

根据以上的分析和视频编码的具体过程,本文将一种具体的贝叶斯代价函数 $C$ 定义为

$$C = D(d,q) + \lambda_1 \times R_m + \lambda_2 \times R_c + \lambda_3 \times R_s \quad (1)$$

其中, $\lambda_1, \lambda_2, \lambda_3$ 均大于零.由式(1)定义的贝叶斯代价函数可看成两部分之和,其第一部分为

$$C_1 = D(d,q) \quad (2)$$

其中, $d = (u, v)$ 是运动矢量, $q$ 为量化因子, $D(d,q)$ 是在运动矢量为 $d$ 和量化因子为 $q$ 的情况下所产生的编码失真,通常这一失真可以用均方误差 $MSE$ 或绝对平均误差 $MAD$ 来衡量,由此可见,这一部分是图象质量在贝叶斯代价中的体现.

其第二部分为

$$C_2 = \lambda_1 \times R_m + \lambda_2 \times R_c + \lambda_3 \times R_s \quad (3)$$

其中, $R_m$ 为对运动矢量 $d$ 进行编码所需要的比特数; $R_c$ 为对量化后的离散余弦变换系数进行编码所需的比特数, $R_s$ 是对辅助信息(例如宏块类型、量化因子等)进行编码所需的比特数; $\lambda_1, \lambda_2, \lambda_3$ 为3个折衷参数,由于其不同的取值,可以得到不同的码率和失真的折衷,因此 $R_m, R_c, R_s$ 可以根据重要程度的不同而取不同的值.另外,由于 $R_c$ 在编码器输出的比特数中只占有较大的比例,并且采用不同的编码策略时,变化也较大,而 $R_m$ 和 $R_s$ 在编码器输出比特数中占较小的比例,而且 $R_m$ 在采用不同的编码策略时,变化比 $R_s$ 大,所以一般取 $\lambda_2 > \lambda_1 > \lambda_3$ ,由此可见,这一部分是输出比特率在贝叶斯代价中的体现.

从以上的分析可以看出,由式(1)定义的贝叶斯代价函数已满足前面所提的3个要求,而且综合考虑了图象质量和输出比特率两个重要因素,所以由式(1)决定的贝叶斯代价函数是合理的.

## 2 宏块类型及其编码方法

本文提出的算法将在通过贝叶斯决策选择编码

策略的同时,确定图象信源的统计特性类型,但要得到正确的图象信源分类,不仅要求给出的贝叶斯代价函数合理,同时要求不同的编码策略可以针对图象信源统计特性的变化来实现有效的压缩编码,因此本文将图象信源分成 4 类,并根据不同的统计特性给出了 6 种编码策略。

根据图象信源不同的统计特性将图象宏块分成如下 4 类:

- (1) 与参考帧相比,没有变化或变化很小;
- (2) 存在以宏块为单位的运动;
- (3) 一个宏块中包含不同方向运动的物体;
- (4) 与参考帧完全不相关。

对于第 1 类,由于宏块与参考帧的相应宏块相比,变化很小,因而可以采用不编码的方法,并在解码时,用参考帧相应宏块来代替当前宏块即可;对于第 2 类以宏块为单位运动的情况,由于运用帧间运动估计可以很好地去掉相关,所以采用帧间编码方式;对于第 3 类一个宏块中包含不同方向运动物体的情况,由于以宏块为单位进行运动估计,不能得到很好的预测效果,所以需将一个宏块分成 4 个块,然后分别进行运动估计;对于第 4 类与参考帧不相关的情况,由于采用运动估计已没有意义,所以采用帧内编码方式。

将运动估计补偿方法与量化因子相结合而给出的 6 种宏块编码方式分别为跳过编码方式(Not coded)、帧内编码方式(INTRA)、帧内+量化因子编码方式(INTRA+q)、帧间编码方式(INTER)、帧间+量化因子编码方式(INTER+q)、高级编码方式(INTER+4V)。这 6 种编码方式与宏块类型的对应关系见表 1。

表 1 宏块类型对应的编码方式

宏块类型	1	2	3	4
编码方式	Not coded	INTER INTER+q	INTER+4V	INTRA INTRA+q

从以上的分析可以看出,这 6 种编码方式可以有效地解决 4 类统计特性图象信源的编码问题。

## 3 基于贝叶斯决策方法的自适应编码算法

### 3.1 算法的计算过程

在给出了贝叶斯代价函数和宏块编码方式后,就需要在 6 种编码方式中,选择代价最小的来对宏

块进行编码。从式(1)可以看出,贝叶斯代价函数的计算量较大,为了减少计算量,可采用以下方法:

①  $\lambda_1, \lambda_2, \lambda_3$  取相同的值,即实际上只有一个折衷参数  $\lambda$ ;

② 通过设定阈值来逐级确定编码方法,即采用阈值法来判定是采用跳过编码方式,还是其它编码方式。如果跳过编码方式的代价小于阈值  $T_n$ ,就采用跳过编码方式,否则再计算其它编码方式的代价。

③ 在帧间编码方式中,由于在编码过程中,运动矢量和量化因子的相关性不大,且将两者分开考虑对最后结果的影响较小,因此可以分别搜索运动矢量  $d$  和量化因子  $q$ ,即首先根据下式估计运动矢量

$$C_m = \sum_{r \in w} \rho(I(r, n) - I(r + d, n - 1)) + \lambda R(d) \quad (4)$$

其中,  $r$  是图象象素的空间坐标,  $n$  为帧序号,  $I(r, n)$  为当前帧当前块的亮度值,  $I(r + d, n - 1)$  为前一重建图象对应块的亮度值,  $w$  为匹配窗口的大小(16×16 或 8×8),  $\rho(\cdot)$  是均方误差(MSE)或绝对误差(MAD),  $R(d)$  为编码运动矢量  $d$  所需要的比特数。这样使式(4)达到最小的运动矢量  $d'$ , 就是我们所要的运动矢量。

然后根据下式确定量化因子

$$C_t = D_r(q_a) + \lambda R(q_a) \quad (5)$$

其中,  $R(q_a)$  是在量化因子附加量为  $q_a$  的情况下,编码离散余弦变换系数所需的比特数,  $D_r(q_a)$  是当前块与前一重建图象中对应块的绝对误差(MAD)。

对于 INTER+4V 编码方式,也是先搜索得到 4 个运动矢量  $d'_1, d'_2, d'_3, d'_4$ , 然后再确定量化因子  $q$ 。

需要指出的是,尽管采取了以上的方法来减少计算量,但算法所需的计算量仍然较大。

宏块编码方法的选择过程如下:

① 计算跳过方式的代价  $C_{\text{cod}=1}$ , 若  $C_{\text{cod}=1} < T_n$ , 则编码方式为跳过方式,即编码方式 Mode 为 Not coded, 然后跳到步骤⑧。

② 分别搜索一个和 4 个运动矢量  $d'$  及  $d'_1, d'_2, d'_3, d'_4$ , 并计算  $C_m(d')$  和  $C_m(d'_1, d'_2, d'_3, d'_4)$ , 如果  $C_m(d'_1, d'_2, d'_3, d'_4) < C_m(d')$ , 则选择高级编码方式对宏块进行量化编码,并计算得到这种编码方式的代价  $C_t(d'_1, d'_2, d'_3, d'_4)$ , 然后,置高级编码方式开关为 1, 并跳到步骤④, 否则,置高级编码方式开关为 0。

③ 在运动矢量  $d'$  的基础上,搜索量化因子的附加量  $q_a$ , 并计算采用帧间编码方式的贝叶斯代价

$C_i(d', q)$ .

④在帧内编码方式中,搜索量化因子的附加量  $q_a$ ,并计算采用帧内编码方式的贝叶斯代价  $C_i(q)$ .

⑤比较  $C_i(d'_1, d'_2, d'_3, d'_4)$  或  $C_i(d', q), C_i(q)$ , 如果  $C_i(d', q)$  或  $C_i(d'_1, d'_2, d'_3, d'_4)$  最小,就跳到步骤⑥; 如果  $C_i(q)$  最小,就跳到步骤⑦.

⑥如果高级编码方式开关为 1,则宏块编码方式 Mode 为高级编码方式 INTER+4V; 如果高级编码方式开关为 0,而且  $q_a=0$ ,则宏块编码方式 Mode 为帧间编码方式 INTER; 如果高级编码方式开关为 0,而且  $q_a \neq 0$ ,则宏块编码方式 Mode 为帧间+量化因子的编码方式 INTER+q.

⑦如果  $q_a=0$ ,则宏块编码方式 Mode 为帧内编码方式 INTRA; 否则,宏块编码方式 Mode 为帧内+量化因子的编码方式 INTRA+q.

⑧按照前面几步所选的编码方式对宏块编码.

### 3.2 折衷参数的选择

从式(1)可以看出,折衷参数  $\lambda$  主要用于在图象质量与输出比特率之间进行折衷, $\lambda$  的大小变化可以调整两者在贝叶斯代价中的比重.若在通信中输出比特率是主要矛盾,则可以使  $\lambda$  较大,这样在编码方式选择过程中就侧重考虑输出码率,而忽略重建图象质量;如果通信中输出比特率较宽裕,则可以使  $\lambda$  较小,这样在编码方式选择过程中就侧重考虑图

象质量,以尽量提高图象的质量.由于输出比特率可以由输出缓存器中的存储量反映出来,因此可根据折衷参数与输出码率之间的关系,给出如下式的一种方法来使折衷参数能根据缓存器的存储量变动而不断更新

$$\lambda_n = \lambda_{n-1} s(t) / s' \quad (6)$$

其中,  $t$  表示编码当前宏块的起始时间;  $\lambda_{n-1}$  为前一个宏块的折衷参数取值;  $\lambda_n$  为编码当前宏块的折衷参数;  $s' = s_{\max} / 2$  是我们希望维持的缓存器的存储量,其中  $s_{\max}$  为缓存器可以达到的最大存储量;  $s(t)$  表示缓存器在时间  $t$  的存储量.

这样通过折衷参数  $\lambda$  随着缓存器当前存储量的不断更新,就可以根据缓存器的当前状态不断调整输出比特率的比重,而使输出比特率尽量平滑,以缓解编码器不断变化的输出比特率与恒定的信道输出比特率之间的矛盾,以达到充分利用信道资源的目的.

## 4 实验结果

本文在 32kbit/s 数码率的情况下对测试图象 Miss American(CIF format 352×288×12 bit/pixel 25 帧/s)做了实验.同时取阈值  $T_n=900$ ,  $\lambda$  的初值取 100,其实验结果如图 1~图 5 所示.

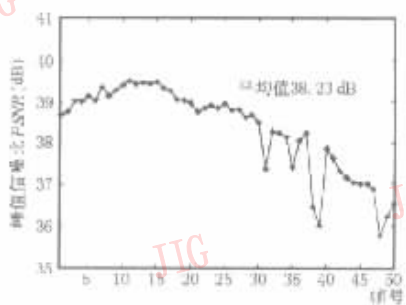


图 1 非自适应编码的信噪比曲线



图 2 非自适应编码的重建图象

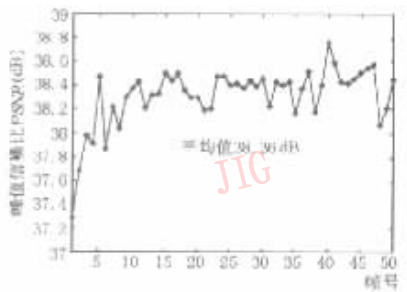


图 3 自适应编码算法的信噪比曲线



图 4 自适应编码的重建图象

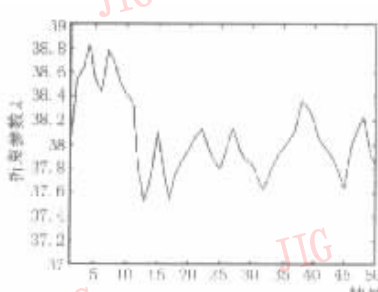


图 5 折衷参数的变化曲线

图 1 和图 3 分别是非自适应编码与自适应编码算法的峰值信噪比曲线,图 2 和图 4 分别是两种算法的重建图象(第 5 帧).从图 1 和图 3 可以看出,基于贝叶斯代价的自适应编码算法的平均峰值信噪比非自适应编码算法,提高了 0.13dB,通过比较图 2 和图 4 可见,重建图象的质量也有所提高.

本文设定缓存器的大小为 10 kbits,从图 5 中可以看到,折衷参数随缓存器的当前状态变化而不断更新调整,并围绕统计所得的均值 100 上下变化,而且在编码过程中没有出现缓存器溢出和为空的现象,从而达到了平滑编码器输出比特率的效果.

## 5 结 论

本文提出的基于贝叶斯决策方法的自适应编码算法,其实质是在综合考虑了图象质量与输出比特率的前提下,通过提出合理的贝叶斯代价函数,并将贝叶斯决策方法引入到视频编码中编码方法的决策过程.然后通过选择贝叶斯代价最小的编码策略来对图象进行编码,以达到选择主观评价最佳的编码策略的目的.同时给出了对图象质量和输出比特率进行折衷的折衷参数  $\lambda$ ,并使  $\lambda$  根据缓存器的存储状态不断更新,从而达到平滑编码器的输出比特率和充分利用信道资源的目的.

### 参 考 文 献

- Ghanbari M. The cross-search algorithm for motion estimation. IEEE Trans. Commun., 1990,38(9):1799~1808.
- Lee L W, Wang L F, Lee J Y *et al.* Dynamic search-window adjustment and interlaced search for block-matching algorithm. IEEE Trans. Circ. Syst. Video Technol., 1993,3(2):85~87.
- Telenor Research. TMN(H.263)encoder/decoder, version 1.4a, TMN(H.263) Codec, May. 1995.
- Wiegand T, Lightstone M, Mukherjee D *et al.* Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard. IEEE Trans. Circ. Syst. Video Technol., 1996,6(2):182~190.
- Kossentini F, Lee Y, Smith M *et al.* Predictive RD optimized motion estimation for very low bit-rate video coding. IEEE Journal on Selected Areas in Communications, 1997,15(9):1752~1763.
- 柳葆芳. 基于码率失真函数的自适应编码算法研究[硕士学位论文]. 郑州:郑州信息工程大学,1999.3.

柳葆芳 1999 年获郑州信息工程学院硕士学位,现为郑州信息工程大学信息技术学院信息科学系博士研究生.主要研究方向为图象处理与分析、图象压缩编码、信息隐藏.

平西建 1982 年获北京航空学院硕士学位,现任郑州信息工程大学信息技术学院信息科学系教授,博士生导师.主要研究兴趣为图象信源编码理论与方法、图象处理与识别、计算机视觉、信息隐藏.

李弼程 1998 年获解放军国防科技大学博士学位,现任郑州信息工程大学信息技术学院信息科学系讲师.主要研究兴趣为图象处理与分析、图象压缩编码、小波、矢量量化.

邵美珍 1988 年获郑州信息工程学院硕士学位,现任郑州信息工程大学信息技术学院信息科学系副教授.主要研究兴趣为模式识别、图象处理与分析、图象信源编码理论与方法.