

一种新的低功耗的图象编 / 解码算法¹

赵乘骥 庞朝阳 朱维乐

(电子科技大学电子工程学院 成都 610054)

摘 要 该文提出了一种改进的低功耗并且只需少量存储空间的图像, 视频编码方案, 它是基于矢量量化的, K. Masselos 等人 1998 年提出的方案使用了一个小的基本码书, 然后通过基本码书中的每一码字进行八种对称变换, 从而把小的基本码书扩展成为一个大的虚码书, 补偿了因为码书的尺寸太小而引起的图象质量下降, 因为它的基本码书小, 从而对存储器的访问次数也就少, 那么功耗也就小. 同原方案相比, 本文做出了如下两个改进: 一是基本码书中存放的是码字的经过量化的 DCT 系数, 寻找最佳匹配码字也放在 DCT 域内进行, 二是改进了 K. Masselos 文中的基本码书的设计方案, 消除了码书中存在的重复. 改进方案具有计算量更小, 操作更灵活的特点.

关键词 矢量量化, 基本码书, 离散余弦变换

中图分类号 TN919.8

1 介 绍

矢量量化是一种高效率的数据压缩方法, 特别是由于它的编 / 解码的不对称性, 因而适合于图像和视频数据的压缩, 它可以看成一个从 k 维欧氏空间 R^k 到它的一个有限子集的一个映射:

$$Q: R^k \rightarrow C$$
$$C = \{c_i: i = 1, 2, \dots, N\} \quad c_i \in R^k \quad (1)$$

其中 C 称为码书, c_i 称为码字, 在码书中找到最为相似的码字后, 我们只存储或传输最佳匹配码字在码书中的索引值, 从而达到图像压缩的目的. 假定一包含有 N 个码字, 码字维数为 k 的码书, 那么常用的全搜索办法编码一个图像块的计算复杂度为:

$$\begin{aligned} \text{编码一个图像块的复杂度} &= kN \text{ 存储器的访问} + N \text{ 距离计算} + N - 1 \text{ 次比较运算} \\ &+ \log_2 N \text{ 输出操作。} \end{aligned}$$

存储器访问是为了计算输入矢量与 N 个码字之间的距离而必须把每个码字的 k 个元素取出来. $\log_2 N$ 次输出操作是对应的最佳匹配码字的索引值的编码比特数. 而解码端的计算复杂度为:

解码一个图像块的复杂度 = k 次存储器的访问 + $\log_2 N$ 输入操作. 矢量量化的编 / 解码的非对称性, 也就是编码端较复杂而解码端极简单, 使矢量量化特别适合于低功耗应用情况, 比如在手提式电脑或用电池供电的终端, 矢量量化被用于视频编码和解码^[1]. 常用的算法的编码端需要大量的存储器用于存储码本, 而访问存储器就是最耗电的操作. 虽然解码端对存储器的访问次数相对较少, 但仍然需要一个大的存储器用于存储码书. 文献 [1] 中提出了一种基于变换的编 / 解码方案 SCTVQ (small codebook transformation vector quantization). 它的框图如图 1. 它通过定义亮度平移变换使输入图像块的平均亮度和码字的平均亮度更为接近以及定义八种对称变换把基本码书扩大为 8 倍基本码书, 再在这个扩大了码书中寻找最佳匹配, 这样把基于存储器的编码转为更多的是依赖与运算的编码. 既减少了片内存储器的数量, 又减少了对存储器的访问次数, 从而就减小了功耗. 本文针对文献 [1] 中的方法提出了改进, 第二部分叙述原文算法和它的码书中可能的重复. 第三部分给出了改进算法. 第四部分为实验结果. 最后给出了结论.

¹ 1998-11-16 收到, 1999-06-17 定稿

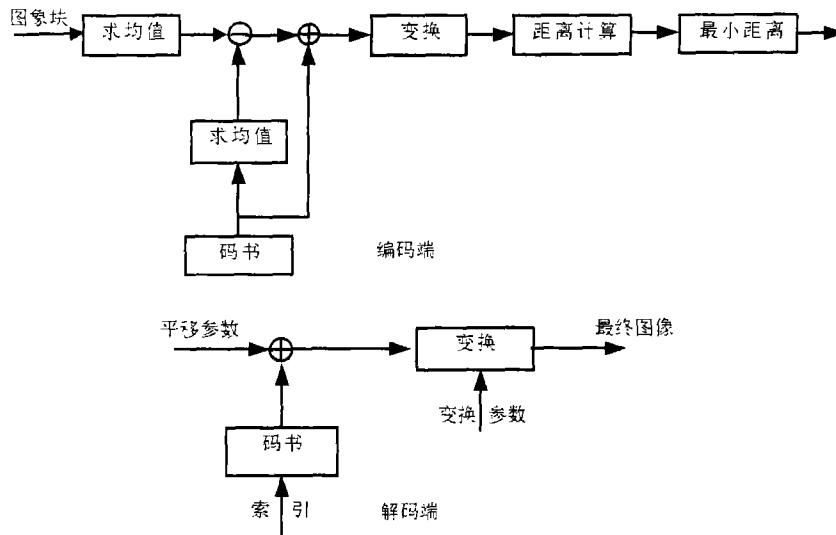


图1 SCTVQ 编 / 解码结构

2 SCTVQ 算法介绍

SCTVQ 算法中定义了亮度平移变换以及八种对称变换, 定义如下:

定义 1 亮度平移变换, 就是把所有基本码书中的码字在变换前先加上一个参数 C .

$C = \frac{1}{B^2} \left[\sum_{i,j=1}^B b_{ij} - \sum_{i,j=1}^B w_{ij} \right]$, 其中 b 是一个 $B \times B$ 点的图像块, w 是一个码字.

定义 2 对称变换定义

- (1) $T_1(w_{ij}) = w_{i,j}$;
- (2) $T_2(w_{ij}) = w_{i,B-1-j}$;
- (3) $T_3(w_{ij}) = w_{B-1-i,j}$;
- (4) $T_4(w_{ij}) = w_{B-1-i,B-1-j}$;
- (5) $T_5(w_{ij}) = w_{j,i}$;
- (6) $T_6(w_{ij}) = w_{B-1-j,i}$;
- (7) $T_7(w_{ij}) = w_{j,B-1-i}$;
- (8) $T_8(w_{ij}) = w_{B-1-j,B-1-i}$;

由于变换过后产生的码书即时产生, 即时使用, 不需要中间存储从而极大的减少了对存储器的访问, 它的编码复杂度为

编码一个子块的复杂度 = kN 次存储器访问 + N 次平移 + $8N$ 次变换
+ $8N$ 次距离计算 + $8N - 1$ 次比较 + $(\log_2 N + 9)$ 次输出操作

解码一个子块的复杂度 = k 次存储器访问 + k 次加法
+ 1次变换 + $(\log_2 N + 9)$ 次输入操作

其中 N 是基本码书的尺寸, 注意这个 N 远小于常规码本的尺寸.

SCTVQ 算法中采用了 LBG 算法来生成基本码书, 除此而外未对码字的选择提出任何要求. 实际上该算法对码字是有限制的, 让我们从二维空间理解一下这个限制, 在基本码书中任

意选一码字 $x_1, x_i, i = 2, 3, \dots, 8$ 是从 x_1 产生的对称码字, 如果码书中的任意其他码字 y_j 等于 x_i , 那么 y_j 是一个重复的码字, 换句话说码书的尺寸由 N 减小为 $N - 1$, 我们可以从图 2 中清楚的看到这一点.

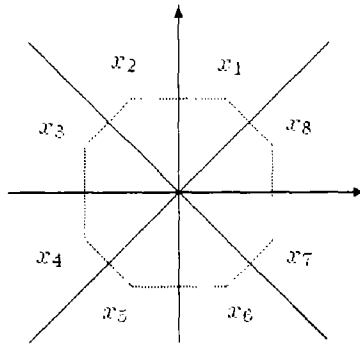


图 2 码字变换示意图

因此必须对基本码书中的码字加以如下限制:

设 C_k, C_l 为基本码书中的两个码字, 则

$$T_i(C_k) \neq C_l, \quad i = 1, 2, \dots, 8 \quad (2)$$

本文的改进算法中将用到这一限制条件.

3 本文算法

假定输入图像块为 V , 码书中的—个码字为 C , 通常的矢量量化算法是在码书中寻找 $\|V - C\|_2^2$ 最小的 C , 而在 SCTVQ 算法中则是寻找 $\|V - T_i[C]\|_2^2$ 最小的 C 以及 T_i , 由于 DCT 变换是正交变换, 因此

$$\|V - T_i[C]\|_2^2 = \|\text{DCT}[V] - \text{DCT}[T_i[C]]\|_2^2 \quad (3)$$

那么在空间域找最小的 C 以及 T_i 完全可以在 DCT 域中找, 结果是一样的. 我们可以很容易的找到 $\text{DCT}[T_i[c]]$ 与 $\text{DCT}[C]$ 之间的关系如下:

设 $\text{DCT}[w] = F(u, v)$, 那么

- (1) $\text{DCT}[T_1[w]] = F(u, v)$;
- (2) $\text{DCT}[T_2[w]] = (-1)^v F(u, v)$;
- (3) $\text{DCT}[T_3[w]] = (-1)^u F(u, v)$;
- (4) $\text{DCT}[T_4[w]] = (-1)^u (-1)^v F(u, v)$;
- (5) $\text{DCT}[T_5[w]] = F(v, u)$;
- (6) $\text{DCT}[T_6[w]] = (-1)^v F(v, u)$;
- (7) $\text{DCT}[T_7[w]] = (-1)^u F(v, u)$;
- (8) $\text{DCT}[T_8[w]] = (-1)^u (-1)^v F(v, u)$; 其中 T_i 的定义如前述. 以下只证明 (2) 式, 其它几式的证明类似.

设 $\text{DCT}[w] = F(u, v)$, $T_2(w_{ij}) = w_{i, B-1-j}$, 那么 $\text{DCT}[T_2[w]] = (-1)^v F(u, v)$.

证明

$$F(u, v) = \frac{2}{B} C(u) C(v) \sum_{i=0}^{B-1} \sum_{j=0}^{B-1} w_{i,j} \cos \frac{2i+1}{2B} u\pi \cos \frac{2j+1}{2B} v\pi$$

$$\text{DCT}[T_2[w]] = \frac{2}{B} C(u) C(v) \sum_{i=0}^{B-1} \sum_{j=0}^{B-1} w_{i,B-1-j} \cos \frac{2i+1}{2B} u\pi \cos \frac{2j+1}{2B} v\pi$$

设 $B - 1 - j = J$, 那么上式可写成

$$\text{DCT}[T_2[w]] = \frac{2}{B} C(u) C(v) \sum_{i=0}^{B-1} \sum_{j=0}^{B-1} w_{i,J} \cos \frac{2i+1}{2B} u\pi \cos \frac{-2J-1+2B}{2B} v\pi$$

$$\text{DCT}[T_2[w]] = (-1)^v F(u, v)$$

证毕

本文算法是基于 (3) 式的, 由于图像像素之间的强烈相关性, 可以只用图像块的 DCT 变换后的部分 DCT 系数代替原来的图像, 对码字也进行相同的处理 (因此我们称这个 DCT 域内的码字为 DCT 码字, 因为存储的是码字的 DCT 系数)。比如只取 3×3 个 DCT 系数代替 4×4 的图像块和码字或 6×6 个系数代替 8×8 的块, 那么计算 $\|\text{DCT}[V] - \text{DCT}[T_i[C]\|_2^2$ 就可大大简化。图 3 画出了本文算法的框图。

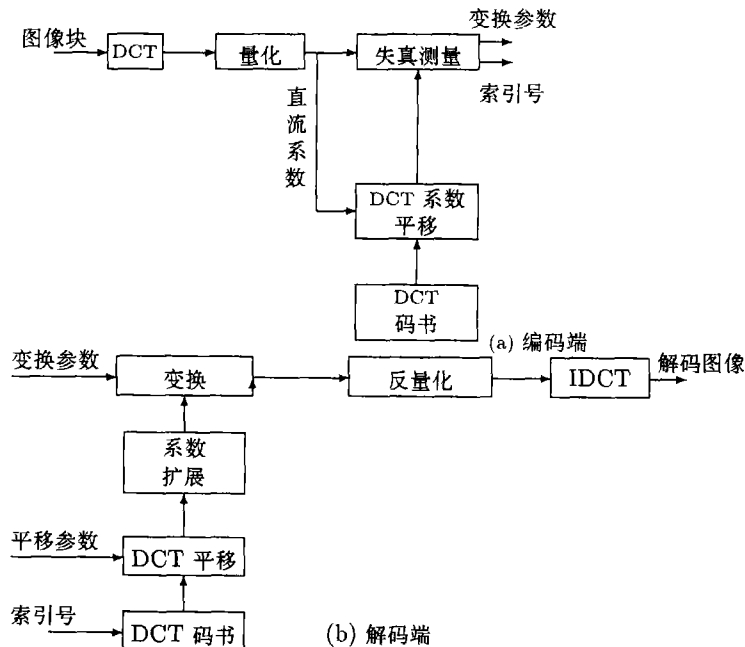


图 3 本文算法编 / 解码结构

根据图 3 可以得出编码, 解码的复杂度为

$$\begin{aligned} \text{编码一个图像块的复杂度} = & 1 \text{次块 DCT 变换} + 1 \text{次块量化} + (k - m)N \text{ 存储器访问} \\ & + N \text{次平移} + 8N \text{ 次变换} + 8N \text{ 次简化的距离计算} \\ & + 8N - 1 \text{比较} + (\log_2 N + 9) \text{ 输出操作} \end{aligned}$$

$$\begin{aligned} \text{解码一个图像块的复杂度} = & (k - m) \text{存储器访问} + 1 \text{次块系数扩展} \\ & + 1 \text{次块 IDCT 变换} + k \text{次加法} \\ & + 1 \text{次变换} + (\log_2 N + 9) \text{ 输入操作} \end{aligned}$$

其中 m 指的是被忽略的系数个数, 容易看出此算法既减少了存储空间及存储器访问次数又减少了计算量。我们可以取成 $k = 16$ (即 4×4 的块), $N = 32$, $m = 7$ (就是只取 4×4 的块的左上 3×3 个系数), 来验证一下。改进算法的计算量如下

$$\begin{aligned} \text{编码一个图像块的复杂度} = & 16 \times 1.25 \text{次乘法} + 1 \text{次块量化} + 9 \times 32 \text{次存储器访问} \\ & + 32 \text{次平移} + 8 \times 32 \text{次变换} + 8 \times 32 \times (9 \text{次乘法} + 8 \text{次加法}) \\ & + 8 \times 32 - 1 \text{次比较} + (5 + 9) \text{次输出操作} \end{aligned}$$

$$\begin{aligned} \text{解码一个图像块的复杂度} = & 9 \text{次存储器访问} + 1 \text{次块系数扩展} \\ & + 16 \times 1.25 \text{次乘法} + 16 \text{次加法} \\ & + 1 \text{次变换} + (5 + 9) \text{次输出操作} \end{aligned}$$

SCTVQ 算法的计算量如下:

$$\begin{aligned} \text{编码一个图像块的复杂度} = & 16 \times 32 \text{次存储器访问} + 32 \text{次平移} + 8 \times 32 \text{次变换} \\ & + 8 \times 32 \times (16 \text{次乘法} + 15 \text{次加法}) + 8 \times 32 - 1 \text{次比较} \\ & + (5 + 9) \text{次输出操作} \end{aligned}$$

$$\begin{aligned} \text{解码一个图像块的复杂度} = & 16 \text{次存储器访问} + 16 \text{次加法} \\ & + 1 \text{次变换} + (5 + 9) \text{次输出操作} \end{aligned}$$

可见 DCT 变换引入的计算量和因为它而节约的计算量相比是微不足道的。考虑到 DCT 变换后的系数每个一般要用 12 比特, 而空域的系数只要 8 比特, 尽管如此 DCT 码书仍然节约了存储空间。

在本文算法中基本码书的生成仍然使用的是 LBG 算法, 不过在每一次迭代后都按 (2) 式进行了检查, 以排除重复的码字。具体产生方法如下:

(1) 用 LBG 算法产生初始码书, 经一次迭代后形成更新过的码书。

(2) 按 (2) 式检查新码书中是否有重复的码字, 如无, 就继续 (1), 如有, 就淘汰其中一个, 并添一新的不重复的码字。继续 (1)。

4 实验结果

在我们的实验中, 采用了两副 512×512 个像素, 256 个灰度等级的测试图像 Lena 和 Boat, 其中 Lena 做为训练图像。码字的维数取了 4×4 和 8×8 两种, 在计算 DCT 距离时只使用了每个 4×4 块的左上 3×3 个 DCT 系数, 码书也一样, 8×8 的块只取了它的左上 6×6 个系数。量化是取了最靠近的整数。实验结果分别如表 1 和表 2。

表1 图像块的大小为 4×4 结果

算法	码书尺寸	PSNR of Lena	PSNR of Boat
SCTVQ	64	32.81	31.27
	32	32.42	30.81
	16	30.20	28.31
本文算法	64	32.94	31.31
	32	32.51	30.86
	16	30.22	28.30

表2 图像块的大小为 8×8 的结果

算法	码书尺寸	PSNR of Lena	PSNR of Boat
SCTVQ	64	27.52	26.22
	32	27.01	25.83
	16	27.00	25.61
本文算法	64	27.61	26.34
	32	27.12	25.95
	16	27.08	25.66

5 结 论

本文提出的算法确实减少了存储空间, 减小了计算量, 降低了功耗。实验结果表明图像质量反而有约 0.1 分贝的提高。

参 考 文 献

- [1] K. Masselos, P. Merakos, T. Stouraitis, C. E. Goutis, A novel algorithm for low-power image and video coding, *IEEE Trans. Circuits and system for video technology*, 1998, CSVT-8(3), 258-263.
- [2] Y. Linde, A. Buzo, R. M. Gray, An algorithm for vector quantizer design, *IEEE Trans. Commun.*, 1980, COM-28(1), 84-95.

A MODIFIED ALGORITHM FOR LOW-POWER IMAGE CODING AND DECODING

Zhao Chengji Pang Chaoyang Zhu Weile

(*University of Electronics Science and Technology of China, Chengdu 610054, China*)

Abstract In this paper, a modified scheme for small memories and low-power image and video coding is presented. It is based on vector quantization. In 1998, K. Masselos uses a small codebook, and by using simple but efficient transformations to the codewords during coding process the small codebook is extended, thus compensated for the quality degradation introduced by the small codebook size. In this way, the power consumption is reduced. Compared to the scheme proposed by K. Masselos, we store the quantized DCT coefficients of codewords in the codebook instead of the codewords itself, and the distortion calculation is performed in the DCT field respectively. Because we use only a part of coefficients to calculate the distances, the distortion calculation is simplified. Furthermore, we modified the way of designing the small codebook; the possible redundancy in K. Masselos's codebook is eliminated.

Key words Vector quantization, Small codebook, DCT transformation

赵乘骥: 男, 1971年生, 博士, 目前从事数字视频技术的研究, 对图像编码, 运动估计较有兴趣。

朱维乐: 男, 1940年生, 教授, 博士生导师。研究兴趣主要有数字视频及 HDTV 系统, 图像序列的三维运动估计和三维结构恢复以及数字视频方向的 VLSI 设计等。