

# 引入 Foveation 视觉模型的 MPEG 视频码流码率转换

代 锋 张勇东 林守勋

(中国科学院计算技术研究所,北京 100080)

**摘 要** 视频码流转换是支持在异质网络环境中进行视频通信的关键技术之一。提出了一种充分利用人类视觉系统特性的 MPEG 视频码流码率转换方法。与原有方法相比,该方法在转换过程中巧妙地使用 Foveation 视觉模型,有效地提高转换效率,产生视觉质量更好的低码率码流,并减少了计算量。

**关键词** 视频转换 Foveation 视觉模型 码率控制

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2005)12-1567-08

## Bit-rate Transcoding of MPEG Video Streams with Foveation Visual Model

DAI Feng, ZHANG Yong-dong, LIN Shou-xun

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

**Abstract** Video transcoding is a key technology to support video communications over heterogeneous networks. In this paper, we propose a new bit-rate transcoding method for MPEG video streams, which can utilize human visual system characteristics adequately. Compared with the conventional method, this method uses skillfully foveation visual model to improve transcoding efficiency, and it can produce new lower bit-streams with better visual quality while achieving computation reduction.

**Keywords** video transcoding, Foveation visual model, rate control

## 1 引 言

随着视频压缩技术和网络技术的发展,各种网络多媒体服务,如多点视频会议、视频点播(VOD)、数字电视等不断出现。为了在异质网络环境中支持各种服务,视频服务器必须具有对已压缩视频码流进行实时转换的功能,从而能够适应客户端与传输信道的异质性<sup>[1]</sup>。码流转换包括码率转换、(空间和时间)分辨率转换、语法转换等。本文主要研究针对 MPEG 视频码流的码率转换,即把已有的 MPEG 视频码流根据传输信道的实际带宽限制转换

为与之相适应的更低码率的码流。

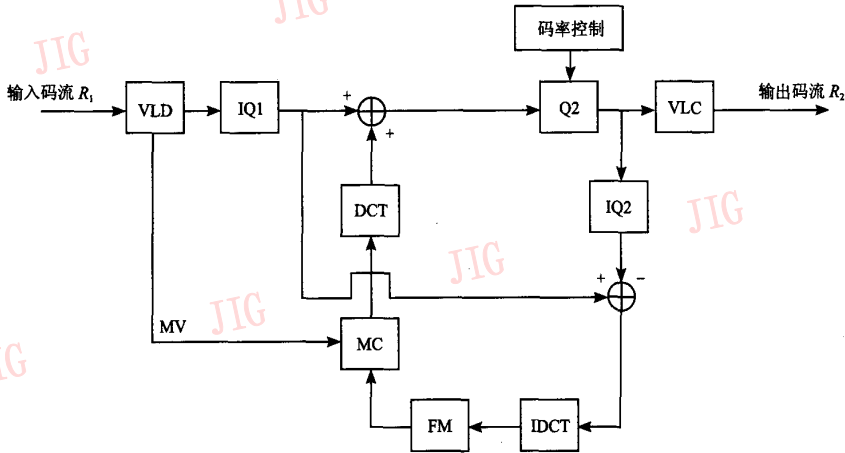
实现码率转换最简单的方式是将输入的视频码流完全解码,然后再压缩成码率更低的码流,这种转换方式可以称为像素域级联转换。这种转换方式具有很大的灵活性,但是计算复杂度很高,如果完全依靠软件来实现,远远不能满足实时应用的需要<sup>[2]</sup>。为此,人们进一步提出了快速像素域级联转换(FCPDT, fast cascaded pixel-domain transcoding)方式<sup>[3]</sup>。如图1所示,该方式通过再使用原有码流中的运动矢量而不重新进行运动估计,使计算复杂度大大降低。对 FCPDT 进一步简化就可以得到另一种转换方式——DDT(DCT 域转换)<sup>[4]</sup>。在 DDT 方式中,运动补偿直

基金项目:国家自然科学基金项目(60302028)

收稿日期:2004-11-26;改回日期:2005-03-21

第一作者简介:代锋(1979~),男,中国科学院计算技术研究所数字化技术研究室博士研究生。研究方向为视频编码与处理。E-mail:

fdai@ict.ac.cn



VLD—变字长解码 VLC—变字长编码 Q—量化 IQ—反量化 FM—帧存储  
DCT—离散余弦变换 IDCT—反离散余弦变换 MV—运动矢量 MC—运动补偿

图 1 像素域快速级联转换结构示意图

Fig. 1 Illustrate of fast pixel-domain transcoding architecture

接在 DCT 域上进行,从而避免了 DCT/IDCT 运算过程,进一步减小了计算量。然而,DDT 方式的灵活性受到限制,当要求改变运动矢量、改变(空间和时间)分辨率等时,用 DDT 很难实现,不易进行扩展。

目前已有的码率转换方法没有很好地利用人类视觉系统(HVS, human visual system)特性,导致所转换成的低码率码流不能很好地与 HVS 特性一致,主观质量较差,转换效率低。当前在视频编码研究领域,随着各种压缩编码技术的日趋完善和成熟,压缩效率已达到较高水平,很难有进一步的提高。为了进一步提高压缩效率满足越来越高的实际需求,世界各国学者都在积极寻找提高压缩效率的新方法和技术,其中受到极大关注的是结合视觉模型的编码方法,充分利用和挖掘 HVS 特性来提高压缩效率,产生主观质量更好的压缩码流<sup>[5]</sup>。如果在码率转换时,引入视觉模型,在带宽有限的条件下根据视觉模型进行码率的合理分配,那么就可以产生主观质量更好的低码率码流,从而有效地提高转换效率。

基于以上考虑,引入 Foveation 视觉模型,提出了一种新的 MPEG 视频流码率转换方法。

## 2 Foveation 视觉模型理论

人视觉系统有这样的生理特点:感光细胞在视网膜上非均匀分布,从而导致人眼对于视觉信息的采样是非均匀的。一般情况下,人眼观看一幅图像时有一

个注视点,可称为 Foveation 点,在该点处人眼具有最高敏感度。以该点为中心,向周围延伸人眼敏感度快速下降。根据生理学试验,可以把 HVS 的对比敏感度描述为视网膜离心率和空域频率的函数<sup>[6,7]</sup>:

$$CT(f, e) = CT_0 \exp\left(\alpha f \frac{e + e_2}{e_2}\right) \quad (1)$$

其中,  $f$  表示空间频率(周期/度),  $e$  是视网膜离心率(度),  $CT_0$  是对比度阈值,  $\alpha$  是空间频率衰减因子,  $e_2$  是清晰度为一半时的离心率。由对比敏感度函数可以得到 HVS 的截止频率函数:

$$f_c(e) = \frac{1}{\alpha} \frac{\ln\left(\frac{CT_c}{CT_0}\right)}{1 + \frac{1}{e_2}} \quad CT_0 \leq CT_c \leq 1.0 \quad (2)$$

其中,  $CT_c = CT(f, e)$ 。这就是所谓的 Foveation 视觉模型。依据 Foveation 视觉模型,在视频图像编码时,对人眼敏感度较高的区域分配较高的码率,给予较高质量编码;对人眼敏感度较低的区域分配较低的码率,给予较低质量编码。从而有效地去除视觉冗余,提高压缩效率。

为了在视频图像编码中有效地结合 Foveation 视觉模型理论,文献[7]给出了在离散空间域上近似 Foveation 视觉模型:给定 Foveation 点,对于图像中的任意一个像素点  $(x, y)$ ,它的截止频率(人眼的最大可感知频率)  $f_c(x, y)$  由下式确定:

$$f_c(x, y) = \min\left\{\frac{i}{8}; d \geq B[i, V], 1 \leq i \leq 8, i \in \mathbf{Z}^+\right\} \quad (3)$$

$$d = (x - x_f)^2 + (y - y_f)^2 \quad (4)$$

$$B[i, V] = \min\{r^2 : [f_c(r, V) \times 8] = i, r \in \mathbf{Z}^+\} \quad (5)$$

$$f_c(r, V) = \frac{1}{1 + k \arctan\left(\frac{r - R}{V}\right)} \quad (6)$$

其中,  $(x_f, y_f)$  代表图像中 Foveation 点坐标,  $V$  代表视点到图像的距离, 模型参数  $k = 13.75$ ,  $R$  代表以 Foveation 点为中心的圆形区域的半径, 对该区域给予最高敏感度 (即  $f_c = 1.0$ ) 的编码。在图像中频率高于截止频率  $f_c(x, y)$  的信息不能被人眼感知。把

一帧图像分为 8 个区域, 每个区域中截止频率相同, 不同的区域截止频率不同, 截止频率取值范围是  $\frac{i}{8} (1 \leq i \leq 8, i \in \mathbf{Z}^+)$ 。

根据 Foveation 视觉模型, 可以设计 Foveation 滤波器对图像进行 Foveation 化<sup>[7]</sup>, Foveation 化后的图像与原始图像在视觉质量上应该是一样的。图 2 显示了 news 原始图像与 Foveation 化后的图像, 其中, Foveation 点在女播音员的面部区域上, 视点到图像的距离  $V = 4$ 。



(a) news 图像



(b) Foveation 化的 news 图像

图 2 图像 Foveation 化的例子

Fig.2 Original and foveated images of news

引入 Foveation 视觉模型的视频图像编码与基于感兴趣区域 (ROI, region-of-interest) 编码有许多相似之处, 它采用同 ROI 相同的思想, 只是采用的生理学模型更符合 HVS, 根据观看的距离、离心率, 把整个图像分为无限多层, 每层的感兴趣程度不同。从这个层面上讲, ROI 是引入 Foveation 视觉模型编码的一个特例。同 ROI 相比, Foveation 视觉模型编码更符合人类视觉, ROI 仅仅把视频数据分为两层, 一层是感兴趣区域; 另一层是非感兴趣区域。Foveation 视觉模型根据 HVS 特性, 把视频数据分为更多层。

### 3 引入 Foveation 视觉模型的 MPEG 视频码流码率转换方法

图 3 为引入 Foveation 视觉模型的 MPEG 视频

码流码率转换方法的结构示意图。鉴于 FCPDT 转换方式具有计算量较小, 结构灵活, 便于扩展的优点, 本方法基于 FCPDT 转换方式的结构, 并依据 Foveation 视觉模型进行了相应的改进。

#### 3.1 部分解码

对输入的 MPEG 视频码流进行变字长解码 (VLC), 之后根据码流中的量化因子信息进行反量化 (IQ1), 得到每个  $8 \times 8$  块的 DCT 系数。

#### 3.2 DCT 系数截断

依据 Foveation 视觉模型, 在  $8 \times 8$  块的 DCT 系数中, 高于截止频率的系数不能被人主观视觉感知, 如果将其去除, 不会影响主观视觉质量, 可以有效地提高转换效率。DCT 系数截断模块就是为实现这一目的而加入的。

可以近似认为一个  $8 \times 8$  块具有唯一截止频率, 一般取  $8 \times 8$  块的中心点为代表, 由它的坐标计算该

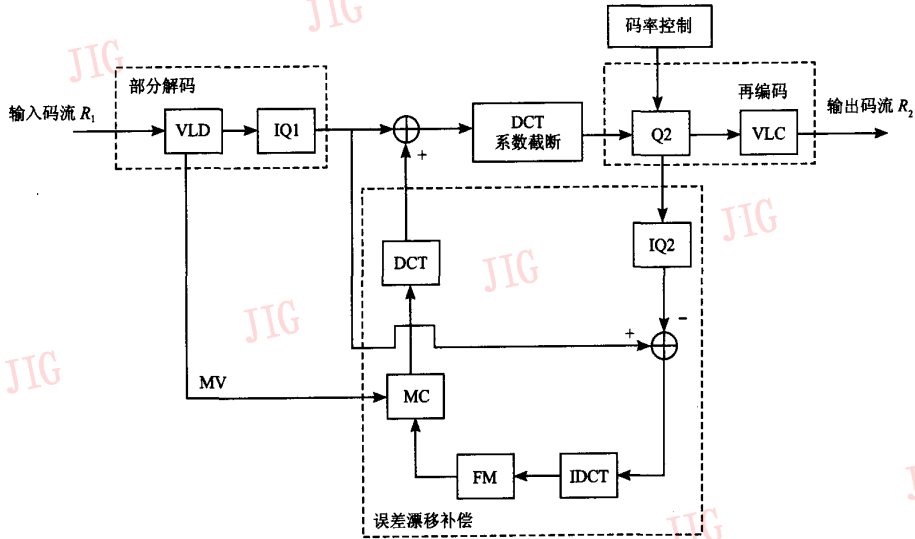


图 3 所提出的转换方法结构示意图

Fig. 3 Illustrate of proposed transcoding architecture

块的截止频率  $f_c$ 。一个  $8 \times 8$  块的 DCT 系数可分成 8 个频带, 构成多分辨率表示, 如图 4 所示。

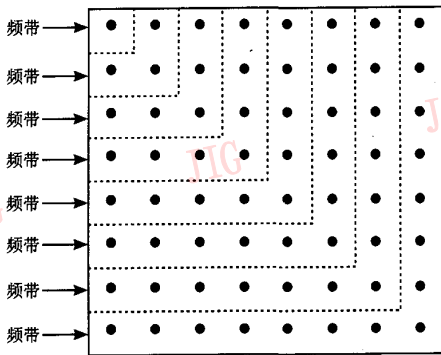


图 4  $8 \times 8$  块 DCT 系数的频带划分

Fig. 4 Division  $8 \times 8$  DCT block into frequency band

对于任意一频带  $m$ , 它的频率  $f(m)$  为  $\frac{m}{8}$  ( $1 \leq m \leq 8, m \in \mathbf{Z}^+$ )。这样基于 Foveation 视觉模型的 DCT 系数截断方法可以如下表述: 给定 Foveation 视觉点, 对于一个  $8 \times 8$  块, 其截止频率为  $f_c$ , 它的一个 DCT 系数为  $F(u, v)$ , 该系数属于频带  $m$ , 那么

$$F(u, v) = \begin{cases} F(u, v) & f(m) \leq f_c \\ 0 & f(m) > f_c \end{cases} \quad (7)$$

### 3.3 码率控制

要把原有 MPEG 视频流的码率降低, 就要运用

码率控制模块重新确定各宏块的量化因子, 根据量化因子对 DCT 系数重新量化。本文依据 Foveation 视觉模型对原有的 MPEG 的 TM5 码率控制方法<sup>[8]</sup>进行改进, 构成新的基于 Foveation 视觉模型的码率控制方法, 其主要步骤如下:

(1) 图像帧级目标编码比特数分配

具体方法与 TM5 方法相同<sup>[8]</sup>。

(2) 宏块级目标编码比特数分配

假设一帧图像的编码比特数为  $R$ , 在此图像中共有  $M$  个宏块, 每个宏块中有  $N$  个  $8 \times 8$  块。原有的 TM5 方法对每个宏块平均分配目标编码比特数, 即对于任一个宏块  $k$ , 它被分配的目标编码比特数为  $r^{(k)} = \frac{R}{M}$ 。而依据 Foveation 视觉模型, 人眼对图像的不同区域具有不同的敏感度, 对于敏感度越高的区域 (其截止频率也就越高), 编码质量应该越高, 也就是说应该对其分配越多的目标编码比特数。基于这样的思想, 根据截止频率的大小比例来分配各宏块的目标编码比特数 (宏块内的截止频率越高, 被分配的目标编码比特数应越多), 即

$$r^{(k)} = \frac{\sum_{j=0}^N (f_c^{(k)}(j))^2}{\sum_{i=0}^{M \times N} (f_c(i))^2} R \quad (8)$$

其中,  $\sum_{j=0}^N (f_c^{(k)}(j))^2$  表示宏块  $k$  内的  $N$  个  $8 \times 8$  块的

截止频率的平方和,  $\sum_{i=0}^{M \times N} (f_c(i))^2$  为图像帧内所有  $8 \times 8$  块的截止频率的平方和。

### (3) 码率控制

根据虚拟缓冲区 (VBV) 的满度, 确定各宏块的参考量化因子  $Q_i$ 。此处采用的方法与 TM5 相同。

### (4) 自适应量化

在 TM5 方法中, 根据宏块的空间活动性来自适应确定它的最终量化因子, 而宏块的空间活动性是该宏块内所有  $8 \times 8$  亮度块空间活动性的最小值, 其中  $8 \times 8$  块空间活动性是由块内的信息变化率  $A$  来确定, 即

$$A = \frac{1}{64} \sum_{i=0}^{64} (p_i - p_{\text{mean}})^2 \quad (9)$$

其中,  $p_{\text{mean}} = \frac{1}{64} \sum_{i=0}^{64} p_i^2$ ,  $p_i$  表示块内第  $i$  个像素的亮度值。在压缩域上这样的信息无法得到, 为此提出了  $8 \times 8$  块在 DCT 域上计算空间活动性  $A_{\text{DCT}}$  的方法:

$$A_{\text{DCT}} = \frac{1}{N} \sum_{i=0}^N |F_i|^2 \quad (10)$$

其中,  $N$  表示该块内低于截止频率的所有交流系数的个数,  $F_i$  表示这  $N$  个系数中的一个值。

根据宏块内所有  $8 \times 8$  亮度块空间活动性的最小值, 确定该宏块的空间活动性 (经规范化后)  $NV_i$ , 那么该宏块的最终量化因子为

$$mq_i = Q_i \times NV_i \quad (11)$$

## 3.4 再编码

根据各宏块的最终量化因子  $mq_i$  对该宏块内的所有 DCT 块的系数进行再量化 (Q2), 之后再继续进行变字长编码 (VLC), 生成满足要求的更低码率的 MPEG 视频。

## 3.5 误差漂移补偿

以上过程即可以实现 MPEG 视频流的码率转换。然而由于对 DCT 系数的再量化 (Q2) 会引起编码端和解码端参考图像的不匹配, 导致误差漂移, 影响转换后所生成码流的图像质量<sup>[3]</sup>。为此需要误差漂移补偿模块来避免误差漂移。

把反量化 (IQ1) 后的 DCT 系数与再量化 (Q2) 后的 DCT 系数的差值进行 IDCT 变换, 得到像素域系数, 送入帧存储器中。然后根据部分解码所得到的运动矢量 (MV) 信息, 在像素域进行运动补偿 (MC), 并将所得的预测值利用 DCT 变换转换成

DCT 系数, 反馈回去与原有预测帧中的残差 DCT 系数相加, 实现误差漂移补偿。

由于要进行 IDCT 与 DCT 运算, 因此与 DDT 转换相比, 本文方法运算量较大。但是根据 Foveation 视觉模型, 对于高于截止频率的 DCT 系数可以不予计算, 据此提出 DCT/IDCT 快速计算方法, 显著减小 DCT/IDCT 计算量。原有的 DCT 及 IDCT 计算公式分别为

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \times \cos\left(\frac{\pi u(2i+1)}{16}\right) \cos\left(\frac{\pi v(2j+1)}{16}\right) \quad (12)$$

$$f(i, j) = \frac{1}{4} \sum_{i=0}^7 \sum_{j=0}^7 C(u) C(v) F(u, v) \times \cos\left(\frac{\pi u(2i+1)}{16}\right) \cos\left(\frac{\pi v(2j+1)}{16}\right) \quad (13)$$

其中,  $u, v$  表示频域坐标,  $i, j$  表示空域坐标。

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0, v = 0 \\ 1 & u \neq 0, v \neq 0 \end{cases}$$

设一个  $8 \times 8$  块的截止频率为  $\frac{t}{8}$  ( $1 \leq t \leq 8, t \in \mathbf{Z}^+$ ) 那么在该块内所有高于截止频率的 DCT 系数不被人眼感知, 可以不予处理, 即赋值为 0。因此在对该块进行 DCT/IDCT 变换时, 只计算低于截止频率的 DCT 系数, 从而 DCT 及 IDCT 计算公式变为

$$F(u, v) = \begin{cases} \frac{1}{4} C(u) C(v) \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \times \cos\left(\frac{\pi u(2i+1)}{16}\right) \cos\left(\frac{\pi v(2j+1)}{16}\right) & u \leq t \text{ 且 } v \leq t \\ 0 & u > t \text{ 或 } v > t \end{cases} \quad (14)$$

$$f(i, j) = \frac{1}{4} \sum_{i=0}^t \sum_{j=0}^t C(u) C(v) F(u, v) \times \cos\left(\frac{\pi u(2i+1)}{16}\right) \cos\left(\frac{\pi v(2j+1)}{16}\right) \quad (15)$$

在整个方法的实现过程, Foveation 点 (视觉关注点) 可以通过以下几种方法获得<sup>[7]</sup>:

- (1) 关注点跟踪器;
- (2) 自动选择;
- (3) 交互模式, 例如通过鼠标点击选择。

通过鼠标交互式选择, 可以由用户自行选择视觉关注点, 实现起来简单可行, 因此这里采用这种方法来确定视觉关注点。

## 4 实验结果

为了评价本文所提出的引入 Foveation 视觉模型的码率转换方法(在此简称为 F-FCPDT),选取两个 MPEG 标准测试序列:Paris 和 Mobile 序列(CIF 格式,帧率 25f/s,长度 300f),用 MPEG-1 编码器把这两个测试序列编码为 2Mbps 的 MPEG-1 码流,用这两个视频流进行实验测试,并把实验结果与 FCPDT 方法相比较。

### 4.1 视觉质量评价指标

在视频图像编码领域,传统的视觉质量评价指标是峰值信噪比(PSNR, peak signal-to-noise ratio),其计算公式如下:

$$\text{PSNR} = 10\log_{10} \frac{255^2}{\frac{1}{N} \sum_{n=1}^N [o(x_n) - c(x_n)]^2} \quad (16)$$

其中, $o(x)$ 和 $c(x)$ 分别代表原始图像和重建图像, $N$ 为图像中的像素数。然而,PSNR 并不能与 HVS 特性完全一致,尤其不能反映人眼对于视觉信息采

样非均匀的特性。为此,在文献[9]中提出了一种新的视觉质量评价指标 FPSNR (foveal peak signal-to-noise ratio),其计算公式如下:

$$\text{FPSNR} = 10\log_{10} \frac{\max[v(x_n)]^2}{\frac{1}{\sum_{n=1}^N f_{cn}^2} \sum_{n=1}^N [v(x_n) - g(x_n)]^2 f_{cn}^2} \quad (17)$$

其中, $v(x)$ 和 $g(x)$ 分别代表原始图像 $o(x)$ 和重建图像 $c(x)$ Foveation 化后的图像, $f_{cn}$ 代表在像素 $x_n$ 处的截止频率。FPSNR 充分考虑到人眼对于视觉信息采样非均匀的特性,用截止频率 $f_c$ 进行加权,能够更好地反映 HVS 特性,因此本文采用 FPSNR 作为视觉质量评价指标。

### 4.2 两种转换方法的视觉质量比较

在实验中,使用 FCPDT 和 F-FCPDT 转换方法(都是由 MPEG-1 编码器经过相应修正而得)把 MPEG-1 测试视频流码率由 2Mbps 分别转换为不同的低码率,在转换过程中 Foveation 点的选定如图 5 所示。



Foveation 点坐标 (100,100)  
(a) Paris 序列



Foveation 点坐标 (250,190)  
(b) Mobile 序列

图 5 转换过程中 Foveation 点的选定位置

Fig. 5 Choosing Foveation point in transcoding

表 1 和表 2 分别给出了 Paris 和 Mobile 序列的平均 FPSNR 比较结果,从中可以看出 F-FCPDT 转换方法较 FCPDT 可以提高 FPSNR 值,尤其对于像 Mobile 那样具有大量运动的图像序列提高程度更加明显(在 2dB 以上)。图 6 显示了对于 Paris 序列将码率由 2Mbps 转换为 0.4Mbps 时,两种方法重建图像比较;图 7 显示了对于 Mobile 序列将码率由

2Mbps 转换为 0.5Mbps 时,两种方法重建图像比较。从图 6 和图 7 中可以看出,对于视觉关注区(图像中圆圈内的区域),F-FCPDT 重建图像较 FCPDT 有更好的主观视觉质量。

以上结果说明 F-FCPDT 方法有效地利用了 Foveation 视觉模型,较原有的 FCPDT 提高了转换效率,能够产生主观质量更好的低码率码流。

表 1 Paris 序列的平均 FPSNR 比较结果

Tab.1 Compare result of Paris FPSNR

单位:dB

转换方法	2Mbps-0.8Mbps	2Mbps-0.6Mbps	2Mbps-0.4Mbps
FCPDT	37.80	36.39	34.09
F-FCPDT	37.95	37.00	35.05

表 2 Mobile 序列的平均 FPSNR 比较结果

Tab.2 Compare result of Mobile FPSNR

单位:dB

转换方法	2Mbps-1Mbps	2Mbps-0.75Mbps	2Mbps-0.5Mbps
FCPDT	31.14	29.79	28.12
F-FCPDT	33.56	32.53	30.55



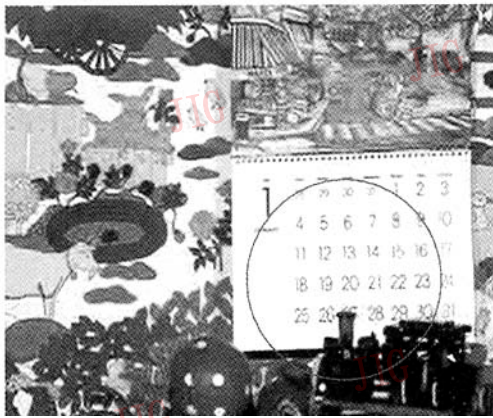
(a) FCPDT



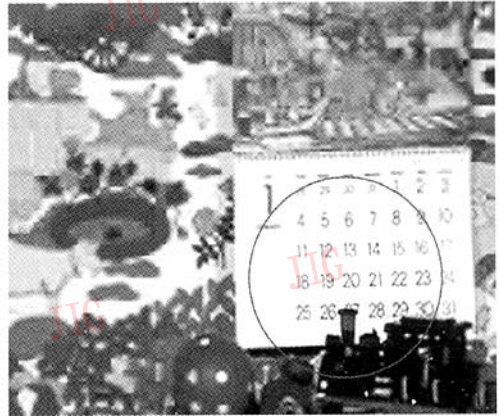
(b) f-FCPDT

图 6 Paris 序列重建图像(第 100 帧)比较

Fig.6 Compare of construed image of Paris(100th frame)



(a) FCPDT



(b) f-FCPDT

图 7 Mobile 序列重建图像(第 100 帧)比较

Fig.7 Compare of construed image of Mobile(100th frame)

### 4.3 两种转换方法的计算复杂度比较

由于引入 Foveation 视觉模型,在 F-FCPDT 方法中,需要计算各  $8 \times 8$  块的截止频率,这是 F-FCPDT 较 FCPDT 主要增加的计算量。但是只有当 Foveation 点位置发生变化时,才需要计算各  $8 \times 8$  块的截止频率;不变化时,不需要计算。一般情况下,

Foveation 点位置在经过许多帧后才发生变化,因此计算 DCT 系数块截止频率的计算量平均起来很小,可以忽略。

DCT/IDCT 运算是转换中计算量较大的部分。但如 3.5 节所述,依据 Foveation 视觉模型可以显著减小 DCT/IDCT 的计算量。把运算一次式(12)称

为 1 次 DCT 运算。在 FCPDT 方法中,每帧图像完成 DCT 运算次数与图像尺寸有关,对于 CIF 格式的图像完成 DCT 运算次数为  $352 \times 288 + 1/2 \times 352 \times 288 = 152\,064$ 。在 F-FCPDT 方法中,每帧图像完成 DCT 运算次数不仅与图像尺寸有关,而且与 Foveation 点的选择有关,对于 CIF 格式的图像,若 Foveation 点选择在图像中央则完成 DCT 运算次数为 23 076,共减少  $152\,064 - 23\,076 = 128\,988$  次 DCT 运算。IDCT 是 DCT 的反运算,F-FCPDT 较 FCPDT 也减少同样数量的 IDCT 运算。基于以上分析可知,与 FCPDT 方法相比,F-FCPDT 明显减少了 DCT/IDCT 运算量。

## 5 结 论

在视频编码研究领域,如何更有效地利用和挖掘 HVS 特性来提高压缩效率是当前各国学者所普遍关注的问题。本文在视频码流码率转换中引入 Foveation 视觉模型,提出了一种新的 MPEG 视频流码率转换方法。该方法基于快速级联像素域转换结构,在转换中巧妙地利用了 Foveation 视觉模型。依据该模型,对原有码流中的 DCT 系数进行截断,去除主观视觉不能感知的 DCT 高频系数,提高转换效率;改进 TM5 码率控制方法,实现有效的基于 HVS 特性的码率分配与控制;进而对 DCT/IDCT 的计算方法进行改进,降低计算量。实验结果表明:与已有转换方法相比,该方法能够有效地提高转换效率,产生主观质量相对更好的低码率码流,并具有较低的

计算量。

下一步,将研究在其他类型的视频码流转换(如空间和时间分辨率转换)中,如何有效地结合 Foveation 视觉模型提高转换效率。

## 参考文献 (References)

- 1 Fung K T, Chan Y L, Siu W C. New architecture for dynamic frame-skipping transcoder [J]. IEEE Transactions on Image Processing, 2002, 11(8):886~900.
- 2 Youn J, Xin J, Sun M T, et al. Video transcoding for multiple clients[A]. In: Proceedings of SPIE: Visual Communication and Image Processing 2000[C]. Sydney, Australia, 2000:76~85.
- 3 Assuncao P A A, Ghanbari M. Transcoding of single-layer MPEG video into lower rates [J]. IEEE Proceedings Vision Image Signal Processing, 1997, 144(6):377~383.
- 4 Assuncao P A A, Ghanbari M. A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit streams [J]. IEEE Transactions on Circuits for Systems and Video Technology, 1998, 8(8):953~967.
- 5 Reid M M, Millar R J, Black N D. Second generation image coding: an overview[J]. ACM Computing Surveys, 1997, 29(1):3~29.
- 6 Geisler W S, Perry J S. A real-time foveated multiresolution system for low-bandwidth video communication [J]. Proceedings of SPIE, 1998, 3299:294~305.
- 7 Lee S. Foveated video compression and visual communications over wireless and wireline networks [D]. Austin, Texas, USA: The University of Texas at Austin, 2000.
- 8 ISO/IEC JTC1/SC29/WG11/93-400, MPEG-2 Video Test Model 5 (TM5)[S].
- 9 Lee S, Pattichis M S, Bovik A C. Foveated video quality assessment [J]. IEEE Transactions on Multimedia, 2002, 4(1):129~132.