

一种改进的可分级视频编码方法及其网络传输研究

张 方 吴成柯 程培星 肖 嵩

(西安电子科技大学综合业务网国家重点实验室 西安 710071)

摘 要: 该文提出了一种改进的可分级视频编码方法。为了适应流媒体的分层传输要求, 该方法通过对 DCT 系数量化残差的位平面编码产生视频流的增强层部分, 其基本层码流由更多的子基本层组成, 各子基本层通过宏块级 DCT 系数重排及 VLC 重组生成。同时, 该文设计了一种针对该分层视频流数据的网络传输自适应不等丢包保护 (AUPLP) 策略, 在估计当前可利用带宽资源的基础上, 实时调整不同层数据的保护力度, 并控制传输截断的层数。仿真结果表明, 与传统方法相比该文方案在低带宽时可获得平均 1.2dB 的编码增益, AUPLP 的应用也大大改善了视频流媒体的传输质量。

关键字: 视频编码方法, 位置重排, MPEG4-FGS, PFGS, AUPLP

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 1009-5896(2005)01-0108-04

Research on an Improved Scalable Video Coding and the Network Transmission

Zhang Fang Wu Cheng-ke Cheng Pei-xing Xiao Song

(National Key Lab of Integrated Service Networks, Xidian Univ., Xi'an 710071, China)

Abstract An improved scalable video coding scheme is proposed in this paper. The method generates the base layer including some sub-base layers by DCT coefficients reordering and VLC reshuffling, and the residues between the original DCT coefficients and the reconstructed DCT coefficients of the base layer are encoded in the enhancement layer with bit-plane coding technology. To transmit the layered streaming video generated by the proposed scheme over the IP network, an Adaptive Unequal Packet Loss Protect (AUPLP) strategy is designed to determine the current available network bandwidth and adjust sending rates according to different situation. Experimental results show that the proposed scheme can improve the average coding efficiency up to 1.2dB compared with conventional methods in low bandwidth, and the AUPLP strategy can further improve the performances of video network transmission system.

Key words Video coding scheme, Data reordering, MPEG4-FGS, Progressive Fine Granularity Scalability (PFGS), Adaptive Unequal Packet Loss Protect (AUPLP)

1 引言

随着 Internet 网络的飞速发展及可用带宽资源的不断增大, 视频流媒体业务也受到越来越多的重视。为了在当前“尽力而为”(Best-effort)的 IP 网络上传输视频流满足业务服务质量 (QoS) 的要求, 具有对时变网络带宽的自适应能力及抗误码性能的视频编码方法及传输控制策略就显得非常重要。最近的研究表明^[1-7], 渐进可分级的视频编码及分层传输控制方案可以有效地解决视频流 IP 网络传输时的大部分问题。从视频编码的角度来看, 渐进可分级使得视频流数据可以适应特殊的网络变化和不同的带宽条件, 在 Internet 上的传输达到尽可能好的业务质量; 从网络结构的角度来看, 分层传输可以使大量不同接入类型、拥有不同可用资源的用户在任何时间接收任何需要的视频流, 即支持单播 (Unicast)

和多播 (Multicast) 业务。

MPEG-4 提出了一种有效的可分级视频编码方法 FGS (Fine Granularity Scalability)^[1,2], 吴枫等人又在其基础上, 通过选择更高质量的增强层码流作为运动估计的参考, 提出了 PFGS (Progressive FGS) 方法^[4-6], 一定程度上提高了系统的编码增益。但遗憾的是 MPEG-4 FGS 及 PFGS 都只能适应小范围的网络带宽变化, 它们都建立在这样一个假设的基础上: 即使在最坏的网络条件下, 视频流的基本层都可以被成功地传输、接收并完全解码。显然, 这在实际应用中是不可能的, 而一旦基本层码流不能被完全接收, 这两种方法就无法重建任何一层的视频数据。针对这一问题, 本文提出了一种改进的可分级视频编码方法, 主要思想是通过 DCT 系数量化残差的位平面编码产生视频流的增强层部分, 基本层码流则由更多的子部分组成, 各子部分通过宏块级 DCT

系数重排和可变长编码 (VLC) 重组生成。同时,在构造并分析 IP 网络信道模型的基础上,本文设计了一种自适应不等重丢包保护策略 (Adaptive Unequal Packet Loss Protect, AUPLP) 来满足可分级视频流用于网络传输时的 QoS 要求。AUPLP 在估计当前可利用网络带宽资源的基础上,实时调整对不同层视频流的纠错保护力度,并控制传输截断的层数以适应网络的拥塞及传输带宽的起伏。试验证明,与传统方法相比,本文在低带宽时可获得平均 1.2dB 编码增益, AUPLP 的应用也大大改善了视频流媒体的传输质量。

2 改进型渐进精细可分级视频编码

现有的视频编码标准已经或多或少地引入了可分级的概念,特别是在 MPEG-4 v4.0 中提出的 FGS,提供了一个有效的 SNR/Rate 可分级编码方案^[1,2]。MPEG-4 FGS 编码器采用相对较大的量化步长来产生基本层视频流,原始 DCT 系数和基本层重建系数之间的差值(残差)被用位平面编码技术来形成更多的增强层比特流^[3]。基于此的一些改进型方案也已经被提出^[4-7],特别是吴枫等人提出的 PFGS 方法可以在较大程度上提高视频流的鲁棒性,在改善编码增益的同时可有效控制误差积累和扩散,从而避免了由于传输误码引起的重建图像错误对后续解码帧的影响。

但是,无论 MPEG-4 FGS 或 PFGS 都只能适应小范围的网络带宽变化。也就是说,这两种方法都是建立在这样一个假设的基础上:即使是最坏的网络条件下,视频流的基本层都可以被成功地传输、接收并完全解码。尽管 PFGS 的基本层码率控制比较灵活,但用于 IP 网络传输的视频流通常都是预先编码后放在视频服务器上的数据,不可能实时调整基本层速率来适应最小可利用带宽,从而就不能满足基本层码流被正确传输并完全解码这一必要条件。而一旦基本层码流不能被完全解码, MPEG-4 FGS 和 PFGS 都无法重建任何一层的视频数据。基于此,本文提出的改进型渐进精细可分级视频编码方案同时采用了 DCT 系数重排、VLC 重整^[8]和 PFGS 3 种方法,将 MPEG-4 的基本层码流更进一步分成多个子层,从而使流媒体能适应更大范围的网络带宽变化。图 1 所示为本文方法的编码流程。

从人眼生理视觉的角度讲, DCT 低频系数对图像质量的影响要大于高频系数对图像质量的影响,即可以认为低频系数的重要性大于高频系数。经过 zigzag 扫描后,系数的重要性沿扫描顺序依次递减,图 2(a)显示了宏块级重要数据信息的分布,颜色越深表明重要性越强。基于此,图 1 中的基本层编码器部分对量化后的 DCT 系数及运动矢量信息进行了位置重排,将所有宏块的宏块头信息、运动矢量、DCT 直流系数排列在一起,形成第一子基本层,经过 zigzag 扫描后的所有宏块的第二个 DCT 系数形成第二子基本层,依次可以形成其他的子基本层,图 2(b)是位置重排后的比特流结构。尽管本文方法 DCT 系数的位置重排是在 VLC 编码之前,但影响变长编码效率的高频系数在重排后具有这样一个规律:即宏块内所有子块相同阶的高频系数依次顺序相连,码流中的连零数据仍然很多,因此 VLC 编码的整体效率并不会降

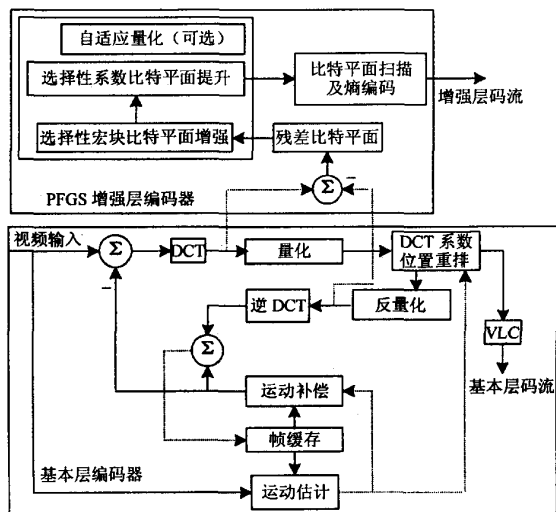


图1 改进型可分级视频编码框图

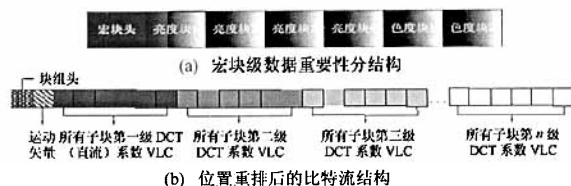


图2

低。本文方法生成的基本层比特流在宏块级也是可分级的,即使基本层的一部分码流因为信道带宽的波动或包丢失不能被准确传输、接收或者解码,接收端仍可以由接收到的部分分子基本层码流来重建获得比较粗糙的图像信息。

本文方法的增强层编码器和 PFGS 基本相同,运动估计选用一些高质量的增强层作为参考 (MPEG-4 FGS 总是选用基本层作为参考),由于增强层重构图像的质量总是要比基本层高,因此就使得运动补偿更有效,从而提高了增强层的编码效率。但同时这也产生了差错扩散,一旦有一帧图像的参考增强层未能完全解码,必然导致后续的所有以此为参考的图像帧增强层无法恢复。为了消除和减少这种差错漂移,规定增强层编码器周期性依据参考帧的基本层对增强层做运动估计/补偿^[4-7]。

3 IP 网络包丢失模型和 AUPLP 策略

在实际应用中,视频流数据通常要在网络中多个服务器、多个用户之间传输,并且每个服务器应能同时响应多个请求,每个客户端也能向不同的服务器提出服务请求。考虑到每个终端用户的带宽不同,而视频流媒体业务对数据的传输延时、包丢失、包失序等要求非常严格,因此需要 QoS 控制模块来实时检测网络拥塞、调整数据流向^[9]。

IP 网络上数据丢包的问题比较复杂,在构造数学模型的基础上可以用两个参数来做定性分析:(1)丢包率 P_L (Packet loss rate); (2) 环回时间 T_{RTT} (Round-trip time)。首先,当 IP 网络路由器或网关检测到网络拥塞时,会丢弃超出当前带宽资源的数据报文,并且 Internet 的不可靠传输 (特别在无

线网络中)也会导致数据的不可靠接收或者丢失;其次,网络拥塞时数据的环回时间会加长。因此,依据 P_L 和 T_{RTT} 的综合分析可以对当前可利用的 IP 网络带宽资源做出较准确的估计,本文采用两状态 Markov 链的 Gilbert-Elliot 模型来分析 IP 网络的丢包率 P_L 。假定任何时刻网络都处于好(G)或坏(B)两状态之一。在 G 状态,信道误码率(P_G)较低,而在 B 状态,信道误码率(P_B)较高,状态之间的转换概率可以计算如下:

$$P_{BG} = n_{BG} / n_B, \quad P_{GB} = n_{GB} / n_G$$

其中 n_B 表示测试时间序列内处于状态 B 的次数; n_G 表示测试时间序列内处于状态 G 的次数; n_{BG} 表示测试时间序列内状态变换组合 B → G 出现的次数; n_{GB} 表示测试时间序列内状态变换组合 G → B 出现的次数。从而,稳态丢包率可定义如下:

$$P_L = \frac{P_{GB}}{P_{BG} + P_{GB}} = \frac{n_{GB}}{n_B n_{GB} + n_G n_{BG}} \quad (1)$$

对 IP 网络数据报文环回时间 T_{RTT} 的估计可以采用如下公式:

$$T_{RTT} = \alpha \times T_{RTT}^{\text{now}} + (1 - \alpha) \times (T_{\text{now}} - T_{\text{send}} - \Delta T) \quad (2)$$

其中 T_{RTT}^{now} 表示当前分组环回时间, T_{now} 为发送方收到应答报文的时刻, T_{send} 为数据被发送的时刻, ΔT 是接收方对报文分组的处理时间, α 是加权系数,通常设定为 0.875。

综合丢包率 P_L 和分组环回时间 T_{RTT} 的作用,可以估算当前可得到的带宽资源为

$$W_{\text{cur}} = C / (T_{RTT} + \sqrt{P_L}) \quad (3)$$

这里的 C 为常数,通常依据接收方采取的应答方式取为 1.22 或者 1.31^[9]。

假设式(3)中 3 个物理量的变化分别为 ΔW_{cur} , ΔT_{RTT} 和 ΔP_L , 则如下方程组成立:

$$\left. \begin{aligned} \Delta W_{\text{cur}}^{\text{RTT}} &= \frac{\partial W_{\text{cur}}}{\partial T_{RTT}} \Delta T_{RTT} = -\frac{1}{(T_{RTT} + \sqrt{P_L})^2} \Delta T_{RTT} \\ \Delta W_{\text{cur}}^{P_L} &= \frac{\partial W_{\text{cur}}}{\partial P_L} \Delta P_L = -\frac{1}{2(T_{RTT} + \sqrt{P_L})^2 \sqrt{P_L}} \Delta P_L \\ \Delta W_{\text{cur}} &= \Delta W_{\text{cur}}^{\text{RTT}} + \Delta W_{\text{cur}}^{P_L} \end{aligned} \right\} \quad (4)$$

其中 $\Delta W_{\text{cur}}^{\text{RTT}}$ 表示由于 ΔT_{RTT} 引起的带宽变化量, $\Delta W_{\text{cur}}^{P_L}$ 表示由于 ΔP_L 引起的带宽变化量。从而当前网络状态可以判断如下:

(1) 当 W_{cur} 下降时,如果 $|\Delta W_{\text{cur}}^{\text{RTT}}| \geq |\Delta W_{\text{cur}}^{P_L}|$, 网络处于拥塞状态;如果 $|\Delta W_{\text{cur}}^{\text{RTT}}| < |\Delta W_{\text{cur}}^{P_L}|$, 则网络处于不可靠传输状态。

(2) 当 W_{cur} 上升时,如果 $|\Delta W_{\text{cur}}^{\text{RTT}}| \geq |\Delta W_{\text{cur}}^{P_L}|$, 与不可靠传输因素相比较而言,网络的拥塞状况有所缓解;如果 $|\Delta W_{\text{cur}}^{\text{RTT}}| < |\Delta W_{\text{cur}}^{P_L}|$, 与上一次可利用带宽的估算相比,网络的不可靠传输(比如误码率)有所改善。

综合前面的分析,本文通过一种新颖的 AUPLP 策略来实现可分级视频流在 IP 网络的 QoS 传输控制。假定 m 为压缩数据分级的层数,接收端参与解码的层数越多,恢复的图像质量就越高。各质量层(可以是子基本层,也可以是增强

层)重要性沿着子基本层 1 → 子基本层 2 → ... 增加层 1 → 增加层 2 ... 的方向递减,表示为

$$I(Q_0) > I(Q_1) > I(Q_2) > \dots > I(Q_{m-1}) \quad (5)$$

视频流数据第 i 质量层长度定义为 $L(i)$, 取 IP 网络带宽变化范围为 $W \in [W_{\min}, W_{\max}]$, 即最坏情况下网络可利用带宽为 $W_{\min}$, 最好情况下网络可利用带宽为 $W_{\max}$ 。在 AUPLP 工作的初始条件中,预设一个标准带宽 W_{std} , 当利用这一资源传送可分级视频码流分组数据时,接收端可以得到满足正常要求的解码重建图像,其传输层数为 m_{std} , 则有 $W_{\text{std}} \in [W_{\min}, W_{\max}]$ 。

AUPLP 根据式(3)估算出当前可利用网络带宽 W_{cur} , 在 $W_{\text{cur}} > W_{\text{std}}$ 时,以一个传输层为单位通过多传或少传来适应网络带宽的时变、起伏;在 $W_{\text{cur}} \leq W_{\text{std}}$ 时,根据对网络处于拥塞状态还是不可靠传输状态的判断,通过前向纠错编码(Forward Error Coding, FEC)提高重要层数据保护力度及增加、减少传输层等几种方法自适应组合的处理,在适应网络带宽变化的同时,尽可能提高数据传输的抗误码及鲁棒性能。为了定性分析当前可利用网络带宽的变化量,将 $[W_{\min}, W_{\max}]$ 分成一个个的小区间,在 $(W_{\text{std}}, W_{\max}]$ 内可表示为

$$S_H(i) = \left[W_{\text{std}} + \sum_{k=0}^i L(m_{\text{std}} + k), W_{\text{std}} + \sum_{k=0}^{i+1} L(m_{\text{std}} + k) \right] \quad (6)$$

其中 $i = 0, 1, \dots, S_{\text{num}} - 1$, S_{num} 为区间数目。在 $[W_{\min}, W_{\text{std}})$ 内划分 m_{std} 个区间,表示为

$$S_L(j) = \left[W_{\text{std}} - \sum_{k=0}^{m_{\text{std}}-j} L(m_{\text{std}} - k), W_{\text{std}} - \sum_{k=0}^{m_{\text{std}}-j+1} L(m_{\text{std}} - k) \right] \quad (7)$$

这里 $j = 0, 1, \dots, m_{\text{std}} - 1$ 。定义 t 为表示时间段的自变量, $f(t)$ 为第 t 个时间段传输的层数, $F(t)$ 为第 t 个时间段传输层中受 FEC 纠错保护的层数,显然 $f(t) \geq F(t)$ 成立。具体算法描述如下:

(1) 根据式(1), (2), (3)分别计算 $P_L(t)$, $T_{RTT}(t)$ 和 $W_{\text{cur}}(t)$;

(2) 如果 $W_{\text{cur}}(t) \geq W_{\text{std}}$ 且 $W_{\text{cur}}(t) \in S_H(i)$, 取 $f(t) = m_{\text{std}} + i$ 且 $F(t) = 0$;

(3) 如果 $W_{\text{cur}}(t) < W_{\text{std}}$ 且 $W_{\text{cur}}(t)$ 与 $W_{\text{cur}}(t-1)$ 不在同一区间, 当 $|\Delta W_{\text{cur}}^{\text{RTT}}(t)| > |\Delta W_{\text{cur}}^{P_L}(t)|$ 时, 取 $f(t) = f(t-1) - 1$ 且 $F(t) = F(t-1)$; 当 $|\Delta W_{\text{cur}}^{\text{RTT}}(t)| \leq |\Delta W_{\text{cur}}^{P_L}(t)|$ 时, 取 $f(t) = f(t-1) - 1$ 且 $F(t) = F(t-1) + 1$, 受 FEC 保护层数的增加是沿着式(5)表示的传输层重要性顺序的方向;

(4) 如果 $W_{\text{cur}}(t) < W_{\text{std}}$ 且 $W_{\text{cur}}(t)$ 与 $W_{\text{cur}}(t-1)$ 不在同一区间, 当 $|\Delta W_{\text{cur}}^{\text{RTT}}(t)| > |\Delta W_{\text{cur}}^{P_L}(t)|$ 时, 取 $f(t) = f(t-1) + 1$ 且 $F(t) = F(t-1)$; 当 $|\Delta W_{\text{cur}}^{\text{RTT}}(t)| \leq |\Delta W_{\text{cur}}^{P_L}(t)|$ 时, 取 $f(t) = f(t-1) + 1$ 且 $F(t) = F(t-1) - 1$, 受 FEC 保护层数的减少是沿着式(5)表示的传输层重要性顺序的反方向;

(5) 回到(1)开始下一个时间段的传输控制。

据此,任意 t 时刻的传送数据分组表示如下:

$$\left\{ \underbrace{L(0) + L_{\text{FEC}}(0)}_{Q_0}, \underbrace{L(1) + L_{\text{FEC}}(1)}_{Q_1}, \dots, \underbrace{L(F(t)-1) + L_{\text{FEC}}(F(t)-1)}_{Q_{F(t)-1}}, \dots, \underbrace{L(f(t)-2)}_{Q_{f(t)-2}}, \underbrace{L(f(t)-1)}_{Q_{f(t)-1}} \right\}$$

其中 $L_{\text{FEC}}(i)$ 表示码长为 $L(i)$ 的传输层经过 FEC 纠错编码后增加的数据冗余，则总的码流长度（对应于传输所需的网络带宽资源）为 $L(t) = \sum_{k=0}^{f(t)-1} L(k) + \sum_{k=0}^{F(t)-1} L_{\text{FEC}}(k)$ 。

4 实验结果与分析

通过软件仿真，我们将本文提出的改进型渐进精细可分级视频编码方案与 MPEG-4 FGS, PFGS 做了比较。测试序列采用 CIF 格式 Foreman 及 Suzie，帧编码速率为 15Hz，从第一帧开始，每隔 15 帧作一次强制 I-VOP 编码，其余均为 P-VOP 编码，基本层采用 TM5 流控算法，网络信道的仿真采用本文第 3 节描述的 GE 模型。

图 3(a)所示为在不同比特率时 3 种视频编码方案的平均 PSNR 曲线，显然在比特率较高时，本文方案与 PFGS 性能相当，但优于 MPEG4-FGS；而随着比特率的降低，本文方法 PSNR 的下降明显比 PFGS 和 MPEG4-FGS 慢，也就是说在低带宽资源时本文方法的重建图像质量要明显优于传统方法。图中横坐标 512kbps 附近时三组仿真结果相似，这是因为在如此高比特率情况下编码效率已接近饱和。为了验证本文提出的 AUPLP 的性能，我们虚拟了一个具有时变可利用带宽资源的仿真信道，并假定包丢失率 P_L 是导致式(3)中网络状态变量 W_{cur} 变换的主要因素。从图 3(b)所示仿真结果来看，本文 AUPLP 策略能够有效地提高整个视频编码传输系统的性能，可分级视频流发送速率的逐层调整使得解码出的视频帧的质量变化能更加平稳，重要层数据的不等重保护也很大程度上改善了视频流的抗误码性能。特别是本文提出的改进型可分级视频编码方案结合 AUPLP 策略后在低带宽的网络条件下可获得最好的编码、传输性能。图 4 所示为 P_L 变化导致可利用网络带宽约为 40kbps 时 Suzie 序列的一组实际图像，这里 AUPLP 中 FEC 采用 (2,1,6) 卷积码，增信删余后编码速率为 3/4。

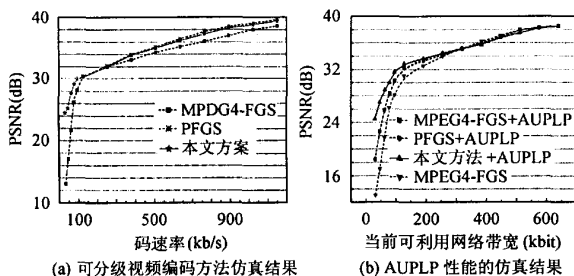


图 3 不同方法的仿真实验结果

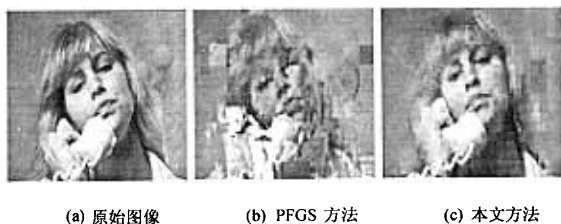


图 4 不同方法的实验结果

5 结论

本文提出了一种改进型渐进精细可分级视频编码方法。为了适应流媒体的分层传输要求和网络视频业务的 QoS 控制，通过对 DCT 系数量化残差的位平面编码产生视频流的增强层部分，其基本层码流由更多的子基本层组成，各子基本层通过宏块级 DCT 系数重排和 VLC 重组生成。同时，在构造并分析网络传输信道模型的基础上，本文设计了一种自适应不等重丢包保护策略 (AUPLP) 来满足 QoS 的要求。AUPLP 在估计当前可利用网络带宽资源的基础上，将当前网络分为不可靠传输及数据拥塞两种状态，并据此分别实时调整对不同层视频流的保护力度，控制传输截断的层数。仿真证明，与传统方法相比，本文方案在低带宽时可获得平均 1.2dB 的编码增益，而且提出的 AUPLP 策略的应用也大大改善了视频流媒体的传输质量。

参考文献

- [1] MPEG video group. FGS VM v3.0. ISO/IEC JTC1/SC29/WG11. N3097, Maui, USA, Dec. 1999.
- [2] Radha H M, Van der Schaar M, Chen Yingwei. The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP. *IEEE Trans. on Multimedia*, 2001, 3(1): 53 – 68.
- [3] Li W. Bitplane coding of DCT coefficients for fine granularity scalability. ISO/IEC JTC1/SC29/WG11, MPEG98/ M3989, Oct. 1998.
- [4] Wu Feng, Li Shipeng, Zhang Y Q. A framework for efficient progressive fine granularity scalable video coding. *IEEE Trans. on CASVT*, 2001, 11(3): 332 – 344.
- [5] Wu Feng, Li Shipeng, Yan Rong, Sun Xiaoyan, Zhang Yaqin. Efficient and universal scalable video coding. *ICIP 2002, Rochester, September 2002*, 2: 37 – 40.
- [6] Sun X, Wu F, Li S, Gao W, Zhang Y Q. Macroblock-based progressive fine granularity scalable video coding. *IEEE International Conference on Multimedia and Expo (ICME)*, Tokyo, Aug. 2001: 461 – 464.
- [7] Peng W S, Chen Y K. Mode-adaptive fine granularity scalability. *ICIP 2001, Greece, Oct. 2001*: 993 – 996.
- [8] Tin Tak Cheung, So M S, Cheng R S K, Leatief K B. Adaptive unequal error protection and VLC reshuffling of image transmission over wireless channels. *Vehicular Technology Conference Proceedings, Tokyo, May 15-18, 2000*, 2: 800 – 804.
- [9] Zhang Qian, Zhu Wenwu, Zhang Y Q. Resource allocation for multimedia streaming over the internet. *IEEE Trans. on Multimedia*, 2001, 3(3): 339 – 355.

张 方：男，1977 年生，博士，主要从事图像压缩编码、网络视频流业务及无线视频等方面的研究工作。

吴成柯：男，1938 年生，教授，博士生导师，主要从事图像处理、图像通信、计算机视觉等方面的研究工作。

程培星：男，1979 年生，硕士生，主要从事图像压缩编码及无线视频等方面的研究。

肖 嵩：女，1977 年生，博士，主要从事图像压缩编码、联合信源信道编码等方面的研究工作。