

采用可分级视频编码和基于 MPEG-21 的视频适配

杨元挺^{1,2}

(1. 福州大学物理与信息工程学院, 福建 福州 350108; 2. 福建信息职业技术学院电子工程系, 福建 福州 350003)

摘要: 开发了一个基于可分级视频编码和 MPEG-21 视频适配的视频流系统, 其主要特征是支持无线通信. 该系统能使无线设备在不破坏当前通信连接的同时在不同无线网络中漫游. 采用的视频流编码方法可以自适应调整由于垂直切换带来的无线连接特性的巨大变化. 提出了一种基于概念证明的方法——分布式数字项适配(DIA), 通过服务器和代理商之间的合作使系统适配更多复杂的情况. 实例说明, 该设计系统成功地实现了执行垂直切换而并不断开网络连接的视频应用.

关键词: 视频适配; MPEG-21 数字项适配; 可分级视频编码; 垂直切换

中图分类号: TN 914

文献标识码: A

文章编号: 0438-0479(2010)06-0806-06

随着无线通信的快速发展和视频编码技术的提高, 无线视频流将会得到广泛应用. 不同的无线网络各自折衷共存形成一个异构无线网, 也称无线覆盖网络^[1]. 应用于不同类型的无线网络间的交换(垂直交换)往往要求保持全局连通性. 现有的大多无线视频流系统设计只针对某个特定的无线网, 不能满足全局通信服务. 通过无线覆盖网络传送视频流是一项具有挑战性的任务. 它不仅要考虑无线网络的弱点, 还要考虑无线覆盖网络中频繁发生的垂直切换. 从应用层来说, 这一任务要求视频编码高度适应不同的情形, 故本文提出了视频适配技术.

视频适配巧妙地视频信号移除冗余信息从而减少端对端资源的需求. 视频适配有 3 种基本实现方法: 1) 为每一种视频序列存储大量的不可分级比特流. 将每个数据流按不同的格式或空间的/时间的/信噪比(SNR)的分辨率进行编码. 该方法简单易行已被广泛应用, 但占用大量存储空间. 2) 实现视频转码, 其内容包括降低空间分辨率和 SNR. 虽然转码灵活且节约存储空间, 但因操作过程复杂, 故不被广泛应用. 3) 使用可分级视频编码(SVC)^[2], 其固有特性使它能适应不同场合. 可分级视频编码为系统提供了简单灵活的适应能力. 适配可以认为是寻找有效资源内容表达的过程. 在异构无线网络和动态无线网络中, 为了最大程度

提高用户的效用, 视频适配必须是内容感知和上下文感知的.

在无线视频流服务中, 传送路径在到达移动终端用户之前很可能要穿越不同网络中的多个结点. 为了提供端到端的服务质量, 需要在传送路径上连续执行视频适配. MPEG-21 已成为一个进行透明传输和充分利用多媒体资源的授权技术, 并通过各种网络和设备为不同的用户服务^[3].

MPEG-21 视频适配和可分级编码是我们设计系统的基础. 本文首先介绍 MPEG-21 数字项适配(DIA)的概念. 然后描述我们开发的可分级视频编解码器(SVC codec), 它是建立在运动补偿时域滤波(MCTF)和 JPEG2000 的基础上. 本文开发的 SVC codec 能够在空间/时间/SNR 领域用高效编码提供完整的适应性, 其技术完全可以和 H. 264 最新发展水平的编码标准相比. 后面提到的垂直切换管理系统能使视频流广泛应用于无线覆盖网络中. 文中还概括了现代 SVC codec 和垂直切换管理系统与 MPEG-21 的 DIA 相整合的方法. 最后提出了一个概念证明的解决方法——分布式 DIA, 它通过服务器和代理商之间的相互合作, 能适应更加复杂的适配环境.

1 MPEG-21 DIA

DIA 的基本思想是定义描述工具, 对媒体比特流的句法结构和内部或外部的使用环境约束进行描述. 这些描述工具不仅仅是接口语言, 还是数字项的一部

分. 对于每一个带有基于适配的位流语法描述的中间节点而言, 只需实现一个 MPEG-21 DIA 的补充适配模块.

为获得最佳 QoS, 媒体资源适配须满足终端设备和网络带来的约束, 而适配 QoS 也指出选择最优参数设置的问题^[4]. 如图 1 示, UED 描述用户特征、终端容量、网络 and 自然环境特征^[5]. UCD 为约束规范建立数学抽象模型. 当达到最佳 QoS 及 ADTE 输出参数被用于转换 Gbsd^[4] 时, ADTE 最优化匹配参数能满足通用环境约束. 注意, ADTE 的实现不是规范的, 且当 ADTE 投入使用时, DIA 并不限制其传输结构.

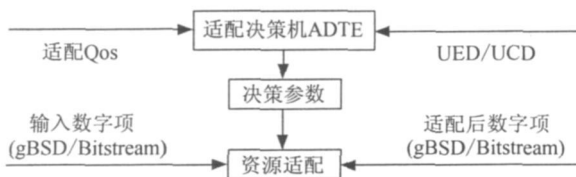


图 1 MPEG-21 DIA 概念

Fig. 1 Digital item adaptation in MPEG-21

MPEG-21 DIA 框架中的视频适应是通过一个比较静态的方式实现的. 根据细节描述的程度, gBSD 可能会很大. 标准 XSLT 处理器要求在转换输入 XML 之前将它完全加载到存储器中^[6]. 这意味着基于 XSLT 的适配不会发生, 除非整个 gBSD 是可用的. 整个 gBSD 文件在会话传送期间, 可能会导致相当数量的启动延迟. 即使 gBSD 已经预先存储了, 改编而成的代理人可能只具备有限的处理能力和记忆能力来处理 gBSD 的实时性. 文献[6]提出 2 种解决方案: 1) 使用一种新的流处理语言 STX, 对基于 SAX 的 XML 事件语言活动进行处理, 并取代 XSLT; 2) 将 gBSD 分割为叫做处理单元 (PU) 的小成分, 它们能被标准 XSLT 独立辨认和处理. 2) 已被批准为 MPEG-21 数字项适配的修正部分.

2 SVC

SVC 视频编码提供了一个有效的、可扩展的方式来表达视频内容, 这有利于进一步提高 MPEG-21 的视频适应性. 运动补偿时域滤波^[7-8]是一种广泛应用的时间分解办法, 因其能有效地检测复用帧冗余, 传统的帧到帧 (MPEG) 或复用帧预测方法 (H. 264) 则很难实现. 最新的图像编码标准, 如 JPEG2000^[9], 提供了一个高度可扩展 (包括成分、质量、空间和位置可扩展性)

和高效的 (根据失真比率 $R-D$ 性能) 图像编码. 它是基于最优切断 (EBCOT) 的嵌入式分块编码^[10]. 在本文中, 为了实现无线视频流的流畅 (简单) 匹配, 我们在整合 MCTF 与 JPEG2000 的基础上设计了一个简单的可分级视频编码方案. 通过探索 MCTF 和 JPEG2000 标准的独特功能, 拟议的编解码器不仅提高了编码的 $R-D$ 性能, 而且还具有较高的可扩展性.

我们开发的编码器的结构如图 2 所示. 整个编码过程包括 3 个组成部分: MCTF、空间编码和最优位截断. 第一步迭代采用 5/3 小波 MCTF 对初始帧 F_k 的每一种颜色组合 (YUV) 进行滤波. 通过 MCTF, 得到了运动矢量和许多 T-bands. 接着用 JPEG2000 将 T-bands 编码成多种质量层, 每一种都有一个 $R-D$ 值. 消除不可行的截断点, 实现最优截断点从而达到了给定的目标码率. 最后形成的视频比特流包含由 MCTF 产生的 MV 信息和每一个 T-band 的 JPEG2000 比特流. 下面详细地介绍编码的组成部分.

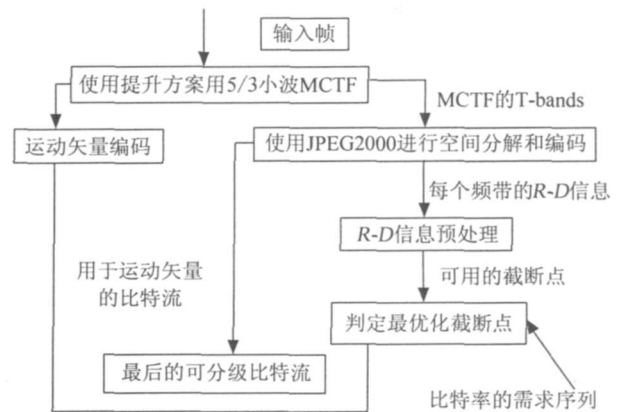


图 2 编码结构图

Fig. 2 Structure of coding

2.1 MCTF

MCTF 通过预测和更新两步提升方案来实施. 在预测中, 利用原始输入帧生成高通带宽 H_k ; 在更新中, 利用原始偶数帧和已获得的高通带宽来生成低通带宽 L_k . 提升步骤可以用数学公式表示为:

$$\begin{aligned}
 H_k[m, n] &= F_{2k+1}[m, n] - 0.5(W_{2k, 2k+1} \\
 &\quad (F_{2k}[m, n]) + W_{2k+2, 2k+1}(F_{2k+2}[m, n])), \\
 L_k[m, n] &= F_{2k}[m, n] + 0.25(W_{2k-1, 2k} \\
 &\quad (H_{k-1}[m, n]) + W_{2k+1, 2k}(H_k[m, n])), \quad (1)
 \end{aligned}$$

其中 $W_{2k, 2k+1}$ 是前一帧的运动补偿映射在后一帧的坐标系的表示. 预测中, 实现了基于块的双边运动评估. 更新中, 利用高通系数 (从每个像素中预测到的权重因

子)来更新每一个像素,这些像素用来在预测和更新之间进行匹配。

2.2 二维分解和空间编码

应用 JPEG2000 标准对 MCTF 生成的 T-band 进行压缩^[9]. 在特殊情况下,首先使用二维离散小波变换(DWT)将每个 T-band 分解成几个时空子带. 然后,将生成的小波系数分成小的码组,每一个码组是位平面到位平面的独立编码. 每个码组比特流将被截断成几个层,每一层都会作用于某一失真等级. 将这些层进行分组,则 JPEG2000 流可以分成若干个质量层. 在实验中,我们为最低时空子带配置了大约 30 个质量层,且有约 20 个质量层用于其它高频带. 这是为了在不同的比特率情况下都能简化最优比特分配,从而为最后的视频流提供有效的质量分级. 后面的“时空子带”都简称为“T-band”。

2.3 最优位截断

最优位截断问题可以总结为:给定一个目标比特率 R ,如何截断每个 T-band JPEG2000 流以达到最优从而使整体失真最小. 每个 T-band JPEG2000 流包含一定数量的质量层,其中相应的比率和失真度已知. 这里, JPEG2000 流的质量层为可用截断点. 我们可以得到最优位截断公式为

$$\min_{L_{B,C}} D_{\text{total}} = \sum_C \sum_B W_C \cdot G_B \cdot D_{L_{B,C},B,C}, \quad (2)$$

$$R_{\text{total}} = \sum_C \sum_B R_{L_{B,C},B,C} < R_{\text{target}}, \quad (3)$$

其中 $L_{B,C}$ 表示在颜色 C 中 T-band B 的质量层 L , $R_{L_{B,C},B,C}$ 和失真 $D_{L_{B,C},B,C}$ 是分别为质量层 $L_{B,C}$ 的比特率和失真度, G_a 是 T-band B 的能量增益; 当要为 YUV 三颜色指定不同的优先级时, W_c 设(默认 $W_c = 1$)为每一种颜色的权重. G_a 的值由滤波系数和 T-band 所在的 MCTF 等级决定.

类似于 JPEG2000^[10] 用到的最优截断, 这个约束极小化问题可以通过引入一个拉格朗日乘数 λ 转变为一个无约束问题, 如下所示:

$$\min_{L_{B,C}} J = D_{\text{total}} + \lambda (R_{\text{total}} - R_{\text{target}}). \quad (4)$$

不是每个 T-band 的所有质量层都是可行的截断点, 因为 JPEG2000 流的失真率曲线并不是严格凸起的. 必须移除不可行的截断点, 这样其余的截断点才会形成一个凸曲线. 要得到给定比特率 R_{target} 值下的最优 λ 值, 我们采用二进制搜索算法, 如图 3 所示.

T-band B 截断点 h 的斜率计算公式为

$$\text{Slope}_B^h = W_C \cdot G_B \cdot \frac{D_B^{h-1} - D_B^h}{r_B^h - r_B^{h-1}}. \quad (5)$$

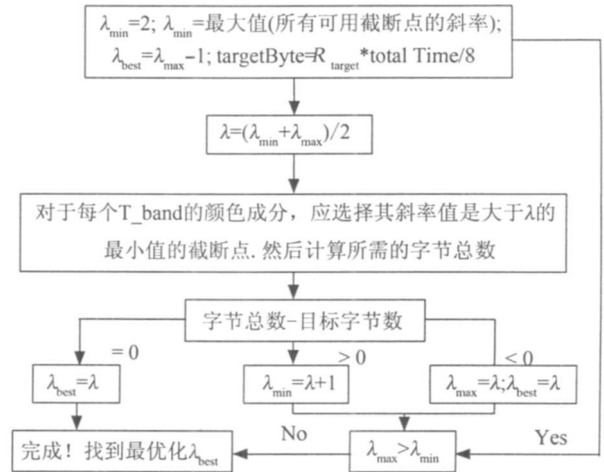


图3 寻找最优截断点的流程图

Fig. 3 Flow chart of seeking optimal truncation point

2.4 编码实现

对拟定的编解码器进行 3 个 CIF 序列测试: akiyo, foreman 和 mobile & calendar. 图 4 就 R - D 性能和 H. 264/AVC1 做了比较. 从拟定的编解码器得到的峰值信噪比 (PSNR), 结果仅从一个单一的比特流 (比如编码器仅执行一次) 计算出来, 而对于 H. 264, 必须为每一个所需的比特率执行整个编码器. 由图 4 可知, 对于低运动序列 (Akiyo) 或高结构序列 (mobile & calendar), 拟定的编解码器的性能非常接近 H. 264. 然而, 对于更高的运动序列 (foreman), 其性能比 H. 264 差. 其原因: 1) H. 264 在每个预测运动方向上允许多个参考帧, 而我们的编解码器的 MCTF 在每个方向仅用了一个参考帧; 2) MCTF 中采用的基于块的运动预测导致 T-bands 产生了很多块效应, 尤其在时域高频带中. 这些块效应会产生纵横交错的纹路, 导致 JPEG2000 性能下降. 因此, 在 JPEG2000 编码之前, 我们通过为 MCTF 引入多个参考值^[8]并在 T-bands 中应用 H. 264/AVC 去块效应滤波器^[11], 提出了可分级视频编解码器, 对于所有类型的视频序列, 它具有和 H. 264 相比拟的功能.

3 垂直切换管理

为了使无处不在的视频流穿过无线覆盖网络, 理想的方法是开发一个智能系统, 在不同的无线网络间垂直切换. 垂直切换管理系统应该能够自动管理移动设备上的多个网络接口, 当移动节点需要切换到其它接口时, 管理系统能够判断出在当前的情况下使用哪个接口最好.

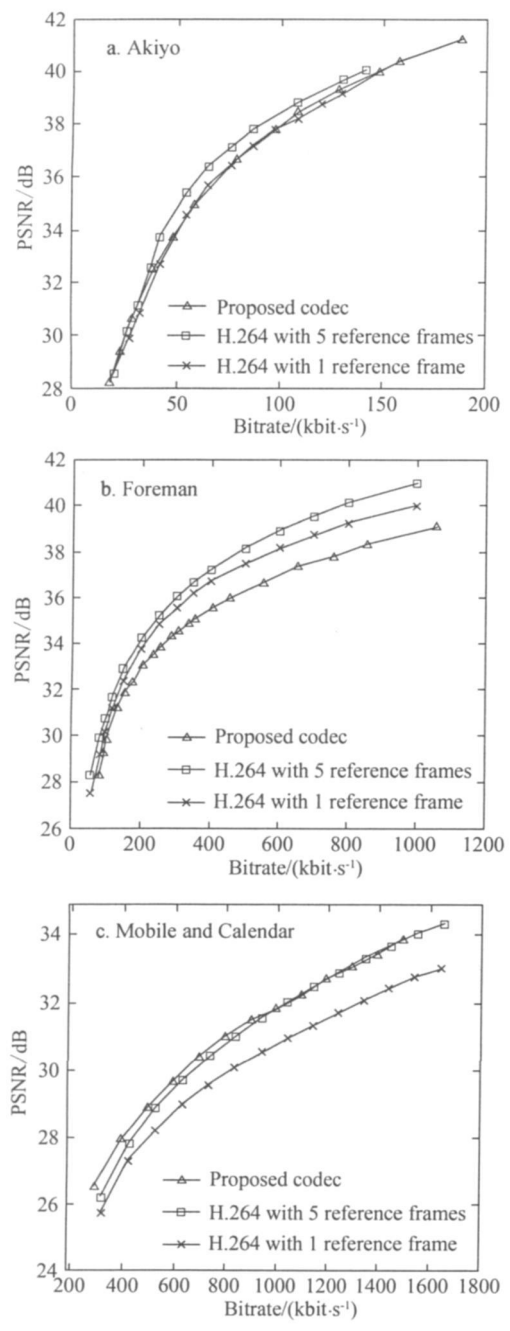


图 4 $R-D$ 参数值
Fig. 4 Parameter value of $R-D$

传统的网络协议(IP)不支持垂直切换, 因为当移动设备连接不同的无线网络时, 通信会被打断. 移动 IP^[12]解决了传统的 IPv4 网络的移动连接问题, 而未来的 IPv6 网络的移动连接将由移动 IPv6 协议来解决. 我们的研究是建立在移动 IPv6 基础上, 使用 MPL 技术实现.

我们的垂直切换管理系统(基于 MIPLv1.0, Linux 内核 2.4.22)见图 5. 网络接口配置数据库存储

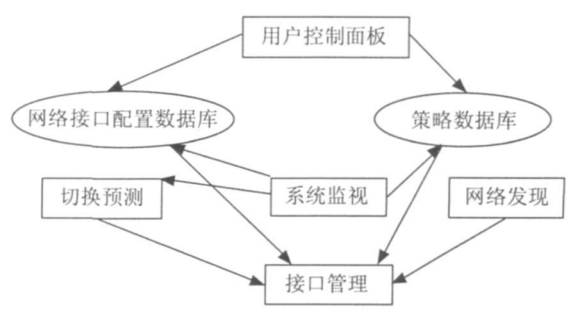


图 5 给定的垂直切换在 Linux 环境下系统的结构
Fig. 5 Structure of given vertical handoff on Linux

了不同类型的网络接口的性能配置参数. 这些参数(功率消耗(P), 可用带宽(B), 最大容量(M), 信号强度(S), 数据包丢失率(L)和使用费用(C))描述了相应的接口类型的特点. 在这个设计系统里, 接口选择基于一个成本函数的策略模式. 在一定时间里的网络使用成本(非使用费用) 用公式表示为

$$f(n)=\omega_b\times\frac{1}{B_n}+\omega_p\times P_n+\omega_c\times C_n. \tag{6}$$

此计算成本用于选择网络接口. 参数 B_n 、 P_n 和 C_n 都取之于网络接口配置数据库. 系数 ω_b 、 ω_p 和 ω_c 是对应参数的权重. 这些权重系数被存储在策略数据库中.

图 5 中, 用户控制面板模型是一个简单的用户空间工具, 可以指定存储在网络接口配置数据库和策略数据库中的各种参数. 系统监控模块测量系统各方面的性能, 包括电力监控, 流通监控和可用带宽监控 3 个子模块. 分别监控移动节点的电池电量, 运行中应用程序使用的交通数据速率和网络的可用带宽. 网络发现模块检测移动节点周围的可用网络, 间歇性地打开闲置的网络接口设备, 寻找无线网络接入点并请求路由器通知. 切换预测模块预测可能与当前网络失去连接的情况的发生, 并通知接口管理模块提前进行切换. 接口管理模块对接口选择和切换操作来说是一个中心模块.

4 基于 SVC 的 MPEGk-21

本节介绍基于 SVC 系统的 MPEG-21 模型, 它可用于无线覆盖网络. 首先, 分析系统架构, 它融合了可分级视频编码模块, 垂直切换管理模块和 MPEG-21 DIA. 然后示范一个案例来证实该系统的功能.

4.1 系统架构

服务器架构见图 6. 基于垂直切换管理模型的移动 IPv6 被部署在网络层. 该模块通过与它在客户端对

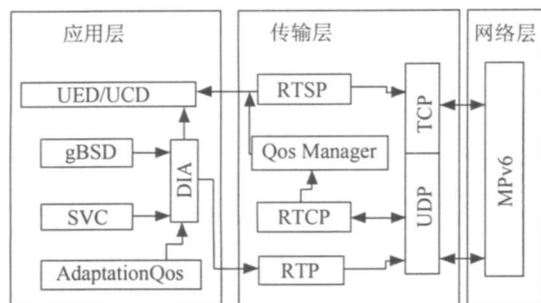


图6 先进的可分级视频流系统的服务器架构

Fig. 6 Server structure of advanced scalability video streaming system

应的模块合作来自动管理移动设备上的多个网络接口,并依当前情形在网络接口选择上做出决定,以便提供无所不在的通讯.传输平台基于 RTP、RTCP 和 UDP 协议. RTSP/TCP^[13] 协议用在配置会话、播放控制和动态环境控制上交流信息.服务质量管理负责感知和分析网络变化.在应用层,一个 SVC 编码视频流连同其 gBSD 形成一个最终被移动用户使用的数据单元.我们开发的 SVC 编码视频依空间、时间、SNR 的组合省力灵活地剪辑,而相应的限制、可行的适应操作和相关功能要预先计算和保存在适应 QoS 中.基于适应 QoS 和 UED, MPEG-21 DIA 模块掌握着对应的适

```

<DIA>
  <Description xsi:type=" UsageEnvironment Type ">
    <UsageEnvironment xsi:type=" Networks Type ">
      <Network>
        <NetworkCharacteristic xsi:type=" NetworkCondition Type ">
          <AvailableBandwidth average=" 58000 "/>
        </NetworkCharacteristic>
      </Network>
    </UsageEnvironment>
  </Description>
</DIA>

```

图7 由 UED 描述的垂直切换之后的网络条件

Fig. 7 Network qualification described by UED

应进程.随着 SVC 和垂直切换管理模块集成进入 MPEG-21 DIA,我们的系统能够为无所不在的通讯提供精确的媒体访问.

4.2 垂直切换的使用实例

下面描述一个从 WLAN 网络垂直切换到蓝牙网络的实例.假设在切换到蓝牙之后的平均带宽是 58 kbit/s,移动用户终端设备在新的带宽条件下将产生一个 UED 并发送到服务器.图 7 为 UED 的一个片段.当服务器接收到新的 UED, MPEG-21 DIA 将开启对应进程.依据网络和终端制约因素, ADTE 决定空间、时间和 SNR 层的优化组合.对于移动设备,一个额外的限制因素就是屏幕尺寸.这种终端限制在 UCD 中的表示为:框架宽度 $\leq$ 显示宽度且框架高度 $\leq$ 显示高度. ADTE 通过在适应 QoS 中开发离散空间来获得最优化的解决方案,在每个空间、时间、SNR 下的适应操作的尺寸都被取样并与视频比特率和相关质量连接在一起,其依据是在 UED 和 UCD 中规定的隐含和明确的限制条件.

从 WLAN 网络转换到蓝牙网络时数据率急剧下

降,大量的 SNR 层被删减以满足新的带宽条件.根据引起突然变化的网络条件的垂直切换情况,系统会自动适应新的带宽条件且对移动用户透明.

5 结 论

本文阐述了无线覆盖网络中基于 SVC 系统的 MPEG-21 的开发经验,详细说明开发的 SVC 编解码器和垂直切换管理系统及它们与 MPEG-21 DIA 结合的方法.通过实例证实了我们设计的系统成功地实现了执行垂直切换而并不断开网络连接的这一功能.根据引发突变的网络条件的垂直切换情况,该系统能自适应新的视频带宽条件且对移动用户透明.

参考文献:

- [1] Stemm M, Katz R H. Vertical handoff in wireless overlay networks[J]. Mobile Networks and Applications, 1998, 3 (4): 335-350.
- [2] Ohm J R. Advances in scalable video coding[J]. Proceedings of the IEEE, 2005, 93: 42-56.

- [3] Bormans J, Gelissen J, Perks A. MPEG-21: the 21st century multimedia framework [J]. IEEE Signal Processing Magazine, 2003, 20(2): 53-62.
- [4] Mukherjee D, Delfosse E, Kim Jae Gon, et al. Optimal adaptation decision taking for terminal and network quality of service [J]. IEEE Transactions on Multimedia, 2005, 7(3): 454-462.
- [5] Burnett I, Pereira F, Van de Walle R, et al. The MPEG-21 book [M]. USA: Wiley, 2006.
- [6] Devillers S, Timmerer C, Heuer J, et al. Bitstream syntax description based adaptation in streaming and constrained environments [J]. IEEE Transactions on Multimedia, 2005, 7(3): 463-470.
- [7] Secker A, Taubman D. Motion compensated highly scalable video compression using an adaptive 3D wavelet transform based on lifting [J]. Proc IEEE Int Conf Image Processing, 2001, 2: 1029-1032.
- [8] Schafer R, Marpe D, Schierl T, et al. MCTF and scalability extension of H.264/AVC and its application to video transmission, storage, and surveillance [J]. Visual Communications and Image Processing, 2005, 5960: 343-354.
- [9] Taubman D S, Marcellin M W. JPEG2000: standard for interactive imaging [J]. Proceedings of the IEEE, 2002, 90: 1336-1357.
- [10] Taubman D. High performance scalable image compression with EBCOT [J]. Proc IEEE Int Conf Image Processing, 2000, 9: 1158-1170.
- [11] Lainema J, Bjontegaard G, List P, et al. Adaptive deblocking filter [J]. IEEE Trans Circuits Syst Video Technol, 2003, 13(7): 614-619.
- [12] Perkins C E. RFC 3344: IP mobility support for IPv4 [EB/OL]. [2003-08]. <http://www.ietf.org/rfc/rfc3344.txt>.
- [13] Wu D, Hou Y T, Zhu W W, et al. On end-to-end architecture for transporting MPEG-4 video over the Internet [J]. IEEE Trans on Circuits and Systems for Video Technology, 2000, 10: 923-941.

MPEG-21 Based Adaptation for Scalable Video Streaming

YANG Yuan ting^{1,2}

(1. College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China;

2. Department of Electronic Engineering, Fujian Polytechnic of Information Technology, Fuzhou 350003, China)

Abstract: In this article, we describe a ubiquitous MPEG-21 based scalable video streaming system. Its main feature is to support seamless communication. The system enables wireless devices to roam among different wireless networks without breaking the ongoing connections. The method of video coding can adaptively adjust to match the significant change of wireless link characteristics due to vertical handoff. A proof of concept solution distributed DIA is proposed, for more complex adaptation scenarios through cooperations among the server and proxies. The system designed in the paper is illustrated successfully to achieved the implementation of the vertical handoff without disconnecting the ongoing connections.

Key words: video adaptation; MPEG-21 digital item adaptation; scalable video coding; vertical handoff