

基于感知分析的立体视频帧重要性模型

章联军, 李少博, 王晓东, 王腾飞

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 由于丢失不同帧对立体视频质量的终端感知影响不同, 因此提出一个面向内容感知的帧重要性区分模型. 根据立体视频序列的时空相关性及视频帧之间错误扩散的特性, 新模型从帧丢失引起的当前帧的错误隐藏失真和帧丢失引起后续帧的错误扩散失真两方面来计算帧重要性. 实验结果表明: 新模型能够在编码端准确地估算各帧对终端感知的重要性, 对不同运动剧烈程度和视差大小的立体视频序列都具有适用性.

关键词: 立体视频; 帧重要性; 内容感知

中图分类号: TP391

文献标志码: A

文章编号: 1001-5132 (2015) 02-0024-04

随着多媒体技术和网络通信技术的迅猛发展, 立体视频相关技术成为目前广泛研究的热点^[1-2]. 为了提高传输效率和终端感知质量, 需要区分立体视频帧的重要性, 建立一个基于内容感知的立体视频帧重要性模型^[3-4].

现有的帧重要区分方法可分为 2 种. 一种是基于解码依赖关系的帧重要性区分方法, 在 1 个 GOP 中, 通常认为 I 帧最重要, P 帧次之, B 帧最不重要^[5]. Yang 等人^[6-8]通过考虑错误扩散因素, 建立 ELEP 模型来度量视频帧的重要性, 模型认为错误发生得越早, 错误扩散积累越严重, 对后续帧的影响越大. 但 ELEP 模型并没有考虑视频内容特性的影响, 不能体现运动剧烈序列和运动缓慢序列帧的重要性区别. 另外一种帧重要性区分方法是基于语法结构的视频内容进行区分. 目前的 H.264/AVC 将视频数据按片编码后分为 3 种不同等级的重要性^[9], 可扩展编码将视频数据分为基本层和增强层, 增强层必须依赖基本层数据解码^[10]. 周茵等人^[11]根据视频码流中基本层和增强层对视频帧所起的作用不同, 认为基本层的重要性高于增强层.

上述 2 种帧重要性区分方法都没有考虑视频内容特性对重要性的影响, 不能准确反映视频帧对终端感知的重要性. 笔者在详细考虑视频内容

特点、时空相关性和视差信息的基础上, 结合帧拷贝的错误隐藏算法建立了一个面向感知的帧重要性区分模型.

1 帧重要性度量方法

视频帧中不同视频内容对终端的感知贡献并不相同. 笔者采用的立体视频编码结构为 IPPP, 第 1 帧为 I 帧, 其他帧为 P 帧, 且文中设定 I 帧为最重要帧, 不做重要性区分. 立体视频 s 视点 t 时刻帧重要测定参数 $I(s, t)$ 可以表示为:

$$I(s, t) = I_c(s, t) + I_p(s, t), \quad (1)$$

式中, $I_c(s, t)$ 是帧丢失给当前帧带来的错误隐藏失真; $I_p(s, t)$ 指帧丢失引起的后续帧的错误扩散失真. 帧重要性度量与具体的错误隐藏方法有关, 文中采用帧拷贝的错误隐藏方法进行数据恢复.

2 立体视频帧重要性区分模型

2.1 右视点帧重要性区分

对于右视点而言, 某帧丢失后直接拷贝左视点同时刻帧, 因此错误隐藏失真可表示为:

$$I_c(s, t) = D(s, t), \quad (2)$$

式中, $D(s, t)$ 指 t 时刻左右视点帧的 MSE.

右视点帧丢失只会对后续右视点帧产生影响,

收稿日期: 2014-07-05.

宁波大学学报(理工版)网址: <http://journalg.nbu.edu.cn/>

基金项目: 国家科技支撑计划 (2012BAH67F01); 浙江省科技厅科研项目 (2012R10009-08); 浙江省教育厅科研项目 (Y201327703); 宁波市科技创新团队研究计划 (2011B81002).

第一作者: 章联军 (1980), 男, 浙江兰溪人, 硕士/实验师, 主要研究方向: 网络通信与传输技术. E-mail: zhanglianjun@nbu.edu.cn

对于后续的右视点帧, 失真只可能由帧间编码宏块产生. 假设后续帧的帧内编码比例为 β , 则某帧丢失对后续第 λ 帧产生的失真为:

$$D(s, t + \lambda) = (1 - \beta)^\lambda \alpha^\lambda D(s, t), \quad (3)$$

式中, α 为错误扩散因子, 物理意义为所有参与预测下一时刻帧的宏块总失真与当前时刻帧失真的比值. α 和序列的运动剧烈程度有关, α 是个大于 0 小于 1 的参数, 根据经验, 可用 $(1 - \beta)^2$ 近似表示. 计算某帧丢失对于后续 λ 帧产生的总失真, 即错误扩散失真可表示为:

$$\begin{aligned} I_p(s, t) &= D(s, t + 1) + D(s, t + 2) + \cdots + D(s, t + \lambda) \\ &= (1 - \beta)\alpha D(s, t) + (1 - \beta)^2 \alpha^2 D(s, t) + \cdots + (1 - \beta)^\lambda \alpha^\lambda D(s, t) = \\ &= \frac{(1 - \beta)\alpha - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - (1 - \beta)\alpha} D(s, t). \end{aligned} \quad (4)$$

综上所述, 右视点帧重要性可表征为:

$$\begin{aligned} I(s, t) &= I_c(s, t) + I_p(s, t) = D(s, t) + \\ &= \frac{(1 - \beta)\alpha - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - (1 - \beta)\alpha} D(s, t) = \\ &= \frac{1 - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - (1 - \beta)\alpha} D(s, t), \end{aligned} \quad (5)$$

其中, 初始化 $D(0, 0) = 0$. 由(5)式可知, 右视点帧重要性和采用的错误隐藏方法、后续帧的帧数、帧内编码比例及视频序列的连续性, 即错误扩散因子 α 有关. 如果知道初始条件和相关参数, 就可求出右视点某帧对于整个立体视频序列的重要性.

2.2 左视点帧重要性区分

左视点的错误隐藏失真和右视点的表示方法相同. 左视点帧丢失会影响所有后续的左右视点帧, 因此错误扩散失真可表示为:

$$I_p(s, t) = I_{PL}(s, t) + I_{PR}(s, t), \quad (6)$$

式中, $I_{PL}(s, t)$ 为左视点帧丢失引起的后续左视点帧的总失真; $I_{PR}(s, t)$ 为左视点帧丢失引起的后续右视点帧的总失真. $I_{PL}(s, t)$ 可表示为:

$$\begin{aligned} I_{PL}(s, t) &= D(s, t + 1) + D(s, t + 2) + \cdots + D(s, t + \lambda) \\ &= (1 - \beta)\alpha D(s, t) + (1 - \beta)^2 \alpha^2 D(s, t) + \cdots + (1 - \beta)^\lambda \alpha^\lambda D(s, t) = \\ &= \frac{(1 - \beta)\alpha - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - (1 - \beta)\alpha} D(s, t). \end{aligned} \quad (7)$$

左视点某帧丢失对后续右视点帧产生的扩散失真可用上一视点同一时刻帧的失真及同一视点

上一时刻帧的失真加权求取, 则右视点后续第 λ 帧失真可以表示为:

$$\begin{aligned} D(s + 1, t + \lambda) &= (1 - \beta)(1 - \gamma)\omega_1 D(s, t + \lambda) + \\ &= (1 - \beta)\gamma\omega_2 D(s + 1, t + \lambda - 1) = \\ &= mD(s, t + \lambda) + nD(s + 1, t + \lambda - 1), \end{aligned} \quad (8)$$

式中, $m = (1 - \beta)(1 - \gamma)\omega_1$, $n = (1 - \beta)\gamma\omega_2$; γ 为视点内预测比例; β 为帧间预测比例; ω_1 和 ω_2 为加权系数, 且二者之和为 1; α 可用 $(1 - \beta)^2$ 近似表示. 所以, 右视点后续 λ 帧的总失真为:

$$\begin{aligned} I_{PR}(s, t) &= D(s + 1, t) + D(s + 1, t + 1) + \cdots + \\ &= D(s + 1, t + \lambda) = m[D(s, t) + D(s, t + 1) + \\ &= D(s, t + 2) + \cdots + D(s, t + \lambda)] + mn[D(s, t) + \\ &= D(s, t + 1) + D(s, t + 2) + \cdots + D(s, t + \lambda - 1)] + \\ &= mn^2[D(s, t) + D(s, t + 1) + D(s, t + 2) + \cdots + \\ &= D(s, t + \lambda - 2)] + \cdots + mn^\lambda D(s, t) = \\ &= m \frac{1 - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - \alpha + \alpha\beta} D(s, t) + nm \cdot \\ &= \frac{1 - (1 - \beta)^\lambda \alpha^\lambda}{1 - \alpha + \alpha\beta} D(s, t) + mn^2 \cdot \\ &= \frac{1 - (1 - \beta)^{\lambda-1} \alpha^{\lambda-1}}{1 - \alpha + \alpha\beta} D(s, t) + \cdots + mn^\lambda D(s, t) = \\ &= (m \frac{1 - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - \alpha + \alpha\beta} + nm \frac{1 - (1 - \beta)^\lambda \alpha^\lambda}{1 - \alpha + \alpha\beta} + \\ &= mn^2 \frac{1 - (1 - \beta)^{\lambda-1} \alpha^{\lambda-1}}{1 - \alpha + \alpha\beta} + \cdots + mn^\lambda) D(s, t), \end{aligned} \quad (9)$$

则左视点帧丢失对后续左右视点产生的总扩散失真可表示为:

$$\begin{aligned} I_p(s, t) &= I_{PL}(s, t) + I_{PR}(s, t) = ((m + 1) \cdot \\ &= \frac{1 - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - \alpha + \alpha\beta} + mn \frac{1 - (1 - \beta)^\lambda \alpha^\lambda}{1 - \alpha + \alpha\beta} + \\ &= mn^2 \frac{1 - (1 - \beta)^{\lambda-1} \alpha^{\lambda-1}}{1 - \alpha + \alpha\beta} + \cdots + mn^\lambda) D(s, t). \end{aligned} \quad (10)$$

综上所述, 左视点帧的重要性模型可表示为:

$$\begin{aligned} I(s, t) &= I_c(s, t) + I_p(s, t) = \\ &= ((m + 1) \frac{1 - (1 - \beta)^{\lambda+1} \alpha^{\lambda+1}}{1 - \alpha + \alpha\beta} + mn \cdot \\ &= \frac{1 - (1 - \beta)^\lambda \alpha^\lambda}{1 - \alpha + \alpha\beta} + mn^2 \frac{1 - (1 - \beta)^{\lambda-1} \alpha^{\lambda-1}}{1 - \alpha + \alpha\beta} + \\ &= \cdots + mn^\lambda + 1) D(s, t). \end{aligned} \quad (11)$$

由(11)式可知, 左视点帧的重要性和后续帧的帧内编码比例、视点内预测比例、相邻时刻帧 MSE

及视频序列连续性有关. 所以如果知道初始条件和相关参数, 即可求出左视点某帧丢失对于整个立体视频序列的重要性.

3 实验验证

为验证提出帧重要性模型的有效性, 在 JM18.2 平台上对 2 个典型的立体视频序列“Flamenco2”和“Ballroom”进行编解码, 帧率为 $30 \text{ f} \cdot \text{s}^{-1}$, QP 取定值 28. 视频丢帧后采用前文所述的错误隐藏方法进行恢复, 图 1 和图 2 分别对应 2 个视频序列的模型值与实际值的 MSE, 重要性大小与 MSE 成正比, MSE 越大, 该帧越重要. 图 3 和图 4 使用 PSNR 来表征 2 个视频序列的帧重要性, 重要性大小与 PSNR 成反比, 即 PSNR 越小, 该帧越重要. 为了减小随机误差, 实验在相同环境下进行多次实验后取平均值.

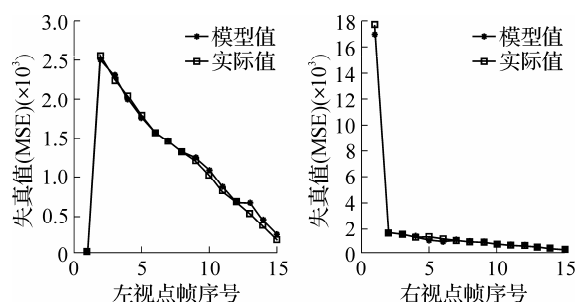


图 1 Flamenco2 每帧丢失对序列带来的总失真

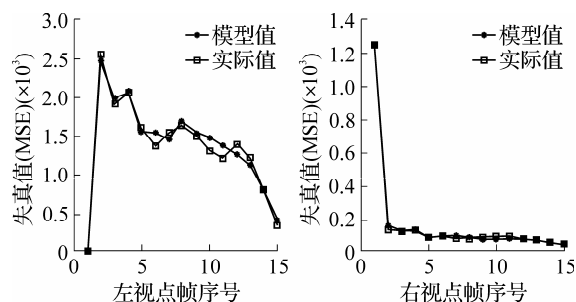


图 2 Ballroom 每帧丢失对序列带来的总失真

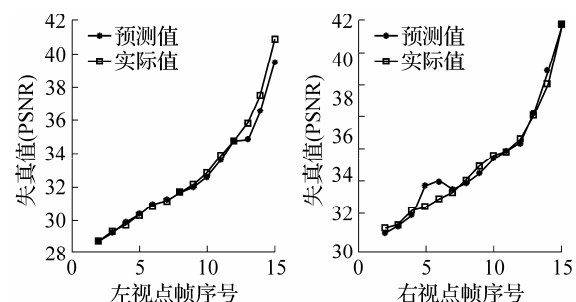


图 3 Flamenco2 某帧丢失后序列的平均 PSNR

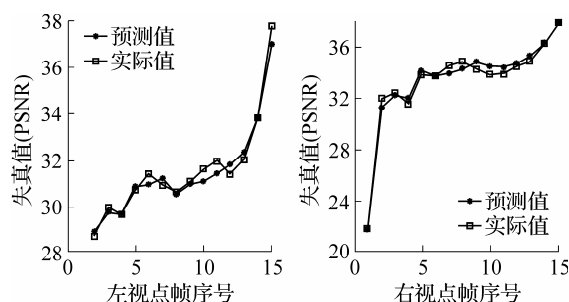


图 4 Ballroom 某帧丢失后序列的平均 PSNR

图中显示两者的曲线轮廓基本符合, 这也证明了提出的帧重要性模型较为准确. 从几个图中还可以看出, 运动剧烈序列 Flamenco2 帧平均失真值大于运动缓慢序列 Ballroom, 这是因为运动缓慢序列的时间相关性更强, 某帧丢失对于本帧和后续帧的影响更小. 2 个序列右视点第 1 帧失真值最大, 且左视点帧的平均重要性大于右视点帧, 这是因为左视点帧丢失会对后续的所有帧产生影响, 而右视点帧丢失只会影响后续右视点帧. 为了客观验证本文模型的准确性, 定义帧重要性平均预测误差:

$$E(s, t) = \frac{\sum_{t=0}^N |D_{\text{sim}}(s, t) - D_{\text{act}}(s, t)|}{\sum_{t=0}^N D_{\text{act}}(s, t)}, \quad (12)$$

式中, $D_{\text{sim}}(s, t)$ 为模型失真值; $D_{\text{act}}(s, t)$ 为实际验证值. 对不同运动剧烈程度序列和视差大小序列进行实验, Rollerblade 和 Crowd 序列为运动剧烈序列, 相机间距分别为 20 cm 和 19.5 cm; Ballroom 为运动缓慢序列, 相机间距为 6.5 cm; Aquarium 为小幅度运动序列, 相机间距为 20 cm. 在相同环境下 100 次试验取平均值得到以上结果. 表 1 则为不同特性序列的帧重要性平均预测误差.

表 1 视频序列帧重要性平均预测误差 %

视点	Ballroom	Crowd	Rollerblade	Aquarium	平均
左视点 (MSE)	5.22	4.78	6.95	4.18	5.28
右视点 (MSE)	8.81	3.26	3.25	7.53	5.71
左视点 (PSNR)	0.94	0.65	0.78	0.70	0.77
右视点 (PSNR)	1.08	0.97	0.72	0.77	0.89

由表 1 可知, 对于以上几种不同运动剧烈程度

和不同视差大小的视频序列, 用 MSE 表征帧重要性时, 笔者提出的帧重要性模型与实际解码得到的真实失真值之间的平均估算误差基本能够控制在 6%以内; 用 PSNR 表征帧重要性时, 平均估算误差可控制在 0.9%之内. 实验结果验明, 文中采用的模型对不同视差大小和运动剧烈程度的视频序列都具有适用性.

4 结论

通过分析立体视频传输过程中的错误隐藏失真和错误扩散失真, 提出了一种面向感知的帧重要性区分模型. 实验结果表明, 该模型适用于不同运动剧烈程度和视差大小的立体视频序列, 能够在编码端准确地估计出各帧丢失对终端感知的重要性, 为确定立体视频编码参数、提高率失真优化性能和指导信道的差错控制和传输的拥塞控制机制等提供了有力保障.

参考文献:

- [1] Gurler C G, Gokemli B, Saygili G, et al. Flexible transport of 3-D video over networks[J]. Proceedings of the IEEE, 2011, 99(4):694-707.
- [2] 周洋, 蒋刚毅, 郁梅, 等. 面向HBP编码格式的立体视频 B 帧整帧丢失分层错误隐藏算法[J]. 电子与信息学报, 2014, 36(2):377-383.
- [3] 冯妮娜, 常义林, 李鹏, 等. 利用 H.264/AVC 数据分割实现 GOP 级视频传输的方法[J]. 西安电子科技大学学报, 2011, 38(3):29-36.
- [4] Lin C H, Wang Y C, Shieh C K, et al. An unequal error protection mechanism for video streaming over IEEE 802.11e WLANs[J]. Computer Networks, 2012, 56(11): 2590-2599.
- [5] 乐德广, 靳勇, 郭东辉, 等. 基于帧重要性认知的网络差错控制策略[J]. 通信学报, 2011, 32(11):11-20.
- [6] Yang X, Zhu C, Li Z G, et al. An unequal packet loss resilience scheme for video over the Internet[J]. IEEE Transactions on Multimedia, 2005, 7(4):753-765.
- [7] 俞伟, 陈曦, 陈学广, 等. 高速无线个域网多视频流信道时间分配算法[J]. 华中科技大学学报, 2008, 36(1):5-8.
- [8] 王曜, 刘文予, 王薇, 等. 一种帧级无线视频传输不平等保护策略[J]. 微型计算机系统, 2005, 26(3):474-477.
- [9] 陈瑞, 张健, 童莹. IEEE 802.11e 中基于失真度量度的视频包传输方法[J]. 计算机应用研究, 2013, 30(6):1813-1816.
- [10] 毛年胜, 卓力. 基于 H.264 SVC 的 IP 网络视频传输系统的实现[J]. 测控技术, 2010, 29(5):5-8.
- [11] 周菡, 于文斌. 视频码流的层次划分研究[J]. 电子科技, 2012, 25(9):154-156.

A Frame Importance Model of Stereoscopic Video Based on Perception Analysis

ZHANG Lian-jun, LI Shao-bo, WANG Xiao-dong, WANG Teng-fei

(Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China)

Abstract: The loss of frames affects the perception quality of stereoscopic video at terminal, to tackle this problem, a model of frame importance is proposed on the basis of content of stereo video. According to the spatial and temporal correlation, and the characteristics of inter-frame error diffusion, the model calculates the frame importance from the error concealment distortion and error diffusion distortion caused by the loss of current frame. The results show that the model can estimate the importance of each frame satisfactorily, and can be adapted to video sequences with different motion intensity and disparity.

Key words: stereoscopic video; frame importance; content perception

(责任编辑 章践立)