

利用运动强度自适应传输视频内容

于俊清¹⁾ 刘冲¹⁾ 何云峰¹⁾ 胡胜红^{1,2)}

¹⁾ (华中科技大学计算机科学与技术学院 武汉 430074)

²⁾ (湖北经济学院计算机学院 武汉 430205)

(yjqing@hust.edu.cn)

摘要 在异构的网络环境下,如何自适应地传输满足多种终端设备和不同用户需求的视频内容具有重要的意义. 以达尔文流媒体服务器为实验平台,提出一套基于运动强度的视频自适应传输策略. 在服务器端通过检测丢包率来控制发送速率等级,避免网络拥塞,并采取一定策略消除因丢帧而产生的马赛克;在自适应策略中加入了运动强度信息,针对不同的运动强度级别进行不同的处理,使得综合视频质量得到提高. 实验结果表明,文中的自适应策略不仅可以避免网络拥塞,而且可以改善视频播放质量.

关键词 运动强度;自适应传输;达尔文流媒体服务器;丢帧;马赛克
中图法分类号 TP391

Adaptive Video Transmission Using Motion Intensity

Yu Junqing¹⁾ Liu Chong¹⁾ He Yunfeng¹⁾ Hu Shenghong^{1,2)}

¹⁾ (School of Computer Science & Technology, Huazhong University of Science & Technology, Wuhan 430074)

²⁾ (Computer School, Hubei University of Economics, Wuhan 430205)

Abstract Adaptive video transmission is dedicated to deliver video content over dynamic and heterogeneous networks and devices. Using Darwin streaming server, a set of adaptive strategy is proposed based on motion intensity. To avoid network congestion, the sending rate level on the server is controlled by detecting the packet loss rate in real time. To eliminate the mosaic effect brought by frame dropping, an algorithm of certain P frame dropping is proposed. As for different motion intensity, the video quality is different when the frame dropping rate is same, so the motion intensity descriptor from MPEG-7 standard is introduced to the adaptive strategy as a parameter. Experimental results demonstrate that the proposed strategies can not only improve the overall video quality, but also eliminate the mosaic effectively.

Key words motion intensity; adaptive transmission; Darwin streaming server; frame dropping; mosaic

随着网络技术和流媒体技术的发展,各种流媒体服务蓬勃涌现,在线点播、视频新闻、视频电话、远程教育等都开始慢慢融入大众生活. 为满足不同用户和各种类型终端设备无缝访问网络上分布广泛的媒体内容,各类自适应传输策略得到了广泛的关注,如自适应编码、网络拥塞控制、服务器端单播和组播

控制、缓冲机制等^[1]. 但这些自适应策略大多没有考虑视频的内容信息. 视频的内容和传输质量是终端用户最为关心的 2 个方面,通常用户较为关心的内容需要较高的视频质量,而用户较不关心的内容可以降低视频质量甚至可以不传输,因此视频内容是指导视频传输的重要信息. 在基于内容的自适应传输

收稿日期:2008-07-16;修回日期:2009-01-12. 基金项目:国家自然科学基金(60703049);湖北省自然科学基金(2007ABA265);武汉市青年科技晨光计划(200850731353). 于俊清,男,1975 生,博士,副教授,CCF 高级会员,主要研究方向为数字媒体处理与检索、多核并行处理. 刘冲,男,1983 年生,硕士,工程师,主要研究方向为流媒体传输. 何云峰,男,1977 年生,博士研究生,讲师,主要研究方向为数字视频处理与检索. 胡胜红,男,1979 年生,博士研究生,讲师,主要研究方向为数字视频处理与传输.

的研究方面,哥伦比亚大学 DVMM 实验室的研究工作具有一定的代表性. Chang 等实现的实时内容自适应视频传输针对网球和篮球等体育视频,提取其中重要的和有意义的信息,有选择地在低带宽网络中传输^[2]. Wang 等实现的基于内容的有效预测及视频自适应主要是为了满足不同的终端设备、网络环境和用户需求,并用来设计视频流的自适应策略^[3]. 本文在实现选帧自适应传输策略的基础上,通过引入运动强度信息,使得选帧更加合理. 该自适应传输策略以 H.264 编码为例,在达尔文流媒体服务器(Darwin streaming server, DSS)上进行了测试. 在 DSS 中首先将网络状况分为 7 级,参考加法增加、乘法减少(additive increase and multiplicative decreasing, AIMD)策略来进行丢帧,并设计算法消除了因 P 帧丢失而产生的马赛克现象;然后在自适应策略中加入运动强度信息作为重要参数,以在相同丢帧率的条件下使用户得到较好的视频质量.

1 运动强度分析

1.1 运动强度的提取

在 MPEG-7 标准中,运动活力描述符用于描述视频的运动程度,其中的运动强度则描述了视频镜头的运动强烈程度^[4],它是按镜头提取的. 但在实际视频中,一方面视频镜头非常短,另一方面一个场景中的多个连续镜头的运动强度比较接近. 为了与视频内容相符合,并方便进行自适应控制,本文针对场景而不是镜头提取运动强度特征. 运动强度的提取可以在 MPEG-1 的压缩域中进行,提取算法步骤如下:

Step1. 使用 MPEG-1 的标准解码器 mpeg2decode 来获得一个 MPEG-1 压缩视频中每个场景包含的所有 P 帧各宏块的运动向量,定义活力矩阵 $C_{mv} = \{R(i, j)\}$. 其中, $R(i, j) = \sqrt{x_{ii}^2 + y_{ii}^2}$,表示 (i, j) 块处的运动活力, x_{ii} 和 y_{ii} 分别表示 (i, j) 块在 x 和 y 方向上的运动向量.

Step2. 计算每个宏块的平均向量幅度

$$C_{mv}^{avg} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J C_{mv}(i, j),$$

其中 I, J 分别代表一个 P 帧中 x 和 y 方向上的宏块数目.

Step3. 计算该 P 帧的运动幅度标准差

$$S = \sqrt{\frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J (C_{mv}^{avg} - C_{mv}(i, j))^2}.$$

Step4. 求运动幅度标准差的期望值 $\bar{S} = E(S)$.

将 \bar{S} 值进行量化就可以得到场景的运动强度,

一般量化为 1~5 级,分别对应超低、低、中等、高和超高运动强度. 从实验的结果考虑,公用中多分辨率格式(common intermediate format, CIF)、25 帧/s 视频的运动强度量化值分别为 50, 80, 100 和 120,这与 MPEG-7 中给定的参考值有些区别.

典型的测试视频帧序列如图 1 所示,从上到下运动强度依次为 1~5 级. 从图 1 中可以直观地看到,视频的运动激烈程度不断增加,第 1 行的视频序列是序幕,字母缓缓出现,测得的运动幅度标准差的期望值为 6.86,属于超低类;第 2 行视频序列是一个人在讲话,移动非常缓慢,测得最终值为 69.50;第 3 行视频序列是一个人在走动,移动速度稍微变快,背景也在变化,测得最终值为 80.13;第 4 行视频序列是一辆车开过来,其中背景、前景都有变化,测得最终值为 110.56;第 5 行视频序列是很多人在熙熙攘攘地下楼梯,整个画面激烈变化,测得最终值为 126.79.



图 1 1~5 级运动强度视频示例图

1.2 运动强度对视频质量的影响

视频的运动强度对传输时视频质量影响很大,下面从 2 个方面来分析运动强度对视频传输质量的影响.

1) 运动强度在丢帧时视频质量的影响

在视频传输中,位速率是变化的,不同运动强度和质量视频的位速率(kbps)的关系如图 2 所示. 从图 2 可以看出,在相同质量条件下,运动强度越高的视频位速率越高,丢帧时可以更有效地解决网络拥塞. 从实验结果来看,假定网络不丢包且发送速率级别一致,在丢帧处理后的播放过程中,视频运动强度

低的视频由于内容变化少,画面一般不会剧烈变化,即使只播放 1 帧,用户也可以清晰地观看视频画面的变化过程.随着视频中运动强度不断增高,变化程度越来越大,视频会因为丢帧而产生越来越多的信息丢失.当运动强度达到第 5 级时,如果只播放 1 帧,看到的将是一幅幅跳动很大的图片,视频中的一些动作被丢失了,视频质量很差.因此在网络不丢包且发送速率级别一致时,运动强度越大,丢帧对视频质量的影响就越大.

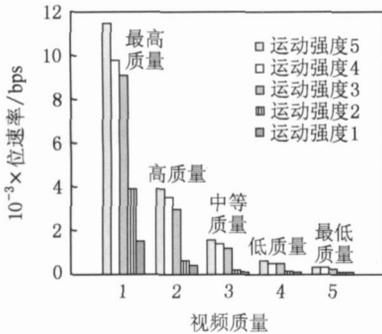


图 2 视频传输码率与视频质量关系示意图

2) 运动强度在丢包时对视频质量的影响

在有网络丢包时,运动强度也在影响视频帧的质量上扮演着重要的角色^[5].用

$$Q(L) = \frac{1}{N} \sum_{n=1}^N \frac{45 - PSNR(n, L)}{45}$$

来计算运动强度为 的视频流在丢包率为 L 时的平均质量损失.其中, $PSNR(n, L)$ 表示第 n 个数据流在处理丢包后的信噪比,45 代表优秀的图像质量, N 表示运动级别为 的数据流总数.这样测得视频损失曲线如图 3 所示,可以看出,在丢包率相同时,随着运动强度升高,视频质量下降率增加.因此,高运动强度的视频质量对丢包率非常敏感,而低运动强度的视频对丢包率不敏感.

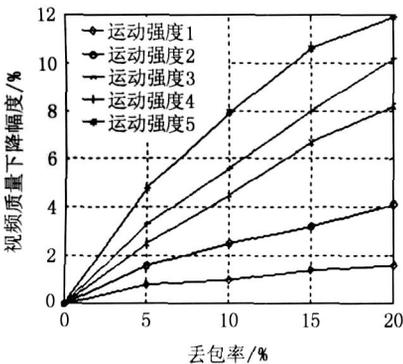


图 3 丢包对不同运动强度视频质量的影响

通过上述分析可以得出结论:运动强度对视频

质量有着非常大的影响,高运动强度的视频应该尽可能少丢帧,而且在网络传输中应尽可能不丢包.

2 基于运动强度的自适应传输

本文通过引入运动强度与视频质量之间的关系,在 DSS 上设计并实现了基于运动强度的视频自适应传输系统.下面对该系统中使用的自适应策略进行系统阐述.

2.1 服务器端的丢帧策略

本文系统传输的视频数据是加入 Hint Track 的 H.264 视频,其中 Hint Track 主要用来协助服务器端打包媒体数据,实现流媒体的传输.通过分析不同运动强度、不同视频质量的视频场景打包后的不同帧的数据包数目所占的比例,可以得出以下结论: 1) I 帧的包所占比例不超过 20%,B 帧和 P 帧的包占绝大多数,而且 P 帧所占比例最多;2) 在低视频质量时,B 帧与 P 帧比例相差不大,当视频质量升高时有先增后减的趋势.

由于 B 帧对其他帧图像没有影响,而 P 帧对 B 帧有影响,因此服务器端的丢帧策略是先丢弃 B 帧,这样对网络带宽的缓解具有明显的效果;丢完 B 帧后可以继续丢弃 P 帧,同样会大幅度缓解网络带宽,直到只剩下 I 帧.本文将视频流的发送速率分为 7 个等级,如表 1 所示.

表 1 视频传输质量分级表

级别	名称	意义
1	kAllPackets	传输所有帧
2	k70PercentBFrames	丢弃 30% 的 B 帧
3	k30PercentBFrames	丢弃 70% 的 B 帧
4	kNoBFrames	丢弃所有 B 帧
5	k70PercentPFrames	丢弃 30% 的 P 帧
6	k30PercentPFrames	丢弃 70% 的 P 帧
7	kKeyFramesOnly	只传输 I 帧

2.2 丢帧优化策略

在传输过程中,I 帧是不能被丢弃的,因为 I 帧的丢弃将导致随后解码的效果急剧恶化,马赛克现象严重,视频将无法观看.B 帧的丢弃不会引起解码错误,因此在丢弃 30% B 帧时应选取相隔较远的头尾 B 帧丢弃,播放时基本感觉不到丢帧;而丢弃 70% B 帧时,则保留相隔较远的头尾 B 帧.P 帧是重要的参考帧,由于随后 P 帧的解码必须依赖于前面的 I 帧或 P 帧,一旦其参考帧丢失,会导致随后的

P 帧解码错误,产生大块的马赛克,严重影响视频质量,因此一些自适应策略经常会一次把 P 帧全部丢弃或者不得不出现马赛克.通过实验可以发现,部分 P 帧的丢失或许就可以使得网络带宽立刻得到缓解;而且 P 帧非常重要,丢弃全部 P 帧后的效果和丢弃部分 P 帧的效果差别非常明显.因此,有选择地丢弃部分 P 帧将具有十分重要的意义.

为了解决丢弃 P 帧时所引起的马赛克现象,本文将需要丢弃的 P 帧集中在 2 个 I 帧之间的后半部分.例如丢弃 30% 的 P 帧时,先读取 I 帧和下一个 I 帧之间的 P 帧数目,顺次读取大约 70% 的 P 帧时,则直接跳到下一个 I 帧.这样使所有的帧都能正常解码,解决了因为丢弃参考帧而引起马赛克的现象,这也是有选择丢弃 P 帧的优点.

2.3 传输中的自适应策略

视频数据在传输时,由于网络传输条件会发生变化,因此需要根据当前的网络传输条件来调整服务器端的发送速率.通过在服务器端检测传输控制协议(real-time transport control protocol, RTP)包中的丢包率来确定当前的网络传输条件,并据此调整视频的发送速率^[6-7].但网络中可能会出现一些突发的波动,此时丢包率会瞬时增加,然后立刻复原,不需要调整发送速率.为了防止这种瞬时突变引起发送速率不必要的大幅变化,需要对丢包率进行平滑处理.本文采用低通滤波方法处理,使用前一次的丢包信息为基准来计算当前的丢包信息,即 $\hat{P}_{loss}(t) = P_{loss}(t) + (1 - \alpha) \hat{P}_{loss}(t - 1)$. 其中, α 取为 1/4 或 1/8; $\hat{P}_{loss}(t)$ 表示平滑处理后的丢包率, $P_{loss}(t)$ 是检测到的丢包率.如果连续 4 次丢包率超过上限,则进行发送速率降低自适应调节;如果连续 6 次接收到丢包率低于下限,则进行发送速率升高自适应调节.

为了避免网络拥塞,保证 TCP 友好,对于实时视频数据传输一般借用 AIMD 的调整思想^[8-9],使速率能够根据探测到的丢包率而加法增加、乘法减少.依据丢包上下限,可以将网络分为轻载、满载和过载 3 种状态.轻载时网络丢包率小于丢包下限 P_1 ,过载时大于丢包上限 P_2 ,满载则处于中间.轻载时网络负载小,可以增加发送速率,过载时必须降速,而满载则继续保持现状^[10].

本文系统使用的策略为

$$k = \begin{cases} k - 1, & P_{loss} < P_1, n \geq 6, k > 1 \\ 4, & P_{loss} > P_2, n \leq 4, k < 4 \\ 7, & P_{loss} > P_2, n \leq 4, k = 4 \end{cases};$$

其中, k 表示当前应该调整到的发送速率级别, n 表示检测到的次数, P_1 为网络轻载的下限, P_2 为网络过载的上限.如果探测到网络轻载,则缓慢地逐步提高速率级别,逐步提高视频质量,使之尽量满足 TCP 友好.如果探测到网络过载,则大幅度降低速率级别:若当前速率级别是 1~3 级时,直接下降到第 4 级,即丢弃所有 B 帧;若当前速率级别是 4~6 级时,直接下降到第 7 级,即只传输 I 帧.

2.4 基于运动强度的传输控制策略

通过前面对运动强度与视频质量之间关系的分析可知,对于运动强度大的视频,既要尽量使得它丢弃帧不太严重,发送级别不能过低,又要保证其在传输时丢包率不会过高.因此,需要根据运动强度来调整网络状态判断的丢包率上限与下限,使得不同运动强度的视频能够进行不同的处理.根据图 3 中对于丢包率与视频质量损失的分析结果,为了使网络在满载状态时的视频最大损失不超过 10%,设定如表 2 所示的丢包率上下限.

表 2 不同运动强度视频丢包率上下限设置表

运动强度	运动强度 1	运动强度 2	运动强度 3	运动强度 4	运动强度 5
下限	1	3	5	7	9
上限	30	25	20	17	15

当运动强度最小时,视频对网络波动不敏感;当网络很差时开始下降速率,或当网络很好时开始提升速率.由于运动强度低的视频本身位速率很低,而且对由网络丢包而产生的视频质量变化也不敏感,所以设置的下限小、上限大,这样的设定符合运动强度低的视频需求.当运动强度最大时,视频对网络波动非常敏感.如果网络丢包率超过一定程度,视频质量损失将非常大,而且运动强度大的视频码率很高,丢帧能更好地解决网络拥塞,所以设置低的网络上限可以保证网络丢包率不会很大.丢包率下限以保证视频质量为主,考虑到高运动强度场景的丢帧对质量的影响比低运动强度场景要大,因此该值不能太小,使得一旦网络状况好转,高运动强度的视频最先提速.

3 实验平台与结果分析

3.1 实验平台的选择

实验平台使用 Apple 公司的开放源代码 DSS 5.5.4 实现.DSS 从整体上分为服务器内核和模块 2

部分,其中内核实现了服务器的运转、线程的管理以及文件打包系统等,模块是具体的处理程序;所有处理功能都可以以模块方式加入.本文选择 RTP 子系统进行二次开发,添加了运动强度描述信息,使得已有的发包例程能根据运动强度选择丢包策略.

3.2 实验结果分析

在测试中,点播的视频质量主观上虽然可以立刻得到非常明显的结论,但是客观上用数据描述却比较困难.本文通过在客户端不断地增加点播服务来消耗网络带宽.未经过自适应传输控制的视频画面会不断恶化,直到停止;而采用自适应策略的视频画面不会出现明显的马赛克,虽然会有些小的中间停顿,但是画面非常清晰,随着网络带宽的缩减,直到最后传输 2 帧/s 到停止.图 4 给出了截取的 10 幅画面的峰值信噪比(peak signal to noise, PSNR).

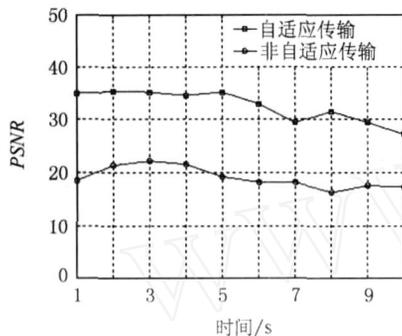


图 4 自适应与非自适应的视频图像质量对比示意图

在发送速率等级均为 k30Percent PFrames 时,对相同的视频进行测试,没有经过处理的视频出现大量马赛克,而经过优化处理的视频基本没有马赛克.截取如图 5 所示的 2 个视频大致相同的时间点,左边为一般策略丢弃 P 帧,右边为选定丢弃 P 帧,可以看到去除马赛克前后的明显效果.同时测得



图 5 消除马赛克前后对比

这 6 幅图相对于原图的 PSNR 值分别为 18.522, 18.382, 17.864 与 23.274, 24.421, 22.704, 如图 6 所示.

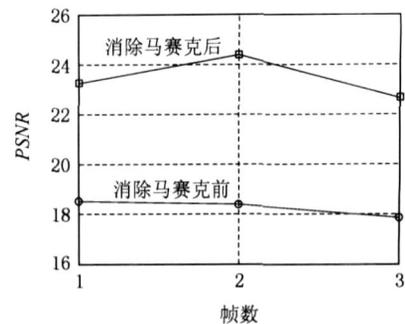


图 6 消除马赛克前后 PSNR 对比示意图

使用运动强度信息后,在客户端同时点播 1~5 级视频.当网络开始阻塞时,第 5 级视频最先开始反应,立刻会因为丢帧而出现播放中短暂跳动,但视频质量依然完好,其他视频依然正常播放;当网络带宽继续被消耗,其他等级的视频依次丢帧,在网络级别传输变化的瞬间会出现马赛克.当网络情况好转时,高运动强度的视频最先好转,其他视频依次好转.与未加入运动强度信息的播放过程相比,在网络开始阻塞时,大多数视频开始同时降低发送速率,影响了本不应该降速的视频,而回升速率时也没有先后次序,有可能使得高运动强度的视频得不到及时回升.

4 结 论

实验结果表明,利用运动强度指导选帧的自适应传输具有很好的实际效果.但是,运动强度只是视频内容的一个方面,其他的视频内容信息,如颜色、摄像机运动、运动轨迹、语义等,如果能够应用到自适应策略中,可能会对有限带宽条件下视频传输质量的提高有益.基于内容的自适应传输的研究还处于不成熟阶段,我们未来将会结合 MPEG-7 和 MPEG-21 对其进行更为系统和深入的研究.

参 考 文 献

- [1] Rejaje R, Handley M, Estrin D. Layered quality adaptation for internet video streaming [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(12): 2530-2543
- [2] Chang S F, Zhong D, Kumar R. Real-time content-based adaptive streaming of sports videos [C] // Proceedings of IEEE Workshop on Content-Based Access of Image and Video Library, Hawaii, 2001: 139-146

- [3] Wang Y, Kim J G, Chang S F. Content-based utility function prediction for real-time MPEG-4 video transcoding [C] //Proceedings of International Conference on Image Processing, Barcelona, 2003: I189-I192
- [4] Peker K A, Alatan A A, Akansu A N. Low-level motion activity features for semantic characterization of video [C] // Proceedings of IEEE International Conference on Multimedia and Expo, Vancouver, 2000: 801-804
- [5] Lotfallah O A, Reisslein M, Panchanathan S. Adaptive video transmission schemes using MPEG-7 motion intensity descriptor [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(8): 929-946
- [6] Cheng W X, Lei Z M. A modified RTP adaptive algorithm [C] //Proceedings of IEEE International Conference on Info-tech and Info-net, Beijing, 2001: 33-38
- [7] Rejaie R, Handley M, Estrin D. RAP: an end-to-end rate-based congestion control mechanism for realtime streams in the internet [C] //Proceedings of the 18th Annual Joint Conference on the IEEE Computer and Communications Societies, New York, 1999: 1337-1345
- [8] Bansal D, Balakrishnan H. TCP-friendly congestion control for real-time streaming applications [R]. Cambridge: Massachusetts Institute of Technology, 2000
- [9] Yan J Y, Katrinis K, May M, *et al.* Media- and TCP-friendly congestion control for scalable video streams [J]. IEEE Transactions on Multimedia, 2006, 8(2): 196-206
- [10] Bertin C, Defré I, Kangasoja J, *et al.* Commerce with MPEG-4 on the internet with QoS [C] //Proceedings of IEEE International Conference on Multimedia Computing and Systems, Florence, 1999: 1043-1045

(上接第 846 页)

- [10] Soucy M, Godin G, Rioux M. A texture-mapping approach for the compression of colored 3D triangulations [J]. The Visual Computer, 1996, 12(10): 503-514
- [11] Garland M, Heckbert P S. Simplifying surfaces with color and texture using quadric error metrics [C] //Proceedings of the Conference on Visualization, Research Triangle Park, 1998: 263-269
- [12] Zhang Liyan, Zhou Rurong, Tang Jie, *et al.* Research on mesh simplification with additional attributes [J]. Journal of Computer-Aided Design & Computer Graphics, 2002, 14(3): 199-203 (in Chinese)
(张丽艳, 周儒荣, 唐杰, 等. 带属性的三角网格模型简化算法研究[J]. 计算机辅助设计与图形学学报, 2002, 14(3): 199-203)
- [13] Hoppe H. Progressive meshes [C] //Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, New Orleans, 1996: 99-108
- [14] Sander P V, Snyder J, Gortler S J, *et al.* Texture mapping progressive meshes [C] //Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Los Angeles, 2001: 409-416
- [15] Garland M, Heckbert P S. Surface simplification using quadric error metrics [C] //Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, Los Angeles, 1997: 209-216
- [16] Dyn N, Levin D, Gregory J A. A butterfly subdivision scheme for surface interpolation with tension control [J]. ACM Transactions on Graphics, 1990, 9(2): 160-169
- [17] Garland M. Multiresolution modeling: survey & future opportunities [C] //Proceedings of the Eurographics, Granada, 1999: 49-65
- [18] van Kaick O M, Pedrini H. A comparative evaluation of metrics for fast mesh simplification [J]. Computer Graphics Forum, 2006, 25(2): 197-210