流媒体视频质量评价（单刺激连续质量评价方法）

2013年12月11日 00:05:22　阅读数：7051

Stefan Winkler等人在论文《Video Quality Evaluation for Internet Streaming Applications》中，介绍了流媒体视频质量评价的方法，以及他们的研究结果，在此记录一下。

注：本文中使用了单刺激连续质量评价方法（SSCQE），很有参考价值。

参与测试的序列如下表所示。注：这些序列可真是够长的！

表1是流媒体（Streaming）格式（360x288）

Table 1. Scenes for streaming content (360×288, 25fps); duration in seconds:frames

Scene #	Description	Characteristics	Duration
A	Letters	Letters with different colors flying in all directions over dark background	10:10
B	News	Male and female speaker in newsroom, almost still	11:23
C	F1 car	Object motion, camera following car, 2 angles	8:20
D	Fast food	Texture, people, fast pans, 2 angles	8:20
E	Coastguard	Two boats crossing on river, medium motion, water motion	11:24
F	Balloons	Amusement park, saturated colors, people, motion	8:05
G	Foreman	Talking head, with pan to construction site, geometric shapes	16:00
H	New York	Slow city flyover, skyscrapers at sunset, detailed texture	10:10
I	Football	Fast camera and object motion, colors	10:10
J	Live concert	Dark scene, spotlights, 3 angles	13:15
K	Cartoon	Characters dancing through scene, with pan	12:14

表2是电影（film）格式（844x360）

Table 2. Scenes for film content (844×360, 25fps); duration in seconds:frames

Scene #	Description	Characteristics	Duration
A	Movie credits	Text on forest flyover towards city skyline	15:03
B	City street	Man leaving shop, walking around building, detail, camera pan	12:01
C	Action	Helicopter crashing into building, explosion, 3 angles	9:12
D	Country road	Camera on car following road, 2 angles	6:17
E	Casino outdoors	Car driving up to casino at night, camera pan, object motion, detail	9:22
F	Casino indoors	People passing through hall, camera follows them	9:22
G	Bridge	Pan to two people crossing bridge, faces	22:12
H	Dinner	Woman talking at dinner table, faces	7:18
I	Living room	Woman in red dress walking down stairs and across room, camera follows her	10:15
J	Desert race	SF race through canyon/desert landscape, several angles	13:02
K	CG movie	Camera pan over characters, very colorful, fade to other scene	9:14

本实验考虑2种损伤：

1.视频压缩损伤

2.网络传输损伤

模型如下图所示。

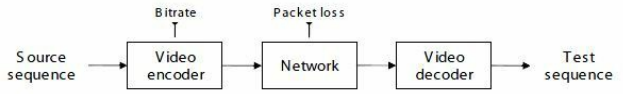


Figure 1. HRC generation chain

本实验考虑了以下编码方式：

- Windows Media Video 8
- Real Video 8
- ISO MPEG-42 (Microsoft implementation)

实验序列的设置如下表所示。注：PLR，Packet Loss Rate，丢包率。

Table 3. HRC's for medium-size streaming content

HRC #	Codec	Bitrate	PLR
1	WM	256 kb/s	–
2	Real	256 kb/s	–
3	Real	256 kb/s	2%
4	WM	512 kb/s	–
5	Real	512 kb/s	–
6	Real	512 kb/s	3%
7	MPEG-4	512 kb/s	–

Table 4. HRC's for high-resolution film content

HRC #	Codec	Bitrate	PLR
1	WM	512 kb/s	–
2	Real	512 kb/s	–
3	Real	512 kb/s	4%
4	MPEG-4	512 kb/s	–
5	WM	1 Mb/s	–
6	Real	1 Mb/s	–
7	Real	1 Mb/s	12%
8	MPEG-4	1 Mb/s	–

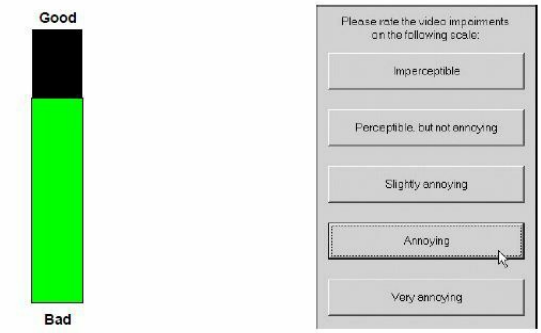
主观质量评价

本实验考虑以下2种主观质量评价方法：

1. 单刺激连续质量评价方法（SSCQE，这个方法比较不常见，但是还是挺重要的）

2.双刺激损伤标度法（DSIS）

质量评价软件打分模块如下图所示。一个是单刺激的，一个是双刺激的。



(a) SSCQE slider (b) DSIS dialog

Figure 2. Voting devices designed for the subjective experiments.

实验安排如下图所示。可见实验量还是挺大的。

Table 5. Summary of experiments

Test	Streaming Content	Film Content
SSCQE session	2 sequences (Table 1) 7 HRCs (Table 3) *14 min.* of test material (with break at half-time)	2 sequences (Table 2) 8 HRCs (Table 4) *16 min.* of test material (with break at half-time)
DSIS session	11 sequences (Table 1) 7 HRCs (Table 3) *28 min.* of test material (with break at half-time)	—

显示其使用15" Sony SDM-S51。其属性如下表所示。

Resolution:	1024×768
Dot pitch:	0.297 mm
Peak luminance:	250 cd/m^2
Contrast ratio:	300:1
Viewing angles:	120° horizontal, 90° vertical
Response times:	10 ms (rise time), 20 ms (fall time)

校准和黑电平调整后，测得如下屏幕属性：

Gamma:	2.2
Color temperature:	6400 K
White luminance:	77 cd/m^2
Video surround:	20 cd/m^2

实验室设置如下图所示：

Figure 3. Laboratory setup

一共20个非专家人员（绝大部分是大学生）参加了实验。

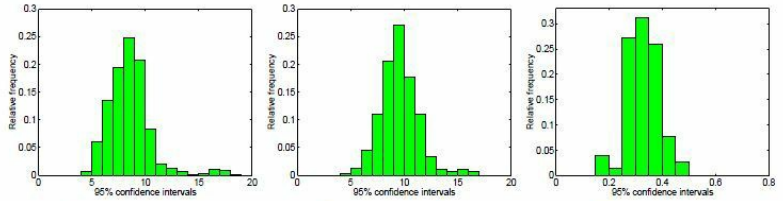
实验数据分析

几种方法的测试区间如下所示：

SSCQE streaming(0-100):8.5

SSCQE film(0-100):9.5

DSIS(0-5):0.33



(a) SSCQE test (streaming) (b) SSCQE test (film) (c) DSIS test (streaming)

SSCQE和DSIS之间的关系如下图所示。

该图可以拟合出一个关系式：SSCQE = 18.9×DSIS − 0.9

Figure 4. Distribution of 95% confidence intervals

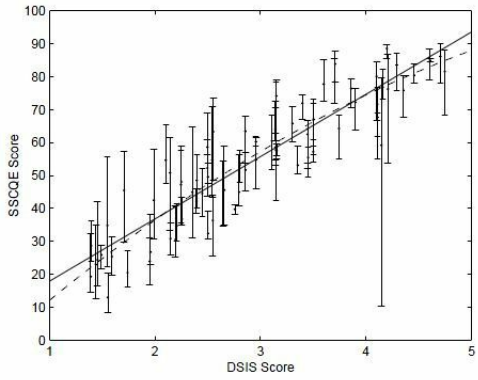


Figure 5. Comparison of DSIS and SSCQE mean scores. The vertical bars indicate the range of SSCQE scores within the corresponding DSIS sequence part; the dots indicate the average over those scores. The solid line represents a linear fit through the data (92% correlation), the dotted line results from a quadratic fit.

SSCQE实验结果如下图所示。横坐标是时间，纵坐标是MOS。灰线外面有一个矩形代表该处有95%的置信区间。

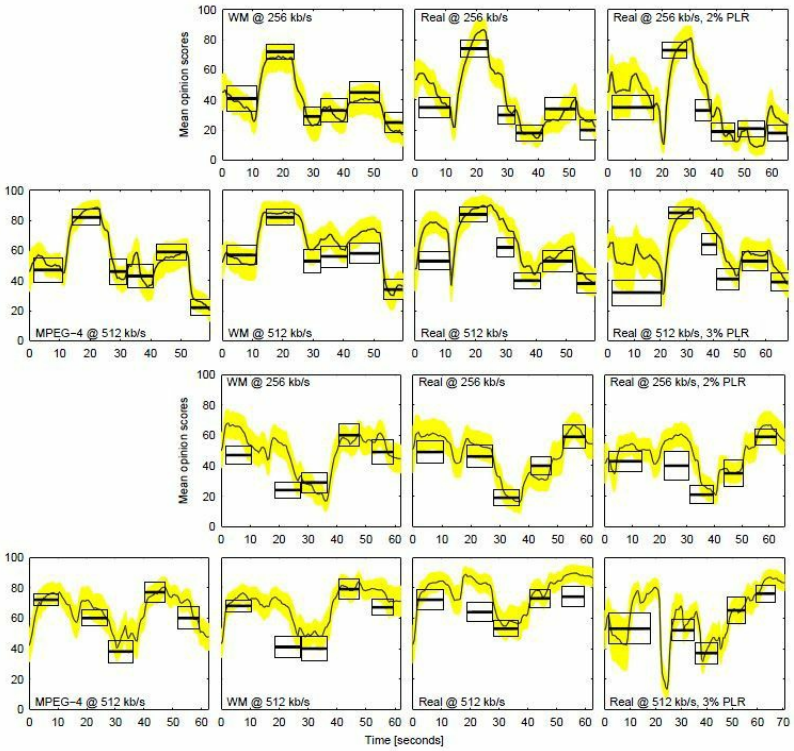
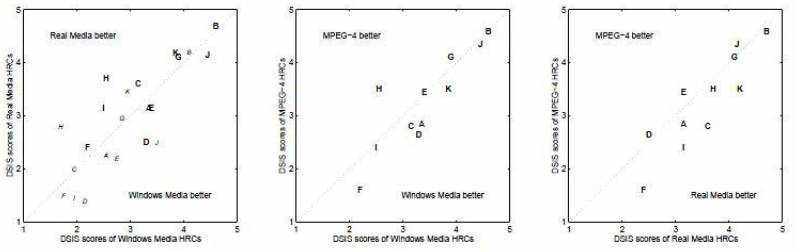


Figure 6. SSCQE ratings (smooth curves) for the 14 streaming test sequences and DSIS ratings (thick line segments) for the corresponding 77 sequence parts. Top two rows: source scenes A–F, bottom two rows: source scenes G–K. The gray bands and the hollow rectangles around the mean values indicate the 95% confidence intervals.

下面两张图显示了编码器比较的结果。横坐标是一种编码器的SSCQE的分数，纵坐标则是另一种编码器的SSCQE的分数。

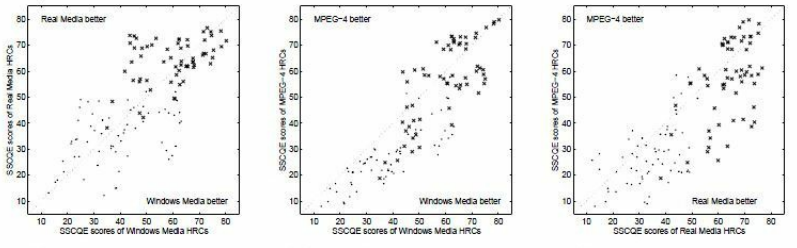
图7是流媒体内容的对比，字母和表1中的序列的字母是相对应的。小的斜体字母代表256kbps码率的结果。大的粗体字代表512kbps码率的结果。

图8中电影内容的对比，点代表512kbps码率的结果，叉代表1Mbps码率的结果。



(a) Real vs. Windows Media (b) MPEG-4 vs. Windows Media (c) MPEG-4 vs. Real Media

Figure 7. Codec comparisons for streaming content. The letters correspond to the scene numbers from Table 1. Small italic letters denote 256 kb/s HRC's, large bold letters denote 512 kb/s HRC's (only HRC's without packet losses are shown).



(a) Real vs. Windows Media (b) MPEG-4 vs. Windows Media (c) MPEG-4 vs. Real Media

Figure 8. Codec comparisons for film content. 512 kb/s HRC's are represented by dots, 1 Mb/s HRC's are represented by crosses (only HRC's without packet losses are shown).

将本次试验的数据用于Genista's Stream PQoS 软件。得到一个无参考视频质量评价方法。该方法主观和客观之间的关系如下图所示。

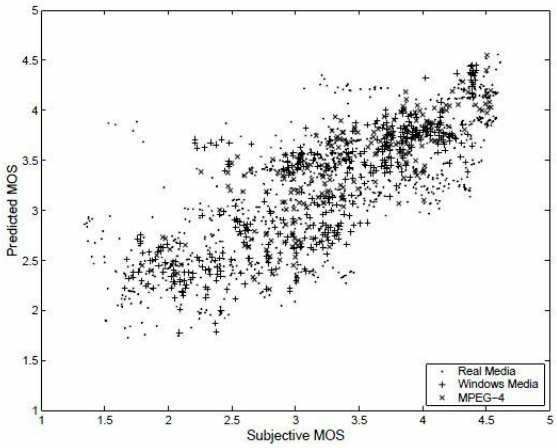


Figure 9. Predicted MOS vs. subjective MOS (both mapped onto the 1-5 scale).

该方法的性能如下所示。

Table 6. MOS prediction performance

	Linear correlation	Rank-order correlation	Prediction error
Real Media	76%	76%	0.54
Real Media (no PL)	84%	83%	0.48
Windows Media	84%	85%	0.41
MPEG-4	83%	84%	0.36
Overall	**78%**	**79%**	**0.48**

文章标签： DSIS SSCQE 视频 质量 流媒体

个人分类： 视频质量评价

所属专栏： 视频质量评价