Towards using entertaining 3D video sequences in entertainment-oriented subjective experiments

Dawid Juszka and Zdzisław Papir
Department of Telecommunications, AGH University of Science and Technology, Kraków, Poland
juszka@kt.agh.edu.pl

1. MOTIVATIONS

Change in users behaviour

- conscious decisions on the video content selection, because of:
  - wide access to VoD services
  - large availability of tag information

Content is a king!

- video quality is not the most important
- subjectively assessed level of content desirability is very important [1]

Source material selection

- crucial for the reliability of test results [4]
- recommendation: to minimize any impact of artistic, aesthetic and storyline aspects on ratings [3]:
  - select the most neutral video content
  - instruct subjects not to base their opinion on the content of the scene or the quality of the acting
- consequently: **entertainment-oriented experiments are conducted on not entertaining source material sequences**
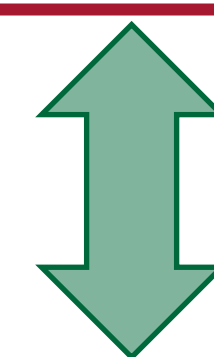
2. HYPOTHESIS

Objective parameters

| SPATIAL PERCEPTUAL INFORMATION (SI) | TEMPORAL PERCEPTUAL INFORMATION (TI) | SCENE-CUT DENSITY (d) |

CORRELATION?

| INTERESTINGNESS (INT) | VISUAL ATTRACTIVENESS (ATR) | 3D EFFECT IMPRESSION (3DE) |

Subjective parameters

4. EXPERIMENT DESIGN AND SETUP

Source content

- Scenes extracted from 9 movie productions (7 feature films and 2 documentaries) recorded on Bluray-3D discs:
  - *The Amazing Spider-Man*
  - *Born to be wild*
  - *Drive Angry*
  - *Hugo*
  - *Man of Steel*
  - *Pirates of the Caribbean: On Stranger Tides*
  - *Polar Bears: Ice Bear*
  - *The Great Gatsby*
  - *The Hobbit: An Unexpected Journey*
- Criterion for scene selection: a logical course of action within the bounds of 30±2s
- **96 scenes selected from the acquired material and 5 sequences for the purpose of training**

Test environment

- ITU-R BT.2021, ITU-R BT.500 and ITU-T P.910
- 42" 3D plasma screen PANASONIC TX-P42GT30 with shutter glasses

Subjects

- 28 subjects (14 men and 14 women)
- pre-checked:
  - RANDOT stereo test
  - Ishihara colour vision test

Methodology description

- Absolute Category Rating (ACR) scenario, but subjects answered questions about their opinions on content
- 15s to answer three questions on the Likert scale (1-5):
  - **How interesting was the presented movie clip?** – opinion about storyline
  - **How visually attractive was the presented movie clip?** – opinion about aesthetics
  - **How intensely did you experience the 3D effect in this movie clip?** – opinion about 3D effect impression
- NASA Task Load Index (NASA-TLX) questionnaire

Objective parameters definitions:

Spatial Perceptual information [5]:

$$SI = \max_{time}\{std_{space}[Sobel(F_n(i,j))]\}$$

Mean SI value for whole sequence:

$$SI_{mean} = \frac{\sum std_{space}[Sobel(F_n(i,j))]}{N}$$

Temporal Perceptual information [5]:

$$TI = \max_{time}\{std_{space}[F_n(i,j) - F_{n-1}(i,j)]\}$$

Mean TI value for whole sequence:

$$TI_{mean} = \frac{\sum std_{space}[F_n(i,j) - F_{n-1}(i,j)]}{N}$$

Scene-cut density:

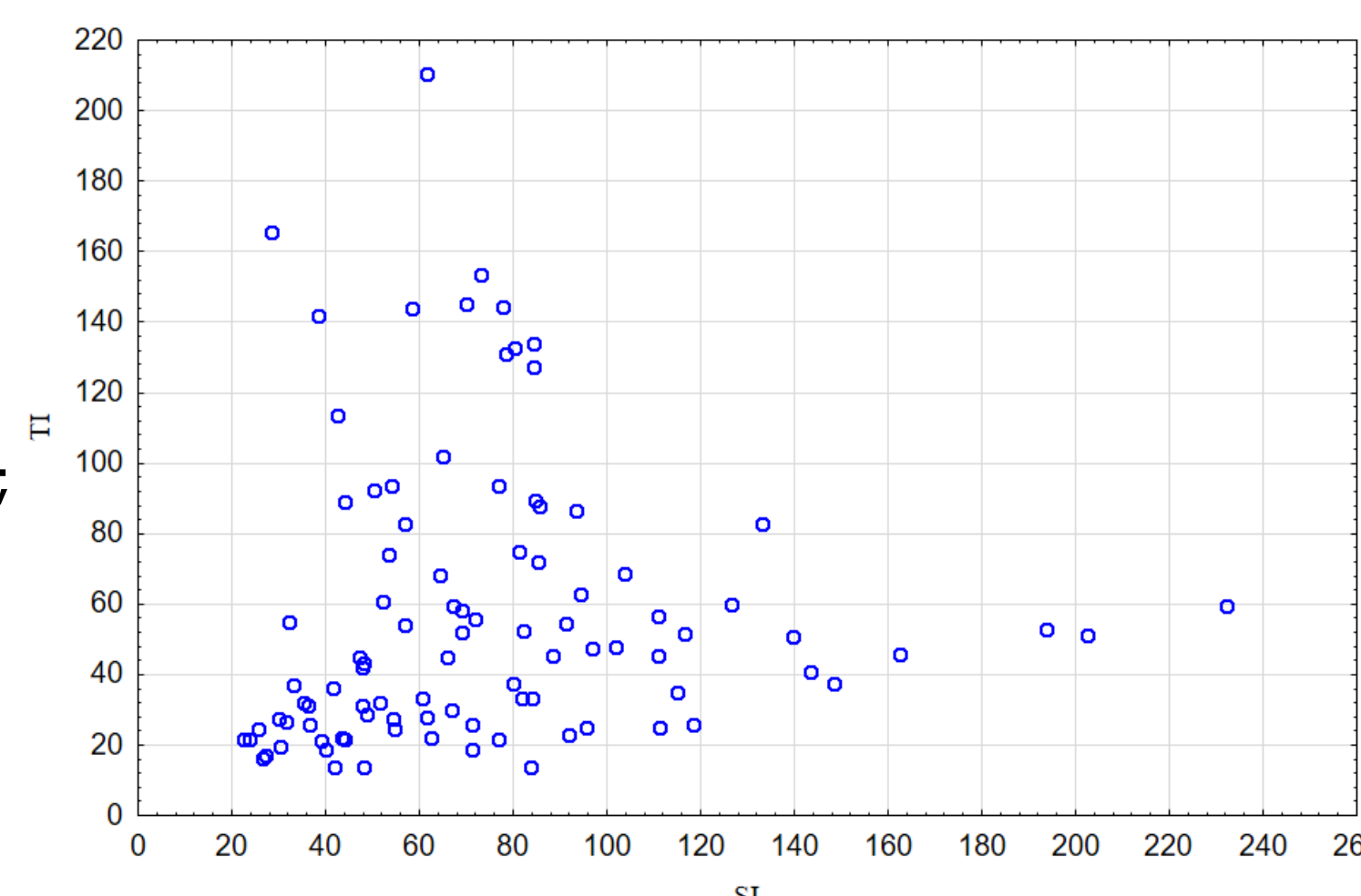$$d = \frac{\sum[r_m - r_{m-1}]}{M}$$

where:
F_n – the luminance plane of video frame at time n;
$std_{space}()$ – standard deviation over the pixels;
$Sobel()$ – a Sobel filter;
r_m – the frame number where m-th scene-cut occurs;
N – total number of frames in a sequence;
M – total number of scene-cuts in a sequence.

5. RESULTS AND CONCLUSIONS

CORRELATION COEFFICIENTS BETWEEN OBJECTIVE PARAMETERS AND SUBJECTIVE EVALUATION

	INT	ATR	3DE	SI	SI_{mean}	TI	TI_{mean}	d
INT	-	**0.55**	**0.36**	-0.04	-0.16	**0.28**	0.19	**-0.27**
ATR	**0.55**	-	**0.83**	**0.47**	**0.42**	**0.23**	**0.42**	**0.26**
3DE	**0.36**	**0.83**	-	**0.54**	**0.46**	**0.34**	**0.54**	**0.31**

*statistically significant values in bold

- objective parameters usually calculated for source video sequence selection needs (**SI** and **TI**) **cannot be used alone** to estimate if video content is visually attractive or interesting for subjects
- similarly, parameters such as **SI_{mean} and TI_{mean}** cannot be used for this purpose
- high correlation between visual attractive scores and 3D effect experience scores - the difficulty in assessing these two experiences separately
- **NASA-TLX questionnaire results**: low scores (mean score approx. 2.7) confirm that the experimental **task was not demanding** for participants and justifies extending the total number of sequences in future work
- **Future work**: the influence of subjective characteristics on QoE assessment scores

REFERENCES

1. P. Kortum and M. Sullivan, "**The Effect of Content Desirability on Subjective Video Quality Ratings**," Human Factors: The Journal of the Human Factors and Ergonomics Society, vol. 52, no. 1, pp. 105–118, may 2010.
2. Y. Zhu, I. Heynderickx, and J. A. Redi, "**Understanding the role of social context and user factors in video Quality of Experience**," Computers in Human Behavior, vol. 49, pp. 412–426, 2015.
3. M. H. Pinson, M. Barkowsky, and P. Le Callet, "**Selecting scenes for 2D and 3D subjective video quality tests**," EURASIP Journal on Image and Video Processing, vol. 2013, p. 50, 2013.
4. M. H. Pinson, L. Janowski, and Z. Papir, "**Video Quality Assessment: Subjective testing of entertainment scenes**," IEEE Signal Processing Magazine, vol. 32, no. 1, pp. 101–114, jan 2015.
5. ITU-T, "**Recommendation ITU-T P.910: Subjective video quality assessment methods for multimedia applications**," 2008.

ACKNOWLEDGMENT