

Redefining ITU-T P.912 Recommendation Requirements for Subjects of Quality Assessments in Recognition Tasks

Mikołaj I. Leszczuk Joel Dumke Lucjan Janowski Artur Koń

Quality Assessment for Recognition Tasks (QART)
Video Quality Experts Group (VQEG)

December 9, 2011

Table of contents

- 1 Introduction
- 2 Experimental Method
 - Source Video Sequences
 - Recognition Test-Plans
- 3 Results
- 4 Conclusion on Standardization

Introduction

- Problems of quality measurements for task-based video partially addressed in Recommendation ITU-T P.912
 - Title: “Subjective Video Quality Assessment Methods for Recognition Tasks”
 - Published: 2008
 - Introducing:
 - Basic definitions
 - Methods of testing
 - Psycho-physical experiments
- Section 7.3 (“Subjects”): “Subjects who are experts in the application field of the target recognition video should be used.”
- Nevertheless, to best authors’ knowledge, expert viewer issue not well verified in specific academic research
- Consequently, we compared groups of subjects assessing video quality for task-based video

Is Subjects' Proficiency Necessary?



Expert subject

- Costly (practitioner):
 - Police officer
 - Doctor
 - Difficult to hire



Non-expert subject

- Cheap (colleague/friend)
 - Student
 - Retired
 - Easy to hire

Figure: *Do I really need to be a security officer in order to participate in a test checking my ability to read license plate numbers in compressed video?*

Table of contents

- 1 Introduction
- 2 **Experimental Method**
 - **Source Video Sequences**
 - Recognition Test-Plans
- 3 Results
- 4 Conclusion on Standardization

Four Categories of Lighting Condition Scenarios

- 1 Outdoor, daytime light
- 2 Indoor, bright with flashing lights
- 3 Indoor, dim with flashing lights
- 4 Indoor, dark with flashing lights

Three Different Distances Used for Clips Creating

- 1 5.2 meters for indoor scenarios
- 2 10.9 meters for outdoor scenarios, objects close
- 3 14.6 meters for outdoor scenarios, objects far

Table of contents

- 1 Introduction
- 2 Experimental Method
 - Source Video Sequences
 - Recognition Test-Plans
- 3 Results
- 4 Conclusion on Standardization

Viewing Conditions of Room Where Test Took Place Following ITU-R BT.500-12 and ITU-T P.910

- Ratio of luminance of inactive screen to peak luminance: ≤ 0.02
- Ratio of the luminance of the screen, when displaying only black level in a completely dark room, to that corresponding to peak white: ≈ 0.01
- Maximum observation angle relative to the normal (this number applies to CRT displays, whereas the appropriate numbers for other displays are under study): 30°
- Ratio of luminance of background behind picture monitor to peak luminance of picture: ≈ 0.15
- Other room illumination: *low*

Arranged Viewing Conditions



Figure: Test environment

Example of User Interface



Figure: User interface for subjective target recognition task test performed

NTIA Test-Plan (1/2)

- NTIA performed the object recognition tests with two groups of viewers
- The Practitioner group
 - All subjects were volunteers and weren't paid for the test
 - Most received invitational travel to Boulder, CO
 - All of them had experience in public safety, including:
 - Police
 - Firemen
 - EMS
 - Forensic Video Analysts
 - Very few were outside the range of 30-60 years old
 - Three had minor color vision problems — their results were not significantly different

NTIA Test-Plan (2/2)

- The Non-Practitioner group
 - Subjects having no experience in image recognition
 - All subjects were paid through a temp agency to take the test
 - None of them had experience in public safety
 - Subjects had a wide variety of ages, but skewed young
 - Two had minor color vision problems — their results were not significantly different

AGH Test-Plan

- Subjects having no experience in image recognition
- All subjects volunteers and weren't paid for their job
- None of them with experience in public safety area

- Almost all subjects 20-25 years old
- One of them with color vision problems — did no worse than other viewers so his results included

Introduction to Results

- 15540 answers were totally given in each experiment
- **Artur, some more slide text here!**

Conclusion on Standardization

- Comparison developed for task-based video
- Specifications amendments for ITU-T P.912 Recommendation developed
- Consequently first sentence of Section 7.3 (“Subjects”) of ITU-T P.912 to get rephrased into: “Subjects who are motivated should be used.”
- Assisting researchers of task-based video quality to identify subjects that will allow them to successfully perform psychophysical experiment required

Acknowledgement

- European Community's Seventh Framework Program (FP7/2007-2013) for received funding for research leading to these results under grant agreement №. 218086 (INDECT)
- U.S. Department of Commerce for received funding for research under Public Safety Communications Research project
- The Department of Homeland Security's Office of Interoperability and Compatibility, and NIST's Office of Law Enforcement Standards