



POLITÉCNICA



# Quality assessment of immersive media: Recent activities within VQEG

Jesús Gutiérrez

*Grupo de Tratamiento de Imágenes (GTI)*

*Universidad Politécnica de Madrid (UPM)*

[jesus.gutierrez@upm.es](mailto:jesus.gutierrez@upm.es)

<https://www.gti.ssr.upm.es>

[www.vqeg.org](http://www.vqeg.org)

- Presentation of VQEG and IMG.
- Subjective Evaluation of Visual Quality and Simulator Sickness of Short 360° Videos: ITU-T Rec. P.919.
- Ongoing and future work:
  - Evaluation of QoE with long 360° Videos.
  - Evaluation of immersive and interactive communication systems.

# Video Quality Experts Group

## VQEG

- **VQEG provides a forum**, via email lists and face-to-face meetings for video quality assessment experts **to exchange information** and **work together** on common goals.
- Main objectives:
  - To **formulate test plans** that clearly and **specifically define the procedures** for performing **subjective assessment tests** and **objective models validations**.
  - To produce **open-source databases** of multimedia material and test results, as well as **software tools**.
  - To **conduct subjective studies** of multimedia and immersive technologies and **provide a place for collaborative model developments**.
- VQEG is open to all interested from **industry, academia, government organizations** and Standard-Developing Organizations (**SDOs**). There are no fees involved, no membership applications and no invitations are needed.
  - Subscription to the main VQEG email list ([ituvidq@vqeg.org](mailto:ituvidq@vqeg.org)) constitutes membership.
  - Check the website for more details: <http://www.vqeg.org/>

- **Mission:** Quality assessment of immersive media.
- **Technologies:**
  - 360-degree content / Virtual Reality.
  - Augmented / Mixed reality.
  - 3D, FVV, multiview technologies, light field.
- **Goals:** Baseline quality assessment of immersive systems.
  - Generate and publish **open-source datasets** of immersive media content.
  - Develop **subjective test methods** and provide **QoE guidelines**.
  - Validation and development of **objective metrics**.
- **Chairs:** Jesús Gutiérrez (UPM), Pablo Pérez (Nokia Bell Labs), Zhenzhong Chen (Wuhan University).
- **Website:** <https://www.its.blrdoc.gov/vqeg/projects/immersive-media-group.aspx>

- **Some works and presentations:**

- 360° video
  - “Project Vertigo: monitoring sickness and discomfort in high-motion 360 video”, Pablo Perez (Nokia Bell-Labs)
  - “Subjective evaluation of 360 degree video quality during head movement”, Glenn Van Wallendael (Ghent University)
  - “A dataset of head and eye movements for 360 degree images”, Jesús Gutiérrez (University of Nantes)
  - “QoE of Omnidirectional (360°) Videos”, Ashutosh Singla (Audiovisual Technology Group, TU Ilmenau)
  - “A framework for adaptive delivery of omnidirectional video”, Christian Timmerer (University of Klagenfurt)
  - “Spherical SSIM for Objective Omnidirectional Video Quality Assessment”, Zhenzhong Chen (Wuhan University)
  - “Video Multimethod Assessment Fusion (VMAF) on 360VR”, Narciso García (Universidad Politécnica de Madrid).
- VR
  - “Quality of Experience for a Virtual Reality simulator”, Kjell Brunnström (RISE Acreo AB)
  - “Evaluation of immersion and presence in VR”, Francesca De Simone (CWI)
  - “Measuring User Quality of Experience in Social VR Systems”, Francesca de Simone (CWI).
- AR
  - “Perceptual Characterization of 3D Graphical Contents based on Visual Attention patterns”, Mona Abid (Univ. of Nantes)
  - “Comparison of Subjective Methods for Quality Assessment”, Yana Nehmé (LIRIS/INSA Lyon)
- FVV
  - “3 datasets for quality assessment in the context of Free Viewpoint Video”, Suiyi Ling (University of Nantes)
  - “Quality Assessment of FTV”, Patrick Le Callet (Université de Nantes).
  - “Considerations on FTV quality assessment”, Narciso García (Universidad Politécnica de Madrid)
- LF
  - “Characterization and selection of light field content for perceptual assessment”, Jesús Gutiérrez (University of Nantes)
  - “No reference quality evaluation of light field content based on structural representation of the epipolar plane image”, Ali Ak (U. Nantes)
- ... Many more in the public ftp (check the meetings): <https://www.its.bldrdoc.gov/vqeg/meetings-home.aspx>

- Presentation of VQEG and IMG.
- **Subjective Evaluation of Visual Quality and Simulator Sickness of Short 360° Videos: ITU-T Rec. P.919.**
- Ongoing and future work:
  - Evaluation of QoE with long 360° Videos.
  - Evaluation of immersive and interactive communication systems.

- Global target of the test plan:
  - Design and execute a **cross-lab test** where we can assess and validate **subjective evaluation methodologies** for 360° video.
  - Contribute to the **standardization** of such methodologies under ITU-T P.919 (Q13/12).
  - Generate a **dataset** of subjectively assessed content for future research.
- What was covered:
  - Assessment of **video quality** and **simulator sickness**...
  - ... of **short** sequences ( $\leq 30$ s), in the spirit of ITU-R BT.500, ITU-T P.910, etc.
- More details in our recently published paper:
  - J. Gutiérrez, P. Pérez, M. Orduna, A. Singla, C. Cortés, P. Mazumdar, I. Viola, K. Brunnström, F. Battisti, N. Cieplińska, D. Juszka, L. Janowski, M. Leszczuk, A. Adeyemi-Ejeye, Y. Hu, Z. Chen, G. Van Wallendael, P. Lambert, C. Díaz, J. Hedlund, O. Hamsis, S. Fremerey, F. Hofmeyer, A. Raake, P. César, M. Carli, N. García, "Subjective evaluation of visual quality and simulator sickness of short 360° videos: ITU-T Rec. P.919", *IEEE Transactions on Multimedia*, Jul. 2021 (Early access).

- **Audiovisual Quality**

- Test methodology
  - ACR vs DCR
- Sequence duration
  - 10s vs 20s
  - 20s vs 30s
  - 10s vs 30s
- Test setup
  - HTC Vive vs Samsung Gear VR vs HTC Vive Pro and Tethered vs Untethered
  - Scoring app vs “Saying out loud”
  - Audio vs no audio

- **Simulator sickness**

- When/how to assess simulator sickness
- Short vs. long questionnaires
  - One single question
  - Reduced SSQ











- Distribution of test conditions and participant labs

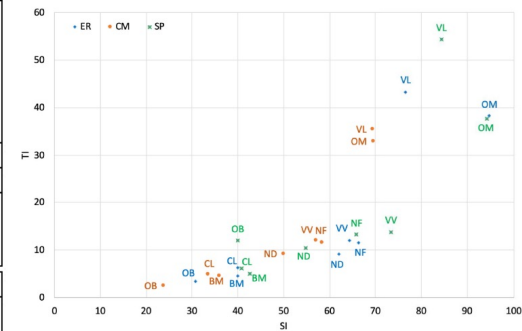
ID	Test Condition	Methodology	Lab	HMDs	Num. of PVSs	PVSs' Length
A	Video duration	ACR	Wuhan	HTC Vive	64	10s & 20s
B	Video duration	ACR	AGH	Oculus Rift	40	20s & 30s
C	Video duration	DCR	Roma3	HTC Vive	40	10s & 20s
D	Video duration	DCR	CWI	Oculus Rift	30	20s & 30s
E	Video duration	ACR	Surrey	HTC Vive	48	10s & 30s
F	Influence of HMD (desktop/mobile, High/low resolution)	ACR	UPM & Nokia	GearVR vs. HTC Vive vs. HTC Vive Pro	48	20s
G	Influence of HMD (Tethered vs. untethered)	ACR	Ghent	HTC Vive Pro	48	20s
H	Influence of audio (Videos with vs. without audio)	ACR	RISE	HTC Vive	48	20s
I	Influence of scoring method (App. vs. voice)	ACR	TU Ilmenau	HTC Vive Pro	48	20s

# Subjective evaluation of short 360° videos

ITU-T Rec. P.919

- Test set: Source sequences

Name (ID)	NokiaDojo (ND)*	NokiaFlamenco (NF)	CheerLeading (CL)*	BrazilMusic (BM)
Screenshot				
Resolution	3840x2160, 30fps	3840x2160, 30fps	4096x2048, 25fps	4096x2048, 25fps
Provider	Nokia	Nokia	TU Ilmenau	TU Ilmenau
Description	Video of an indoor sport course, with ambient audio. Contains stitching artifacts.	Indoor dance course, with ambient audio. Contains stitching artifacts.	Cheerleading session indoors, with ambient audio.	Indoor scene of a band playing Brazilian music. With audio.
Name (ID)	VSenseLuther (VL)	VSenseVaude (VV)	OculusMotion (OM)	OculusBeach (OB)*
Screenshot				
Resolution	4096x2048, 30fps	4096x2048, 30fps	3840x1920, 30fps	3840x1920, 30fps
Provider	VSense	VSense	Oculus	Oculus
Description	Video with animation content and a main character. Contains various shots (indoors and outdoors) and audio.	Video where a girl speaks to the camera. Contains audio and various indoor and outdoor shots.	Camera moving in a city. Contains music and two shots: one in daylight and one at night.	Scene with music of a beach at sunset with people dancing and moving.



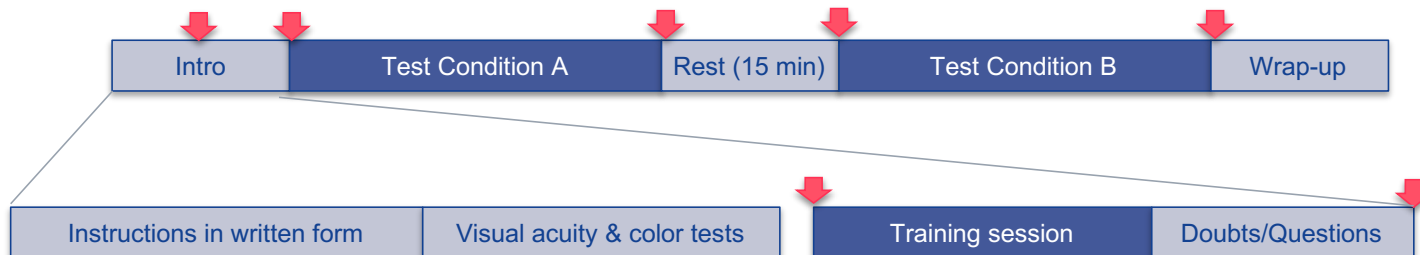
- Wide range of content characteristics.
- SI/TI computed over equirectangular (ER), cube-map (CM) and spherical (SP), to account for possible projection distortions [De Simone, 2019].

- Test set: Encoded sequences
  - 8 HRCs (including Hidden Reference):
    - Kvazaar encoder, HEVC, fixed QP.
    - 4 homogeneous QPs: 15 (HR), 22, 32, 42.
    - 4 non-homogeneous QPs (tiles).

#Tiles	Transition	ROI	QPs							
8x5**	Smooth	90°	42	37	32	22	22	32	37	42
6x3	Smooth	120°	42	32	22	22	32	42		
8x5**	Abrupt	180°	42	42	22	22	22	22	42	42
6x3	Abrupt	120°	37	37	22	22	37	37		

- Session structure and duration:
  - Each subject evaluates the same test clip under two different conditions:
    - Use within-subject statistics for better statistical power.
  - Each condition is tested in one active period:
    - After each active session: rest period of 15 min.
  - Red arrows: score simulator sickness.
  - App Miro360: Developed, tested, and validated subjective assessment SW:
    - For desktop (HTC and Oculus) and mobile (GearVR and Daydream).

C. Cortés, P. Pérez, and N. García,  
“Unity3D-based app for 360VR  
subjective quality assessment with  
customizable questionnaires”, IEEE  
ICCE 2019.  
<https://git.gti.ssr.upm.es/pub/Miro360>



### • Simulator Sickness:

- In each measurement, subjects scored both:
  - First, a single question (1-5 scale): always the same (A or B), randomly assigned
  - Second, SSQ [Kennedy *et al.*, 1993]

A. Are you feeling any sickness or discomfort now?  
(Vertigo scale) [Pérez *et al.*, 2018]

No problem	No perceptible effect, natural feeling
Light effects	Slight discomfort, but no sickness
Uncomfortable	Moderate discomfort, but tolerable for a while
Unpleasant	Strong discomfort or sickness, but can continue the test
Unbearable	Strong discomfort or sickness, and want to stop test

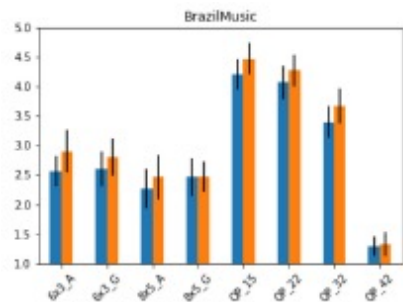
B. How is your level of dizziness or nausea? (Short-SSQ) [Tran *et al.*, 2017]

Very dizzy
Dizzy
Slightly dizzy
Not dizzy
Absolutely not dizzy

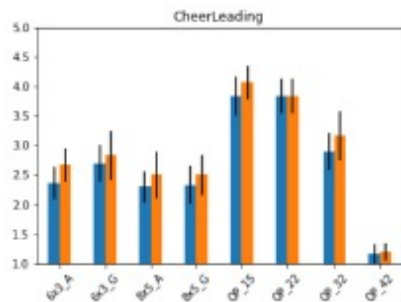
- Observers
  - At least 28 participants per lab.
  - Visual screening: Visual acuity and color vision.
  - Subjects were sitting in a swivel chair.

Lab	Test ID	Number of Observers			Age			Experience with VR Headsets				
		Total	Female	Male	Min	Max	Avg	Times=1	Times<5	5<Times<20	Times> 20	Every day
Wuhan	A	30	15	15	20	30	24.5	8	15	7	0	0
AGH	B	40	13	27	18	79	28.5	13	17	8	2	0
Roma3	C	30	8	22	21	57	30.6	7	10	2	8	2
CWI	D	28	14	14	21	60	27.6	2	12	5	6	3
Surrey	E	31	10	21	19	44	25.9	13	12	3	2	1
UPM & Nokia	F	60	25	35	20	31	23.2	18	32	9	1	0
Ghent	G	30	4	26	23	45	31.6	3	14	7	5	1
RISE	H	28	16	12	22	66	41.6	3	16	8	1	0
TU Ilmenau	I	29	14	15	20	37	25.9	4	18	4	3	0
Total		306	119	187	18	79	28.8	71	146	53	28	7
			38.9%	61.1%				23.20%	47.71%	17.32%	9.15%	2.29%

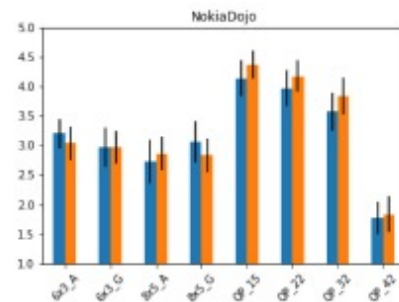
- Results of audiovisual quality: Wuhan, ACR, videos of 10s (blue) and 20s (orange).



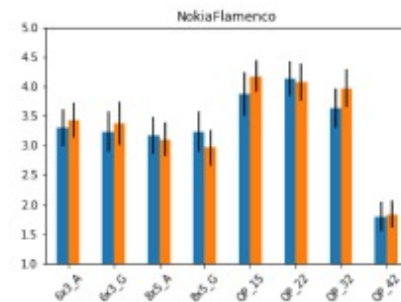
(a)



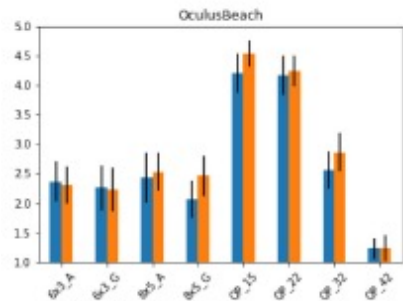
(b)



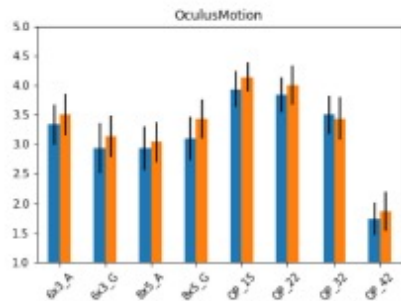
(c)



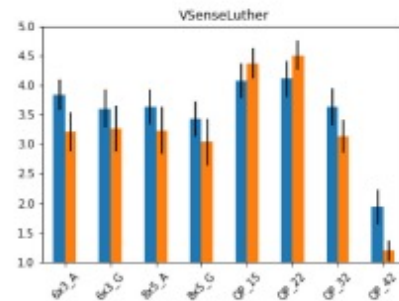
(d)



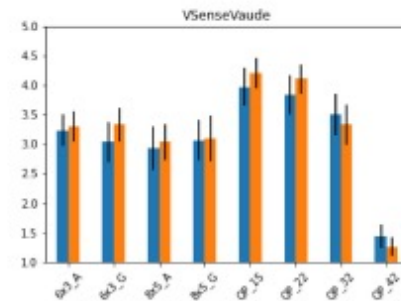
(e)



(f)



(g)

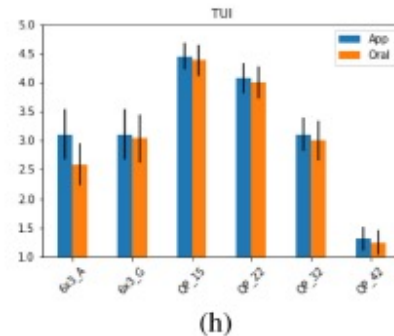
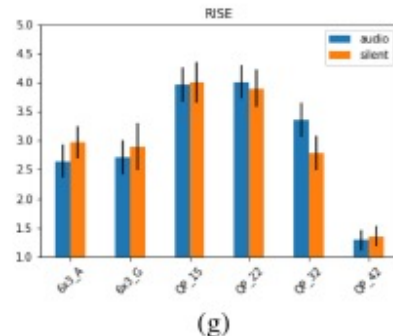
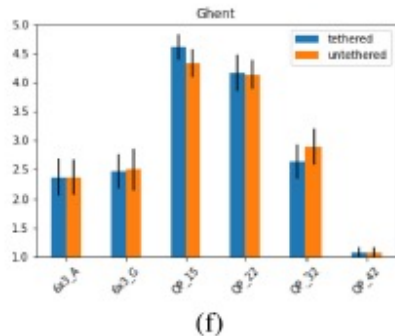
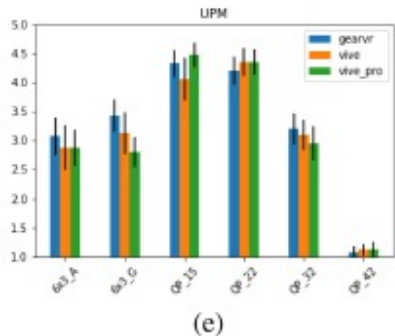
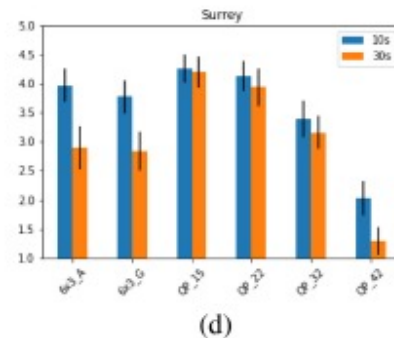
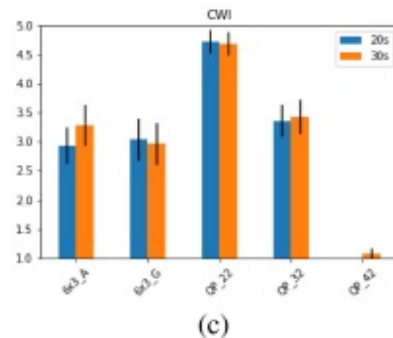
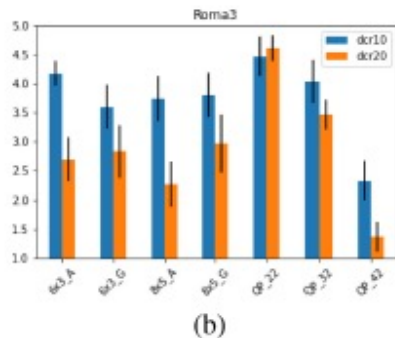
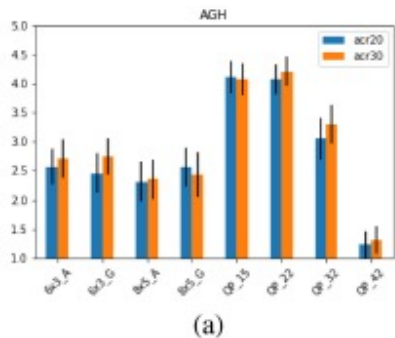


(h)

# Subjective evaluation of short 360° videos

ITU-T Rec. P.919

- Results of audiovisual quality: All laboratories (all tested conditions), VSenseLuther.

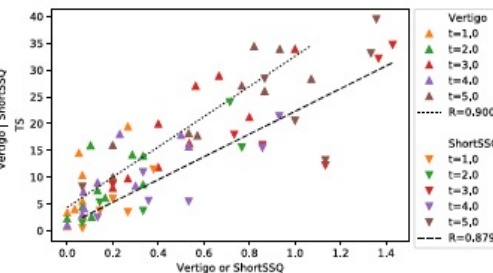
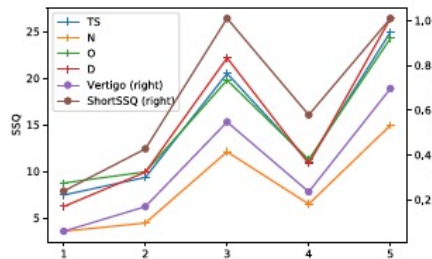
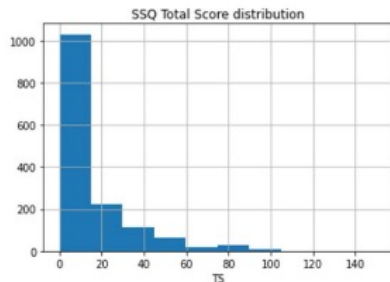




# Subjective evaluation of short 360° videos

ITU-T Rec. P.919

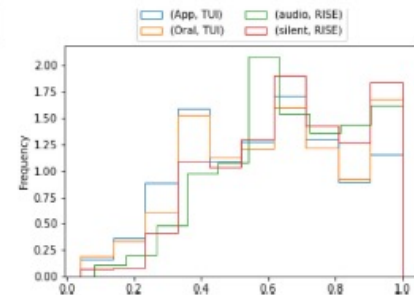
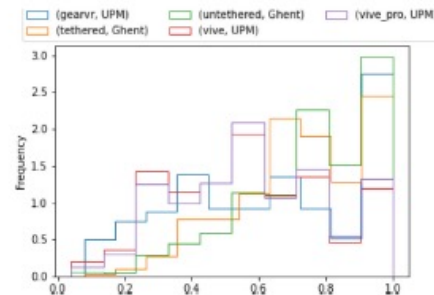
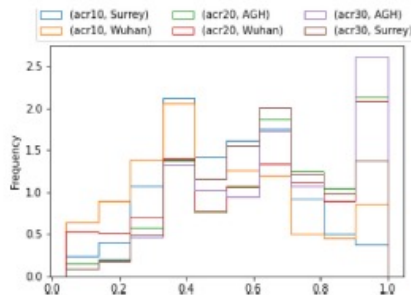
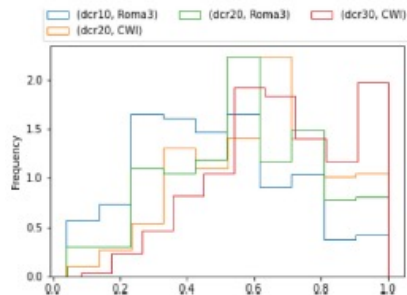
## Results of SSQ:



Can we use a subset of the SSQ?  
PLCC among different SSQ

Questionnaire	SSQ	VRSQ	CSQ	New-FA
SSQ	1.000	0.958	0.918	0.951
VRSQ	0.958	1.000	0.870	0.905
CSQ	0.918	0.870	1.000	0.878
New-FA	0.951	0.905	0.878	1.000

## Exploration results:



- Outcomes included in the recommendation ITU-T P.919:
  - Methodology: Both **ACR** and **DCR** can be used for assessment of coding quality.
  - Duration of videos: subjective tests of coding degradations with 360° videos can be done with sequences of **10 seconds**.
  - HMD: any commercial HMD (tethered or untethered) can be used in visual quality tests (providing enough resolution and refresh rate)
  - Audio: it is possible to use test stimuli either with or without audio to evaluate visual quality (spatial audio was not considered).
  - Method to collect ratings: Voting interfaces and verbal voting.
  - Minimum number of observers: 28 participants.
  - SSQ: both Vertigo scale and VRSQ [Kim *et al.*, 2018] can be alternatives to SSQ.
- **Dataset will be available soon** in <https://www.its.bldrdoc.gov/vqeg/video-datasets-and-organizations.aspx>

- Presentation of VQEG and IMG.
- Subjective Evaluation of Visual Quality and Simulator Sickness of Short 360° Videos: ITU-T Rec. P.919.
- Ongoing and future work:
  - **Evaluation of QoE with long 360° Videos.**
  - Evaluation of immersive and interactive communication systems.

# Ongoing and future work

## Evaluation of QoE with long 360° Videos

- After studying the evaluation of short videos → Long 360° videos.

- Definition: a short (2-10 minutes) full content item.
- Requires content immersive evaluation [Pinson *et al.*, 2014].

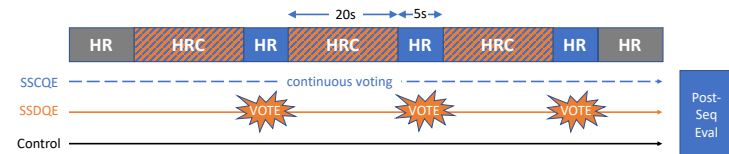
M. Orduna, P. Pérez, J. Gutiérrez, N. García,  
“Methodology to Assess Quality, Presence,  
Empathy, Attitude, and Attention in Social  
VR: International Experiences Use Case”,  
arXiv:2103.02550

- Evaluation:

- Single-Stimulus Continuous Quality Evaluation [ITU-T P.910] → In-sequence
- Single-Stimulus Discrete Quality Evaluation* (SSDQE) [Gutiérrez *et al.*, 2012] → In-sequence
- ACR → Post-sequence

- Outcomes:

- SSDQE provides similar results to what ACR would provide.
- SSCQE → to be analyzed.
- ACR is significantly affected by in-sequence evaluation.
- Quality evaluation does not affect sense of presence.



SSDQE / SSCQE (?) / ACR are valid  
for long sequences

- Presentation of VQEG and IMG.
- Subjective Evaluation of Visual Quality and Simulator Sickness of Short 360° Videos: ITU-T Rec. P.919.
- Ongoing and future work:
  - Evaluation of QoE with long 360° Videos.
  - **Evaluation of immersive and interactive communication systems.**

## Evaluation of immersive and interactive communication systems

- How can we test bi-directional immersive communication systems?
  - Evaluate effect of **technical factors** (e.g., latency / bitrate / etc.) in **QoE** (e.g., immersion, interactivity, etc.)
  - Compare with other systems / experiments.
- Related recommendations:
  - ITU-T P.920 - Interactive test methods for audiovisual communications.
    - Some tasks proposed to evaluate effect of technical factors of video-conference services (05/2000):
      - E.g.: one of the subjects shows and describes a plastic building block and the other one is required to reproduce it.
  - ITU-T P.13xx series: telemeeting assessment.
  - ITU-T P.QXM - QoE Assessment of eXtended Reality (XR) Meeting.
    - Best practices for QoE assessment of tele-meetings with XR elements.

## Evaluation of immersive and interactive communication systems

- Test plan:
  - Run a cross-lab experiment using any available collaborative immersive technology:
    - “The same” experiment in completely different setups
  - Gather a set of immersive communication systems:
    - Real-time 360° / Free-viewpoint video telepresence.
    - Social VR with pointcloud transmission / with avatars.
    - XR collaboration.
    - ...
  - Create an experiment that covers all basic functionalities:
    - Conversation between people.
    - Discussion about objects in the immersive space.
    - Interaction with (local / remote / virtual) objects in the immersive space.

# Ongoing and future work

## Evaluation of immersive and interactive communication systems

- Identified technologies:



### FVV Live (GTI-UPM)

Narciso García

[narciso@gti.ssr.upm.es](mailto:narciso@gti.ssr.upm.es)



### VRTogether (CWI)

Pablo César

[p.s.cesar@cwi.nl](mailto:p.s.cesar@cwi.nl)



### The Owl (Nokia Bell Labs)

Pablo Pérez

[pablo.perez@nokia-bell-labs.com](mailto:pablo.perez@nokia-bell-labs.com)



## Evaluation of immersive and interactive communication systems

- Interested groups:
  - UPM (Spain)
  - Nokia Bell-Labs (Spain)
  - CWI (The Netherlands)
  - TU Ilmenau (Germany)
  - Ghent University (Belgium)
  - RISE (Sweden)
  - Wuhan University (China)
  - University of Surrey (United Kingdom)
  - AGH University of Science and Technology (Poland)
  - ...
  - **You?**



# Acknowledgment and references

- Pablo Pérez
- Marta Orduna
- Ashutosh Singla
- Carlos Cortés
- Pramit Mazumdar
- Irene Viola
- Kjell Brunnström
- Federica Battisti
- Natalia Cieplinska
- Dawid Juszka
- Lucjan Janowski
- Mikolaj Igor Leszczuk
- Anthony Adeyemi-Ejeye
- Yaosi Hu
- Zhengzhong Chen
- Glenn Van Wallendael
- Peter Lambert
- Cesar Diaz
- John Hedlund
- Omar Hamsis
- Stephan Fremerey
- Frank Hofmeyer
- Alexander Raake
- Pablo Cesar
- Marco Carli
- Narciso García
- Philip Corriveau
- Patrick Le Callet
- ...
- F. De Simone, **J. Gutiérrez**, P. Le Callet, "Complexity measurement and characterization of 360-degree content", *Human Vision and Electronic Imaging Conference*, Jan. 2019.
- **J. Gutiérrez**, P. Pérez, M. Orduna, A. Singla, C. Cortés, P. Mazumdar, I. Viola, K. Brunnström, F. Battisti, N. Cieplińska, D. Juszka, L. Janowski, M. Leszczuk, A. Adeyemi-Ejeye, Y. Hu, Z. Chen, G. Van Wallendael, P. Lambert, C. Díaz, J. Hedlund, O. Hamsis, S. Fremerey, F. Hofmeyer, A. Raake, P. César, M. Carli, N. García, "Subjective evaluation of visual quality and simulator sickness of short 360° videos: ITU-T Rec. P.919", *IEEE Transactions on Multimedia*, Jul. 2021 (Early access).
- R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness", *Int. J. Aviation Psychol.*, vol. 3, no. 3, pp. 203–220, Nov. 1993.
- P. Pérez, N. Oyaga, J. J. Ruiz, and A. Villegas, "Towards systematic analysis of cybersickness in high motion omnidirectional video", *Int. Conf. Qual. Multimedia Experience*, 2018.
- H. T. T. Tran, N. P. Ngoc, C. T. Pham, Y. J. Jung, and T. C. Thang, "A subjective study on QoE of 360 video for VR communication", *IEEE Int. Workshop Multimedia Signal Process.*, 2017.
- H. K. Kim, J. Park, Y. Choi, and M. Choe, "Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a VR environment", *Appl. Ergonom.*, vol. 69, pp. 66–73, May 2018.
- M. H. Pinson, M. Sullivan, and A. Catellier, "A New Method for Immersive Audiovisual Subjective Testing", *Int. Work. Video Process. Qual. Metrics Consum. Electron.*, Jan. 2014.
- **J. Gutierrez**, P. Perez, F. Jaureguizar, J. Cabrera, and N. Garcia, "Validation of a novel approach to subjective quality evaluation of conventional and 3D broadcasted video services", *International Workshop on Quality of Multimedia Experience*, Jul. 2012.
- M. Orduna, P. Pérez, **J. Gutiérrez**, N. García, "Methodology to Assess Quality, Presence, Empathy, Attitude, and Attention in Social VR: International Experiences Use Case", arXiv:2103.02550.
- P. Carballeira, C. Carmona, C. Díaz, D. Berjón, D. Corregidor, J. Cabrera, F. Morán, C. Doblado, S. Arnaldo, M.M. Martín, N. García, "FVV Live: A real-time free-viewpoint video system with consumer electronics hardware", *IEEE Transactions on Multimedia*, 2021.
- D. Berjón, P. Carballeira, J. Cabrera, C. Carmona, D. Corregidor, C. Díaz, F. Morán, N. García, "FVV Live: real-time, low-cost free viewpoint video", *IEEE Int. Conf. on Multimedia and Expo Workshops*, Jul. 2020.
- R. Kachach, M. Orduna, J. Rodríguez, P. Pérez, A. Villegas, J. Cabrera, N. Garcia, "Immersive Telepresence in Remote Education", *International Workshop on Immersive Mixed and Virtual Environment Systems*, Sep. 2021.
- R. Kachach, D. González-Morín, F. Pereira, P. Pérez, I. Benito, J. Ruiz, E. González, A. Villegas, "A Multi-Peer, Low Cost Immersive Communication System for Pandemic Times", *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, Mar. 2021.
- S. Rossi, I. Viola, J. Jansen, S. Subramanyam, L. Toni, and P. Cesar, "Influence of Narrative Elements on User Behaviour in Photorealistic Social VR", *International Workshop on Immersive Mixed and Virtual Environment Systems*, Sep. 2021.
- R. Mekuria, K. Blom, and P. Cesar Design, "Implementation and Evaluation of a Point Cloud Codec for Tele-Immersive Video", *IEEE Transactions on Circuits and Systems for Video Technology*, 27(4), pp. 828-842, Apr. 2017.



POLITÉCNICA

VQEG

# Quality assessment of immersive media: Recent activities within VQEG

Jesús Gutiérrez

*Grupo de Tratamiento de Imágenes (GTI)*

*Universidad Politécnica de Madrid (UPM)*

[jesus.gutierrez@upm.es](mailto:jesus.gutierrez@upm.es)

<https://www.gti.ssr.upm.es>

[www.vqeg.org](http://www.vqeg.org)