Complexity measurement and characterization of 360-degree content

VQEG Meeting, March 2019, Berlin

Francesca De Simone
Distributed & Interactive Systems Group
Centrum Wiskunde & Informatica
Amsterdam, the Netherlands

Jesús Gutiérrez
Patrick Le Callet
Image, Interaction, Perception Group (IPI)
Laboratoire des Sciences du Numérique de Nantes (LS2N)
Ecole Polytechnique de l’Université de Nantes

Paper at IS&T Electronic Imaging 2019 Symposium
Content characterization

Appropriate content selection is crucial to:

- **Validate and benchmark** image/video compression/streaming/processing/etc. algorithms
- Design “useful” **subjective quality** assessment tests
- Analyze the results from **objective quality** metrics

“Subjective video quality assessment is a well-understood field, yet **scene selection is often driven by convenience or content availability.**” from “Selecting scenes for 2D and 3D subjective video quality tests,” M. Pinson et al., EURASIP J. Image Video Process., 2013

Content selection should be done according to:

- Application under study
- Purpose of the task (perceptual evaluation, system benchmarking, etc.)
- **Representative content characteristics**!
Related work on characterization of 2D content:

- Type of content:
  - Natural vs. animation
  - Realistic application (e.g., movies, sports, etc.) vs. system-performance evaluation (e.g., coding development and standardization)

- Duration, resolution, framerate…

- Complexity properties:
  - Spatial complexity: Spatial indicator (SI) [ITU-T P.910]
  - Temporal complexity: Temporal indicator (TI) [ITU-T P.910]
  - Colorfulness, contrast, etc.

Related work on characterization of Stereoscopic 3D content:

- Horizontal disparities.
- Depth budget.
- Motion in depth plane.
- Disparity changes (e.g., DTI).
- Occlusions.
- Visual comfort and fatigue.

“Efficient measurement of stereoscopic 3D video content issues”, S. Winkler, EI 2014.


perceived depth in natural images and study of its relation with monocular and binocular depth cues, P. Lebreton et al. EI 2014
Related work on characterization of Light Field content and Free Viewpoint TV:

- Occluded Pixels
- Refocusing range


- Critical Trajectories: Hypothetical Rendering Trajectories

A Study on the Impact of Visualization Techniques on Light Field Perception, Battisti et al. EUSIPCO 2018
Prediction of the Influence of Navigation Scan-path on Perceived Quality of Free-Viewpoint Videos, Li et al. IEEE JETCAS.
Related work on characterization of HDR and WCG content:

- « perceptual difference » due to Tone Mapping Operator

- « perceptual difference » due to gamut reduction
  A Perception-Based Framework for Wide Color Gamut Content Selection, Lee et al. ICIP 2017.
Characterization of 360-degree content

- Omnidirectional signal acquisition:

- Applications:
  - Robotics
  - Virtual Reality (puts viewer in a virtual scene)
Characterization of 360-degree content

New technical and perceptual factors must be considered
- Geometrical domain of the signal
- Interactive navigation by users

Previous work in the SoA (service provider viewpoint):

Open questions:
- Content characterization to be performed in which domain?
- Can content be characterized based on visual attention and navigation patterns?
Characterization of 360-degree content

Open questions:
- **Content characterization to be performed in which domain?**
- Can content be characterized based on visual attention and navigation patterns?
Spherical versus planar representation

- To use existing algorithms it is convenient to map the spherical signal in planar domain, for example:
Spherical versus planar representation

Sampling?
Spherical versus planar representation

Any projection introduces distortions or discontinuities, so the signal is planar but its characteristics may vary… Does this influence content characterization?
Our case study

Salient360! Dataset
- 85 equirectangular images
- Processed data from **head and eye movements**:
  - Head saliency maps.
  - Head-Eye saliency maps.
  - Head-Eye scanpaths
  - Head scanpaths/trajectories
- To access it: [https://salient360.ls2n.fr](https://salient360.ls2n.fr)
- Used for the ICME Grand Challenges 2017 and 2018.
Same measure, different domains & dependency on viewing direction: example

- Spatial perceptual information: $SI = \text{std}[\text{Sobel}(\text{image}_Y)]$

ITU-T P.910 “Subjective video quality assessment methods for multimedia applications”

Equirectangular
Overall SI = 15.8

Cubemap
Overall SI = 11.6
- UpFace SI = 10.3
- FrontFaceSI = 10.1
- LeftFaceSI = 9.9
- RightFaceSI = 10.4
- BackFaceSI = 7.6
- BottomFaceSI = 21.2
Same measure, different domains

Scatter plot of SI value computed on equirectangular vs cubemap

Example of inconsistent ranking

- Equirectangular SI:
  Lower complexity - Higher complexity

- Cubemap SI:
  Lower complexity - Higher complexity
Dependency with viewing direction

- One single value might not be very informative to characterize the entire 360-images

Boxplot of SI values per cube face on entire set of images

Boxplot of SI values per image on all cube faces
Dependency with viewing direction

- One single value might not be very informative to characterize the entire 360-images

Image in the set with lowest SI variability across faces

Image in the set with highest SI variability across faces

Boxplot of SI values per image on all cube faces
Characterization of 360-degree content

Open questions:
- Content characterization to be performed in which domain?
- Can content be characterized based on visual attention and navigation patterns?
Related work on characterization of 2D content:

Visual attention complexity:

- **Inter-Observer Congruency/Agreement (IOC):**
  “Prediction of the inter-observer visual congruency (IOVC) and application to image ranking,” O. Le Meur and al. *ACM MM 2011*.
  - Requires fixations per observer!
  - ↓IOC, nothing stands out → **Exploratory** images
  - ↑IOC, things attracting attention → **Focused** images

  - ↓Entropy, saliency is concentrated → **Focused** images
  - ↑Entropy, saliency is diffused → **Exploratory** images

- Generally good correlation between both measures
Using ground-truth eye+head tracking data

IOC on scanpaths and entropy on the head+eye saliency maps (sphere sampling)

2nd most **focused** image by Entropy, 3rd by IOC.

Most **exploratory** image by Entropy, 2nd by IOC.

SROCC = -0.7

But some inconsistencies...

Ranking difference of 34 between Entropy and IOC.
Using ground-truth eye+head tracking data

IOC on head trajectories and entropy on the head saliency maps (sphere sampling):
- Good correlation between rankings from Head+Eye and Head with IOC
- Worse performance for entropy

SROCC = 0.91
SROCC = -0.49
SROCC = -0.68
Using saliency models

Can we characterize without the ground-truth tracking data (i.e., without previous exploration tests)?

- Is entropy of the predicted saliency maps a good estimator?

- Considered models: Best models in GC Salient360! 2017/18* with online available code:
  - Prediction of Head+Eye Saliency:
  
  - Prediction of Head Saliency:
  
  *https://salient360.ls2n.fr
Using saliency models

- Models can differentiate between clearly “focused” / “exploratory” images
  - Head+eye saliency predictions are more sensitive.
- Predictions of saliency models still improvable: limited use for direct ranking

Focused image:
- Entropy HE saliency model: 4.8
- Entropy H saliency model: 5.4

Exploratory image:
- Entropy HE saliency model: 7
- Entropy H saliency model: 6.3
Using saliency models

- Prediction saliency distribution in latitude and longitude: Saliency models are useful!
Conclusions

- **Geometrical domain** of the signal and **interactive navigation** by users must be considered when characterizing 360-degree content

- **Visual attention complexity** in addition to other characteristics

- **On-going work**: clustering based on Visual attention, characterization of 360-degree videos
Complexity measurement and characterization of 360-degree content

VQEG Meeting, March 2019, Berlin

Francesca De Simone
Distributed & Interactive Systems Group
Centrum Wiskunde & Informatica
Amsterdam, the Netherlands

Jesús Gutiérrez
Image, Interaction, Perception Group (IPI)
Laboratoire des Sciences du Numérique de Nantes (LS2N)
Ecole Polytechnique de l’Université de Nantes

Patrick Le Callet

Paper at IS&T Electronic Imaging 2019 Symposium