Video Multimethod Assessment Fusion (VMAF) on 360VR contents

VQEG Mountain View 2018

Marta Orduna*, César Díaz*, Lara Muñoz*, Pablo Pérez†, Ignacio Benito†, and Narciso García*

* Grupo de Tratamiento de Imágenes (GTI)
  Universidad Politécnica de Madrid

† Nokia Bell Labs
  Madrid

narciso@gti.ssr.upm.es
Presentation scheme

- Introduction
- Review of quality metrics on 360VR contents
- Work approach
- Test material
- VMAF computation
- Subjective assessment
- Results
- Conclusions
Introduction

• Main challenge:
  • to provide omnidirectional content guaranteeing an immersive experience and saving bit rate

• Main solutions:
  • Definition of different perceptible levels of quality
  • Efficient delivery schemes
  • Users’ behavior → Attention maps
  • Exploitation of peculiarities of the type of projection

• All these solutions require a quality metric
Introduction – 360VR (omnidirectional) video
Introduction – Video tiling
Review of quality metrics on 360VR contents

- Spherical - PSNR (S-PSNR)
- Weighted to Spherically - PSNR (WS-PSNR)
- Craster Parabolic Projection - PSNR (CPP-PSNR)
- Uniformly Sampled Spherical - USS-PSNR
- Multi-Scale SSIM - MS-SSIM
- VMAF
- SpatioTemporal - VMAF (ST-VMAF)
Work approach

• VMAF has provided significantly good results on different types of non-immersive contents and viewing conditions

• Research question: can VMAF be applied to omnidirectional content \textit{without making any specific adjustments}\textit{?}

• Underlying hypothesis:
  There is a monotonic relationship between 2D-VMAF and 360VR-VMAF (non-existing)

• If it is true, we can avoid:
  • generating a large and rich specific 360VR video dataset
  • carrying out numerous subjective quality assessments
  • performing the corresponding training and testing stage
Work approach

• The validation of VMAF on 360VR contents is carried out in two steps:

  • VMAF application to omnidirectional sequences encoded with constant QP in the whole range of possible values to obtain the variation of the score with the encoding parameter

  • VMAF scores validation through a subjective assessment
    VMAF-vs-QP curve is monotonically decreasing by the nature of the encoding
    → adjustment with a finite number of key operating points
Test material

- A wide range of contents selected with different features in terms of color, texture, camera motion, composition, and content in the scenes.
VMAF computation

| Number of reference videos | 9 |
|----------------------------|---|
| Duration                   | 10 seconds |
| Encoding                   | H.265/HEVC |
| Resolution                 | 4K (3840x1920) |
| Hypothetical Reference Circuits (HRCs) | QP range (1-51) |
| Framerate                  | 25 fps |

Total number of videos: 459

No temporal pooling challenge
4K throughout the process
Quality degradation vs QP

QP = 29
QP = 36
QP = 40
QP = 46
QP = 50
Subjective assessment – Test material

**DATASET CHARACTERISTICS**

| Feature                  | Value            |
|--------------------------|------------------|
| Number of source videos  | 9                |
| Duration                 | 10 seconds       |
| Encoding                 | H.265/HEVC       |
| Resolution               | 4K (3840x1920)   |
| Number of QP values      | 6                |

Total of videos: **54**

VQEG 2018 – VMAF on VR360 contents - #
Subjective assessment – Test material

- VMAF-vs-QP curve is monotonically decreasing by the nature of the encoding → VMAF can be adjusted with a finite number of QPs, which correspond to anchor VMAF scores in the curve for all the used contents

QP (A) chosen to obtain a bitrate similar to that of the original video available in database

Check: QP (A) < QP (B)
Subjective assessment – Test session

Methodology

**ACR-HR**

| Grade | Description |
|-------|-------------|
| 5     | Excellent   |
| 4     | Good        |
| 3     | Fair        |
| 2     | Poor        |
| 1     | Bad         |

- **No training session** (no reference given about max/min quality)
- All videos viewed and scored by each subject
- Duration around **15 minutes** (assuming 5 seconds for evaluation)
- **24 observers** (age between 21 and 36, average age of 26)
- All observers with normal or corrected vision
- No subject removal because of being considered an outlier

Equipment & Environment

Different randomization for each session

VQEG 2018 – VMAF on VR360 contents - #
Experimental results - MOS
Experimental results - DMOS
VMAF adjustment for 360VR contents

Good fit

Bad fit
# PLCC and RMSE between VMAF and DMOS

| CONTENT          | PEARSON (QB, QC, QD, QE, QF) | PEARSON (QB, QC, QD, QE) | RMSE (QB, QC, QD, QE, QF) | RMSE (QB, QC, QD, QE) |
|------------------|-------------------------------|---------------------------|---------------------------|-----------------------|
| AbandonedBuilding | 0.995                         | 0.997                     | 3.433                     | 1.983                 |
| Alaska           | 0.992                         | 0.994                     | 5.661                     | 2.488                 |
| Beach            | 0.992                         | 0.991                     | 4.213                     | 2.470                 |
| Caribbean Vacation | 0.961                      | 0.997                     | 6.982                     | 6.787                 |
| FemaleBasket     | 0.984                         | 1.000                     | 7.097                     | 1.764                 |
| Happyland        | 0.940                         | 0.979                     | 9.338                     | 9.991                 |
| Lions            | 0.987                         | 0.997                     | 4.029                     | 4.446                 |
| Sunset           | 0.996                         | 0.998                     | 5.016                     | 5.490                 |
| Waterfall        | 0.996                         | 0.990                     | 5.511                     | 4.295                 |
| AVERAGE          | 0.983                         | 0.994                     | 5.698                     | 4.413                 |
Conclusions

• Exhaustive study on the feasibility of VMAF on 360VR contents

• VMAF works sufficiently correctly with omnidirectional contents, without performing any particular adjustments

• The creation of a 360VR dataset can be avoided, thus saving computing and time resources
Video Multimethod Assessment Fusion (VMAF) on 360VR contents

Questions – Discussion – Debate - ...