武漢大學

VQEG

Towards Subjective Quality Assessment for Panoramic Video

VQEG Imersive Media Group (IMG)

Wuhan University

Zhenzhong Chen, Yingxue Zhang

zzchen@whu.edu.cn
grace@whu.edu.cn

武漢大學

We propose a subjective panoramic video quality assessment protocol for coding applications, which specially considers the display of the video via HMD. Based on the proposed protocol, a subjective video quality database for panoramic videos is established.

CONTENTS

Diplay of Panoramic Sequences for Subjective Quality Assessment 1

Subjective Quality Assessment Test 2

武漢大學

1 Diplay of Panoramic Sequences for Subjective Quality Assessment

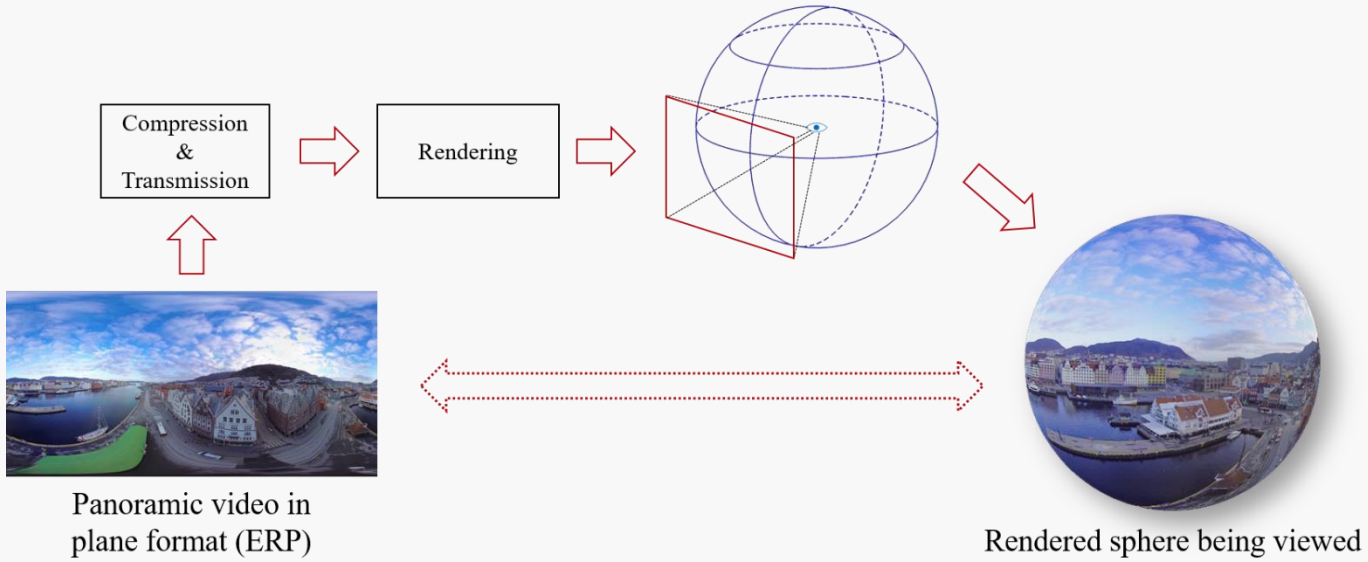
Display of Panoramic videos with HMDs

Problem on Subjective Quality Assessment of Panoramic Videos

Obstacles to the Per-pixel Display

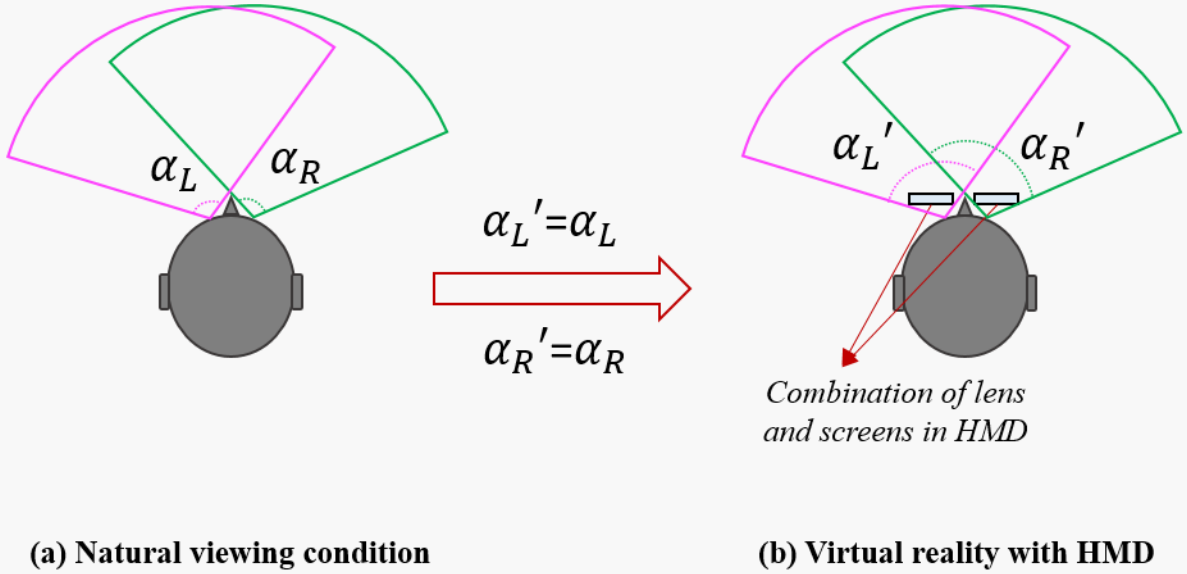
Finding the Optimal Display Resolution for Per-pixel Display

Display of Panoramic videos with HMDs



- The compression before display makes up the main motivation of our subjective quality assessment.
- Since existing coding systems cannot be applied to videos in sphere format, the panoramic videos must first be mapped onto a plane in accordance with certain geometric transformation rules, e.g., Equi-rectangular projection (ERP), Cube Map projection (CMP), Icosahedral projection (ISP).
- The compressed plane video will again be rendered into a sphere while displaying to viewers.

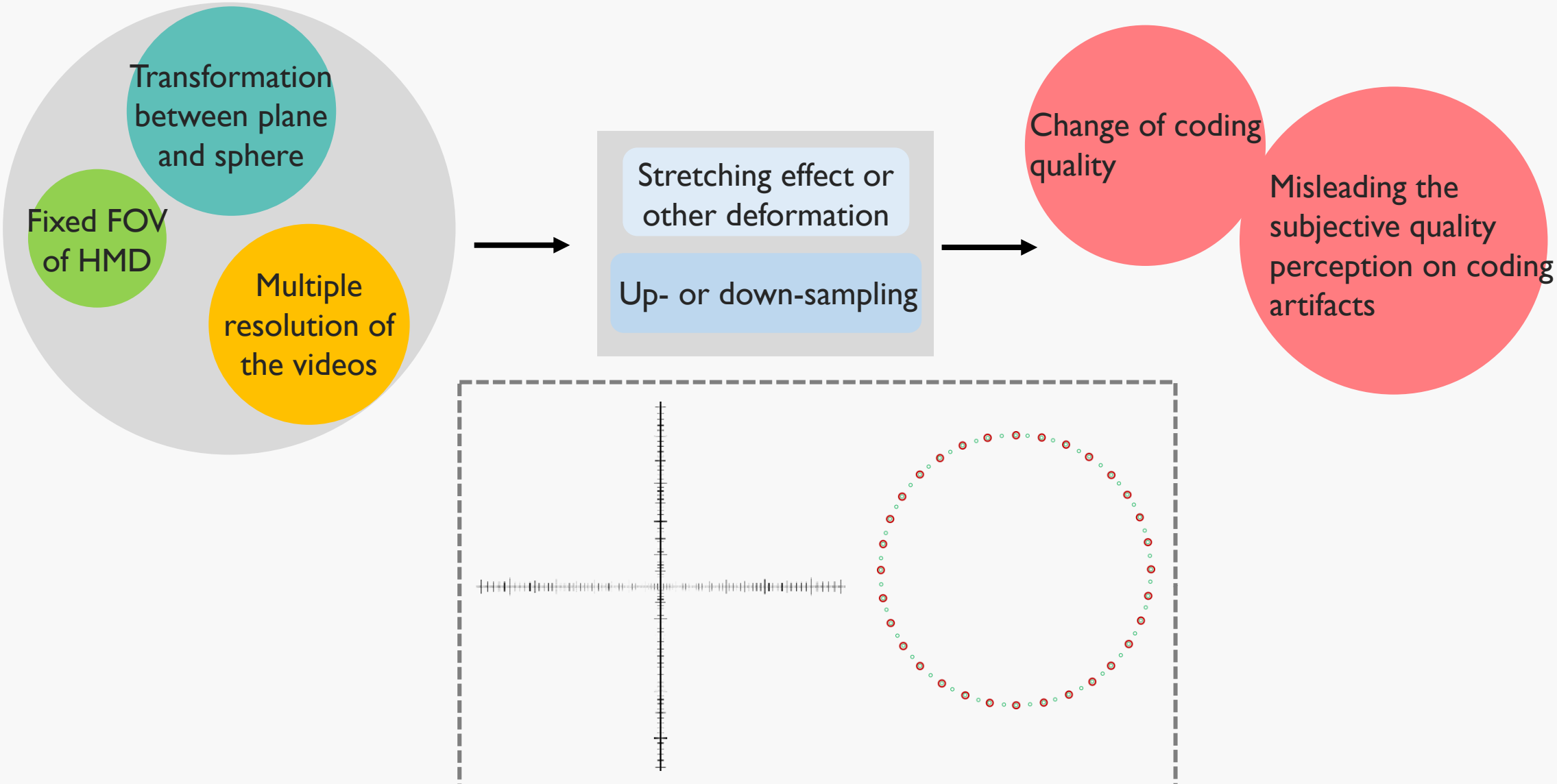
Problem on Subjective Quality Assessment of Panoramic Videos



(a) Natural viewing condition **(b) Virtual reality with HMD**

- Immersion requires that the virtual content can fill the entire FOV in HMD.
- In order to bring immersive experience to the viewers, the FOV αL' and αR' of the HMD must keep fixed and consistent with that of the human eyes (shown as αL and αR).
- The screen size of HMD is limited while the panoramic videos are commonly of high resolution.

Obstacles to the Per-pixel Display



The Optimal Display Resolution for Subjective Quality Assessment

武漢大學

The sampling problem will interfere the subjects' opinion on the video quality in terms of coding system being evaluated.

We propose to sample the original videos to an optimal resolution with respect to a certain HMD before introducing coding artifacts.

The optimal resolution

Guarantees least sampling degradation while displaying the video by keeping a maximized area on the center to be presented in a per-pixel manner.

Finding the Optimal Display Resolution

The sampling points

The cluster of lines connecting the left eye and each pixel intersects the equator on a set of sampling points that will finally be projected onto the integral pixel positions of the screen.

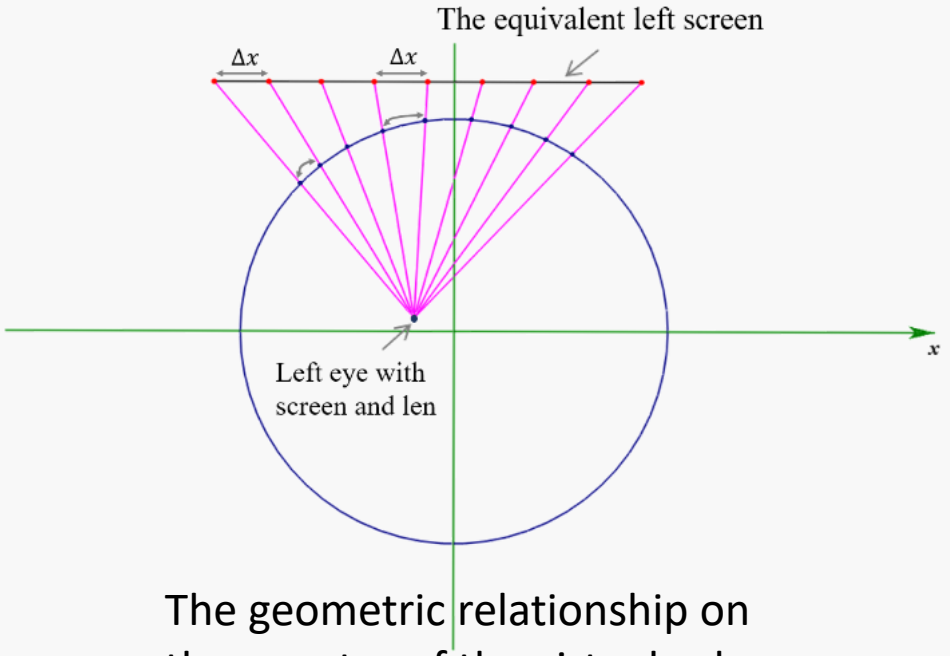
$$Y = \frac{y_o - y_l}{m\Delta x + x_o - x_l}(X - x_l) + y_l$$

(x_o, y_o): *the left end of the screen* (x_l, y_l): *the position of left eye*

Δx: *the constant interval between adjacent pixels.*

***X* and *Y*:** the coordinates of intersection points on the sphere:

$$X^2 + Y^2 = r^2$$



The geometric relationship on the equator of the virtual sphere

r: the radius of the sphere and is empirically set to 12.915 with the criteria to make the range of per-pixel display as large as possible considering the HMD used in the test.

The points whose vertical coordinate is greater than zero are determined to be the positions of the sampling points on the equator.

Finding the Optimal Display Resolution

The distribution of sampling points ->optimal resolution

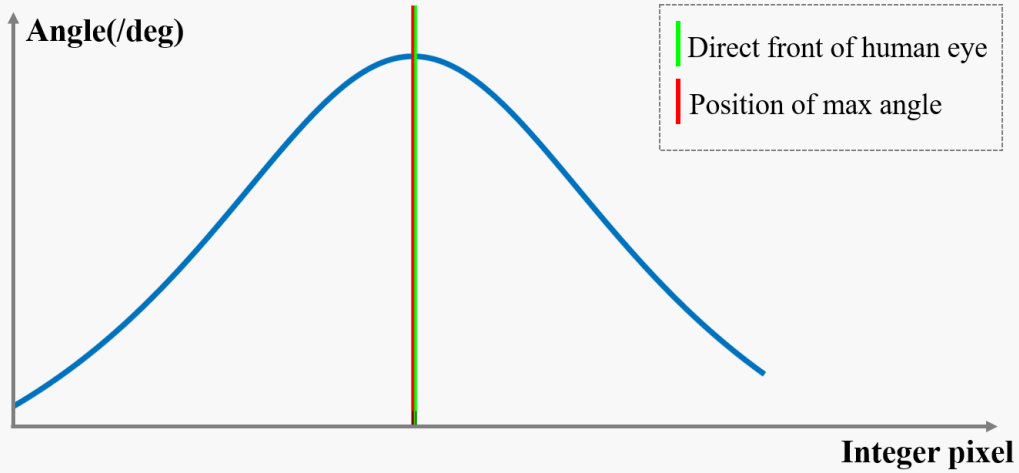
After obtaining the coordinates of the sampling points, the angle between the lines crossing zero point and the nth sampling point is calculated with the radius and the horizontal coordinate X_n:

$$\alpha_n = \sin^{-1}(\frac{X_n}{r})$$

The angle $\Delta\alpha$ between the adjacent sampling points is figured out:

$$\Delta\alpha = \alpha_n - \alpha_{n-1}$$

the distribution of the Δα on the screen



The sampling points on the sphere are not uniformly mapped onto the screen due to perspective projection.

Considering the observers' visual tendency towards the center area, the optimal horizontal resolution is defined as:

$$W = \frac{360}{\Delta\alpha_{mid}}$$

W: the optimal horizontal resolution, and the vertical resolution can be calculated with the constraints of specific coding system.

$\Delta\alpha_{mid}$*:* the angle between center point on the screen and its adjacent one.

武漢大學

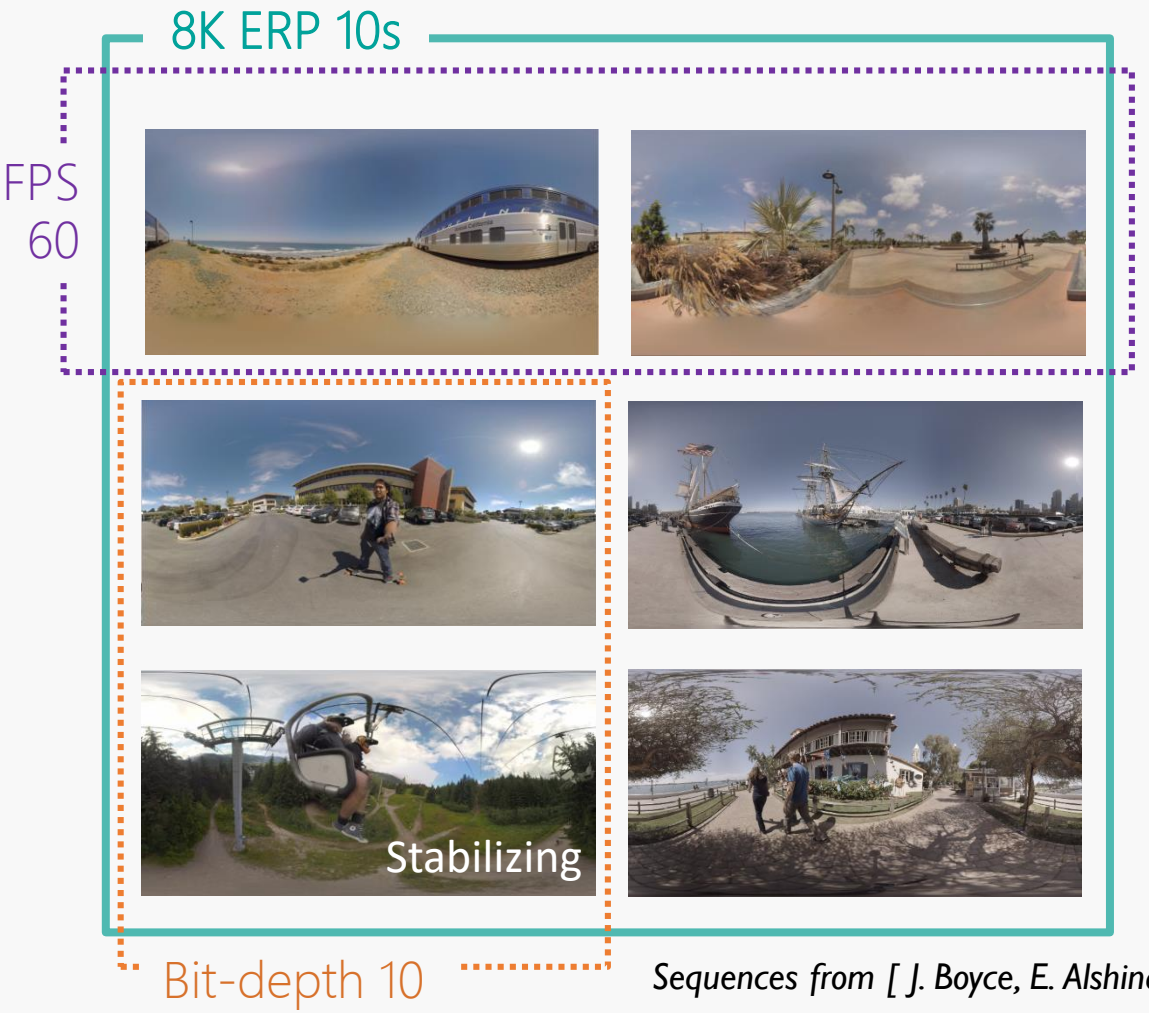
2 Subjective Quality Assessment Test

Sequences

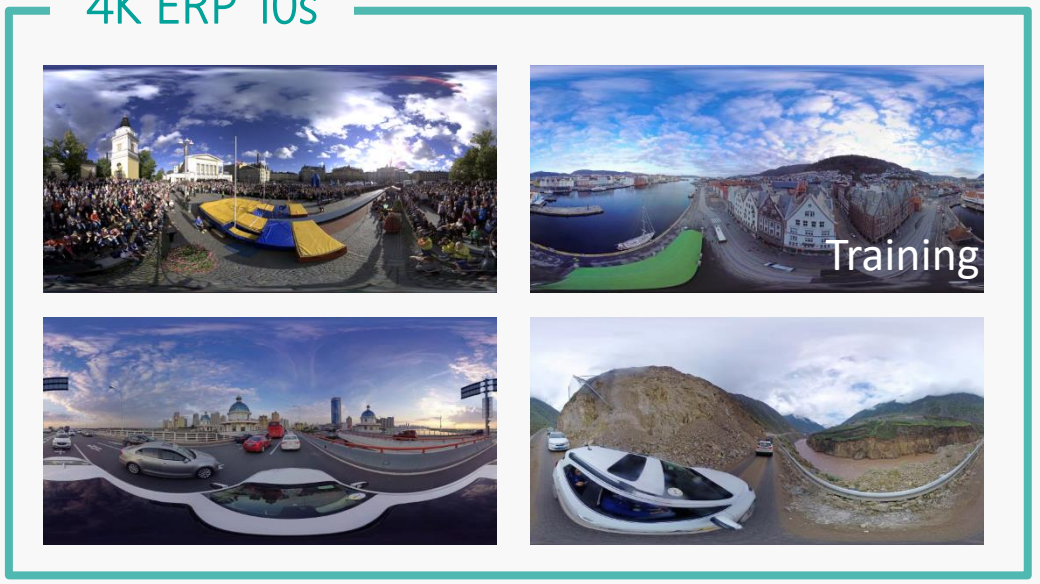
Test design

Rating data processing and analysis

Sequences



- Coding impairments to generate test sequences:
  - HM-16.14 with 360-Lib
  - 5 QP values, i. e., 22, 27, 32, 37, 42
- 60 sequences are generated from the 10 references
  - 6 for training, 3 for stabilizing, 48 for testing.



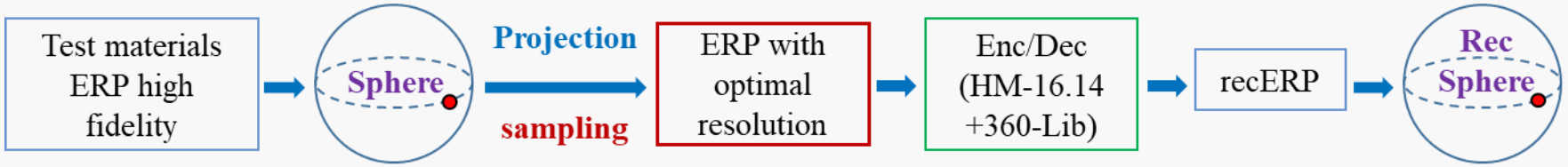
Sequences from [J. Boyce, E. Alshina, A. Abbas, Y. Ye, "JVET common test conditions and evaluation procedures for 360° video", Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-D1030, 4th Meeting, Oct. 2016.]

Sequences

Sampling of the Original Sequences

The original sequences are first sampled to the optimal resolution (3600 × 1800 for HTC VIVE) before coding.

The optimally sampled sequences are used as references. Coding artifacts are introduced based on the sampled references to make a fair comparison between references and test sequences.



Sequences

Coding Artifacts

(a) Reference

(b) QP=22

(c) QP=27

(d) QP=32

(e) QP=37

(f) QP=42

Test Design

Subjects, devices and assessment procedure

- Subject: **30**

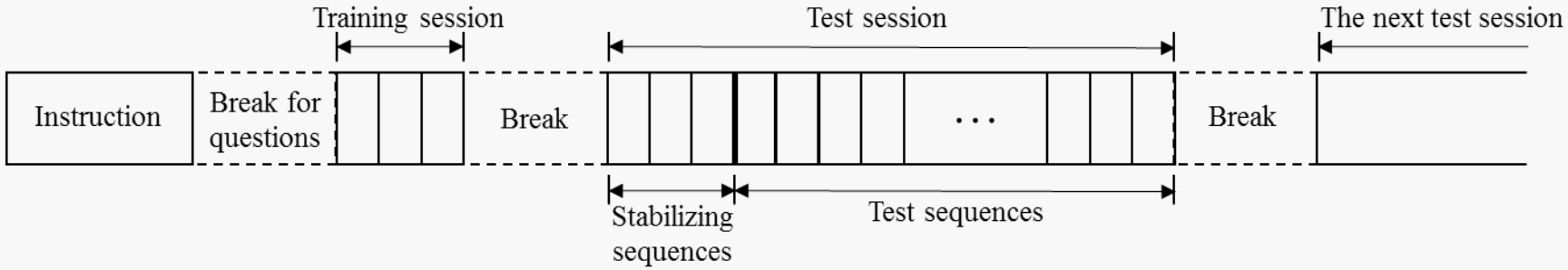
17 MALES 13 FEMALES

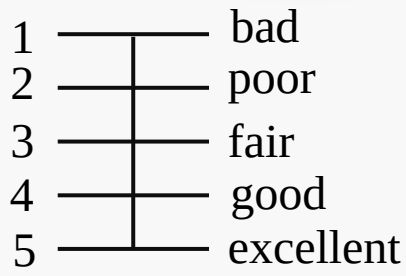
- **Non-expert**
- **undergraduate and graduate students**
- **Normal or corrected-to-normal visual acuity**

- Devices: HTC VIVE

- view all directions freely

- Testing procedure with ACR-HR

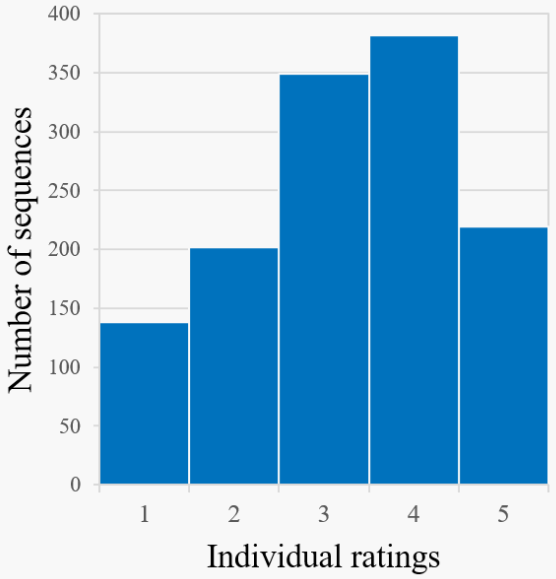




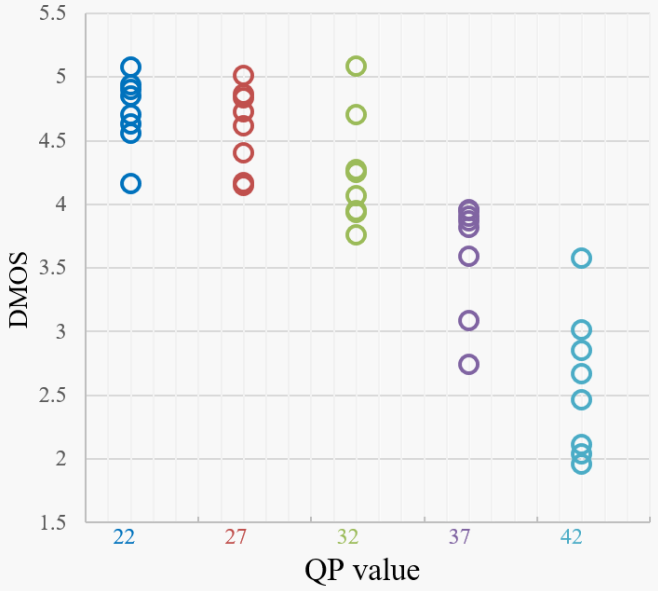
Rating Data Processing and Analysis

As specified in [ITU-R BT. 500], A subject will be discarded if more than a specific percentage of his/her rating scores are out of the expected normal range.

The ratings of **3** subjects are discarded by the post-experiment screening process setting the percentage to **5%.**

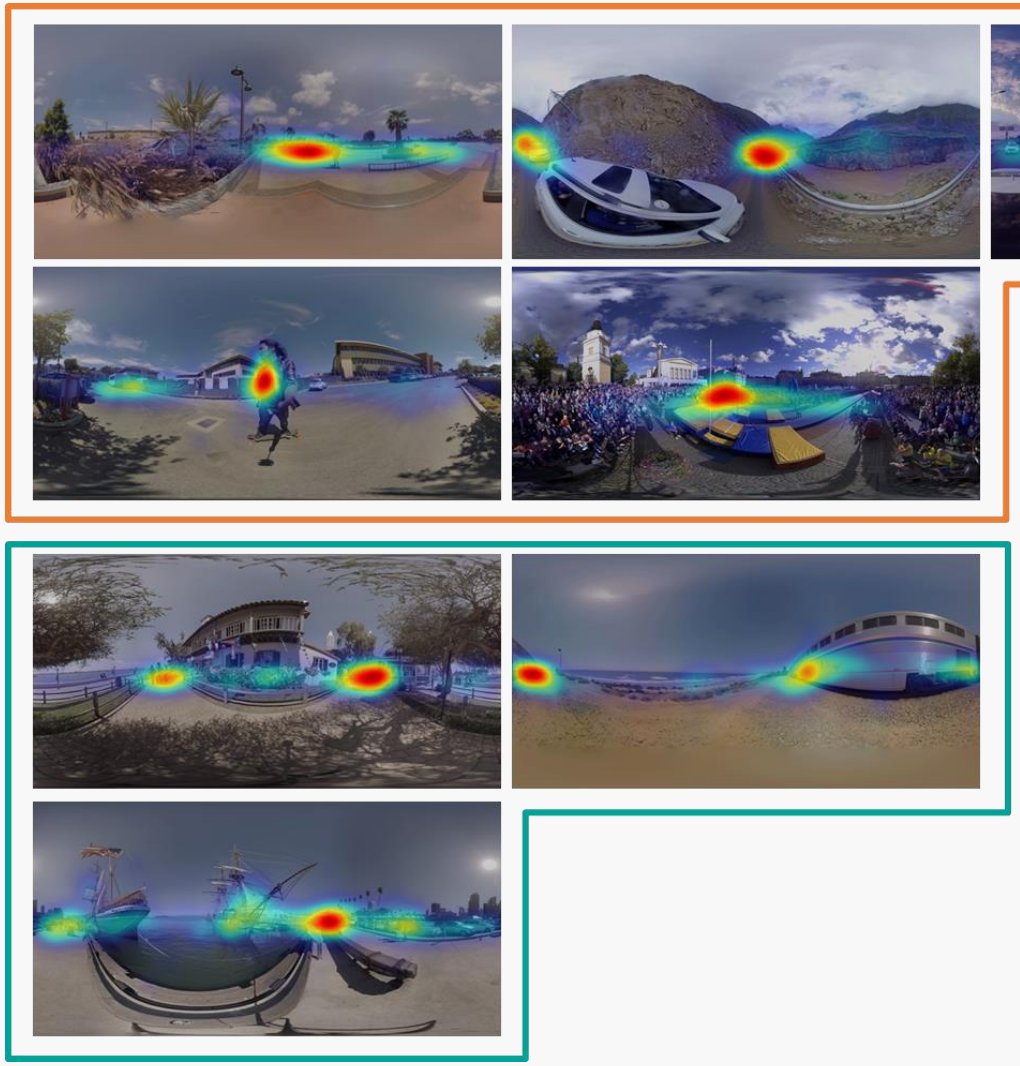


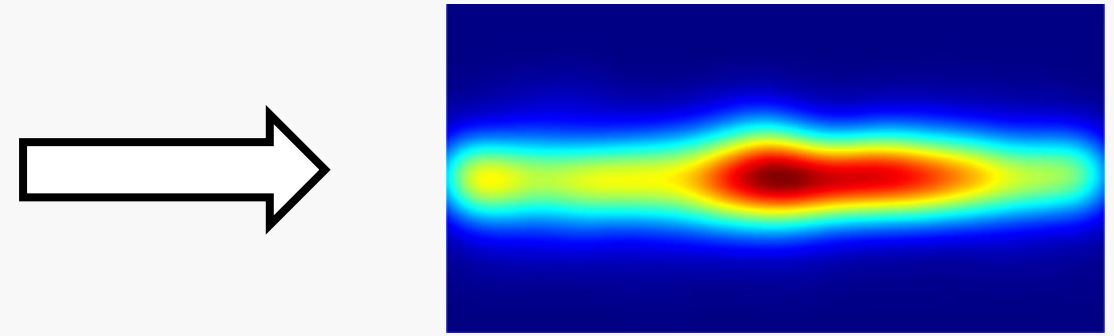
Distribution of the individual rating scores of all the subjects on all the sequences.



Distribution of DMOS over the five compression levels

Viewing Consistency





Conclusion

- a subjective panoramic video quality assessment protocol is proposed for coding applications
  - Considering the projection and the resolution limitation of HMDs, the method of sampling the video sequence to an optimal resolution before coding is proposed first.
  - With the optimal display resolution, a maximized range of per-pixel display on the center area of the video can be guaranteed, alleviating degradations caused by sampling of the HMDs and thus making the assessment more reliable.
- A subjective quality database for panoramic videos is established based on the proposed protocol.


Thanks!
iip.whu.edu.cn