VQEG 3DTV Group

Test Plan for establishing a
Ground Truth for Quality of Experience in 3D
for assessment methodologies in
3D Video Quality Assessment
GroTruQoE3D1

Draft Version 1.02, July 2013

Editors’ note: A blue highlight will occur before proposals that require explanation (e.g., text to be deleted).

Tracked changes are used to identify proposals that have not been agreed upon.

Changes that have been agreed upon are not marked. The wording of agreements made occurred during audio calls may need to be adjusted.

Contact: Marcus Barkowsky Tel: +33 2 40 68 30 22 Email: Marcus.Barkowsky@univ-nantes.fr

 Second author

 Third author

**Editorial History**

|  |  |  |
| --- | --- | --- |
| **Version** | **Date** | **Nature of the modification** |
| 1.0 | December3, 2012 | Initial Draft, edited by Yao Koudota and Marcus Barkowsky |
| 1.1 | December, 2012 | Singapore Meeting |
| 1.2 | July, 2013 | Ghent Meeting |

**Table of Contents**

1. Introduction 8

1.1. Scope 8

2. Overview: Expectations, Division of Labor and Ownership 9

2.1. Participation 9

2.2. Release of Video Sequences, Subjective Data and the Official Data Analysis 9

2.3. Using The Data in Publications 9

2.4. Release of Video Sequences 9

3. Video sequence data set 10

3.1. Reference Source Sequences (SRC) 10

3.2. Hypothetical Reference Circuits (HRC) 13

3.3. Availability of the SRC and PVS sequences 17

4. Subjective Rating Tests 18

4.1. Subjective assessment variables 18

4.2. Interaction with participants 18

4.3. Subjective Test Method: Paired Comparison 18

4.4. Length of sessions 19

4.5. Playback system 19

4.6. Subjects and Subjective Test Control 19

4.7. Instructions for Subjects and Failure to Follow Instructions 19

4.8. Subjective Data File Format 19

4.9. Division of video sequences into lab evaluations 21

4.10. Documentation 21

5. Data Analysis 21

5.1. Verification 21

5.2. Conversion to scale values 21

6. Appendix 21

6.1. Annex 1 21

6.2. Annex 2 26

6.3. Annex 3 26

6.4. Annex 4 26

List of Acronyms

ACR-HRR Absolute Category Rating with Hidden Reference Removal

ANOVA ANalysis Of VAriance

CCIR Comite Consultatif International des Radiocommunications

CODEC Coder-Decoder

FR Full Reference

GOP Group of Pictures

HD High Definition (television)

HRC Hypothetical Reference Circuit

ITU International Telecommunications Union

ITU-R ITU RadiocommunicationsStandardization Sector

ITU-T ITU Telecommunications Standardization Sector

MM Multimedia

MOS Mean Opinion Score

MPEG Motion Pictures Expert Group

NR No (or Zero) Reference

NTSC National Television Standard Committee (60-Hz TV, used mainly in US and Canada)

OS Opinion Score – a single subject’s answer

PAL Phase Alternating Line (50-Hz TV, used in Europe and elsewhere)

PVS Processed Video Sequence

RR Reduced Reference

SMPTE Society of Motion Picture and Television Engineers

SRC Source Reference Channel or Circuit

VQEG Video Quality Experts Group

List of Definitions

Picture quality refers to the quality of 3D images in terms of degradations of the pictorial quality. In most cases, this may also be perceived even on the left and on the right view separately. Typical examples are: blockiness and blurriness.

Visual discomfort is a negative sensation of the observer. It should be seen as a perceived state of the observer, therefore requiring a questionnaire for evaluating its presence and strength. It may be hinted by several symptoms (reported by the observer) and clinical signs (measured objectively). Usually it is supposed to have a steep rise time and a steep fall time, it occurs when watching a particular 3D scene and disappears immediately after the viewing is finished, however, the recovery time may be long (e.g. some of the symptoms and clinical signs may be observed for a longer period).

Visual fatigue shall be defined in this context as a syndrome in the medical sense. Its presence is assessed by the observation of zero, one or several symptoms (reported by the observer) and zero, one or several clinical signs (measured objectively). Usually it is supposed to have a longer rise time and a longer fall time than visual discomfort, it is not instantaneously diagnosed in conjunction with a certain 3D stimulus and remains a certain time after the 3D viewing has finished.

Physical symptoms are defined as indicated by the observation of muscular pain. Examples are shoulder pain, neck pain, <to be continued, we need a clear distinction between observation and interpretation>

Depth quantity is defined in this context as the *amount of depth* that the observer perceives in a 3D scene representation. This combines monocular and binocular cues and may therefore not necessarily be correlated to the amount of disparity in the 3D video.

Depth realism (Depth naturalness) is defined in this context as the *experience of the reconstruction of the 3D depth* that is present in the real 3D scene that has been captured by cameras in a 3D video.

View-asymmetric conditions are defined as processing steps which result in 3D view asymmetries. This usually results in a difference in video quality or in geometrical degradations that may lead to binocular rivalry and may induce a combination of visual discomfort, reduced depth perception, and reduced picture quality.

---

Intended frame rate is defined as the number of video frames per second physically stored for some representation of a video sequence. The intended frame rate may be constant or may change with time. Two examples of *constant*intended frame rates are a BetacamSP tape containing 25 fps and a VQEG FR-TV Phase I compliant 625-line YUV file containing 25 fps; these both have an absolute frame rate of 25 fps. One example of a *variable*absolute frame rate is a computer file containing only new frames; in this case the intended frame rate exactly matches the effective frame rate. The content of video frames is not considered when determining intended frame rate.

Anomalous frame repetition is defined as an event where the HRC outputs a single frame repeatedly in response to an unusual or out of the ordinary event. Anomalous frame repetition includes but is not limited to the following types of events: an error in the transmission channel, a change in the delay through the transmission channel, limited computer resources impacting the decoder’s performance, and limited computer resources impacting the display of the video signal.

Constant frame skipping is defined as an event where the HRC outputs frames with updated content at an effective frame rate that is fixed and less than the source frame rate.

Effective frame rate is defined as the number of unique frames (i.e., total frames – repeated frames) per second.

Frame rate is the number of (progressive) frames displayed per second (fps).

Live Network Conditions are defined as errors imposed upon the digital video bit stream as a result of live network conditions. Examples of error sources include packet loss due to heavy network traffic, increased delay due to transmission route changes, multi-path on a broadcast signal, and fingerprints on a DVD. Live network conditions tend to be unpredictable and unrepeatable.

Pausing with skipping (formerly frame skipping) is defined as events where the video pauses for some period of time and then restarts with some loss of video information. In pausing with skipping, the temporal delay through the system will vary about an average system delay, sometimes increasing and sometimes decreasing. One example of pausing with skipping is a pair of IP Videophones, where heavy network traffic causes the IP Videophone display to freeze briefly; when the IP Videophone display continues, some content has been lost. Another example is a videoconferencing system that performs constant frame skipping or variable frame skipping. Constant frame skipping and variable frame skipping are subset of pausing with skipping. A processed video sequence containing pausing with skipping will be approximately the same duration as the associated original video sequence.

Pausing without skipping (formerly frame freeze) is defined as any event where the video pauses for some period of time and then restarts without losing any video information. Hence, the temporal delay through the system must increase. One example of pausing without skipping is a computer simultaneously downloading and playing an AVI file, where heavy network traffic causes the player to pause briefly and then continue playing. A processed video sequence containing pausing without skipping events will always be longer in duration than the associated original video sequence.

Refresh rate is defined as the rate at which the computer monitor is updated.

Rewinding is defined as an event where the HRC playback jumps backwards in time. Rewinding can occur immediately after a pause. Given the reference sequence (A B C D E F G H I), two example processed sequence containing rewinding are (A B C D B C D E F) and (A B C CCC A B C). Rewinding can occur as a response to transmission error; for example, a video player encounters a transmission error, pauses while it conceals the error internally, and then resumes by playing video prior to the frame displayed when the transmission distortion was encountered. Rewinding is different from variable frame skipping because the subjects see the same content again and the motion is much more jumpy.

Simulated transmission errors are defined as errors imposed upon the digital video bit stream in a highly controlled environment. Examples include simulated packet loss rates and simulated bit errors. Parameters used to control simulated transmission errors are well defined.

Source frame rate (SFR) is the intended frame rate of the original source video sequences. The source frame rate is constant.

Transmission errors are defined as any error resulting from sending the video data over a transmission channel. Examples of transmission errors are corrupted data (bit errors) and lost packets / lost frames. Such errors may be generated in live network conditions or through simulation.

Variable frame skipping is defined as an event where the HRC outputs frames with updated content at an effective frame rate that changes with time. The temporal delay through the system will increase and decrease with time, varying about an average system delay. A processed video sequence containing variable frame skipping will be approximately the same duration as the associated original video sequence.

# Introduction

This document defines the conditions of an evaluation test aiming towards the standardization of the subjective assessment methodologies for different degradations seen in 3DTV visualization conducted by the Video Quality Experts Group (VQEG). It describes the subjective dataset that was used in this evaluation, the different kinds of setups that were used in the participating labs, the immediate results that were obtained from the observers and the statistical analysis that was performed on the acquired data. The text is based on discussions and decisions from meetings of the VQEG 3DTV working group (3DTV) at the periodic face-to-face meetings as well as on conference calls and in email discussion.

The outcome of the project will be a subjective dataset that has been evaluated by the Paired Comparison method. It is therefore independent on any scale. Various assessment methods shall be tried on the same video dataset in order to learn about their prediction quality. These tests may be conducted during the evaluation of the GroTruQoE3D1 or after it has been finished and analyzed. The goal is to evaluate and establish a list of verified assessment methods for 3DTV for standardization in ITU Recommendations. It shall be noted that the validity of a particular assessment methodology may be restricted to a certain scope such as image quality evaluation.

## Scope

The goal of this project is to establish a ground truth database for Quality of Experience in 3DTV (GroTruQoE3D) measurement methodologies. It has been agreed within the Video Quality Experts Group that observers are capable of voting in a Paired Comparison experiment for their preferred condition in 3D. It is further assumed that this preference will provide their opinion related to their own internal and individual scale. This voting process can therefore be performed without taking into consideration the influence of external scales that may be used in other experiments such as naturalness, depth quality, depth quantity, image quality and so on. The experiment will be conducted in a collaborative way in several laboratories and the results will be compared using a common set of sequences. Careful examination of the combination of the subjective experiments will follow. Finally, scale values will be provided for further research using the Bradley-Terry and Thurstone-Mosteller models.

The results of the GroTruQoE3D evaluation may be used to verify the performance of existing subjective quality assessment methods, the impact of different perceptual measurement scales, the influence of observer training on the results, etc. New methodologies may be developed based on the results.

# Overview: Expectations, Division of Labor and Ownership

## Participation

The participation to the project is open for any subjective assessment lab.

## Release of Video Sequences, Subjective Data and the Official Data Analysis

The video sequences are made publicly available by VQEG for download by any partner who wants to perform a subjective assessment in the context of this work or any associated study.

VQEG will make available each individual viewer’s scores (i.e., including rejected viewers) and the associated data as provided by the participating labs. This viewer data will not include any indication of the viewer’s identity, and should indicate the following data: (1) whether the viewer was rejected, (2) country of origin, which indicates frame rate that the viewer typically views, (3) gender (male or female), (4) age of viewer (rounded to the nearest decade would be fine), (5) experience of the viewer for 3D, (6) type of video that the viewer typically views (e.g., standard definition television, HDTV, IPTV, Video Conferencing, mobile TV, iPod, cell phone).

A standard questionnaire may be established that lists the questions asked of all viewers. This questionnaire may include other questions, and must take no longer than 5 minutes to complete. If possible, the questionnaire should be automated and (after translation) be used by all viewers.

## Using The Data in Publications

Publications based on the collected data are highly appreciated. Appropriate citations and eventual co-authorship should be respected and welcomed.

## Release of Video Sequences

All of the video sequences from at least 3 datasets will be made public. Most of the video sequences in these datasets will be available for research and development purposes only (e.g., not for trade shows or other commercial purposes). This same usage restriction will likely apply to the HDTV datasets that are made public.

All of the video sequences from at least 1 dataset will be kept private (i.e., only shared between HDTV ILG and proponents who submit one or more models).

# Video sequence data set

The video sequences used in this subjective assessment have been particularly designed for this purpose. The source sequences have been shot and selected to cover a wide range of different content features while providing a comfortable viewing in the viewing conditions that were fixed for this test plan. The selection of the Hypothetical Reference Circuits (HRC) has been performed on a large set of degradations that have been evaluated by experts on the three scale “picture quality”, “visual discomfort”, and “depth quantity” (to be verified, Excel file to be added in appendix).

## Reference Source Sequences (SRC)

Sequences derived from footage filmed by a stereoscopic camera or animated films. The format of the sequences is yuv422 stored in an AVI container at 25 fps and a resolution of Full-HD 1920x1080 pixels for each of the two views. The sequences are selected in such a way that they exhibit different properties such as coding complexity, motion, brightness, 3D effect and maximum disparity range. Most of the sequences feature 400 frames, corresponding to 16 seconds, only the “Umbrella” sequence is shorter with 325 frames (13 seconds).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SRC | Preview | Name & Description | Frame / Duration | Origin |
| 1 | A description... | Car and barrier gate | 1080p25/400 | University of Nantes |
| 2 | A description... | old man dancing on platform | 1080p25/400 | http://www.elephantsdream.org |
| 3 | A description... | A woman standing in the foreground and a fountain in the background | 1080p25/400 | National Institute of Information and Communication Technology |
| 4 | A description... | Tree leaves and wind | 1080p25/400 | University of Nantes |
| 5 | A description... | person flying a drone | 1080p25/400 | University of Nantes |
| 6 | A description... | Basket ball training | 1080p25/400 | University of Nantes |
| 7 | A description... | News report mimic | 1080p25/400 | University of Nantes |
| 8 | A description... | 1. Boxer warms up2. Boxing training(fade in/fade out effect) | 1080p25/400 | University of Nantes |
| 9 | A description... | two people attacked by birdstwo people in a capsule projected by a mechanical arm | 1080p25/400 | http://www.elephantsdream.org |
| 10 | A description... | Persons meeting in a hall. | 1080p25/400 | University of Nantes |
|  |  |  |  |  |
|  |  |  |  |  |
|  | Additional sequences |  |  |  |
|  | A description... | Women take pictured in the park(replacement) | 1080p25/400 | National Institute of Information and Communication Technology |
|  | A description... | Person playing with an umbrella | 1080p25/325 | University of Nantes |
|  | A description... | two people attacked by birds | 1080p25/400 | http://www.elephantsdream.org |
|  | A description... | 1. 2 players score2. Goal keeping | 1080p25/400 | University of Nantes |
|  | A description... | Two lab assistants working | 1080p25/400 | University of Nantes |
|  | A description... | Phone call in an office | 1080p25/400 | University of Nantes |
|  | A description... | A woman standing in the foreground and a fountain in the background | 1080p25/400 | National Institute of Information and Communication Technology |

## Hypothetical Reference Circuits (HRC)

(remark: Please let me know when HRC references are wrong)



**HRC 0: Reference 3D**

 This is the undistorted reference sequence which is shown to the observers in HRC0 in order to create a hidden reference. These sequences were used in order to create different PVS.

**HRC 1: Reference 2D**

 The reference sequences for the 2D viewing case were created by replacing the right view with the left view, i.e. the viewer perceives two times the same view. This sets the displayed disparity to zero and therefore, the image appears as 2D.

**HRC 2: Reduction of spatial resolution in horizontal and vertical direction by a factor of 4**

The reference frames are downsampled by a Lanczos-3 filter to a quarter of their original width and a quarter of their original height. Then the images are upscaled again to Full-HD resolution using the same filter.

**HRC 3: Frame rate reduction by a factor of 3**

 Reduction of the frame rate by a factor of three using a 0 order temporal filter. One frame is repeated three times, while the two following frames are removed. No additional degradation has been added.



Figure 2: Frame rate reduction by a factor 3

**HRC 4: Asymmetric modification of brightness**

The right view is processed in such a way that it has a brightness offset to the left view. This may be experienced in an incorrect luminance calibration of the cameras. The change is performed in the Y Cb Cr color space by adding a constant value (v) to each pixel of the Y component of this view.

Y = Y + v

The left view is not processed, it is directly taken from the reference HRC0.

**HRC 5: Asymmetric modification of gamma**

The right view is processed in such a way that it has a gamma offset to the left view. An exponential luminance curve processing, often also referred to as gamma value processing, is performed on the Y component of the Y Cb Cr color space. The value that was used equals 0.5 and therefore brighter parts of the image are emphasized by appearing even brighter while the black level value remains unchanged (as compared to the processing of HRC4).

**HRC 6-7: Horizontal disparity offset**

A horizontal offset of the left and the right view with opposite sign produces a shift of the content in depth due to the change of disparity. The border pixels that were introduced in each image were filled with black color. The shift of -30 pixels for HRC6 and 30 pixels for HRC7 has been introduced symmetrically, i.e. the total disparity offset was -60 for HRC6 and 60 for HRC7.



Figure 6: Scheme of horizontal disparity offset

**HRC 8: Vertical disparity offset**

The same processing as in HRC6 is applied for HRC8 but in vertical direction as opposed to the horizontal direction used in HRC6. This simulates a wrong camera calibration and the left and the right view are more difficult to fuse. The shift of -20 pixels has been introduced symmetrically, i.e. the total disparity offset was -40 pixels.



Figure 7: Scheme of vertical disparity offset

**HRC 9: Geometric Distortion**

The left view is processed using the program Stirmark. This program may add geometric deformations to images in order to test watermarking algorithms. In this HRC, it was used to simulate strong camera calibration errors. The parameter act on the maximum distance a corner can move inwards, for this hrc the maximum is 80. The right view is not processed, it is directly taken from the reference HRC0.

**HRC 10: 2D to 3D**

A 3D effect for 2D videos may be obtained by using a geometrical stretching transformation. A commonly applied simple algorithm which has been suggested in "Tiefenbehandlung, 2D-Videos in 3D abspielen (in german)":Dr. Volker Zota, Jan-Keno Janssen; c't Magazin fuer Computer Technik, Heise Verlag, page 116, Volume 6, 2010 is used in this HRC. It start from the left view and is performed in three steps:



Figure 4: Processing chain for HRC 10

AVISynth: AVISynth Script (Annex 3) with virtualdub to produce a second view from the given video sequence, i.e. convert 2D video to 3D.

FPS reduction: As the output is 50fps, a frame rate reduction is implemented skipping one frame



Figure 5: Frame rate divide by factor 2

Video cropping: Separation of left and right view

**HRC 11-12: Coding with H.264**

 These sequences were encoded in H.264 format. They were first converted from the original AVI (yuv 422) files to YUV (YUV 420) files using the lanczos-3 filter for color downsampling, then the encoding and decoding was performed with the ITU reference encoder JM18.2. The encoder configuration file can be found in the Annex 1. The decoding provided a sequence in YUV full HD format showing typical H.264 artifacts. The encoding was performed at a fixed QP setting of 32 for HRC11 and 44 for HRC12.



Figure 1: scheme of HRC 2 and 3

**HRC 13: Asymmetric coding condition**

This HRC uses the result of HRC11 (H.264 encoding at a fixed QP of 32) for the left view and the result of HRC12 (H.264 encoding at a fixed QP of 44) for the right view. The result is a strong asymmetrical encoding condition.

**HRC 14: H.264 Encoding and Edge enhancement at 40%**

The process of encoding and decoding the sequences in H.264 format using a QP of 32 is identical to that used for HRC 11. An additional edge enhancement filter has been applied in order to enhance the high frequency components and to visually sharpen the contours.

**HRC 15: Error concealment with H.264**

Transmission errors are introduced in the encoded H.264 bitstream of HRC11 (H.264 encoding at fixed QP32) before decoding. The following chain is used:



Figure 8: scheme of HRC 15

AVI to YUV: Converting AVI 422 to YUV 420 sequences, using Lanczos-3 filter

JM18.2 H264 encoding: Encoding the sequence with JM 18.2 in H.264 Annex-B bitstream format

Strong packet losses with Siranon: Simulation of a network with packet loss, this part is divided further into 3 steps, the Transmitter (Annex 4), the Packet Loss Simulator (Annex 5), and finally the Receiver (Annex 6). The software Sirannon in its version 1.0 was used to create the impacted bitstream.

H.264 decoding with JM18.2: Decoding the sequence with JM 18.2 H.264, the configuration file is detailed in Annex 2

YUV to AVI: Converting YUV 420 to AVI 422 sequences, using a Lanczos-3 filter for the color upsampling

**HRC 16: 2D video encoded in H.264**

This HRC uses the left view of the result of HRC12 (H.264 encoding at a fixed QP of 44) for the left view and right view.

**HRC 17: Encoding with JPEG2000**

The encoding is performed using a JPEG 2000 encoder / decoder. In particular, the implementation of kakadu v6.3 was used. The bitrate was fixed at 2Mbps. In order to implement this HRC, six operations were performed in succession:

Figure 3: scheme of HRC 17

AVI to YUV: Converting AVI 422 to YUV 420 sequences, using a Lanczos-3 filter

YUV to ppm: Converting the YUV sequence to PPM image format.

Encoding JPEG2000: Encoding ppm images to image j2c with kakadu encoder v6.3

Decoding JPEG2000: decoding j2c images to image j2c, with kakadu decoder v6.3 using the bitrate as indicated.

ppm to YUV: Converting the PPM format images to a YUV sequence.

YUV to AVI: Converting YUV 420 to AVI 422 sequences , using the Lanczos-3 filter

## Availability of the SRC and PVS sequences

The sequences are available for free download at the following location:

<to be inserted>

# Subjective Rating Tests

Subjective tests will be performed using as conditions in the lab that are as close as possible.

## Subjective assessment variables

The subjective assessment environment specified by the ITU-T BT.500 will be followed as close as possible with the following parameters:

* Voting methodology: Paired Comparison (2-3 labs time-parallel experiments, 5-6 time-sequential experiments)
* Viewing distance: A viewing distance of 3H will be used for active shutter glasses, for passive glasses, a viewing distance of 4.5H will be used as a compromise between vertical and horizontal resolution (where H = Picture Height (picture is defined as the size of the video window, not the physical display.)
* Background lighting: The exact value shall be reported. It is allowed to use no background lighting
* Room illumination: low, not interfering with active glasses, shall be reported: at eye position towards the screen with the screen showing gray, a photo without flash of the setup is required
* Voting interface: Paper, on-screen or separate screen (to be documented)
* Stimulus in between presentations: Preferred as a 3D structured gray stimulus, for example gray squares at different depth, gray screen at Y=80 possible (to be documented). This may also apply to the voting screen interface, i.e. a 3D voting screen may be preferred
* Number of observers in parallel for a single screen: Maximum 3 (to be documented)
* (Section voted for in Ghent 2013)

## Interaction with participants

The observer screening needs to contain at least:

* Depth acuity (Randot Stereo test or equivalent)
* Far vision (Snellen Chart or equivalent)
* Near vision (Snellen Chart or equivalent)
* Color vision (Ishihara Plates or equivalent)

Further ophthalmologic tests may be conducted in addition such as:

* Vergence facility (prism test)
* Phoriae
* Extent of Panum’s Area
* Break-off point in near
* …

A questionnaire regarding the experience of the subjects with 3D viewing experience shall be asked to the subjects. The questions are listed in Annex A. (to be provided by IRCCyN)

Further information about the health state of the subjects prior and post-test may be acquired in an additional questionnaire. An example is provided in Annex B. (to be provided by IRCCyN)

## Subjective Test Method: Paired Comparison

The subjective assessment method that shall be used by all labs in the Ground Truth experiments is Paired Comparison. Additional experiments may be conducted using different methods.

<insert description of PC>

The Paired Comparison setup may be performed either time parallel (using two screens side by side) or time sequential (using a single screen showing the two stimuli one after the other).

## Length of sessions

The time of actively viewing videos and voting will be limited to 50 minutes per session.

## Playback system

The playback of the video sequences shall provide the highest possible quality in each lab. Ideally, the sequences are played back on a broadcast quality 3D shutter glasses screen in Full-HD resolution per view.

Depending on the availability of the appropriate equipment, a subjective assessment with a Blu-Ray3D player with slightly compressed videos and a consumer type 3D display is acceptable because the sequences show a wide range of 3D quality conditions. For polarized displays, 2D Blu-Ray discs should be preferred to eliminate the influence of the processing of the 3D display.

## Subjects and Subjective Test Control

Each test will require 40 subjects.

Only non-expert viewers will participate. The term non-expert is used in the sense that the viewers’ work does not involve video picture quality and they are not experienced assessors. They must not have participated in a subjective quality test over a period of six months. All viewers will be screened prior to participation for the following:

* normal (20/30) visual acuity with or without corrective glasses (per Snellen test or equivalent).
* normal color vision (per Ishihara test or equivalent).
* familiarity with the language sufficient to comprehend instruction and to provide valid responses using the semantic judgment terms expressed in that language.

## Instructions for Subjects and Failure to Follow Instructions

For many labs, obtaining a reasonably representative sample of subjects is difficult. Therefore, obtaining and retaining a valid data set from each subject is important. The following procedures are highly recommended to ensure valid subjective data:

* Write out a set of instructions that the experimenter will read to each test subject. The instructions should clearly explain why the test is being run, what the subject will see, and what the subject should do. Pre-test the instructions with non-experts to make sure they are clear; revise as necessary.
* Explain that it is important for subjects to pay attention to the video on each trial.
* There are no “correct” ratings. The instructions should not suggest that there is a correct rating or provide any feedback as to the “correctness” of any response. The instructions should emphasize that the test is being conducted to learn viewers’ judgments of the quality of the samples, and that it is the subject’s opinion that determines the appropriate rating.
* Paying subjects helps keep them motivated.
* Subjects should be instructed to watch the entire video sequence before voting. The screen should say when to vote (e.g., “vote now”).

If it is suspected that a subject is not responding to the video stimuli or is responding in a manner contrary to the instructions, their data may be discarded and a replacement subject can be tested. The experimenter will report the number of subjects’ datasets discarded and the criteria for doing so. Example criteria for discarding subjective data sets are:

* The same rating is used for all or most of the PVSs.

## Subjective Data File Format

Subjective data should NOT be submitted in archival form (i.e., every piece of data possible in one file). The working file should be a spreadsheet listing only the following necessary information:

* Experiment ID
* Source ID Number
* HRC ID Number
* Video File
* Each Viewer’s Rating in a separate column (Viewer ID identified in header row)

All other information should be in a separate file that can later be merged for archiving (if desired). This second file should have all the other "nice to know" information indexed to the subjectIDs: date, demographics of subject, eye exam results, etc. A third file, possibly also indexed to lab or subject, should have ACCURATE information about the design of the HRCs and possible something about the SRCs.

An example table is shown below (where HRC “0” is the original video sequence).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | Viewer ID | Viewer ID | Viewer ID | Viewer ID | … | Viewer ID |
| Experiment | SRC Num | HRC Num | File | 1 | 2 | 3 | 4 | … | 24 |
| XYZ | 1 | 1 | xyz\_src1\_hrc1.avi | 5 | 4 | 5 | 5 | … | 4 |
| XYZ | 2 | 1 | xyz\_src2\_hrc1.avi | 3 | 2 | 4 | 3 | … | 3 |
| XYZ | 1 | 7 | xyz\_src1\_hrc7.avi | 1 | 1 | 2 | 1 | … | 2 |
| XYZ | 3 | 0 | xyz\_src3\_hrc0.avi | 5 | 4 | 5 | 5 | … | 5 |

<needs to be adapted for Paired Comparison>

## Division of video sequences into lab evaluations

Each participating lab evaluates a fixed number of pairs using the Paired Comparison method. The pairs for each lab are randomly selected. A number of constraints are applied. A common set of pairs are evaluated by all labs.

The division into pair groups for the different labs is performed as follows: (verification: Jing Li)

* Rank order all pairs based on their quality using either prior knowledge or based on a pretest. (Decision: Ordering by mean or ordering by maximum quality of the two sequences in the pair)
* Split the pairs equally in terms of quality for each lab:
	+ For n labs, select iteratively n sequences starting from x=1 [x:x+n]
	+ Permute the n sequences randomly and assign them to the labs in ascending order
	+ Increase x by n and continue with the next group of n sequences
* This selection process guarantees that each lab has a set of video sequences that are spaced evenly over the whole evaluation space

Remark: This is not an optimal choice for performing Paired Comparison experiments. If only one lab would perform all experiments, performing the pairwise comparison by groups of quality, i.e. one experiment for only high quality, one experiment for only low quality would lead to a higher precision because the observers compare more often “difficult” conditions.

Remark2: The common set uses the a subset of the Optimized Square Design method providing stabilizing results over the sequences.

## Documentation

A detailed documentation of the experiment shall be provided by each lab. It shall include as much information of the viewing setup and environment as possible. Photos of the environment may help understanding the setup.

# Data Analysis

The data analysis serves several purposes. The first goal is the evaluating the validity of the acquired data in the different subjective assessment labs, thus allowing the creation of a large common dataset. As Paired Comparison is used as assessment methodology, scale adaptation problems do not arise. The second goal is to establish a scale value for each PVS that was evaluated in Paired Comparison methodology. This eases the comparison of the results with assessment methods that use direct scales such as Absolute Category Rating or Double Stimulus Continuous Quality Scale.

## Verification

<to bediscussed / written>

## Conversion to scale values

<to bediscussed / written>

# Participating labs

Acreo : Later in spring, Passive or active technology (or passive projector), sequential Paired Comparison, using PC

NTIA: Spring, active technology, 55” inch, sequential Paired Comparison, using PC

Vittorio: Passive 47” inch LG display, 47LM620S and can participate during spring

Intel: maybe

KDDI: maybe after April, 100” active shutter or Sony 42” passive glasses display, using either PC or specific playback

AGH: 42” Polarized or Active, one experiment, using PC, experiment before next meeting possible (also helping with data analysis)

Yonsei University: no information

Orange Labs: Maybe

UPM: maybe

Next steps:

* Player Software
* Check the display conditions
* Work on the subjective assessment questionnaires
* Algorithm for splitting the experiment

# Appendix

## Annex 1

Configuration file for encoding H.264 using the ITU reference encoder implementation JM 18.2

# New Input File Format is as follows

# <ParameterName> = <ParameterValue> # Comment

#

# See configfile.h for a list of supported ParameterNames

#

# For bug reporting and known issues see:

# https://ipbt.hhi.fraunhofer.de

##########################################################################################

# Files

##########################################################################################

InputFile= ../../JEG264HMIX1\_src01\_original\_s1920x1080p25n250v0.yuv # Input sequence

InputHeaderLength = 0 # If the inputfile has a header, state it's length in byte here

StartFrame = 0 # Start frame for encoding. (0-N)

FramesToBeEncoded = 250 # Number of frames to be coded

FrameRate = 25 # Frame Rate per second (0.1-100.0)

SourceWidth = 1920 # Source frame width

SourceHeight = 1080 # Source frame height

SourceResize = 0 # Resize source size for output

OutputWidth = 1920 # Output frame width

OutputHeight = 1080 # Output frame height

TraceFile = JEG264HMIX1\_src01\_original\_ip32\_qi38\_qp38\_qb38\_hier0\_s1920x1080p25n250v0\_trace.txt # Trace file

ReconFile = JEG264HMIX1\_src01\_original\_ip32\_qi38\_qp38\_qb38\_hier0\_s1920x1080p25n250v0.yuv # Recontruction YUV file

OutputFile = JEG264HMIX1\_src01\_original\_ip32\_qi38\_qp38\_qb38\_hier0\_s1920x1080p25n250v0.264 # Bitstream

StatsFile = JEG264HMIX1\_src01\_original\_ip32\_qi38\_qp38\_qb38\_hier0\_s1920x1080p25n250v0\_stat.txt # Coding statistics file

##########################################################################################

# Encoder Control

##########################################################################################

ProfileIDC = 77 # Profile IDC (66=baseline, 77=main, 88=extended; FREXT Profiles: 100=High, 110 =High 10, 122=High 4:2:2, 244=High 4:4:4, 44=CAVLC 4:4:4 Intra)

IntraProfile = 0 # Activate Intra Profile for FRExt (0: false, 1: true)

# (e.g. ProfileIDC=110, IntraProfile=1 => High 10 Intra Profile)

LevelIDC = 50 # Level IDC (e.g. 20 = level 2.0)

IntraPeriod = 32 # Period of I-pictures (0=only first)

IDRPeriod = 32 # Period of IDR pictures (0=only first)

AdaptiveIntraPeriod = 1 # Adaptive intra period

AdaptiveIDRPeriod = 0 # Adaptive IDR period

IntraDelay = 0 # Intra (IDR) picture delay (i.e. coding structure of PPIPPP... )

EnableIDRGOP = 0 # Support for IDR closed GOPs (0: disabled, 1: enabled)

EnableOpenGOP = 0 # Support for open GOPs (0: disabled, 1: enabled)

QPISlice = 38 # Quant. param for I Slices (0-51)

QPPSlice = 38 # Quant. param for P Slices (0-51)

FrameSkip = 0 # Number of frames to be skipped in input (e.g 2 will code every third frame).

# Note that this now excludes intermediate (i.e. B) coded pictures

ChromaQPOffset = 0 # Chroma QP offset (-51..51)

DisableSubpelME = 0 # Disable Subpixel Motion Estimation (0=off/default, 1=on)

SearchRange = 64 # Max search range

MEDistortionFPel = 0 # Select error metric for Full-Pel ME (0: SAD, 1: SSE, 2: Hadamard SAD)

MEDistortionHPel = 2 # Select error metric for Half-Pel ME (0: SAD, 1: SSE, 2: Hadamard SAD)

MEDistortionQPel = 2 # Select error metric for Quarter-Pel ME (0: SAD, 1: SSE, 2: Hadamard SAD)

MDDistortion = 2 # Select error metric for Mode Decision (0: SAD, 1: SSE, 2: Hadamard SAD)

SkipDeBlockNonRef = 0 # Skip Deblocking (regardless of DFParametersFlag) for non-reference frames (0: off, 1: on)

ChromaMCBuffer = 1 # Calculate Color component interpolated values in advance and store them.

# Provides a trade-off between memory and computational complexity

# (0: disabled/default, 1: enabled)

ChromaMEEnable = 0 # Take into account Color component information during ME

# (0: only first component/default,

# 1: All Color components - Integer refinement only

# 2: All Color components - All refinements)

ChromaMEWeight = 1 # Weighting for chroma components. This parameter should have a relationship with color format.

NumberReferenceFrames = 5 # Number of previous frames used for inter motion search (0-16)

PList0References = 0 # P slice List 0 reference override (0 disable, N <= NumberReferenceFrames)

Log2MaxFNumMinus4 = 0 # Sets log2\_max\_frame\_num\_minus4 (-1 : based on FramesToBeEncoded/Auto, >=0 : Log2MaxFNumMinus4)

Log2MaxPOCLsbMinus4 = -1 # Sets log2\_max\_pic\_order\_cnt\_lsb\_minus4 (-1 : Auto, >=0 : Log2MaxPOCLsbMinus4)

GenerateMultiplePPS = 0 # Transmit multiple parameter sets. Currently parameters basically enable all WP modes (0: diabled, 1: enabled)

ResendPPS = 0 # Resend PPS (with pic\_parameter\_set\_id 0) for every coded Frame/Field pair (0: disabled, 1: enabled)

MbLineIntraUpdate = 0 # Error robustness(extra intra macro block updates)(0=off, N: One GOB every N frames are intra coded)

RandomIntraMBRefresh = 0 # Forced intra MBs per picture

##########################################################################################

# PSlice Mode types

##########################################################################################

PSliceSkip = 1 # P-Slice Skip mode consideration (0=disable, 1=enable)

PSliceSearch16x16 = 1 # P-Slice Inter block search 16x16 (0=disable, 1=enable)

PSliceSearch16x8 = 1 # P-Slice Inter block search 16x8 (0=disable, 1=enable)

PSliceSearch8x16 = 1 # P-Slice Inter block search 8x16 (0=disable, 1=enable)

PSliceSearch8x8 = 1 # P-Slice Inter block search 8x8 (0=disable, 1=enable)

PSliceSearch8x4 = 1 # P-Slice Inter block search 8x4 (0=disable, 1=enable)

PSliceSearch4x8 = 1 # P-Slice Inter block search 4x8 (0=disable, 1=enable)

PSliceSearch4x4 = 1 # P-Slice Inter block search 4x4 (0=disable, 1=enable)

DisableIntra4x4 = 0 # Disable Intra 4x4 modes

DisableIntra16x16 = 0 # Disable Intra 16x16 modes

DisableIntraInInter = 0 # Disable Intra modes for inter slices

IntraDisableInterOnly = 0 # Apply Disabling Intra conditions only to Inter Slices (0:disable/default,1: enable)

Intra4x4ParDisable = 0 # Disable Vertical & Horizontal 4x4

Intra4x4DiagDisable = 0 # Disable Diagonal 45degree 4x4

Intra4x4DirDisable = 0 # Disable Other Diagonal 4x4

Intra16x16ParDisable = 0 # Disable Vertical & Horizontal 16x16

Intra16x16PlaneDisable = 0 # Disable Planar 16x16

ChromaIntraDisable = 0 # Disable Intra Chroma modes other than DC

EnableIPCM = 1 # Enable IPCM macroblock mode

DisposableP = 0 # Enable Disposable P slices in the primary layer (0: disable/default, 1: enable)

DispPQPOffset = 0 #Quantizer offset for disposable P slices (0: default)

PreferDispOrder = 1 # Prefer display order when building the prediction structure as opposed to coding order (affects intra and IDR periodic insertion, among others)

PreferPowerOfTwo = 0 # Prefer prediction structures that have lengths expressed as powers of two

FrmStructBufferLength = 16 # Length of the frame structure unit buffer; it can be overriden for certain cases

ChangeQPFrame = 0 # Frame in display order from which to apply the Change QP offsets

ChangeQPI = 0 # Change QP offset value for I\_SLICE

ChangeQPP = 0 # Change QP offset value for P\_SLICE

ChangeQPB = 0 # Change QP offset value for B\_SLICE

ChangeQPSI = 0 # Change QP offset value for SI\_SLICE

ChangeQPSP = 0 # Change QP offset value for SP\_SLICE

##########################################################################################

# B Slices

##########################################################################################

NumberBFrames = 2 # Number of B coded frames inserted (0=not used)

PReplaceBSlice = 0 # Replace B-coded slices with P-coded slices when NumberBFrames>0

QPBSlice = 38 # Quant. param for B slices (0-51)

BRefPicQPOffset = -1 # Quantization offset for reference B coded pictures (-51..51)

DirectModeType = 1 # Direct Mode Type (0:Temporal 1:Spatial)

DirectInferenceFlag = 1 # Direct Inference Flag (0: Disable 1: Enable)

BList0References = 0 # B slice List 0 reference override (0 disable, N <= NumberReferenceFrames)

BList1References = 1 # B slice List 1 reference override (0 disable, N <= NumberReferenceFrames)

# 1 List1 reference is usually recommended for normal GOP Structures.

# A larger value is usually more appropriate if a more flexible

# structure is used (i.e. using HierarchicalCoding)

BReferencePictures= 0 # Referenced B coded pictures (0=off, 1=B references for secondary layer, 2=B references for primary layer)

HierarchicalCoding= 0 # B hierarchical coding (0= off, 1= 2 layers, 2= 2 full hierarchy, 3 = explicit)

HierarchyLevelQPEnable = 1 # Adjust QP based on hierarchy level (in increments of 1). Overrides BRefPicQPOffset behavior.(0=off, 1=on)

ExplicitHierarchyFormat = "b1r0b3r0b2e2b0e2b4r2" # Explicit Enhancement GOP. Format is {FrameDisplay\_orderReferenceQP}.

# Valid values for reference type is r:reference, e:non reference.

ExplicitSeqCoding= 0 # Enable support for explicit sequence coding

ExplicitSeqFile= "explicit\_seq.cfg"

LowDelay= 0 # Apply HierarchicalCoding without delay (i.e., encode in the captured/display order)

ReferenceReorder= 1 # Reorder References according to Poc distance for HierarchicalCoding (0=off, 1=enable, 2=use when LowDelay is set)

PocMemoryManagement= 1 # Memory management based on Poc Distances for HierarchicalCoding (0=off, 1=on, 2=use when LowDelay is set)

SetFirstAsLongTerm= 0 # Set first frame as long term

BiPredMotionEstimation = 1 # Enable Bipredictive based Motion Estimation (0:disabled, 1:enabled)

BiPredMERefinements = 3 # Bipredictive ME extra refinements (0: single, N: N extra refinements (1 default)

BiPredMESearchRange = 64 #Bipredictive ME Search range (8 default). Note that range is halved for every extra refinement.

BiPredMESubPel = 2 # Bipredictive ME Subpixel Consideration (0: disabled, 1: single level, 2: dual level)

##########################################################################################

# Output Control, NALs

##########################################################################################

SymbolMode = 1 # 0=CAVLC 1= CABAC

OutFileMode = 0 # Output file mode, 0:Annex B, 1:RTP

##########################################################################################

# Picture based Multi-pass encoding

#########################################################################################

RDPictureDecision= 0 # Perform multiple pass coding and make RD optimal decision among them

RDPSliceBTest= 0 # Perform Slice level RD decision between P and B slices.

RDPictureMaxPassISlice= 1 # Max number of coding passes for I slices, valid values [1,3], default is 1

RDPictureMaxPassPSlice= 2 # Max number of coding passes for P slices, valid values [1,6], default is 2

RDPictureMaxPassBSlice= 3 # Max number of coding passes for B slices, valid values [1,6], default is 3

RDPictureFrameQPPSlice= 0 # Perform additional frame level QP check (QP+/-1) for P slices, 0: disabled (default), 1: enabled

RDPictureFrameQPBSlice= 0 # Perform additional frame level QP check (QP+/-1) for B slices, 0: disabled, 1: enabled (default)

RDPictureDeblocking= 0 # Perform another coding pass to check non-deblocked picture, 0: disabled (default), 1: enabled

RDPictureDirectMode= 0 # Perform another coding pass to check the alternative direct mode for B slices, , 0: disabled (default), 1: enabled

##########################################################################################

# Deblocking filter parameters

##########################################################################################

DFParametersFlag = 0 # Configure deblocking filter (0=parameters below ignored, 1=parameters sent)

# Note that for pictures with multiple slice types,

# only the type of the first slice will be considered.

DFDisableRefISlice = 0 # Disable deblocking filter in reference I coded pictures (0=Filter, 1=No Filter).

DFAlphaRefISlice = 0 # Reference I coded pictures Alpha offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFBetaRefISlice = 0 # Reference I coded pictures Beta offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFDisableNRefISlice = 0 # Disable deblocking filter in non reference I coded pictures (0=Filter, 1=No Filter).

DFAlphaNRefISlice = 0 # Non Reference I coded pictures Alpha offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFBetaNRefISlice = 0 # Non Reference I coded pictures Beta offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFDisableRefPSlice = 0 # Disable deblocking filter in reference P coded pictures (0=Filter, 1=No Filter).

DFAlphaRefPSlice = 0 # Reference P coded pictures Alpha offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFBetaRefPSlice = 0 # Reference P coded pictures Beta offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFDisableNRefPSlice = 0 # Disable deblocking filter in non reference P coded pictures (0=Filter, 1=No Filter).

DFAlphaNRefPSlice = 0 # Non Reference P coded pictures Alpha offset div. 2, {-6, -5, ... 0, +1, .. +6}

DFBetaNRefPSlice = 0 # Non Reference P coded pictures Beta offset div. 2, {-6, -5, ... 0, +1, .. +6}

##########################################################################################

# Error Resilience / Slices

##########################################################################################

SliceMode = 1 # Slice mode (0=off 1=fixed #mb in slice 2=fixed #bytes in slice 3=use callback)

SliceArgument = 120 # Slice argument (Arguments to modes 1 and 2 above)

num\_slice\_groups\_minus1 = 0 # Number of Slice Groups Minus 1, 0 == no FMO, 1 == two slice groups, etc.

#slice\_group\_map\_type = 0 # 0: Interleave, 1: Dispersed, 2: Foreground with left-over,

# 3: Box-out, 4: Raster Scan 5: Wipe

# 6: Explicit, slice\_group\_id read from SliceGroupConfigFileName

#slice\_group\_change\_direction\_flag = 0 # 0: box-out clockwise, raster scan or wipe right,

# 1: box-out counter clockwise, reverse raster scan or wipe left

#slice\_group\_change\_rate\_minus1 = 85 #

#SliceGroupConfigFileName = "0006\_sg0conf.cfg" # Used for slice\_group\_map\_type 0, 2, 6

UseRedundantPicture = 0 # 0: not used, 1: enabled

NumRedundantHierarchy = 1 # 0-4

PrimaryGOPLength = 10 # GOP length for redundant allocation (1-16)

# NumberReferenceFrames must be no less than PrimaryGOPLength when redundant slice enabled

NumRefPrimary = 1 # Actually used number of references for primary slices (1-16)

##########################################################################################

# Search Range Restriction / RD Optimization

##########################################################################################

RestrictSearchRange= 2 # restriction for (0: blocks and ref, 1: ref, 2: no restrictions)

RDOptimization= 1 # rd-optimized mode decision

# 0: RD-off (Low complexity mode)

# 1: RD-on (High complexity mode)

# 2: RD-on (Fast high complexity mode - not work in FREX Profiles)

# 3: with losses

I16RDOpt = 0 # perform rd-optimized mode decision for Intra 16x16 MB

# 0: SAD-based mode decision for Intra 16x16 MB

# 1: RD-based mode decision for Intra 16x16 MB

SubMBCodingState= 1 # submacroblock coding state

# 0: lowest complexity, do not store or reset coding state during sub-MB mode decision

# 1: medium complexity, reset to master coding state (for current mode) during sub-MB mode decision

# 2: highest complexity, store and reset coding state during sub-MB mode decision

DistortionSSIM= 0 # Compute SSIM distortion. (0: disabled/default, 1: enabled)

DistortionMS\_SSIM = 0 # ComputeMultiscale SSIM distortion. (0: disabled/default, 1: enabled)

SSIMOverlapSize= 8 # Overlap size to calculate SSIM distortion (1: pixel by pixel, 8: no overlap)

DistortionYUVtoRGB= 0 # Calculate distortion in RGB domain after conversion from YCbCr (0:off, 1:on)

CtxAdptLagrangeMult= 0 # Context Adaptive Lagrange Multiplier

# 0: disabled (default)

# 1: enabled (works best when RDOptimization=0)

FastCrIntraDecision= 1 # Fast Chroma intra mode decision (0:off, 1:on)

DisableThresholding= 0 # Disable Thresholding of Transform Coefficients (0:off, 1:on)

SkipIntraInInterSlices= 0 # Skips Intra mode checking in inter slices if certain mode decisions are satisfied (0: off, 1: on)

WeightY= 1 # Luma weight for RDO

WeightCb= 1 # Cb weight for RDO

WeightCr= 1 # Cr weight for RDO

##########################################################################################

# Explicit Lambda Usage

##########################################################################################

UseExplicitLambdaParams = 0 # Use explicit lambda scaling parameters (0:disabled, 1:enable lambda weight, 2: use explicit lambda value)

UpdateLambdaChromaME= 0 # Update lambda given Chroma ME consideration

FixedLambdaISlice= 0.1 # Fixed Lambda value for I slices

FixedLambdaPSlice= 0.1 # Fixed Lambda value for P slices

LambdaWeightISlice= 0.65 # scaling param for I slices. This will be used as a multiplier i.e. lambda=LambdaWeightISlice \* 2^((QP-12)/3)

LambdaWeightPSlice= 0.68 # scaling param for P slices. This will be used as a multiplier i.e. lambda=LambdaWeightPSlice \* 2^((QP-12)/3)

LossRateA= 5 # expected packet loss rate of the channel for the first partition, only valid if RDOptimization = 3

LossRateB= 0 # expected packet loss rate of the channel for the second partition, only valid if RDOptimization = 3

LossRateC= 0 # expected packet loss rate of the channel for the third partition, only valid if RDOptimization = 3

FirstFrameCorrect = 0 # If 1, the first frame is encoded under the assumption that it is always correctly received.

NumberOfDecoders = 30 # Numbers of decoders used to simulate the channel, only valid if RDOptimization = 3

RestrictRefFrames= 0 # Doesnt allow reference to areas that have been intra updated in a later frame.

##########################################################################################

# Additional Stuff

#########################################################################################

UseConstrainedIntraPred = 0 # If 1, Inter pixels are not used for Intra macroblock prediction.

NumberofLeakyBuckets= 8 # Number of Leaky Bucket values

LeakyBucketRateFile= "leakybucketrate.cfg" # File from which encoder derives rate values

LeakyBucketParamFile= "leakybucketparam.cfg" # File where encoder stores leakybucketparams

NumFramesInELayerSubSeq = 0 # number of frames in the Enhanced Scalability Layer(0: no Enhanced Layer)

SparePictureOption= 0 # (0: no spare picture info, 1: spare picture available)

SparePictureDetectionThr = 6 # Threshold for spare reference pictures detection

SparePicturePercentageThr = 92 # Threshold for the spare macroblock percentage

PicOrderCntType = 0 # (0: POC mode 0, 1: POC mode 1, 2: POC mode 2)

########################################################################################

#Rate control

########################################################################################

RateControlEnable = 0 # 0 Disable, 1 Enable

Bitrate = 45020 # Bitrate(bps)

InitialQP = 0 # Initial Quantization Parameter for the first I frame

# InitialQp depends on two values: Bits Per Picture,

# and the GOP length

BasicUnit = 0 # Number of MBs in the basic unit

# should be a fraction of the total number

# of MBs in a frame ("0" sets a BU equal to a frame)

ChannelType = 0 # type of channel( 1=time varying channel; 0=Constant channel)

RCUpdateMode = 0 # Rate Control type. Modes supported :

# 0 = original JM rate control,

# 1 = rate control that is applied to all frames regardless of the slice type,

# 2 = original plus intelligent QP selection for I and B slices (including Hierarchical),

# 3 = original + hybrid quadratic rate control for I and B slice using bit rate statistics

#

RCISliceBitRatio = 1.0 # target ratio of bits for I-coded pictures compared to P-coded Pictures (for RCUpdateMode=3)

RCBSliceBitRatio0 = 0.5 # target ratio of bits for B-coded pictures compared to P-coded Pictures - temporal level 0 (for RCUpdateMode=3)

RCBSliceBitRatio1 = 0.25 # target ratio of bits for B-coded pictures compared to P-coded Pictures - temporal level 1 (for RCUpdateMode=3)

RCBSliceBitRatio2 = 0.25 # target ratio of bits for B-coded pictures compared to P-coded Pictures - temporal level 2 (for RCUpdateMode=3)

RCBSliceBitRatio3 = 0.25 # target ratio of bits for B-coded pictures compared to P-coded Pictures - temporal level 3 (for RCUpdateMode=3)

RCBSliceBitRatio4 = 0.25 # target ratio of bits for B-coded pictures compared to P-coded Pictures - temporal level 4 (for RCUpdateMode=3)

RCBoverPRatio = 0.45 # ratio of bit rate usage of a B-coded picture over a P-coded picture for the SAME QP (for RCUpdateMode=3)

RCIoverPRatio = 3.80 # ratio of bit rate usage of an I-coded picture over a P-coded picture for the SAME QP (for RCUpdateMode=3)

RCMinQPPSlice= 8 # minimum P Slice QP value for rate control

RCMaxQPPSlice = 44 # maximum P Slice QP value for rate control

RCMinQPISlice= 8 # minimum I Slice QP value for rate control

RCMaxQPISlice = 36 # maximum I Slice QP value for rate control

########################################################################################

#Fast Mode Decision

########################################################################################

EarlySkipEnable = 0 # Early skip detection (0: Disable 1: Enable)

SelectiveIntraEnable = 0 # Selective Intra mode decision (0: Disable 1: Enable)

ReportFrameStats = 0 # (0:Disable Frame Statistics 1: Enable)

DisplayEncParams = 0 # (0:Disable Display of Encoder Params 1: Enable)

Verbose = 1 # level of display verboseness

# 0: short, 1: normal (default), 2: detailed, 3: detailed/nvb

# 4: with additional MB level lambda info

########################################################################################

#Rounding Offset control

########################################################################################

OffsetMatrixPresentFlag = 0 # Enable Explicit Offset Quantization Matrices (0: disable 1: enable)

QOffsetMatrixFile = "q\_offset.cfg" # Explicit Quantization Matrices file

AdaptiveRounding = 1 # Enable Adaptive Rounding based on JVT-N011 (0: disable, 1: enable)

AdaptRoundingFixed = 1 # Enable Global Adaptive rounding for all qps (0: disable, 1: enable - default/old)

AdaptRndPeriod = 16 # Period in terms of MBs for updating rounding offsets.

# 0 performs update at the picture level. Default is 16. 1 is as in JVT-N011.

AdaptRndChroma = 1 # Enables coefficient rounding adaptation for chroma

AdaptRndWFactorIRef = 4 # Adaptive Rounding Weight for I/SI slices in reference pictures /4096

AdaptRndWFactorPRef = 4 # Adaptive Rounding Weight for P/SP slices in reference pictures /4096

AdaptRndWFactorINRef = 4 # Adaptive Rounding Weight for I/SI slices in non reference pictures /4096

AdaptRndWFactorPNRef = 4 # Adaptive Rounding Weight for P/SP slices in non reference pictures /4096

AdaptRndCrWFactorIRef = 4 # Chroma Adaptive Rounding Weight for I/SI slices in reference pictures /4096

AdaptRndCrWFactorPRef = 4 # Chroma Adaptive Rounding Weight for P/SP slices in reference pictures /4096

AdaptRndCrWFactorINRef = 4 # Chroma Adaptive Rounding Weight for I/SI slices in non reference pictures /4096

AdaptRndCrWFactorPNRef = 4 # Chroma Adaptive Rounding Weight for P/SP slices in non reference pictures /4096

########################################################################################

#Fast Motion Estimation Control Parameters

########################################################################################

SearchMode = 1 # Motion estimation mode

# -1 = Full Search

# 0 = Fast Full Search (default)

# 1 = UMHexagon Search

# 2 = Simplified UMHexagon Search

# 3 = Enhanced Predictive Zonal Search (EPZS)

UMHexDSR = 1 # Use Search Range Prediction. Only for UMHexagonS method

# (0:disable, 1:enabled/default)

UMHexScale = 3 # Use Scale\_factor for different image sizes. Only for UMHexagonS method

# (0:disable, 3:/default)

# Increasing value can speed up Motion Search.

EPZSPattern = 2 # Select EPZS primary refinement pattern.

# (0: small diamond, 1: square, 2: extended diamond/default,

# 3: large diamond, 4: SBP Large Diamond,

# 5: PMVFAST )

EPZSDualRefinement = 3 # Enables secondary refinement pattern.

# (0:disabled, 1: small diamond, 2: square,

# 3: extended diamond/default, 4: large diamond,

# 5: SBP Large Diamond, 6: PMVFAST )

EPZSFixedPredictors = 2 # Enables Window based predictors

# (0:disabled, 1: P only, 2: P and B/default)

EPZSTemporal = 1 # Enables temporal predictors

# (0: disabled, 1: enabled/default)

EPZSSpatialMem = 1 # Enables spatial memory predictors

# (0: disabled, 1: enabled/default)

EPZSBlockType = 1 # Enables block type Predictors

# (0: disabled, 1: enabled/default)

EPZSMinThresScale = 0 # Scaler for EPZS minimum threshold (0 default).

# Increasing value can speed up encoding.

EPZSMedThresScale = 1 # Scaler for EPZS median threshold (1 default).

# Increasing value can speed up encoding.

EPZSMaxThresScale = 2 # Scaler for EPZS maximum threshold (1 default).

# Increasing value can speed up encoding.

EPZSSubPelME = 1 # EPZS Subpel ME consideration

EPZSSubPelMEBiPred = 1 # EPZS Subpel ME consideration for BiPred partitions

EPZSSubPelThresScale = 2 # EPZS Subpel ME Threshold scaler

EPZSSubPelGrid = 1 # Perform EPZS using a subpixel grid

########################################################################################

# SEI Parameters

########################################################################################

GenerateSEIMessage = 0 # Generate an SEI Text Message

SEIMessageText = "H.264/AVC Encoder" # Text SEI Message

UseMVLimits = 0 # Use MV Limits

SetMVXLimit = 512 # Horizontal MV Limit (in integer units)

SetMVYLimit = 512 # Vertical MV Limit (in integer units)

########################################################################################

# VUI Parameters

########################################################################################

# the variables below do not affect encoding and decoding

# (many are dummy variables but others can be useful when supported by the decoder)

EnableVUISupport = 0 # Enable VUI Parameters

## Annex 2

Configuration file for H.264 decoding for JM 18.2

DroneSeq2\_ip64\_qi32\_qp32\_qb32\_hier0\_pcap\_s1920x1080p25n400v0.264 ........H.264/AVC coded bitstream

DroneSeq2\_ip64\_qi32\_qp32\_qb32\_hier0\_pcap\_s1920x1080p25n400v0.yuv ........Output file, YUV/RGB

test\_rec.yuv ........Ref sequence (for SNR)

1 ........Write 4:2:0 chroma components for monochrome streams

0 ........NAL mode (0=Annex B, 1: RTP packets)

0 ........SNR computation offset

2 ........Poc Scale (1 or 2)

500000 ........Rate\_Decoder

104000 ........B\_decoder

73000 ........F\_decoder

leakybucketparam.cfg ........LeakyBucketParams

1 ........Err Concealment(0:Off,1:Frame Copy,2:Motion Copy)

2 ........Reference POC gap (2: IPP (Default), 4: IbP / IpP)

2 ........POC gap (2: IPP /IbP/IpP (Default), 4: IPP with frame skip = 1 etc.)

0 ........Silent decode

1 ........Enable Deblocking filter in intra only profiles (0=disable, 1=filter according to SPS parameters)

DroneSeq2\_ip64\_qi32\_qp32\_qb32\_hier0\_pcap\_s1920x1080p25n400v02.xml ........XML trace

4 ........XML trace log level

## Annex 3

Avisynth script used for “converting 2D and 3D”, i.e. applying a geometrical transformation

offset = 1

scale = 10.0

LoadPlugin("C:\Program Files\AviSynth 2.5\plugins\MT.dll")

setMTmode(2,0)

AviSource("Z:/home/hrc/Documents/hrc\_scripts/encode\_JM18\_2Dto3D/src/./DroneSeq2\_s1920x1080p25n400v0.avi")

ConvertToYUY2()

super = MSuper(pel=1,hpad=0,vpad=0,chroma=false)

backward\_vec = MAnalyse(super,blksize=32,isb=true,chroma=false,searchparam=1,search=0)

forward\_vec = MAnalyse(super,blksize=32,isb=false,chroma=false,searchparam=1,search=0)

MFlowFps(super,backward\_vec,forward\_vec,num=2\*FramerateNumerator(last),den=FramerateDenominator(last),mask=0)

width\_scaled = width()\*(1.0+scale/100)

cropper = int((width\_scaled - width())\*0.5)

cropper = cropper + cropper%2

videoR = last

videoL = Trim(videoR, offset, FrameCount(videoR))

videoR = videoR.crop(cropper,0,-cropper,0).lanczosresize(width(videoR),height(videoR))

StackHorizontal(videoL,Trim(videoR,0,FrameCount(videoR)-1))

## Annex 4

Simulation of packet transmission with Sirannon



Annex 5:

Simulation of packet loss with Sirannon



Annex 6:

Simulation of a packet receiver using Sirannon

