# Quality in Wireless Transmissions of 3D Video with the HEVC Standard

## Jesús D. Quintero Polanco\*, Martin D. Bravo Obando and Diego F. Sendoya

Department of Electronic Engineering, Surcolombiana University, Colombia; diego.sendoya@usco.edu.co, jdavid@usco.edu.co, martin.bravo@usco.edu.co

# Abstract

**Objectives:** In this work, the assessment of 3D video quality is performed with the HEVC (High Efficiency Video Coding) standard over a wireless transmission considering the diversity MIMO schemes. Methods/Statistical Analysis: The video sequences in both Multiview Extension and 3D-Extension with its depth map using the HEVC standard were encoded. In addition, three different diversity MIMO schemes and a turbo codec in the system were considered. The BER results obtained using the Montecarlo simulation for packet loss were implemented to evaluate the 3D video sequence at the output of the system, it based in the objective and subjective assessment. Findings: In 3D video encoding with multiple views using HEVC standard with GOP All-Intra coding structure, packets of different texture views can be lost, but in 3D video encoding with DIBR and depth data, you cannot miss any package of the texture views, because the video cannot be decoded. The coding structures of GOP type IPBB and IPPP is more susceptible to the package loss, so the transmission protocol TCP is the suitable. Therefore, this scheme in real- time transmission is not recommended. The increase of diversity in the MIMO schemes decreases the package loss, in this way, the increasing of the video quality that is verified using the objective assessment, but it with a cost of velocity transmission due that is required more intervals of transmission in the MIMO scheme i.e. (with redundant information). The packet loss in wireless transmissions with MIMO 3D video schemes encoded with HEVC standard it has greater impact on the degradation of the video sequences of high mobility than the low mobility. Application/Improvements: The assessment of 3D Video quality using the HEVC standard and its transmission through a wireless channel is using the diversity MIMO schemes which takes advantage of spatial dimension.

Keywords: Video Quality, 3D Video, HEVC, MIMO

# 1. Introduction

Video technology has evolved from its beginnings and with it the need for multimedia transmission in wired and wireless mode. In the last two decades, there has been a great development in the quality of the image provided to the user, leading now to have HD and Ultra HD quality<sup>1</sup>, creating requirements of high demand to the systems of encoding and transmission in terms of packet loss and distortion, that image compression is optimal and bandwidth decreases.

The appearance of video applications in three dimensions generates greater expectations for the user in terms of services, achieving greater demands in the coding and video transmission standard. That is why with the standard of encoding HEVC ("High Efficiency Video Coding")<sup>2</sup> its extension 3D-Extension HEVC<sup>3</sup>, focused on video coding MVC and DIBR ("Depth Image Based Rendering<sup>4</sup> in constant development by experts from around the world, the project is aimed at evaluating the quality of 3D videos in terms of their condition, transmission and decoding, guiding this evaluation to the loss of information packages.

The main objective of the project is to evaluate the quality of 3D video, using the extensions of the HEVC coding standard, on wireless channels with considered flat frequency.

The content of this article is organized in the following manner. Section 1 describes the configuration of the HEVC scheme, 3D video coding, data transport and packaging, wireless transmission<sup>5</sup>, the estimation of the loss of packages and the acquisition of objective and subjective parameters of quality<sup>6</sup>. In section 2, the results of the transmission schemes, the HEVC scheme and the final comparison between the objective and subjective quality parameters are discussed.

# 2. Methodology

To carry out the research it was necessary to divide the project into stages with constant feedback between them. The stages that were developed were

## Stage 1: Configuration of the Scheme HEVC

This stage consisted in making the configuration of the HEVC scheme that encodes 3D video taking advantage of the correlation that exists between the different views. For the investigative work, the JMVC extension of the H.264 recommendations was studied to configure 3D-Extension of H.265. These recommendations allow an efficient coding of sequences captured simultaneously from several video cameras, to be able to predict images that are not only based on the temporal relationship between frames of the same sequence but are based on frames of images sequences of adjacent cameras.

For the configuration of the HEVC scheme, there was collaboration of the mailing list of the "Joint Collaborative Team on Video Coding (JCT-VC)", which gathers and contacts video coding experts around the world<sup>7</sup>.

## Stage 2: 3D Video Coding

In this stage, videos developed in laboratories were used under controlled parameters that are apt to encode in both Multiview Extension (MV-HEVC) and 3D-Extension (3D-HEVC) with its depth map and its subsequent synthesis of views (rendering) for displays or 3D screens.

With the video sequences of Table 1, the coding tests were carried out. It specifies its resolution, the number

Table 1.	3D video sequences
I HOIV II	ob flaco sequences

Sequence of video	Resolution	Frame to encode	Frame/s
Newspaper	1024 X 768	300	30
Poznan_ Street	1920 X 1088	250	25
Undo_ Dancer	1920 X 1088	250	25

of frames that each video has and the number of frames that are played per second in YUV 4: 2: 0 format<sup>2</sup>. The Newspaper sequence is of standard resolution, the Poznan\_Street sequence is UltraHD resolution with low mobility, and the Undo\_Dancer sequence is UltraHD resolution with high mobility.

In the cases of coding that were proposed for the different groups of video views, it focused on the modification and efficiency of the GOP (Group of Pictures) for transmission with packet loss. Each GOP must start with an I image, which are images that only use intraframe compression where each image is processed independently of the others and contain all the information necessary for its reconstruction. The P images or prediction images are generated from the closest I image that precede it. The B images are generated from previous type I or P images. Table 2 refers to the different types of GOP (Group of Pictures) of size 8 and search and matching 64 that were worked for the IPPP and IPBB structures, whereas the GOP IIII (All-Intra) structures are of size 1 and rank of search and coincidence 64.

Fable 2.	Coding test	cases
----------	-------------	-------

	IPBB	
2 Views	IIII	
	IPPP	
	IPBB	
2 Views + Depth	IIII	
	IPPP	
	IPBB	Size of GOP
3 Views	IIII	
	IPPP	
	IPBB	
3 Views + Depth	IIII	
	IPPP	

An example of a coding structure of Table 2 and how it is created is seen in Figure 1.

## Stage 3: Transport and Data Packaging

The packaging of the coding is a very important part of the HEVC standard, since, with it, a significant improvement in the transmission and distortion rate is obtained. The data encoded with HEVC is encapsulated in NAL (Network Abstraction Layer) which contains an integer



G

#### Figure 1. A GOP structure.

number of bits. To understand how NALs are encapsulated in H.265, the ITU-T recommendations of the telecommunications standard are studied and analysed<sup>8</sup>. The HEVC bit stream format can be categorized as a stream of NAL units or a stream of bytes. The format of NAL units is conceptually the most basic, where the sequences of NALs are organized in order of decoding, having imposed certain limitations. The byte stream format can be constructed from the stream format of NALs units in order of decoding and by prefixing each NAL unit with a start code prefix with bytes of value 0 to form a stream of bytes.

In this stage, a software component was developed in Python3 language that analyzes the correlation between the layer of depth and texture proposed as objective, identifying the number of NALs, the size of the packages, the NAL type and the view to the which belongs.

#### **Stage 4: Wireless Transmission**

This stage consists of two parts:

#### a) Verification of Transmission Schemes

Three different MIMO schemes were implemented, which must be verified first. The first scheme is the traditional Alamouti 2X2, with a maximum order of diversity of 2N\_rx, where N\_rx is the number of receiving antennas, and block rate of 1, since it transmits 2 symbols every 2 instants of time. The block code of this scheme is expressed in Equation 1:

$$G2 = \begin{pmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{pmatrix}$$
(1)

The second scheme is a 4X4 MIMO with a block rate of  $\frac{1}{2}$  with a maximum order of diversity of 4N\_rx, allowing the transmission of 4 symbols every 8 instants of time. Your block code is expressed in Equation 2.

$$G4 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ -s_2 & s_1 & -s_4 & s_3 \\ -s_3 & s_4 & s_1 & -s_2 \\ -s_4 & -s_3 & s_2 & s_1 \\ s_1^* & s_2^* & s_3^* & s_4^* \\ -s_2^* & s_1^* & -s_4^* & s_3^* \\ -s_3^* & s_4^* & s_1^* & -s_2^* \\ -s_4^* & -s_3^* & s_2^* & s_1^* \end{pmatrix}$$
(2)

The third scheme is an 8X8 MIMO with a block rate of ½ with a maximum order of diversity of 8N\_rx, transmitting 8 symbols every 16 instants of time. Your block code is expressed in Equation 3:

$$16 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ -s_2 & s_1 & s_4 & -s_3 & s_6 & -s_5 & -s_8 & s_7 \\ -s_3 & -s_4 & s_1 & s_2 & s_7 & s_8 & -s_5 & -s_6 \\ -s_4 & s_3 & -s_2 & s_1 & s_8 & -s_7 & s_6 & -s_5 \\ -s_5 & -s_6 & -s_7 & -s_8 & s_1 & s_2 & s_3 & s_4 \\ -s_6 & s_5 & -s_8 & s_7 & -s_2 & s_1 & -s_4 & s_3 \\ -s_7 & s_8 & s_5 & -s_6 & -s_3 & s_4 & s_1 & -s_2 \\ -s_8 & -s_7 & s_6 & s_5 & -s_4 & -s_3 & s_2 & s_1 \\ s_1^* & s_2^* & s_3^* & s_4^* & s_5^* & s_6^* & s_7^* & s_8^* \\ -s_2^* & s_1^* & s_4^* & -s_3^* & s_6^* & -s_5^* & -s_8^* & s_7^* \\ -s_3^* & -s_4^* & s_1^* & s_2^* & s_7^* & s_8^* & -s_7^* & s_6^* & -s_5^* \\ -s_5^* & -s_6^* & -s_7^* & -s_8^* & s_1^* & s_2^* & s_3^* & s_4^* \\ -s_6^* & s_5^* & -s_8^* & s_7^* & -s_2^* & s_1^* & -s_4^* & s_3^* \\ -s_6^* & s_5^* & -s_8^* & s_7^* & -s_2^* & s_1^* & -s_4^* & s_3^* \\ -s_6^* & s_5^* & -s_8^* & s_7^* & -s_2^* & s_1^* & -s_4^* & s_3^* \\ -s_8^* & -s_7^* & s_6^* & s_5^* & -s_4^* & -s_3^* & s_2^* & s_1^* \end{pmatrix}$$

$$(30)$$

The block diagrams are formed by the symbols s that are transmitted and their respective complex conjugate s ^ \*. For the 2X2 MIMO scheme, the distribution of the symbols is of the OSTBC nature, and the MIMO 4X4 and 8X8 schemes are STBC in nature.

The 3 schemes were implemented in C ++ language using IT ++ communication libraries<sup>9</sup>, performing QAM modulation of order 16 at different values 💷 of signal-to-noise ratio



Figure 2. Verification of the schemes.

(SNR) and using the Monte Carlo method, as suggested in other works<sup>10</sup>, to have as a final value the bit error rate (BER), parameter necessary for the analysis of packet loss for a real case of 3D video transmission. The verification results of the schemes are shown in Figure 2. It is observed that BER decreases as the complexity of the scheme increases. This is due to the diversity gain of each scheme, which increases according to the number of antennas increases.

#### b) Channel Coding

Once the transmission schemes have been verified, the channel encoder is implemented. The Turbo Códec was configured, since it is the indicated one to obtain favorable results based on the signal to noise ratio. This channel encoder was implemented in the transmission schemes with the IT ++ library.

#### **Stage 5: Loss of Packages**

With the BER results from the previous stage, the packet loss is formulated in a MIMO environment, [PER] \_ MIMO. For the estimation of [PER] \_MIMO each bit within the packet has the same BER and the bit errors are not correlated, then [PER] \_MIMO can be related to the BER through a packet containing L bits<sup>4</sup>.

$$PER_{MIMO} = 1 - \left(1 - BER\right)^{L} \tag{4}$$

With this value of PERMIME, the identification of the number of NALs of each bitstream, that the size of the NAL units is variable and taking as reference that the maximum size of TCP packing of an MSS (Maximum Segment Size) according to the RFC 6691 standard is 525 bytes<sup>11</sup>, the number of lost packets is estimated:

$$Packet_{Lost} = \frac{N_{NALs} * \frac{M_{NALs}}{525} * PER_{MIMO}}{100}$$
(5)

Where N\_NALs is the total number of NAL units in the bitstream, M\_NALs is the average size of the NALs in the bitstream being analyzed and [PER] \_MIMO is the packet loss in the MIMO environment. MIMO. Alamouti 2X2 (a), MIMO 4X4 (b) y MIMO 8X8 (c).

#### **Stage 6: Objective Evaluation**

To carry out the objective evaluation, bitstreams are decoded with the different patterns of packet loss, process

in which the video is obtained in .YUV format and the aim is to determine the signal to noise ratio (PSNR) for the frames that make up the video original and the degraded video and thus determine the quality level when comparing the two values.

To calculate the PSNR, we first have to determine the Mean Square Error (MSE) of each frame by Equation 6.

$$MSE = \frac{1}{MNT} \sum_{i=1}^{M} \sum_{i=1}^{N} \sum_{i=1}^{T} \left[ x(m,n,t) - y(m,n,t) \right]^{2}$$
(6)

Where the image has M \* N pixels and T frames; x, y corresponds to the pixels of the original signal and the degraded signal respectively.

In this work, when this measure is already taken for a couple of frames (original - degraded) in a plane (R, G or B), it is only necessary to perform the same calculation for the other pairs of frames and then averaging them. Then the same is done for the other two planes.

$$MSE_{v} = \frac{1}{3F} \left( \sum_{j=1}^{F} MSE_{R}(j) + \sum_{j=1}^{F} MSE_{G}(j) + \sum_{j=1}^{F} MSE_{B}(j) \right)$$
(7)

$$PSNR = 10 \log_{10} \left( \frac{L^2}{MSE_v} \right)$$
(8)

Where F is the number of frames and MSER (j); MSEG (j); MSEB (j) is the value of the squared error calculated for the j-frame in the plane R, G,  $B^{12}$ . Finally, Table 3 that relates the PSNR with the MOS is exposed<sup>12</sup>.

 Table 3.
 Relationship between PSNR and MOS values

PSNR (dB)	MOS
> 37	5
31 - 37	4
25 - 31	3
20 - 25	2
< 20	1

During the development of the project we worked with a software component programmed in C ++ with the help of the OpenCV video and image treatment library, and

the PSNR of the frames was evaluated, but for reasons of agility in the evaluation process, the tests were they did with the h.264 JMVC software.

#### **Stage 7: Subjective Evaluation**

The objective of this experiment is to obtain the subjective assessment of the quality of a certain number of video sequences by viewers, since these being the final receivers, more safety and reliability in the determination of the quality of the sequences is achieved.

Before beginning the evaluation, a Snellen test is carried out as suggested in recommendation ITU-T 500-11<sup>13</sup>, then a scenario is adapted to certain requirements and a physical format is delivered so that the viewer would give his opinion of the projected sequences.

The tests consisted of the screening of 5 sessions, each with 2 3D video sequences in such a way that they did not exceed 20 minutes<sup>14,15</sup>. The projections were made in a controlled space whose background lighting conditions were less than 20 lux. In each session, the original 3D video sequence "Undo\_Dancer" was projected in UltraHD, which lasts 9 seconds, followed by a 5 second relaxation period to finally project the decoded video sequence to be evaluated. After the screening of this last sequence, an estimated 40 seconds were given for the viewers to rate the video sequence according to the indications given in the introduction by one of the project's members.

For the execution of the subjective test, the DSIS (Degradation Scale with Double Stimulus) method was taken due to its low complexity and allows the degradation level to be qualified between the original and the coded video. The total sample at the end of the test is 18 users.

The presentation structure of the subjective test sessions is shown in Figure 3. The time T1 corresponds to the duration of the reference video; T2 is the relaxation period between the reference video and the video to be evaluated; T3 is the length of the encoded and degraded video which the viewers will give their subjective rating comparing it with the reference video; and T4 corresponds to the space assigned to qualify the video in the evaluation format delivered in physics.

The DSIS method has a default category of values for MOS interpretation where viewers choose one of five options, as shown in Table 4.



Figure 3. Presentation structure of the test materials.

Table 4.	Categories	of values	□□ for	DSIS
----------	------------	-----------	--------	------

Description	MOS
The degradation is imperceptible	5
The degradation is imperceptible, but does not annoying	4
The degradation is slightly annoying	3
The degradation is annoying	2
The degradation is very annoying	1

# 3. Results and Discussion

# 3.1 Transmission Schemes

The results obtained from the wireless transmission channels that were designed with the implementation of the channel encoder were examined.

This simulation was carried out with an input signal generated with bits of random form whose size was equal to or greater than 1.04MB (1,048,576 bytes), the modulation was the same proposed for the verification of the systems (16QAM) and ten were obtained one thousand (10,000) iterations of BER for each estimated SNR value.

These iterations obey the Monte Carlo method, in order to calculate a precise value of BER in each process. The results of this simulation are shown in Figure 4.

From the results shown in Figure 4, we can determine that:

- 1. The complexity of the schemes continues to play an important role, since, as the diversity gain increases, the bit error rate decreases.
- 2. By implementing the channel encoder in any of the schemes, the bit error rate is improved by considerably reducing the signal to noise (SNR) of the system.
- 3. The 3D video transmission requires channel coding, and it works with a bit error rate of 10-4 with a





**Figure 4.** Results MIMO Schematics with Turbo Codec. (a) MIMO 2X2, (b) 4X4 and (c) 8X8.

relatively low signal to noise ratio, which is why it is considered as the best channels for the transmission of the 4X4 and 8X8.

Once having the values DD of BER of each scheme, the analysis of the loss of packages for a real case of transmission of 3D video is made.

## 3.2 Standard HEVC

In this section, coding tests focused on working Group of Pictures (GOP) structures that were not affected by packet loss were performed. Taking into account the experiences with 2D video coding, it was concluded that the best structure to encode the video and minimize packet losses was the IPPP. This type of structure is represented in Figure 5. Because two or more views are worked on in the 3D video, always having a base or main view, tests of coding schemes with GOP structures of 3 views were made as shown in Figure 6.



Figure 5. Structure IPPP.



**Figure 6.** First GOP structure of 3 views.

Where each line represents the GOP of each view, in which, view B is the base or main view. During the tests, it was possible to demonstrate that with this type of structure, no packet or bit can be lost that is going to be transmitted and the only way to transmit with this type of GOP structure is oriented in the TCP protocol. A second GOP structure was also tested for three views shown in Figure 7.

It was evidenced that if the GOP structure of Figure 6 of hierarchy P is used instead of hierarchy B a large amount of losses for 3D video arise since the distance between the objective image and the reference image increases, and this increase does not provide no coding benefit by bi-prediction of encoded images subsequently. However, it is remembered that simply using the IPPP structure in 2D video generates acceptable losses.

For the tests of this work oriented to the loss of packages it was concluded that the best GOP structure to use was All-Intra. Its structure is represented in Figure 8. This structure is from a cfg file for a 3D video view.



Figure 7. Second GOP structure of 3 views.

Ι	II	III	IV	V	VI	VII	VIII	IX	Х	XI	XII	XIII
FrameI:	Ι	0	0	0.442	0	0	0	0	0	0	0	1
Frame1:	В	1	1	0.442	0	0	0	4	0	0	0	1

Figure 8. Structure All-Intra.

Two video sequences from Table 1 of high and low mobility with All-Intra GOP structure were coded, and the packet loss formula named in Equation (5) was applied for each BER value, of each signal-to-noise ratio., resulting from the previous section.

## 3.3 Video 3D of 2 View

For each BER value, 40 iterations of packet loss were made and each of them was decoded and the PSNR measurement of each view was made.

The relationship of SNR versus PSNR is shown in Figure 9 for the case of the Poznan\_Street video sequence, and in Figure 10 for the case of the Undo\_Dancer video sequence, where the PSNR value is the measure of the 40 iterations with each BER value of the 3 transmission schemes worked.

The graphs show that in a real case of wireless transmission of 3D video of multiple views, the greater the complexities of the transmission channels, the lower are the packet losses. This loss to the same error of BER affects more the quality of the high mobility video Undo\_Dancer than the low mobility Poznan\_ Street. In the highest values III of SNR there are no packet losses as it is in the case of the 8x8 scheme, where at positive SNR values IIII the quality is excellent.

When losing packets in the transmission of 3D video, the loss is more influential on one of the views than on the rest, so that the overall quality of the video is not affected



**Figure 9.** Relationship SNR against PSNR Poznan\_ Street of 2 views.

to a large extent, and taking into account that the structure GOP worked is All-Intra, the losses in one of the views does not greatly affect the remaining views, as evidenced in Figure 11, where we can see that the luminance



Figure 10. SNR vs PSNR Undo\_Dancer of 2 view.

component (Y) is the most affected by the losses and components of blue and red chrominance (Cb and Cr) are affected to a lesser extent, since the structure .YUV of the videos worked is 4: 2: 0. Each point represents an



Figure 11. Sequence data sample Undo\_ Dancer.

iteration of the data obtained in the 8x8 scheme with the sequence Undo\_Dancer.

# 3.4 Video of 2 Views Plus Depth and Rendering

The Poznan\_ Street video sequence of two views was coded, each with its corresponding depth layer, resulting in 4 views for the transmission with All-Intra GOP structure. During the tests it was possible to demonstrate that in this type of video sequences only data or packages corresponding to the depth views can be lost, since when losing information of the texture views, the decoding process is unsuccessful and all data is lost information. The results of this test are shown in Figure 12.

Texture views remain constant in the excellent quality range while depth layers lose information. Once this information is decoded, the rendering process is performed, through which 3 texture views are synthesized with the information of the depth layers and the cfg file of the original video cameras. The results of the evaluation



Figure 12. SNR vs PSNR Video 3D of 2 view.

of the video rendering with losses after transmission are shown in Figure 13.

It is noted that the views I and V correspond to the texture views transmitted without loss whereas the views II, III and IV are those that resulted from the rendering. These views have losses since they were created from the information of depth layers with packet losses. It is worth mentioning that the losses are not significant to affect the quality of the video.



Figure 13. Rendering.

## 3.5 Standard Coding Quality HEVC

To evaluate the quality at a different bit rate of the HEVC coding standard, the parameter QP (Quantization Parameter) was varied, which is inversely proportional to the quality of the video. The parameter can be modified in integers ranging from 21 to 54. The results for

the Newspaper video sequence with the specifications in Table 1 with GOP IBBB structure are shown in Figure 14. Taking into account Figure 14, it can be said that the higher the bit rate, the higher the coding quality.



Figure 14.PSNR Newspaper.

# 3.6 Subjective Evaluation

In this section, the results obtained from the subjective test are examined. As a first point to be discussed, all the data and information obtained from the 18 users who participated in the subjective test and their respective MOS scale scores were tabulated and sorted in a spreadsheet.

Each video sequence has a different PSNR value. These values III are found in Table 5. We can see that the sessions were projected from lower to higher quality, comparing them with the quality of the original video.

Taking as a reference Table 3 where the PSNR values are related to the MOS scale and Table 4 that discusses the value scheme for the DSIS evaluation, it was determined that the 3D video sequences evaluated in sessions 1 and 2 correspond to the MOS scale of order 2 ("The degradation is annoying"), the sequence in session 3 corresponds to the MOS scale of order 3 ("The degradation is slightly annoying"), and the sequences of sessions 4 and 5 to the MOS scale of order 4 ("The degradation is noticeable but does not bother"). The original 3D video used as a reference for the subjective rating corresponds to the MOS scale of order 5 ("The degradation is imperceptible"). This information is displayed in Figure 15. After having an average MOS measurement for each qualifying session subjectively, we proceeded to graph the results as shown in Figure 16.



Figure 15. Ideal MOS measurement.



Figure 16. Measure MOS subjective test.

Figure 17 shows the comparison and relationship between Figures 15 and 16. Based on this, we can determine that:

1. The 3D video sequence evaluated in the fifth session was the best valued by the viewers, which is acceptable since the PSNR value of the views of the 3D video projected in this session were the highest.

Table 5.PSNR	values DD of 3D	video sequences
--------------	-----------------	-----------------

View	Session 1(dB)	Session 2(dB)	Session 3(dB)	Session 4(dB)	Session 5(dB)	Original (dB)
1st	21,69	24,24	28,95	33,87	36,71	38,44
2nd	22,13	24,16	27,88	34,15	36,8	38,44

- 2. In the same way, it can be seen that the viewers gave the worst ratings for the video sequences projected in sessions 1 and 2, which is also acceptable when having their views, the lowest PSNR values II worked.
- 3. You can see in the graph that the scale from lowest to highest quality matches the order of the sessions of the projected video sequences.
- 4. The subjective qualification of sessions 1 and 2 are similar although they are approximated to the MOS scale of order 3. This is due to the so different ratings of the spectators for their visual perception. However, the sequences with the lowest scores of all the evaluated ones remain.
- 5. The objective of this subjective test was to demonstrate how the loss of packets based on the Bit Error Rate (BER) of a MIMO wireless transmission system affects the quality of the 3D video of two views in the perception of a real user.



**Figure 17.** Comparison between ideal value of MOS and average MOS of subjective evaluation.

# 4. Conclusions

In 3D video encoding with multiple views, following the recommendations of the HEVC standard with GOP All-Intra coding structure, packets of the different texture views can be lost; but in the 3D video encoding with DIBR and depth data, you cannot miss any package of the texture views, but of the depth views. In case you lose information about the texture views, the video cannot be decoded.

3D video encoding with the recommendations of the HEVC standard with GOP coding structures of type IPBB and IPPP is very susceptible to information loss, so the

most recommended transmission protocol for transmission is TCP. The implementation of these schemes in realtime transmission is not recommended.

The increase in diversity in the MIMO wireless transmission schemes decreases transmission loss by increasing the quality in the objective measurements of the transmission and the perception of the user in the information transmitted.

The implementation and standardization of the packaging of NAL units, in coding and transmission of 3D video with HEVC, increases and benefits the transmission and decreases the loss of packages, because by the recommendations of the HEVC standard, a classification of NALs with different priority levels, in such a way that they are presented in different sizes depending on the frame section of the video to which they belong and their level of importance within the GOP of the video encoding.

The loss of information in the transmission of 3D video in wireless schemes affects more significantly one of the views, in comparison with the others, so that if taking advantage of the correlation between the coded views, the information of the view with loss, when obtaining and taking advantage of the MV (movement vectors) created during coding.

The loss of information in the videos transmitted in the YUV 4: 2: 0 formats was presented to a greater extent in the luminance components and little in the chrominance components, since human vision is more sensitive to the luminance components. The packet loss in wireless transmissions with MIMO 3D video schemes encoded with the recommendation of the HEVC standard affects more the quality of the high mobility video than the low mobility video.

# 5. References

- 1. TCP Options and Maximum Segment Size (MSS). [Internet]. [cited 2012 Jul 23]. Available from:https://tools. ietf.org/html/rfc6691.
- Bross B, Han W-J, Sullivan GJ, Ohm J-R, Wiegand T. High Efficiency Video Coding (HEVC). Text specification draft 9, document JCTVC-K1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC); 2012
- 3. Culturación. Diferencias entre HD, Ultra HD y 4K. [Internet]. [cited 2014 Mar 12]. Available from:http://culturacion.com/diferencias-entre-hd-ultra-hd-y-4k/.

- 4. Dorado S, Santacruz J, Pino F. Propuesta de Estructuras con Tensores para la Implementación de Esquemas MIMO. Revista S&T. 2014; 12(29):85–102.
- 5. Haykin S, Moher M. Modern wireless communications. International Edition. Pearson Prentice Hall; 2005. p. 357–63.
- ITU-R. Recommendation BT 500-11: Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union Recommendation Assembly; 2002.
- TTU-T H.264. Series H: Audiovisual and multimedia systems, infrastructure of audiovisual services-coding of moving video, advanced video coding for generic audiovisual services. The International Telecommunication Union; 2011. p. 1–674.
- López S. Transmisión de Video 3D sobre Redes P2P. Proyecto Fin de Carrera Ingeniería Técnica de Telecomunicaciones, Especialidad en Telemática. Universitat Politécnica De Catalunya; 2012. p. 1–58.
- Ndjiki-Nya P, Köppel M, Doshkov D, Lakshman H, Merkle P, Müller K, Wiegand T. Depth image-based rendering with advanced texture synthesis for 3-D video. IEEE Transactions On Multimedia. 2011; 13(3):453–65. https:// doi.org/10.1109/TMM.2011.2128862
- 10. Neto AV, Da Silva HW, Cerqueira EC, De Souza JN, Pirmez L, Gomes DG, Aguiar RL. Enhancing smart grid with session-

oriented communication system to truly support reliability and robustness. Smart Grid and Renewable Energy. 2012; 3(4):260–5. https://doi.org/10.4236/sgre.2012.34036

- Obando M, Freitas W, Cavalcanti F. Switching between hybrid MIMO structures for video transmission based on distortion model. IEEE 72nd Vehicular Technology Conference; 2010. p. 1–5. https://doi.org/10.1109/ VETECF.2010.5594528
- 12. Souceforge. IT++ Documentation [Internet]. [cited 2013 Jul 06]. Available from: http://itpp.sourceforge.net/4.3.1.
- Suarez J, Valencia O. Dise-o de un modelo objetivo para evaluar la calidad de video 3D anaglifo. Proyecto Fin de Carrera Ingeniería Electrónica. Universidad Surcolombiana; 2016. p. 25–9.
- 14. Sullivan G, Ohm J-R, Han W-J, Wiegand T. Overview of the high Efficiency Video Coding (HEVC) Standard. IEEE Transactions on Circuits and Systems for Video Technology. 2012; 22(12):1649–68. https://doi.org/10.1109/ TCSVT.2012.2221191
- Gerhard T, Ying C, Karsten Müller; Jens-Rainer Ohm; Anthony Vetro; Ye-Kui Wang overview of the multiview and 3D extensions of high efficiency video coding. Transactions on Circuits and Systems for Video Technology. 2016; 26(1):35–49. https://doi.org/10.1109/TCSVT.2015.2477935