<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *15th International Conference on Signal Processing and Communication Systems, ICSPCS 2021, Virtual, Online, 13 December 2021 through 15 December 2021*.

Citation for the original published paper:

Tran, D N., Zepernick, H-J., Chu, T M. (2021)

Viewing Direction Weighted Bit Plane LSB Data Hiding in 360° Videos

In: Wysocki T.A., Wysocki B.J. (ed.), *2021 15th International Conference on Signal Processing and Communication Systems, ICSPCS 2021 - Proceedings* Institute of Electrical and Electronics Engineers Inc.

<https://doi.org/10.1109/ICSPCS53099.2021.9660281>

N.B. When citing this work, cite the original published paper.

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:bth-22675>

Viewing Direction Weighted Bit Plane LSB Data Hiding in 360° Videos

Dang Ninh Tran, Hans-Jürgen Zepernick, and Thi My Chinh Chu
Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden
E-mail: {dnt, hjz, cch}@bth.se

Abstract—In this paper, we propose a viewing direction weighted bit plane least significant bit (LSB) data hiding method for 360° videos. Normalized Gaussian mixture models (GMMs) of the distributions of the relative viewing direction frequency for latitude, longitude, and both latitude and longitude are used to describe the human viewing behavior of 360° videos on head-mounted displays (HMDs). Apart from formulating data hiding weight functions to control the amount of secret data to be hidden at given latitudes and longitudes, the GMMs are used in this work also to define viewing direction weighted bit plane functions. These viewing direction weighted bit plane functions specify the number of bit planes to be used for data hiding depending on the latitude and longitude. As humans pay more attention to the central front view of 360° videos when watched on HMDs, the number of engaged bit planes is decreased accordingly from the poles to the equator, and from the west and the east to the front view. A performance assessment of the proposed viewing direction weighted bit plane LSB data hiding method is conducted, showing the trade-offs between capacity and objective perceptual video quality. The results indicate that it is preferable to first utilize a given bit plane for data hiding in an entire 360° video frame before engaging the next higher bit plane subject to the viewing direction weighted bit plane functions.

Index Terms—LSB data hiding, 360° videos, viewing direction, bit plane, capacity, non-content-based perceptual SSIM.

I. INTRODUCTION

Recent years have seen a rapid development of immersive digital media such as virtual reality (VR) and augmented reality (AR). In [1], it has been forecast that 90% of VR content will be in the form of 360° videos offering $360^\circ \times 180^\circ$ viewing directions in terms of latitude and longitude. A more immersive experience compared to traditional videos can be achieved by means of a head-mounted display (HMD) which allows three degrees of freedom in head movements for VR content exploration along with a large field of view. To ensure good visual quality of 360° videos when being viewed on an HMD, 4K to 8K resolutions are typically used with current technology.

The large resolutions associated with each frame of a 360° video may also be exploited as a resource for hiding secret data. Regarding networked digital value platforms for the delivery of advanced extended reality applications, e.g., considering 5G and 6G mobile networks [2], [3], data hiding techniques for 360° videos shall preferably have low computational load with little impact on link latency. As such, least significant bit (LSB) or *b*-rightmost LSB data hiding may be used to conceal the secret data in the words of depth *B* of a cover object. A survey on LSB data hiding in digital images

and videos is provided in [4] which gives a classification of these techniques with respect to the spatial, transform, and quantum domains of digital images, and the raw and compressed domains of digital videos.

Regarding data hiding for digital visual media, characteristics of the human visual system (HVS) are taken into account to reduce the perceptibility of data hiding. In particular, the HVS is less sensitive to sharp intensity changes in the edge regions of visual stimuli compared to other regions. Related works exploring this characteristic are focused on finding such preferable edge regions for data hiding, e.g., the pixel-differencing method proposed in [5] and improvements of this method using edge detectors [6]–[12]. Applications of edge detectors together with LSB data hiding in images are provided in [13], [14]. An increase in data hiding capacity can be obtained by using morphological operations such as dilation and closing to engage pixels around the edges for data hiding [15]. Morphological operations have been applied for LSB data hiding in high-definition images [16] and for spherical LSB data hiding in 360° videos [17].

Furthermore, it has been shown that humans statistically focus more on the equator region and the front region compared to the poles and peripheral regions when watching 360° videos on HMDs. For example, in [18], average statistics on the relative frequency of pixel accesses versus latitude are provided indicating that users give more attention to the equator region compared to the poles when viewing natural scenes. This finding was deduced from head motion trajectories that were gathered in a subjective experiment in which 10 participants watched 10 different natural 360° video scenes on an HMD while standing. The results about the relative frequency of pixel access versus latitude provided in [18] were used in [19] to propose a visual attention based LSB data hiding method for 360° videos. In [20], aiming at 360° video quality assessment and modeling, an extensive subjective experiment on viewing directions of 360° videos on HMDs has been reported. The 40 participants in this experiment were seated on a swivel chair supporting free scene exploration on an HMD. The recorded Euler angles were used to calculate each participants' viewing directions in terms of latitude and longitude. It has been shown that the distributions of viewing direction in both latitude and longitude can be formulated as Gaussian mixture models (GMMs). In view of the findings in [20], a viewing direction based LSB data hiding method for 360° videos has been proposed in [21]. Normalized GMMs

representing the distributions of viewing direction frequency for latitude and longitude are used to control the amount of secret data to be hidden at the latitude, longitude, or both latitude and longitude of 360° videos. Analytical expressions for the data hiding capacity of this method are derived. The fidelity and visual quality of the proposed method are assessed in terms of the peak signal-to-noise ratio (PSNR), weighted-to-spherically uniform PSNR (WS-PSNR), non-content-based perceptual PSNR (NCP-PSNR), structural similarity (SSIM) index, and the non-content-based perceptual SSIM (NCP-SSIM). Furthermore, it has been shown that NCP-SSIM is the most effective metric among the considered metrics with respect to assessing the visual quality of this viewing direction based LSB data hiding method.

Motivated by all of the above, in this paper, a viewing direction weighted bit plane LSB data hiding method for 360° videos is proposed. In this method, not only the amount of secret data to be hidden in the coordinates of 360° videos is controlled subject to the distributions of the viewing directions as in [21] but also the number of bit planes to be used at the respective coordinates. For this purpose, viewing direction weighted bit plane functions are defined. In contrast to [21], which performs data hiding from the south pole to the north pole or/and from the west to the east, in this work, data hiding is performed simultaneously commencing from both poles to the equator or/and from the west and the east to the front view. The main contributions of this paper are summarized as follows:

- A viewing direction weighted bit plane LSB data hiding method for 360° videos along with two strategies of scheduling data hiding in the bit planes of 360° video frames are proposed.
- Viewing direction weighted bit plane functions with respect to the latitude, longitude, and both latitude and longitude are introduced.
- Strategy I performs data hiding first in all bit planes of a given segment of a video frame, i.e., latitude or/and longitude, and then moves to the next video frame segment.
- Strategy II performs data hiding first in an entire video frame for a given bit plane and then moves to the next higher bit plane.
- The progressions for the actual number of bit planes used for data hiding in the 360° video frames with respect to the viewing direction weighted bit plane functions are provided for a wide range of experimental settings.
- Numerical results are provided to illustrate the performance trade-offs between capacity and NCP-SSIM for different scenarios.

The remainder of this paper is organized as follows. Section II describes the models that are used for viewing direction weighted bit plane LSB data hiding, i.e., normalized GMMs for normalized relative viewing direction frequencies, data hiding weight functions, and viewing direction weighted bit plane functions. Two novel strategies of viewing direction weighted bit plane LSB data hiding in 360° videos are

proposed in Section III. The performance measures used for assessing the proposed strategies are presented in Section IV. Section V provides numerical results about the performance of the proposed viewing direction weighted bit plane LSB data hiding method. Conclusions are given in Section VI.

II. MODELS FOR VIEWING DIRECTION WEIGHTED BIT PLANE LSB DATA HIDING

In [20], as a result of a comprehensive subjective experiment, it was shown that the distributions of the viewing directions that humans watch 360° videos on an HMD can be represented by GMMs. This finding led to GMMs of viewing directions as functions of latitude (south pole to north pole, i.e., $\phi = -90^\circ$ to $\phi = 90^\circ$), longitude (west to east, i.e., $\theta = -180^\circ$ to $\theta = 180^\circ$), and both latitude and longitude. In the context of LSB data hiding, these GMMs can be used to derive normalized GMMs that allow to formulate viewing direction based data hiding weight functions and viewing direction weighted bit plane functions as follows.

A. Normalized Relative Viewing Direction Frequency

In view of the results reported in [20], normalized relative viewing direction frequencies and their related models in terms of normalized GMMs were derived in [21] to define the weights for data hiding in latitude, longitude, and both latitude and longitude. The normalized GMMs are given as sums of three Gaussian components as follows [21]:

1) *Normalized GMM for the Latitude:* The normalized GMM with respect to the maximum relative latitude viewing direction frequency is given as

$$\hat{f}(\phi) = \sum_{k=1}^3 \hat{a}_k \exp \left[- \left(\frac{\phi - b_k}{c_k} \right)^2 \right] \quad (1)$$

where the peak height $\hat{a}_k$, center position b_k , and standard deviation c_k of the components of $\hat{f}(\phi)$ are provided in Table I.

TABLE I
PARAMETERS OF THE NORMALIZED GMM $\hat{f}(\phi)$ FOR LATITUDE ϕ

k	$\hat{a}_k$	b_k	c_k
1	0.2272	-2.3738	6.6437
2	0.6333	1.8260	14.8171
3	0.1727	1.4618	36.1311

2) *Normalized GMM for the Longitude:* The normalized GMM with respect to the maximum relative longitude viewing direction frequency has been obtained as

$$\hat{g}(\theta) = \sum_{l=1}^3 \hat{a}_l \exp \left[- \left(\frac{\theta - b_l}{c_l} \right)^2 \right] \quad (2)$$

where $\hat{a}_l$, b_l , and c_l of $\hat{g}(\theta)$ are given in Table II.

TABLE II
PARAMETERS OF THE NORMALIZED GMM $\hat{g}(\theta)$ FOR LONGITUDE θ

l	$\hat{a}_l$	$\hat{b}_l$	$\hat{c}_l$
1	0.1988	-0.1549	4.6740
2	0.6198	1.5140	18.51
3	0.1871	6.3670	110.5

3) *Normalized GMM for the Latitude and Longitude*: Given (1) and (2), the normalized GMM accounting for both latitude and longitude viewing directions can be expressed as

$$\hat{u}(\phi, \theta) = \hat{f}(\phi) \cdot \hat{g}(\theta) \quad (3)$$

B. Data Hiding Weight Functions

The normalized GMMs $\hat{f}(\phi)$, $\hat{g}(\theta)$, and $\hat{u}(\phi, \theta)$ given in (1), (2), and (3), respectively, can be applied to specify the related viewing direction based data hiding weight functions $w_1(\phi)$, $w_2(\theta)$, and $w_3(\phi, \theta)$. In particular, based on the normalized GMMs providing models for the normalized relative viewing direction frequencies, visual impairments to a 360° video due to data hiding should be kept small for directions that users frequently view. The data hiding weight functions may therefore be defined as complement of the respective normalized relative viewing direction frequencies [see also Figs. 1(a)-(c)]:

$$w_1(\phi) = 1 - \hat{f}(\phi) \quad (4)$$

$$w_2(\theta) = 1 - \hat{g}(\theta) \quad (5)$$

$$w_3(\phi, \theta) = 1 - \hat{u}(\phi, \theta) \quad (6)$$

C. Viewing Direction Weighted Bit Plane Functions

Apart from viewing direction based weighting of the amount of data to be hidden in a frame of a 360° video, in this paper, the number of bit planes to be used for data hiding is also weighted as a function of the latitude, longitude, and both latitude and longitude. As such, the number of bit planes used for data hiding is decreased from the poles to the equator, and from the west and the east to the front view of a 360° video subject to the respective data hiding weight functions (4), (5), and (6). In particular, viewing direction weighted bit plane functions may be defined as

$$b_1(\phi) = \lfloor w_1(\phi) \cdot \text{BP}_{max} \rfloor \quad (7)$$

$$b_2(\theta) = \lfloor w_2(\theta) \cdot \text{BP}_{max} \rfloor \quad (8)$$

$$b_3(\phi, \theta) = \lfloor w_3(\phi, \theta) \cdot \text{BP}_{max} \rfloor \quad (9)$$

where BP_{max} denotes the maximum number of bit planes to be used for data hiding. The round function $\lfloor x \rfloor$ [22] returns the nearest integer to the argument x . In agreement with our earlier work on LSB data hiding, e.g. [21], we have set the maximum number of bit planes used for data hiding to $\text{BP}_{max} = 6$. Engaging even higher bit planes than this already large value would severely degrade the visual quality of 360° videos.

Figs. 2(a)-(c) show the progressions of the viewing direction weighted bit plane functions $b_1(\phi)$, $b_2(\theta)$, and $b_3(\phi, \theta)$, respectively. As for $b_1(\phi)$ depicted in Fig. 2(a) as a function of latitude ϕ , the number of bit planes allowed for data hiding

Fig. 1. Viewing direction based data hiding weight functions: (a) Latitude, $w_1(\phi)$, (b) Longitude, $w_2(\theta)$, (c) Latitude and longitude, $w_3(\phi, \theta)$.

decreases when moving from the poles to the equator. A similar progression is observed in Fig. 2(b) for $b_2(\theta)$ as a function of longitude θ for which the number of bit planes allowed for data hiding decreases when moving from the west and the east to the front view. Regarding $b_3(\phi, \theta)$ as a function of both latitude and longitude shown in Fig. 2(c), the number of bit planes allowed for data hiding decreases when moving toward the center front view.

III. VIEWING DIRECTION WEIGHTED BIT PLANE LSB DATA HIDING STRATEGIES

This section proposes two strategies of viewing direction weighted bit plane LSB data hiding in 360° videos. In particular, the viewing direction weighted bit plane functions

Fig. 2. Viewing direction weighted bit plane functions for $BP_{max} = 6$: (a) Latitude, $b_1(\phi)$, (b) Longitude, $b_2(\theta)$, (c) Latitude and longitude, $b_3(\phi, \theta)$.

proposed in (7), (8), and (9), control the number of bit planes to be used for data hiding. To preserve visual quality, the engaged number of bit planes is decreased toward the equator and the front view of 360° videos.

A. LSB Data Hiding

The term least significant bit refers to a bit in a multiple-bit binary number or word of bit width B that has the lowest order. The different positions with reference to the bit width of a word are also called bit planes. In the context of data hiding, the expression “insignificant” also relates to the property of a bit to be changed without noticeable impact on the quality of the source. A modification of the LSBs of a sequence of words of a 360° cover video typically has little effect on the

visual quality. Accordingly, LSB data hiding may be extended to higher bit planes, which increases capacity at the expense of visual quality.

B. Viewing Direction Based LSB Data Hiding

A viewing direction based LSB data hiding method for 360° videos has been proposed in [21]. This method uses the data hiding weight functions given in (4), (5), and (6) for controlling the amount of secret data to be hidden at the latitudes and longitudes. As such, the finding that humans view the front region near the equator much more often than other regions of 360° videos is accounted for. Further, LSB or b -rightmost LSB substitution is performed randomly in the selected segments of the latitude, longitude, or both latitude and longitude. The placement of secret data at random positions can be realized by pseudo-random number generators which are widely used for stream ciphers. Hardware or software implementations of linear feedback shift registers (LFSRs) have been widely used as pseudo-random number generators [23]. To allow the recovery of the secret data from the stego-object, the coefficients of the chosen feedback polynomial describing the feedback taps of the LFSR may be stored at the receiver together with an agreed upon seed.

The performance assessment of this method has shown that NCP-SSIM is the most effective metric among the considered metrics with respect to assessing the visual quality of the proposed viewing direction based LSB data hiding method. In particular, NCP-SSIM turned out to accommodate the distribution of viewing direction frequencies as well as spatial perceptual information into the visual quality assessment. In the following, viewing direction based LSB data hiding is used along with viewing direction weighted bit plane functions while NCP-SSIM is used for visual quality assessment.

C. Data Hiding Strategies Using Viewing Direction Weighted Bit Plane Functions

Figs. 3(a)-(b) show the flow diagrams of the two proposed viewing direction weighted bit plane LSB data hiding strategies, i.e., Strategy I and Strategy II for the case of data hiding in the latitude. Both strategies commence with the same initialization steps: (1) Reading the 360° cover video and secret data, (2) Segmentation of the 360° cover video frames with respect to the latitude into intervals $[\phi, \phi \pm \Delta)$ with a selected width or step size Δ , (3) Reading of the data hiding weight function $w_1(\phi)$, (4) Calculation of the viewing direction weighted bit plane function $b_1(\phi)$ for a selected maximum number of bit planes BP_{max} to be used for data hiding. Furthermore, the initial latitudes are set to $\phi_1 = -90^\circ$ (south pole) and $\phi_2 = 90^\circ$ (north pole), and the initial bit plane is set to the LSB, i.e., $b = 1$. It should be mentioned that the step size used for the segmentation of the frames with respect to the latitude is given in degrees, e.g., $\Delta = 1^\circ$. The quality threshold ϵ is set with respect to the visual quality metric used for visual quality assessment (VQA). After these common initialization steps, Strategy I [see Fig. 3(a)] and Strategy II

Fig. 3. Flow diagrams of the proposed viewing direction weighted bit plane LSB data hiding strategies with respect to the latitude: (a) Strategy I, (b) Strategy II.

[see Fig. 3(b)] continue with different processing steps and loops as follows.

1) *Strategy I*: Data hiding is performed in each segment given by the intervals of latitudes $[\phi_1, \phi_1 + \Delta)$ and $[\phi_2, \phi_2 - \Delta)$ until the maximum allowable bit plane defined by $b_1(\phi_1)$ and $b_1(\phi_2)$ for these segments is reached before moving to the subsequent segments $[\phi_1 + \Delta, \phi_1 + 2\Delta)$ and $[\phi_2 - \Delta, \phi_2 - 2\Delta)$. It should be mentioned that each segment for a given latitude contains a number of horizontal lines of a 360° cover video frame that depends on the selected width or step size Δ . In particular, using $w_1(\phi)$, viewing direction based LSB data hiding at the b -th bit plane is performed for the segments of a 360° cover video frame commencing from the latitudes ϕ_1 and ϕ_2 with reference to the north pole and the south pole. If the VQA returns a value that is greater than the selected quality threshold ϵ and the current bit plane b is smaller than the value of the weighted bit plane function $b_1(\phi_1)$ and $b_1(\phi_2)$, then the next bit plane $b := b + 1$ is also used for data hiding in the current segments, i.e., intervals of latitudes $[\phi_1, \phi_1 + \Delta)$ and $[\phi_2, \phi_2 - \Delta)$. If $VQA > \epsilon$ holds, this loop is followed until the number of engaged bit planes b reaches or exceeds the value defined by $b_1(\phi_1)$ and $b_1(\phi_2)$ for the current segments. Then, the next segment of width Δ with $\phi_1 := \phi_1 + \Delta$ and

$\phi_2 := \phi_2 - \Delta$ are used for data hiding commencing again with bit plane $b = 1$. The above process continues until the VQA returns a value that falls below the selected quality threshold ϵ or all segments of the considered 360° cover video frame have been used for data hiding.

2) *Strategy II*: An eligible b -th bit plane with respect to the viewing direction weighted bit plane function $b_1(\phi)$ is used for data hiding throughout all segments of the entire 360° video frame before moving to the next higher bit plane. In other words, the latitude sweeps from the south pole and the north pole with step size Δ toward the equator while the bit plane is kept fixed. Specifically, using $w_1(\phi)$, viewing direction based LSB data hiding at the b -th bit plane is performed throughout the entire 360° video frame as long as the current bit plane b is smaller than or equal to the value bounded by the weighted bit plane function $b_1(\phi)$, and the VQA gives a value that is higher than the quality threshold ϵ . Once the b -th bit plane has been used for the entire 360° video frame, then the next bit plane $b := b + 1$ is engaged for viewing direction based LSB data hiding. This process ends once all possible bit planes with respect to $b_1(\phi)$ are consumed or the VQA returns a value that falls below the selected quality threshold ϵ .

Strategy I and II can be easily applied to the longitude θ

using the viewing direction weighted bit plane function $b_2(\theta)$. The same applies to LSB data hiding in both latitude ϕ and longitude θ for which the viewing direction weighted bit plane function $b_3(\phi, \theta)$ is used and the decisions in the flow diagrams need to account for the latitude and longitude.

IV. PERFORMANCE MEASURES

The performance of the viewing direction weighted bit plane LSB data hiding method is assessed in terms of data hiding capacity, or capacity for brevity, and NCP-SSIM regarding visual quality. This allows to study the trade-off between capacity and objective perceptual 360° video quality.

A. Capacity

Capacity refers to the number of secret bits that can be embedded in a cover object, e.g., a 360° cover video in the considered application. In particular, for 360° cover videos, capacity can be defined as the average number of secret bits concealed in each 360° cover video frame and is measured in bits per pixel per frame (bpp/frame). In contrast to [21], where capacity can be analytically expressed, the capacity of the proposed LSB data hiding strategies is experimentally obtained by accumulating the number of secret bits embedded in each 360° cover video frame. This is because of the video quality assessment condition $VQA > \epsilon$ (see Fig. 3). The outcome whether this condition holds is not known beforehand but only after the random LSB substitution has been performed.

B. NCP-SSIM

NCP-SSIM is an objective visual quality metric for 360° videos that has recently been proposed in [20] as follows. Consider a 360° video of resolution $N \times M$ (width×height) given in equirectangular projection (ERP) format. The probability of a pixel with coordinates (s, t) being in the viewing direction during a given frame is obtained as

$$v(s, t) = u\left(-180\left(\frac{t-1}{M-1} - \frac{1}{2}\right), -360\left(\frac{s-1}{N-1} - \frac{1}{2}\right)\right) \quad (10)$$

with pixel coordinates being in the intervals $1 \leq s \leq N$ and $1 \leq t \leq M$, and $u(\phi, \theta)$ denotes the distribution of the viewing direction frequency. Then, the normalized non-content-based weight map $\hat{w}(s, t)$ is defined as

$$\hat{w}(s, t) = \frac{w(s, t)}{\sum_{s, t} w(s, t)} \quad (11)$$

where the non-content-based weight map $w(s, t)$ is defined using (10) as

$$w(s, t) = \max_{s', t' \in \mathbf{V}_{s, t}} v(s', t') \quad (12)$$

and $\mathbf{V}_{s, t}$ denotes the collection of the viewing directions of all viewports that include the pixel (s, t) .

Then, using the normalized non-content-based weight map $\hat{w}(s, t)$ given in (11), the NCP-SSIM is defined as

$$\text{NCP-SSIM} = \sum_{s, t} m_{\text{SSIM}}(s, t) \cdot \hat{w}(s, t) \quad (13)$$

where $m_{\text{SSIM}}(s, t)$ denotes the SSIM map with local SSIM value for pixel (s, t) [24].

V. NUMERICAL RESULTS

In this section, a performance assessment of the viewing direction weighted bit plane LSB data hiding in 360° videos for Strategy I and Strategy II is provided. In particular, the progressions of the actual number of bit planes used for data hiding are presented subject to the viewing direction weighted bit plane functions $b_1(\phi)$ for latitude, $b_2(\theta)$ for longitude, and $b_3(\phi, \theta)$ for latitude and longitude. The relationships between capacity and NCP-SSIM are provided for different scenarios.

A. Cover Video

The 360° video scene “Salon”, which is publicly available at the VQA-ODV database [25], [26], is used as cover video. This video has a resolution of 4096×2048 pixels, is of 10 s duration, and runs at a frame rate of 30 frames per second. Fig. 4 shows a sample frame of the 360° video scene “Salon” in ERP format with uniformly distributed samples of latitude and longitude. In the ERP format, the vertical coordinate is the latitude ϕ of the sphere and the horizontal coordinate is the longitude θ of the sphere. Furthermore, we use the YV12 color coding model which comprises of an 8 bit Y plane (luminance component) of resolution $N \times M$ followed by 8 bit 2×2 subsampled $N/2 \times M/2$ V and U planes (chrominance components) [27].

Fig. 4. Sample frame of the 360° cover video “Salon” in ERP format with uniformly spaced latitude and longitude.

B. Data Hiding in the Latitude

In this section, the performance of the two proposed viewing direction weighted bit plane LSB data hiding strategies is compared for the case of segmenting the 360° video with respect to the latitude. The step size $\Delta = 1^\circ$ has been used for incrementing and decrementing the latitude ϕ_1 and ϕ_2 , respectively. Accordingly, a 360° video frame is organized into segments of width Δ with respect to the latitude. The maximum number of bit planes to be used for data hiding in these segments is set to $\text{BP}_{\max} = 6$. LSB and b -rightmost LSB substitution has been applied to the Y component of the YV12 color encoded 360° cover video. It should be mentioned that modifying only the Y component of the YV12 format reduces the impact of data hiding on the visual quality as the pixels are considered individually while groups of four pixels would

Fig. 5. Progression of the number of bit planes used for data hiding in the 360° video frames as a function of latitude: (a)-(c) Strategy I, (d)-(e) Strategy II.

Fig. 6. Capacity versus NCP-SSIM with the strategies applied to the latitude: (a) Strategy I, (b) Strategy II.

need to be modified in the U and V components. In addition, the size of the Y component for a given frame is four times higher compared to the size of the U and V components.

Figs. 5(a)-(f) show the actual number of bit planes used at each latitude of a 360° video frame for different quality

thresholds. The results obtained for Strategy I in Figs. 5(a)-(c) illustrate that data hiding uses the selected maximum number of bit planes in clusters around the pole regions and drops steeply when the quality threshold is reached. The results for Strategy II, on the other hand, show that utilizing an entire bit plane before moving to the next higher bit plane gives a step-like progression from the poles to the equator as shown in Figs. 5(d)-(f). In particular, for the quality threshold of $\epsilon = 0.9$, the number of bit planes used for data hiding does not exceed $b = 4$. In the cases that the demands on visual quality are lower, i.e., $\epsilon = 0.8$ and 0.7 , up to $b = 5$ bit planes are used increasingly with the decrease of the quality threshold. If the quality threshold were to be further reduced, the selected maximum number of bit planes $BP_{max} = 6$ would be engaged commencing from the poles before step-wise being decremented toward the equator.

Figs. 6(a)-(b) compare the capacity offered by Strategy I and Strategy II for data hiding as a function of NCP-SSIM. The capacity smoothly increases with the decrease of NCP-SSIM for Strategy I while the capacity progresses over several plateaus for Strategy II. It should be mentioned that the number of plateaus observed for Strategy II in Fig. 6(b) corresponds to the actual maximum number of bit planes used in Figs. 5(d)-(f). For example, a maximum of four bit planes is used for the quality thresholds $\epsilon = 0.9$ while five bit planes are used for $\epsilon = 0.8$ and 0.7 , which leads to four and five plateaus in the progression of the capacity versus NCP-SSIM. Although the progression of capacity for Strategy I and Strategy II differs, the differences at NCP-SSIM = 0.9, 0.8, and 0.7 are minor.

To shed further light on the characteristics of Strategy I and Strategy II, sample stego-video frames in ERP format are provided in Figs. 7(a)-(f) for visual inspection. As for

Fig. 7. Sample 360° stego-video frames in ERP format for different quality thresholds $\epsilon = 0.9, 0.8, 0.7$: (a)-(c) Strategy I, (d)-(f) Strategy II.

the stego-video frames in Figs. 7(a)-(c) generated by Strategy I, impairments are clearly visible at the ceiling and floor structures commencing from the poles to the equator. Furthermore, the region around the equator that is free of data hiding becomes narrower with decreasing quality threshold [see also Figs. 5(a)-(c)]. In contrast, the stego-video frames in Figs. 7(d)-(f) obtained with Strategy II do not suffer from noticeable visual quality degradation for the considered values of the quality threshold. The finding that Strategy II outperforms Strategy I in terms of visual quality of sample stego-video frames has also been obtained through visual inspection of stego-frames for data hiding as function of longitude, and both latitude and longitude. As such, focus is given to Strategy II in the remainder of this section.

C. Data Hiding in the Longitude

Figs. 8(a)-(c) show the progression of the actual number of bit planes used for data hiding as a function of the longitude subject to different quality thresholds. Similar as for the latitude, a step-wise decrease of the number of bit planes is observed when sweeping the longitude from the west and the east toward the front view ($\theta = 0^\circ$). In the case of the high quality threshold $\epsilon = 0.9$, neither bit plane $b = 6$ nor $b = 5$ are used but only up to $b = 4$ bit planes are engaged for data hiding. In other words, instead of utilizing the selected maximum number of bit planes $BP_{max} = 6$ for a certain range of longitudes commencing from the poles, a given bit plane is considered first for data hiding throughout an entire frame before moving to the next higher bit plane. Furthermore, for the lower quality thresholds of $\epsilon = 0.8$ and $\epsilon = 0.7$, bit plane $b = 5$ is also used for data hiding while spanning over a smaller range of longitudes for $\epsilon = 0.8$ compared to $\epsilon = 0.7$.

Fig. 9 presents the capacity versus NCP-SSIM that was obtained with Strategy II applied to the longitude. It can be observed that the capacity increases at the expense of lower

visual quality in terms of NCP-SSIM, and reaches similar values as for the case that Strategy II is applied to the latitude. Similarly, the capacity progresses through a number of plateaus with increased NCP-SSIM corresponding to the number of bit planes used for data hiding. However, the capacity increases less rapidly at the beginning of each of the five plateaus compared to the curve in Fig. 6(b) obtained with Strategy II applied to the latitude. This behavior is due to the larger interval for the longitude spanning over 360° compared to the interval for the latitude spanning over 180°.

D. Data Hiding in the Latitude and Longitude

Figs. 10(a)-(c) provide heatmaps of the number of bit planes used for data hiding in each 360° video frame as a function of both latitude and longitude subject to different quality thresholds. The lowest color (dark blue) and highest color (dark red) of the colormap correspond to $b = 0$ and $b = 6$ bit planes. The heatmaps show that the center front view is kept free of data hiding while an increased number of bit planes are utilized with the increase of the latitudes toward the poles, and longitudes toward the west and the east. Furthermore, the actual maximum number of bit planes used and its spatial portion of a frame increase with the decrease of the quality threshold, i.e., $b = 4$ (yellow) for $\epsilon = 0.9$, $b = 5$ (red) for $\epsilon = 0.8$, and again $b = 5$ for $\epsilon = 0.7$ but for a larger spatial portion of a frame. However, the maximum number of bit planes $BP_{max} = 6$ is not reached in these examples giving the opportunity of increasing capacity if further visual quality reduction below $\epsilon = 0.7$ was tolerable.

Fig. 11 presents the capacity versus NCP-SSIM for Strategy II applied to the latitude and longitude. Clearly, as only a small region in the center front view is kept free of data hiding, the highest capacity is obtained for this case compared to the results shown in Fig. 6 and Fig. 9. Again, the progression of the capacity over several plateaus with increasing NCP-

Fig. 8. Progression of the number of bit planes used for data hiding in the 360° video frames as a function of longitude using Strategy II: (a) $\epsilon = 0.9$, (b) $\epsilon = 0.8$, (c) $\epsilon = 0.7$.

SSIM is clearly visible, illustrating the quality level at which an additional bit plane is engaged in data hiding.

VI. CONCLUSIONS

In this paper, we have proposed and assessed a viewing direction weighted bit plane LSB data hiding method for 360° videos. This method takes into account the human behavior of focusing more on the central front view when watching 360° videos on an HMD compared to the poles and peripheral regions toward the west and the east. In particular, normalized GMMs have been used to represent the distributions of the relative viewing direction frequency for latitude, longitude, and both latitude and longitude. These normalized GMMs have been used to not only formulate data hiding weight

Fig. 9. Capacity versus NCP-SSIM for Strategy II applied to the longitude.

functions that control the amount of secret data to be hidden at given latitudes and longitudes, but also to define viewing direction weighted bit plane functions. Based on the viewing direction weighted bit plane functions, the number of bit planes to be used for data hiding is selected according to the human viewing behavior of 360° videos on HMDs. In particular, the number of bit planes used for data hiding is decreased accordingly from the poles to the equator, and the west and the east to the front view. A performance assessment of the proposed viewing direction weighted bit plane LSB data hiding method has been conducted revealing the trade-offs between capacity and visual quality in terms of NCP-SSIM. The results have shown that utilizing a given bit plane for data hiding in an entire 360° video frame before engaging the next higher bit plane (Strategy II) gives superior visual quality of the 360° stego-video compared to using all allowable bit planes for a given 360° video frame segment before attending the next segment (Strategy I). Future directions may exploit the fact that HMDs provide a limited viewport rather than showing the entire sphere at a given time. As such, content-based-perceptual approaches including more advanced visual attention models may be considered together with adaptive LSB data hiding that accounts for viewport dynamics.

ACKNOWLEDGMENTS

This work has been supported in part by The Knowledge Foundation, Sweden, through the ViaTech project (Contract 20170056). Dang Ninh Tran has been supported by a VIED scholarship awarded by the Vietnamese Government.

REFERENCES

- [1] Huawei iLab, *VR Big Data Report*, Huawei Technologies Co., Ltd., Shenzhen, P. R. China, 2017.
- [2] 3GPP TR 26.918 V16.0.0, *Virtual Reality (VR) Media Services Over 3GPP (Release 16)*, 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, Dec. 2018.
- [3] S. Yrjölä, P. Ahokangas, and M. Matinmikko-Blue (Eds.), *White Paper on Business of 6G. (6G Research Visions, No. 3)*, University of Oulu, Finland, Jun. 2020.
- [4] D. N. Tran, *On LSB Data Hiding in New Digital Media*, Licentiate Dissertation, Blekinge Inst. of Technol., Karlskrona, Sweden, May 2020.
- [5] D.-C. Wu and W.-H. Tsai, "A Steganography Method for Images by Pixel-Value Differencing," *Pattern Recognit. Lett.*, vol. 24, no. 9, pp. 1613–1626, Sep. 2003.

Fig. 10. Heatmaps of the number of bit planes used for data hiding in each 360° video frame as a function of both latitude and longitude: (a) $\epsilon = 0.9$, (b) $\epsilon = 0.8$, (c) $\epsilon = 0.7$.

Fig. 11. Capacity versus NCP-SSIM for Strategy II applied to the latitude and longitude.

- [6] W. J. Chen, C. C. Chang, and T. H. N. Le, "High Payload Steganography Mechanism Using Hybrid Edge Detector," *Expert Systems with Appl.*, vol. 37, no. 4, pp. 3292–3301, Apr. 2010.
- [7] A. K. Pal and T. Pramanik, "Design of an Edge Detection Based Image Steganography with High Embedding Capacity," in *Proc. Quality, Reliability, Security and Robustness in Heterogeneous Networks*, Greder Noida, India, Jan. 2013, pp. 794–800.
- [8] H. Tseng and H. Leng, "High-Payload Block-Based Data Hiding Scheme Using Hybrid Edge Detector With Minimal Distortion," *IET Image Process.*, vol. 8, no. 11, pp. 647–654, Nov. 2014.
- [9] S. Sun, "A Novel Edge Based Image Steganography With 2^k Correction and Huffman Encoding," *Information Process. Lett.*, vol. 116, no. 2, pp. 93–99, Feb. 2016.
- [10] H. Dadgostar and F. Afsari, "Image Steganography Based on Interval-Valued Intuitionistic Fuzzy Edge Detection and Modified LSB," *J. of Inf. Security and Appl.*, vol. 30, pp. 94–204, Oct. 2016.
- [11] J. Bai, C. C. Chang, T. S. Nguyen, C. Zhu, and Y. Liu, "A High Payload Steganographic Algorithm Based on Edge Detection," *Displays*, vol. 46, pp. 42–51, Jan. 2017.
- [12] S. K. Ghosal, J. K. Mandal, and R. Sarkar, "High Payload Image Steganography Based on Laplacian of Gaussian (LoG) Edge Detector," *Multimedia Tools and Appl.*, vol. 77, no. 23, pp. 30403–30418, Dec. 2018.
- [13] W. Luo, F. Huang, and J. Huang, "Edge Adaptive Image Steganography Based on LSB Matching Revisited," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 201–214, Jun. 2010.
- [14] S. Chakraborty, A. Jalal, and C. Bhatnagar, "LSB Based Non Blind Predictive Edge Adaptive Image Steganography," *Multimedia Tools and Appl.*, vol. 76, no. 6, pp. 7973–7987, Mar. 2017.
- [15] K. Gaurav and U. Ghanekar, "Image Steganography Based on Canny Edge Detection, Dilation Operator and Hybrid Coding," *J. of Inf. Security and Appl.*, vol. 41, pp. 41–51, Aug. 2018.
- [16] D. N. Tran and H.-J. Zepernick, "On LSB Data Hiding in High-Definition Images Using Morphological Operations," in *Proc. Int. Symp. on Commun. and Inf. Technol.*, Ho Chi Minh City, Vietnam, Sep. 2019, pp. 386–391.
- [17] —, "Spherical LSB Data Hiding in 360° Videos Using Morphological Operations," in *Proc. Int. Conf. on Signal Process. and Commun. Systems*, Gold Coast, Australia, Dec. 2019, pp. 573–582.
- [18] M. Yu, H. Laksman, and B. Girod, "Framework to Evaluate Omnidirectional Video Coding Schemes," in *Proc. IEEE Int. Symp. on Mixed and Augmented Reality*, Fukuoka, Japan, Sep./Oct. 2015, pp. 31–36.
- [19] D. N. Tran, H.-J. Zepernick, and T. M. C. Chu, "Visual Attention Based LSB Data Hiding in 360° Videos," in *Proc. IEEE Int. Symp. on Signal Process. and Commun. Systems*, Adelaide, Australia, Dec. 2020, pp. 1–8.
- [20] M. Xu, C. Li, Z. Chen, Z. Wang, and Z. Guan, "Assessing Visual Quality of Omnidirectional Videos," *IEEE Trans. Circuits and Systems for Video Technol.*, vol. 29, no. 12, pp. 3516–3530, Dec. 2019.
- [21] D. N. Tran, H.-J. Zepernick, and T. M. C. Chu, "Viewing Direction Based LSB Data Hiding in 360° Videos," *Electronics*, vol. 10, no. 13, pp. 1–32, Jun. 2021.
- [22] J. Hastad, B. Just, J. C. Lagarias, and C. P. Schnorr, "Polynomial Time Algorithms for Finding Integer Relations Among Real Numbers," *SIAM J. Comput.*, vol. 18, pp. 859–881, 1988.
- [23] H.-J. Zepernick and A. Finger, *Pseudo Random Signal Processing: Theory and Application*. Chichester: John Wiley & Sons, 2005.
- [24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment From Error Visibility to Structural Similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [25] VQA-ODV, Beihang University, School of Electronic and Information Engineering, Beijing, China, 2018 (accessed August 8, 2021). [Online]. Available: <https://github.com/Archer-Tatsu/VQA-ODV>
- [26] L. Chen, X. Mai, D. Xinzhe, and W. Zulin, "Bridge the Gap Between VQA and Human Behavior on Omnidirectional Video: A Large-Scale Dataset and a Deep Learning Model," in *Proc. ACM Int. Conf. on Multimedia*, Seoul, Republic of Korea, Oct. 2018, pp. 932–940.
- [27] Bjarne, *YUV Pixel Formats*, 2019, (accessed May 8, 2021). [Online]. Available: <https://www.fourcc.org/yuv.php>