FREE-VIEW POINT TV WATERMARKING EVALUATED ON GENERATED ARBITRARY VIEWS

Evlambios E. Apostolidis and Georgios A. Triantafyllidis
Informatics and Telematics Institute, Thessaloniki, Greece

Keywords: Free-view point TV, watermarking.

Abstract: The recent advances in Image Based Rendering (IBR) has pioneered a new technology, free view point television, in which TV-viewers select freely the viewing position and angle by the application of IBR on the transmitted multi-view video. In this paper, exhaustive tests were carried out to conclude to the best possible free-view point TV watermarking evaluated on arbitrary views. The watermark should not only be extracted from a generated arbitrary view, it should also be resistant to common video processing and multi-view video processing operations.

1 INTRODUCTION

Image based rendering (IBR) has been developed in the last ten years as an alternative to the traditional geometry based rendering techniques. IBR aims to produce a projection of a 3D scene for an arbitrary view point by using a number of original camera views of the scene. This approach contains the original effects already in the original camera views and consequently, yields more natural views compared to the traditional geometry based methods, which are modeled with one texture and additional supporting textures that sometimes lacks natural appearance. There are several advantages to this approach:

- The display algorithms for image-based rendering require modest computational resources and are thus suitable for real-time implementation on workstations and personal computers.
- The cost of interactively viewing the scene is independent of scene complexity.
- The source of the pre-acquired images can be from a real or virtual environment, i.e. from digitized photographs or from rendered models. In fact, the two can be mixed together.

Moreover, IBR is more preferable, since images are easier to obtain and simpler to handle compared to describing a geometric model, a texture and a texture map in the traditional approach (Zhang and Chen, 2004). Due to these advantages, IBR has attracted much attention from researchers in vision and signal processing and shown a great progress in the last decade. Yet, it is possible to see real-time demonstration of free-view TV, where TV-viewers select freely the viewing position and angle on the transmitted multi-view point video.

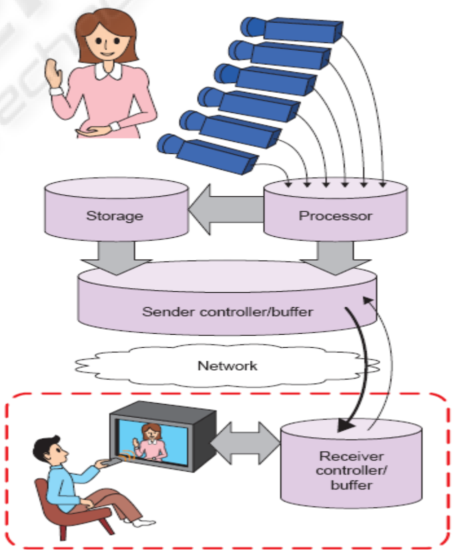


Figure 1: A general scheme for Free-Viewpoint Television.

Free-view point video (see Fig. 1) is expected to be a next-generation visual application (MPEG Meeting, 2003). It provides the user with realistic impressions by means of high interactivity and photorealistic image quality. It lets the user freely change his/her viewpoint (i.e., viewing position and

Apostolidis E. and Triantafyllidis G. (2008).
FREE-VIEW POINT TV WATERMARKING EVALUATED ON GENERATED ARBITRARY VIEWS.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 146-151
DOI: 10.5220/0001084201460151

viewing direction) and enjoy more photorealistic 3D images. With these functionalities, it can be used for various services, such as broadcasting, visual communication, and education.

In this context, it is apparent that copyright protection problems also exist and should be solved for free-view TV. Among many alternative rights management methods, the copyright problem for visual data is also approached by means of embedding hidden imperceptible information, called watermark, into the image and video content (Koz and Alatan, 2005). Hence, watermarking can also be a good candidate for the solution of the copyright problem for the free-view TV, as well. However, the problem is more complicated compared to image and mono-view video case.

First of all, concerning with the robustness requirement, the watermark should not only be resistant to common video processing and multi-view video processing operations, it should also be extracted from a generated arbitrary view (Fig. 2). In order to extract the watermark from such a rendered view, the watermark detection scheme should involve an estimation procedure for the imagery camera position and orientation, where the rendered view is generated. In addition, the watermark should also be invisible and survive from image based rendering operations, such as frame interpolation between neighbour cameras and pixel interpolation inside each camera frame

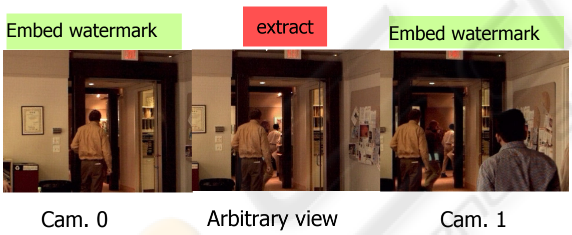


Figure 2: The watermarking problem for free view television.

2 WATERMARKING EMBEDDING AND DETECTION

In the literature, the most well known and useful IBR representation is the light field, due to its simplicity that only the original images are used to construct the imagery views. Therefore, the proposed watermarking method is specially tailored for the aim of watermark extraction from the imagery views, which are generated by using light field rendering.

In light field rendering (LFR), a light ray is indexed as (u_0,v_0,s_0,t_0), where (u_0,v_0) and (s_0,t_0) are the intersections of the light ray with the two parallel planes namely, camera (uv) plane and focal (st) plane. The planes are discretized so that a finite number of light rays are recorded. If all the discretized points from the focal plane are connected to one point on the camera plane, an image (2D array of light fields) is resulted. Actually, this resulting image becomes sheared perspective projection of the camera frame at that point (Levoy and Hanrahan, 1996). 4D representation of light field can also be interpreted as a 2D image array, as it is shown in Fig. 3. The watermark is embedded to each image of this 2D image array.

The proposed method embeds the watermark into each image of the light field slab (Fig. 3) by exploiting spatial sensitivity of HVS. For that purpose, the watermark is modulated with the resulting output image after filtering each light field image by a high pass filter and spatially added to that image.

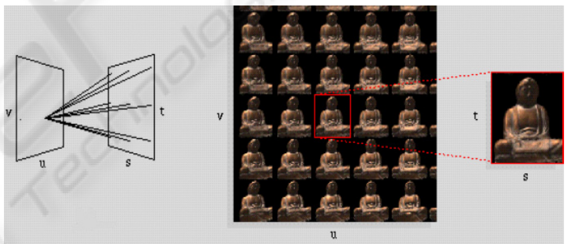


Figure 3: A sample light field image array: Buddha light field.

More specifically, the method applies the same watermarking operation for each light field image as follows:

$$I_{uv}^{*}(s,t) = I_{uv}(s,t) + \alpha . H_{uv}(s,t) . W(s,t) \qquad (1)$$

where I_{uv} is the light field image corresponding to the camera at the (u,v) position, H_{uv} is the output image after high pass filtering, α is the global scaling factor to adjust the watermark strength, W is the watermark sequence, I_{uv}^{*} is the watermarked light field image. A correlation-based scheme is proposed for watermark detection as depicted in Fig. 4.

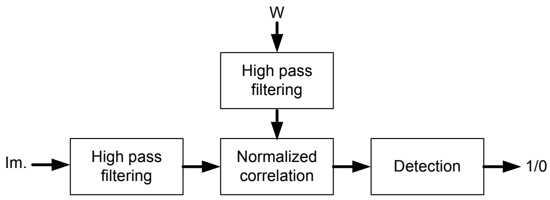


Figure 4: Watermark detection.

3 EXPERIMENTAL RESULTS

Buddha light fields (The Stanford Light Field Archive, 2007) are used during the simulations. We tested various watermarking sequences (as shown in Table 1) in order to conclude to the one with the best performance.

Table 1: Watermark test sequences.

No	DISTRIBUTION	MATLAB CODE
1	Chi-Square	random('chi2' , v , 256 , 256)
2	Exponential	random('exp' , μ , 256 , 256)
3	Geometric	random('geo' , p , 256 , 256)
4	Poisson	random('poiss' , λ , 256 , 256)
5	Rayleigh	random('rayl' , b , 256 , 256)
6	Beta	random('beta' , a , b , 256 , 256)
7	Binomial	random('bino', n , p , 256 , 256)
8	Extreme Value	random('ev', μ , σ , 256 , 256)
9	Gamma	random('gam' , a , b , 256 , 256)
10	Negative Binomial	random('nbin' , p , r , 256 , 256)
11	Normal (Gaussian)	random('norm' , μ , σ , 256 , 256)
12	Uniform	random('unif' , a , b , 256 , 256)
13	Weibull	random('wbl', a , b , 256 , 256)
14	Generalized Pareto	random('gp', κ , σ , θ , 256 , 256)
15	Hypergeometric	random('hyge', M , K , n , 256,256)

We also tested the sequences: Extreme Value, F, Non-central F, Lognormal, Student's distribution and Non-central T distributions which however did not yield acceptable results and as a result are not considered further.

As it was stated in the introduction, the main problem in watermarking of free view point video is the successful extraction of the watermark from a random generated view. So, we tested the watermarks efficiency on imaginary views created by rendering procedures on the source data of Buddha's images. The creation of a rendered view from the light field of Buddha's images requires the determination of some characteristics:

First thing is the interpolation method for the construction of the rendered view. We selected to have two choices:

1. Bilinear Interpolation.
2. Nearest Neighborhood Interpolation.

Secondly is the viewing position or the determination of the imagery camera's spot. This spot is determined by two elements:

1. The coordinates of the spot of virtual camera.
2. The orientation, as a vector with coordinates, for the virtual camera.

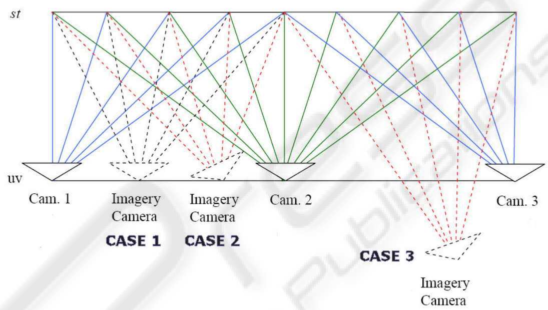
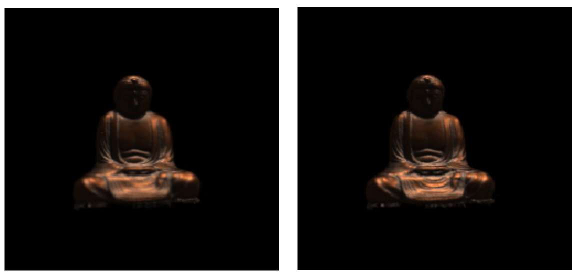


Figure 5: Configurations for the imagery camera position and rotation.

In (Apostolidis, Koz and Triantafyllidis, 2007), tests are carried out for the cases 1 and 2 (see Fig. 5) where the imagery camera is located in the camera plane. In this paper, we focus on case 3 where the imagery camera is in an arbitrary position and its rotation is not unity. More specifically we select 4 positions of the imagery camera (within case 3) in order to include all the possible viewing positions of case 3 (zoom in and out, rotation left or right). Figs 6,7,8,9 illustrate the four data sets that are used for our experimental results.



(a) (b)

Figure 6: Test 1: Imagery camera rendered views with zoom in.

(a): Nearest Neighborhood
(b): Bilinear
Virtual camera's adjustment:
Camera's center coordinates: [0, 0, 1.8]
Orientation Vector: [0, 0, 1]

(a) (b)

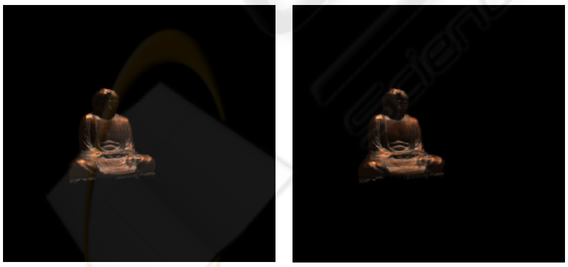
Figure 7: Test 2: Imagery camera rendered views with zoom out.

(a): Nearest Neighborhood
(b): Bilinear
Virtual camera's adjustment:
Camera's center coordinates: [0, 0, 2.5]
Orientation Vector: [0, 0, 1]

(a) (b)

Figure 8: Test 3: Imagery camera rendered views with zoom in and left rotation.

(a): Nearest Neighborhood
(b): Bilinear
Virtual camera's adjustment:
Camera's center coordinates: [0.4, 0, 1.8]
Orientation Vector: [0, 0, 1]



(a) (b)

Figure 9: Test 4: Imagery camera rendered views with zoom out and right rotation.

(a): Nearest Neighborhood
(b): Bilinear
Virtual camera's adjustment:
Camera's center coordinates: [-0.5, 0, 2.5]
Orientation Vector: [-0.3, 0, 1]

We evaluated the performance of the watermark sequences in terms of their robustness and imperceptibility using a blind detection scheme. The watermarks were tested in cases of zooming and rotation in order to conclude to the most robust watermarks. Obviously, the best possible watermark sequence must show the following properties:

- Big PSNR of the watermarked view to ensure high image quality.
- Big Correlation of the watermark with the watermarked view to ensure easy (blind) detection.

Our aim is to find a satisfactory balance between PSNR and Correlation that combines robustness of watermark and good quality of the view. We tested the 15 distributions (of Table 1) as watermarks using the watermarking procedure described in Section 2.

Taking into account the results reported in (Apostolidis, Koz and Triantafyllidis, 2007) for the cases 1 and 2 (see Fig. 5) where the imagery camera is located in the camera plane, we narrowed our tests to the following distributions that succeed better in combining robustness of watermark and good quality of the view:

- Beta Distribution
- Extreme Value Distribution
- Normal (Gaussian) Distribution
- Exponential Distribution
- Negative Binomial Distribution

The next step is to determine all the parameters for each of these five distributions in order to produce the best possible results in terms of robustness and imperceptibility for the case 3. Five data sets for these "best" parameters of each of the five distributions are listed in Table 2.

In Table 3, the results are given for the four tests mentioned above according to:

- The interpolation method for the construction of the rendered view (Nearest Neighborhood / Bilinear Interpolation).
- The data sets of the parameters for each of the five selected distributions.

For each of these tests, we calculated the PSNR and correlation values in order to evaluate the data sets and distributions in terms of robustness and imperceptibility using blind detection. The signs "+" and "–" indicate the efficiency of the corresponding data set of the distribution.

Table 2: For each distribution, five data sets of the parameters that produce the best results.

	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5
Beta					
Par. 1	0,25	0,5	0,75	1	1,5
Par. 2	0,75	1,25	0,25	0,75	1,25
Filter	0,7	0,85	0,9	0,9	0,9
Scalar	0,1	0,15	0,15	0,1	0,15
Extreme Value					
Par. 1	0,25	0,5	0,5	0,75	1,25
Par. 2	0,25	0,25	0,75	0,25	0,5
Filter	0,9	0,9	0,9	0,9	0,9
Scalar	0,25	0,1	0,1	0,1	0,1
Normal / Gaussian					
Par. 1	0	0	0,75	1	1
Par. 2	0,25	0,5	0,25	0,25	0,5
Filter	0,75	0,7	0,8	0,8	0,9
Scalar	0,15	0,1	0,1	0,1	0,1
Exponential					
Par. 1	0,5	0,3	0,2	0,2	0,1
Filter	0,9	0,75	0,7	0,75	0,85
Scalar	0,15	0,15	0,15	0,2	0,25
Negative Binomial					
Par. 1	0,1	0,3	0,7	0,7	0,9
Par. 2	0,7	0,7	0,9	0,9	0,9
Filter	0,9	0,85	0,7	0,9	0,85
Scalar	0,15	0,2	0,2	0,25	0,15

Table 3: Results for the five selected distributions and the corresponding five selected data sets. Please note that "+" means more efficient result in terms of robustness and imperceptibility using blind detection, while "-" means that the result is less sufficient (a predefined threshold has been used for the PSNR and correlation values).

Beta Distribution				
	Nearest Neighborhood / Bilinear Interpolation			
	Test 1	Test 2	Test 3	Test 4
Data Set1	+ / -	+ / -	+ / +	- / -
Data Set2	+ / +	- / +	+ / -	- / +
Data Set3	- / -	+ / +	+ / +	- / -
Data Set4	+ / -	- / -	- / -	+ / +
Data Set5	+ / +	+ / -	- / +	+ / -
Extreme Value Distribution				
	Nearest Neighborhood / Bilinear Interpolation			
	Test 1	Test 2	Test 3	Test 4
Data Set1	+ / -	- / -	- / +	+ / +
Data Set2	+ / -	- / -	+ / +	- / -
Data Set3	- / -	+ / +	+ / +	- / +
Data Set4	+ / -	- / -	- / +	- / +
Data Set5	+ / +	- / +	+ / -	+ / +

Normal / Gaussian Distribution				
	Nearest Neighborhood / Bilinear Interpolation			
	Test 1	Test 2	Test 3	Test 4
Data Set1	+ / +	+ / +	- / +	- / -
Data Set2	- / -	- / -	+ / +	- / +
Data Set3	+ / -	+ / +	- / +	- / +
Data Set4	+ / -	+ / +	+ / -	+ / +
Data Set5	- / +	- / +	+ / +	+ / +

Exponential Distribution				
	Nearest Neighborhood / Bilinear Interpolation			
	Test 1	Test 2	Test 3	Test 4
Data Set1	+ / +	+ / +	- / +	+ / +
Data Set2	- / -	- / +	- / +	- / -
Data Set3	+ / +	+ / +	+ / -	- / +
Data Set4	+ / +	+ / -	- / -	- / -
Data Set5	- / +	- / +	- / +	+ / -

Negative Binomial Distribution				
	Nearest Neighborhood / Bilinear Interpolation			
	Test 1	Test 2	Test 3	Test 4
Data Set1	- / -	+ / -	- / -	- / -
Data Set2	- / -	- / -	+ / -	+ / +
Data Set3	+ / -	- / -	- / -	+ / -
Data Set4	- / +	- / -	- / +	- / -
Data Set5	+ / +	+ / -	- / -	+ / +

Taking into account the "+" and "–" for the data sets of the distribution, we can conclude to the best possible data sets as listed in Table 4:

Table 4: "Winning" data sets and distributions.

Normal/ Gaussian distribution	**Data Set 5** Parameter a=1 Parameter b=0.5 Filter Cutoff=0.9 Scalar Factor=0.1
Exponential distribution	**Data Set 1** Parameter a=0.1 Filter Cutoff=0.85 Scalar Factor=0.25

4 CONCLUSIONS

In free view point video, the user might record a personal video for an arbitrarily selected view and misuse the content, so it is apparent that copyright protection problems should be solved for free-view TV. In this paper we employed several distributions as watermark sequences and we tested them in terms of robustness and imperceptibility.

ACKNOWLEDGEMENTS

This work was supported by the General Secretariat of Research and Technology Hellas under the InfoSoc "TRAVIS: Traffic VISual monitoring" project and the EC under the FP6 IST Network of Excellence: "3DTV-Integrated Three-Dimensional Television - Capture, Transmission, and Display" (contract FP6-511568). The authors should like to thank A. Koz for his valuable help.

REFERENCES

Zhang, C., and Chen, T., 2004. A survey on image based rendering-representation, sampling and compression. In *Signal Processing: Image Communications, vol. 19, pp 1-28*.

MPEG Meeting, 2003. *Applications and Requirements for 3DAV*, document N5877, Trondheim, Norway.

Koz, A., and Alatan A.A, 2005. Oblivious Video Watermarking Using Temporal Sensitivity of HVS. In *IEEE International Conference on Image Processing, Vol. 1, pp 961 – 964*.

Levoy, M., Hanrahan, P., 1996. Light field rendering. In *Computer Graphics (SIGGRAPH'96), New Orleans, pp. 31–42*.

Koz, A., Çıgla C., and Alatan, A.A., 2006. Free-view Watermarking for Free-view Television. In *IEEE International Conference on Image Processing 2006, Atlanta, GA, USA*.

The Stanford Light Field Archive, 2007. http://graphics.stanford.edu/software/lightpack/lifs.html

Apostolidis, E., Koz A., and Triantafyllidis G.A., 2007. Best Watermarking Selection for Free-View Point Television. In *14th International Conference on systems, Signals and Image Processing IWSSIP 2007 and 6th EURASIP Conference Focused on Speech and Image Processing, Multimedia Communications and Services EC-SIPMCS 2007, Maribor, Slovenia*.