Plenoptic image compression comparison between JPEG, JPEG2000 and SPITH

Rogério Seiji Higa, Roger Fredy Larico Chavez, Ricardo Barroso Leite, Rangel Arthur and Yuzo Iano

***Abstract*— Plenoptic image is a novel visual representation that contains more information than traditional images. Different focal planes and different perspectives can be recovered by a rendering algorithm. In this work existing state of the art compression schemes are tested on plenoptic images, and its behavior is analyzed. The analysis is made not on the plenoptic image reconstruction, but on the rendered views. In this way the compression artifacts are analyzed in terms of the interested rendered view. The results showed that the existing compression schemes can be used with good results to compress plenoptic images, in general the JPEG and JPEG2000 had the best performance but for low bitrates the SPIHT had the advantage.**

***Index Terms*—light field, plenoptic image, image compression, refocusing, JPEG, JPEG2000, SPIHT**

I. Introduction

The digital light field camera was introduced by Ren Ng in [1], where a microlens array was used to sample angular information about the light rays. Images captured by a light field or plenoptic camera provide more information about the scene than traditional images, like alternative points of view and focus planes.

The effective resolution of the plenoptic camera is very low compared to the image captured [2-5]. This is because it sacrifices spatial resolution to capture information about the angle of the light ray, which is useful in many applications. Many works attempts to increase the effective resolution using super-resolution techniques [6-7].

Different focal planes and points of view can be recovered by a rendering algorithm. This increase in the information however requires more storage space in relation to the effective image resolution.

There are a few researches in compression of this new data representation [8-11], but the path to standards devoted to this specific type of image is a long road. The question that arises is: can we use the existing compression standards for plenoptic images? What are the effects of the compression on the rendered images?

Rogério Seiji Higa, Roger Fredy Larico Chavez, Ricardo Barroso Leite, and Yuzo Iano are with the State University of Campinas, Campinas, SP 13083-852 Brazil, e-mail: rhiga@decom.fee.unicamp.br, rlarico@decom.fee.unicamp.br, rleite@decom.fee.unicamp.br, yuzo@decom.fee.unicamp.br.

Rangel Arthur is with the State University of Campinas, Limeira, SP, Brazil, email: rangel@ft.unicamp.br

Standards like JPEG [12] and JPEG2000 [13] are very well established, including hardware and image editing software support. Using them is advantageous on the commercial point of view, since adapting them are cheaper than establishing a new standard. There is also the SPIHT, which is a fast algorithm with better results than JPEG, but it does not have any commercial version [14-15].

The main contribution of this work is the study of the plenoptic image compression behavior. It is used existing compression schemes and existing rendering algorithms without modification. This replicates a real scenario where commercial cameras are modified to capture plenoptic images, and seeks to answer if the application would be not hindered by the compression.

In section II the details on plenoptic images are given, in the section III it is described the compression schemes used, and in section IV the results are presented.

II. Plenoptic Image

What we call plenoptic image is the image captured by a plenoptic camera that uses a microlens array to sample the light field or plenoptic function [16-18].

A. Plenoptic Function

The plenoptic function describes the distribution of the light at each point of the space. It is a 7D function in its full form. Levoy presented in [18] a two plane parameterization that reduced its dimensionality to a 4D function, shown in Fig. 1

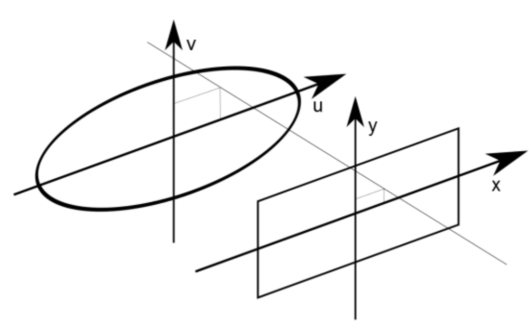

Fig. 1 Plenoptic function parameterization by two planes.

For a plenoptic camera, the first plane maps the points of the sensor and the other plane maps angles of the light ray.

B. Plenoptic Camera

The microlens array is placed in front of the sensor as shown in Fig. 2. The one used in our tests is manufactured by Raytrix, and the microlens array has an hexagonal pattern that increases the useful area of the sensor;

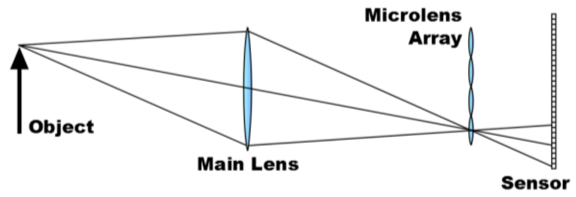

Fig. 2. Plenoptic camera design, the microlens array separate the light rays before hitting the sensor.

The center of each microlens can be considered a point in the plane x-y, and the main lens the u-v plane.

C. *Plenoptic Image Structure*

The light rays that pass though the main lens are separated by the microlens array before hitting the sensor. Each microlens converts the angular information to positional coordinates. The design of a plenoptic camera is shown in Fig. 3. The captured image by the camera is shown in Fig. Notice how the microimages are organized in a hexagonal pattern, which increases the utilization of the sensor.

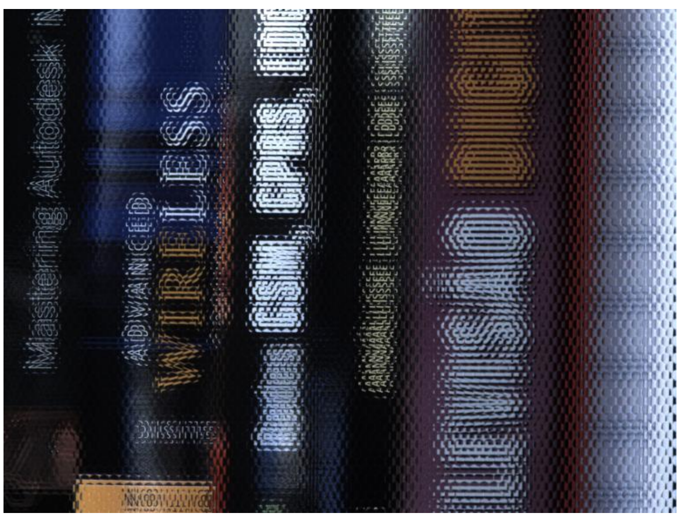

Fig. 3. Example of plenoptic image.

The plenoptic image is composed by a series of microimages. The same feature can be perceived in various microimages, in slightly different positions. If the focus on the main lens is manually adjusted, the feature can move or change multiplicity across the micro-images.

The compression is applied on this image, and then recovered to be used for the rendering algorithm.

D. *View Rendering*

The rendering algorithm is the same as described in [3]. A patch from each microimages is taken from the plenoptic image, and then they are tiled together to form a rendered view. The Fig. 4 illustrates the process. The size of the patch is what determines where the focal plane is.

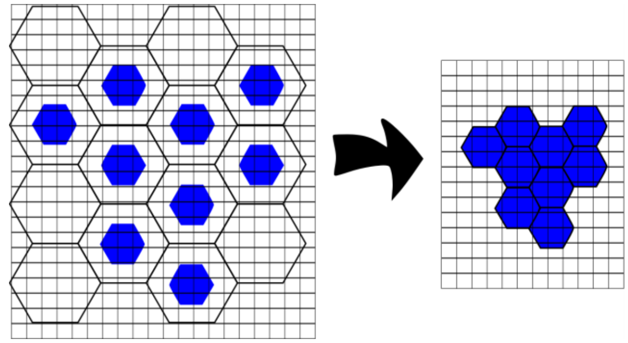

Fig. 4. Rendering process to generate a focus image from the plenoptic image.

III. COMPRESSION SCHEMES

A. *JPEG*

The JPEG standard [12], illustrated in Fig. 5, is based on the use of the DCT in 8×8 blocks, followed by quantization and entropy coding. The block approach leads to a better correlation of the input data, enhancing the algorithm performance. But this leads to the blocking artifacts where discontinuities appear between the blocks.

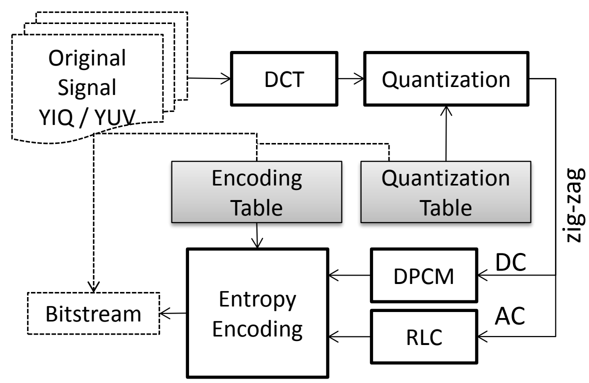

Fig. 5. Block diagram of the standard JPEG encoder.

B. *SPIHT*

The SPIHT is based on the wavelet transform, and uses a tree representation for the coefficients. It was introduced by Said and Pearlman in [14-15] as a refined version of the EZW. The energy is concentrated near the root node, as shown in Fig. 6, so the descending nodes usually have less energy than its parent. The compression is achieved by partitioning the tree, pruning nodes below a certain threshold. Advantages of this scheme are a fast progressive image transmission and an exact bitrate control.

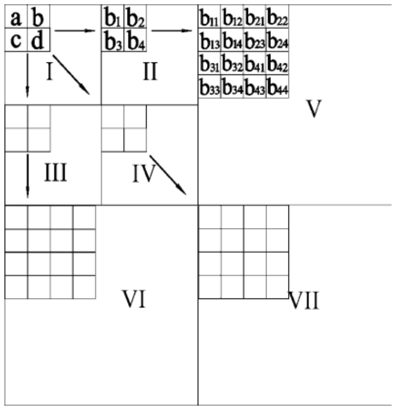

Fig. 6. Wavelet decomposition tree structure used by the SPIHT

C. *JPEG2000*

The JPEG2000 is also based on the wavelet transform, and can achieve up to 90% of compression without loss of quality. It uses the EBCOT (Embedded Block Coding with Optimized Truncation) [19] and is capable of lossy and lossless compression. Fig. 7 shows how the JPEG2000 works.

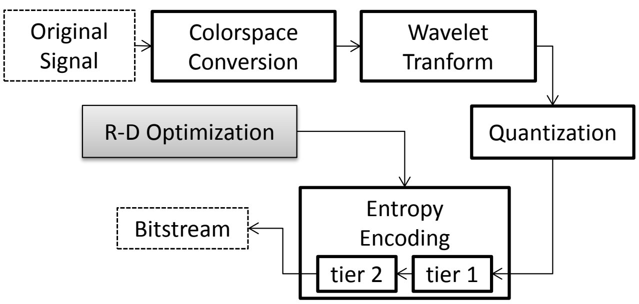

Fig. 7. Block diagram of the standard JPEG2000 encoder

The JPEG2000 works very similar to the SPIHT as it also uses the wavelet transform, but the main difference is the use of an adaptive arithmetic coding that explores the intensity of the wavelet coefficients to generate a number of lower magnitudes based on the neighbor coefficients.

IV. RESULTS

The framework of the tests is shown in Fig. 8. The plenoptic image is compressed with one of the methods described earlier. Then we compare the rendered images using PSNR and SSIM.

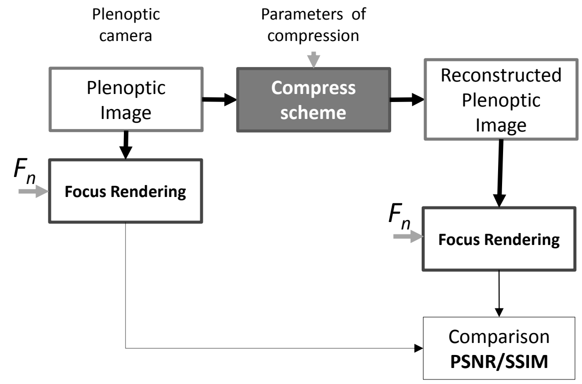

Fig. 8. Compression analysis framework.

The plenoptic images used here are shown in Fig. 16, Fig. 17, Fig. 18, and Fig. 19.

Fig. 9. Plenoptic image: books.

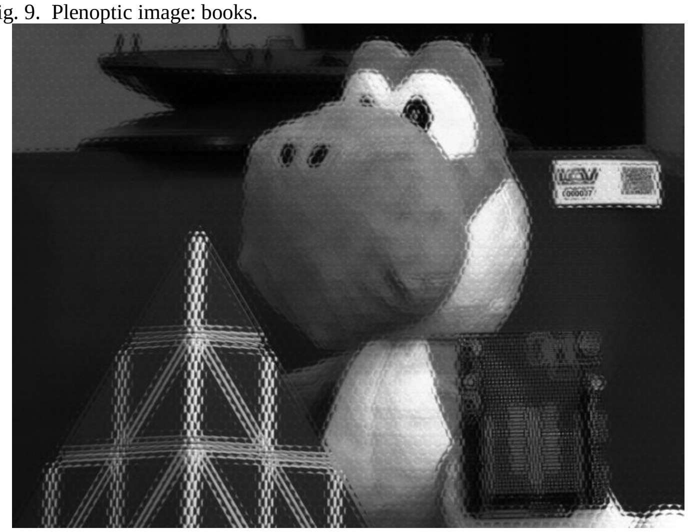

Fig. 10. Plenoptic image: doll.

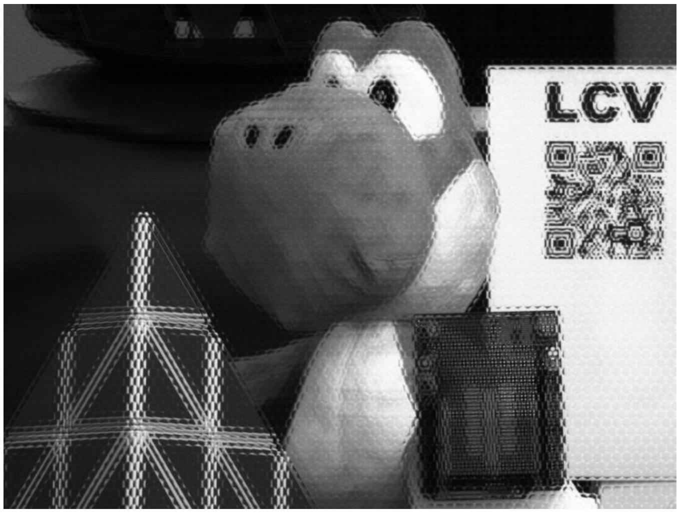

Fig. 11. Plenoptic image:QR doll.

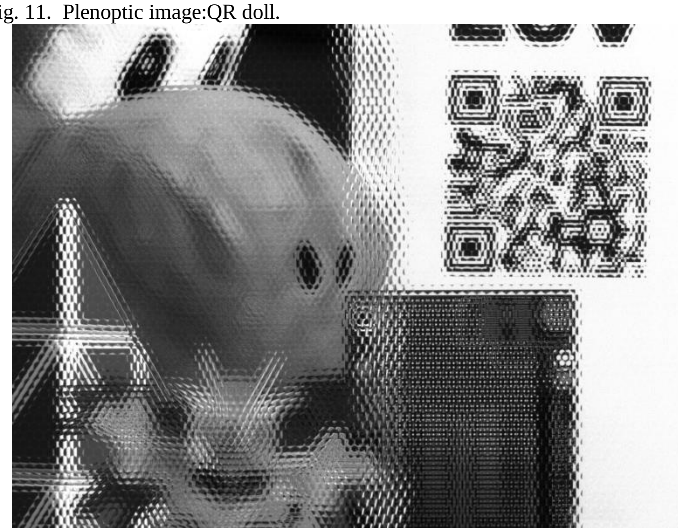

Fig. 12. Plenoptic image: wire.

Fig. 13 shows another angle of the scene of Fig. 12, observe that the objects are at different depths.

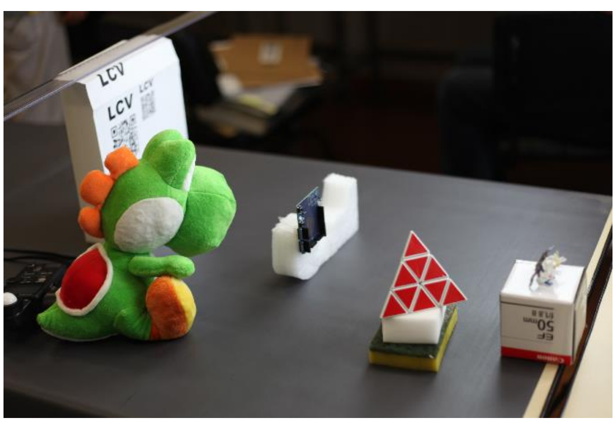

Fig. 13. Lateral view of the objects captured by the plenoptic camera, taken with a normal camera.

The Fig. 14 is a rendered view on the farthest focal plane, observe how the QR code is sharp. The Fig. 15 is focused at the wire, this object is not visible in the plenoptic image, see Fig. 12.

Fig. 14. Rendering of wire focused at the QR code.

The blocking effect in Fig. 15 is caused by the rendering algorithm, the tilling process is the responsible for it. If the patch has a suitable size this blocking effects disapears, that is, the focus is at the plane where there is no blocking effects.

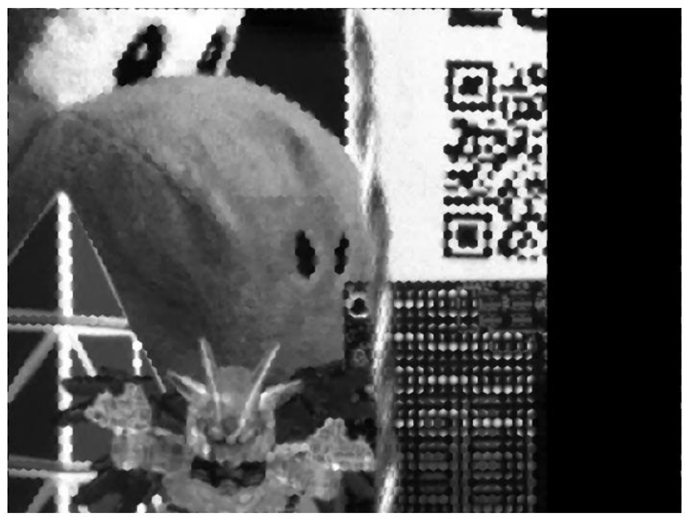

Fig. 15. Rendering of wire focused at nearest object.

The Fig. 16, Fig. 17, Fig. 18 and Fig. 19 show the PSNR performance of the compression schemes. They are the average of the PSNR for each focus plane rendered from the plenoptic images. The JPEG and JPEG2000 have overall a better performance than SPIHT, with an edge of advantage for the JPEG2000. But for low bitrates the SPIHT has a better performance.

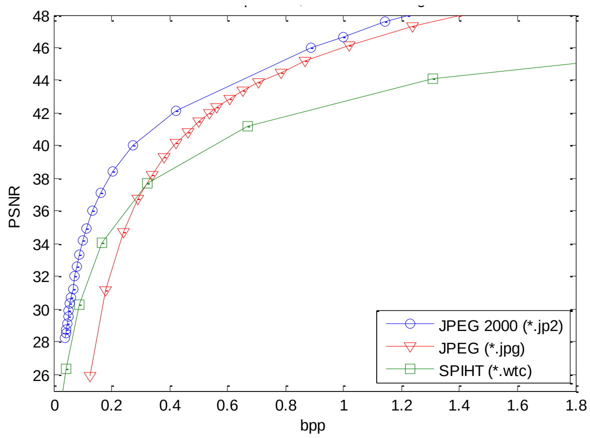

Fig. 16. PSNR results for books.

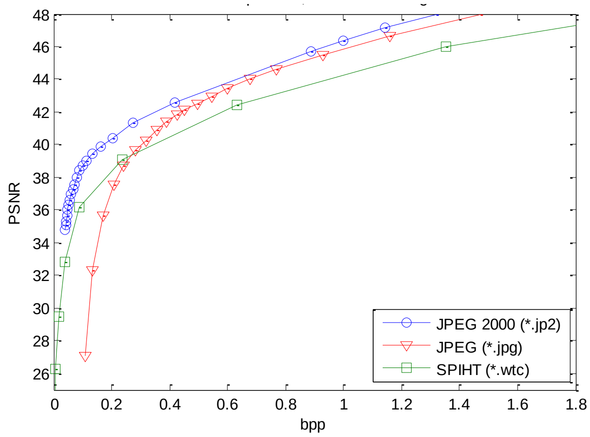

Fig. 17. PSNR results for doll.

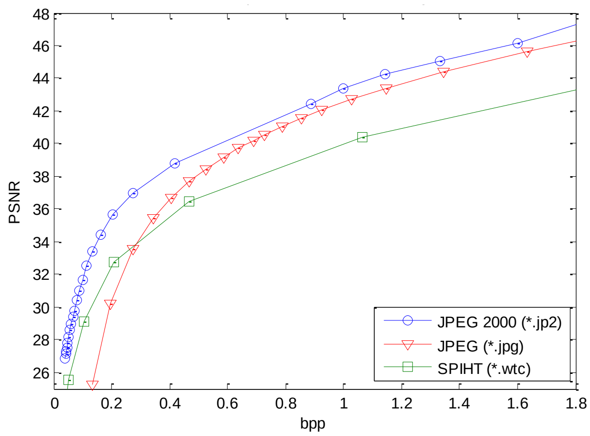

Fig. 18. PSNR results for QR-doll.

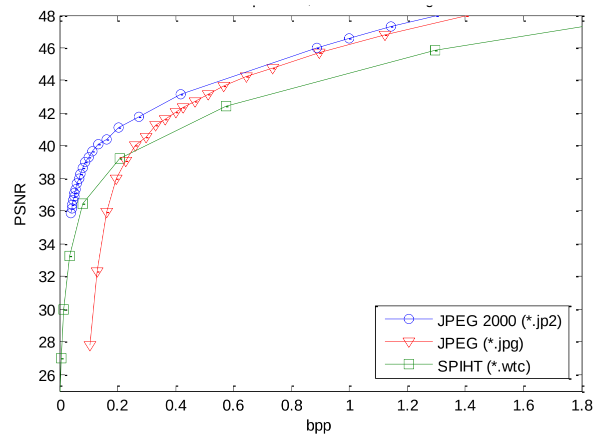

Fig. 19. PSNR results for wire.

The Fig. 20, Fig. 21,Fig. 22 and Fig. 23 shows the results analyzed with the SSIM. The SPIHT shows a slight better performance using this metric.

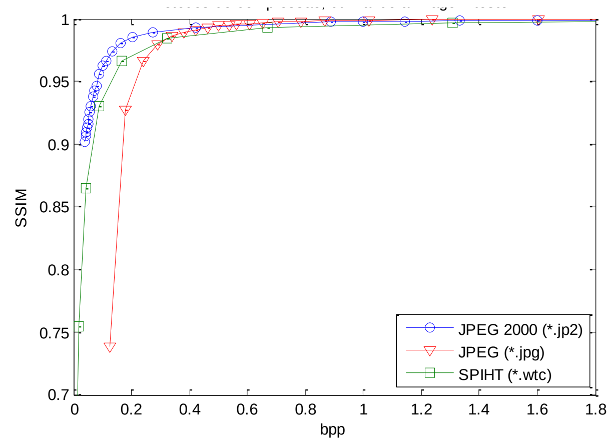

Fig. 20. SSIM results for books.

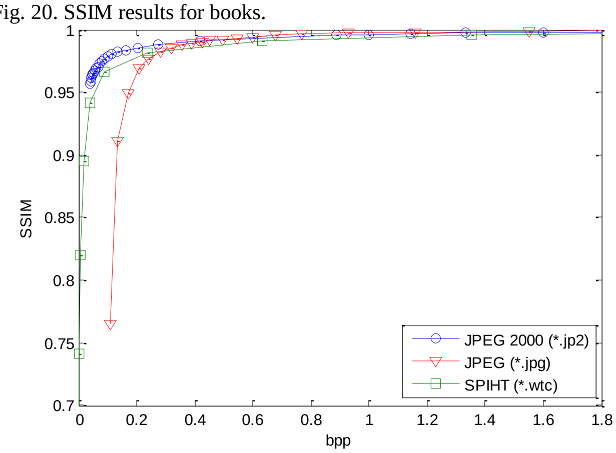

Fig. 21. SSIM results for doll.

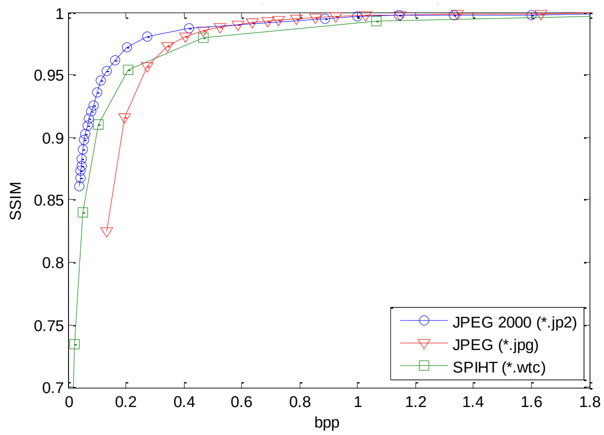

Fig. 22. SSIM results for QR-doll.

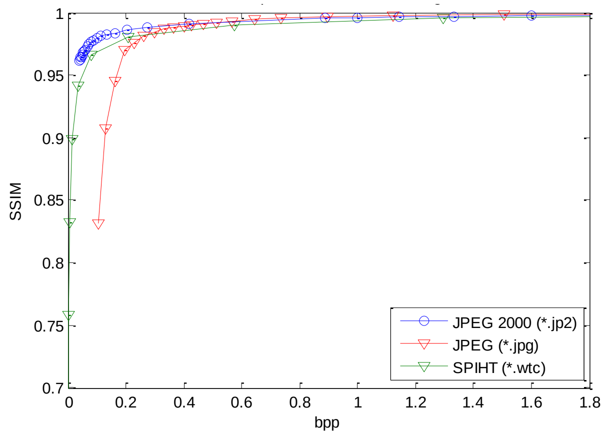

Fig. 23. SSIM results for wire.

V. Conclusions

The main objective of this work was to establish the best compression algorithm for applications that uses the plenoptic images. It is known that the best compression performance for common images belongs to the JPEG2000, which is matured over the years. This knowledge holds true for plenoptic images as well.

As seen in the performance curves, SPIHT is a good alternative for high compression ratios compared to the JPEG, and has a lower complexity than JPEG2000. The SPIHT is a good alternative for applications with limited bandwidth or in remote locations. In general the JPEG and JPEG2000 have a better performance than SPIHT, this holds true for both PSNR and SSIM metrics.

A detailed analysis of the geometry and rendering algorithm can lead to a better performance in modified version of these compression algorithms. For example, the compression blocks could be set to be exactly as large as the microimages, avoiding blocking artifacts. Future works should explore these plenoptic image characteristics.

Acknowledgment

The authors would like to thank the following institutions for making this work possible: CAPES-RHTVD, CAPES, CTIC-RNP, CNPq, Fapesp, Faepex-Unicamp.

References

[1] R. Ng, “Digital light field photography,” Stanford, 2006.
[2] T. Georgiev, A. Lumsdaine, and G. Chunev, “Focused plenoptic camera and rendering,” Camera, vol. 19, no. April, p. 021106, Jan. 2009.
[3] A. Lumsdaine and T. Georgiev, “Full Resolution Lightfield Rendering,” 2008.
[4] R. Ng, M. Levoy, G. Duval, M. Horowitz, and P. Hanrahan, “Light Field Photography with a Hand-held Plenoptic Camera,” 2005.
[5] T. Georgiev and a. Lumsdaine, “Reducing Plenoptic Camera Artifacts,” Computer Graphics Forum, vol. 29, no. 6, pp. 1955–1968, Sep. 2010.
[6] T. Georgiev and A. Lumsdaine, “Superresolution with plenoptic 2.0 cameras,” Signal Recovery and Synthesis, no. x, pp. 2–4, 2009.
[7] T. E. Bishop, S. Zanetti, and P. Favaro, “Light field superresolution,” 2009 IEEE International Conference on Computational Photography (ICCP), pp. 1–9, Apr. 2009.
[8] N. Gehrig and P. L. Dragotti, “Geometry-driven distributed compression of the plenoptic function: performance bounds and constructive algorithms.,” IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, vol. 18, no. 3, pp. 457–70, Mar. 2009.
[9] Shing-Chow Chan; King-To Ng; Zhi-Feng Gan; Kin-Lok Chan; Heung-Yeung Shum, "The plenoptic videos: capturing, rendering and compression," *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on* , vol.3, no., pp.III-905,III-908 Vol.3, 23-26 May 2004
[10] Jim Schwiegerling, Gabriel C. Birch and J. Scott Tyo, “Analysis and compression of plenoptic camera images with Zernike polynomials,” Proc. SPIE 8487, Novel Optical Systems Design and Optimization XV, 84870G (October 19, 2012); doi:10.1117/12.929720.
[11] Andriy Gelman, Pier Luigi Dragotti and Vladan Velisavljević, “Centralized and interactive compression of multiview images,” Proc. SPIE 8135, Applications of Digital Image Processing XXXIV, 81350J (September 23, 2011); doi:10.1117/12.895982
[12] W. B. Pennebaker and J. L. Mitchell, JPEG still image data compression standard, vol. 34. Van Nostrand Reinhold, 1993, p. xviii, 638 p.
[13] D. S. Taubman and M. W. Marcellin, JPEG2000: Image Compression Fundamentals, Standards and Practice, vol. 11, no. 2. Springer, 2002, p. 286.
[14] A. Said and W. A. Pearlman, “A new, fast, and efficient image codec based on set partitioning in hierarchical trees,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, no. 3, pp. 243–250, 1996.
[15] W. A. Pearlman and A. Said, “Set partition coding: Part i of set partition coding and image wavelet coding systems,” *Found. Trends Signal Process.*, vol. 2, pp. 95–180, February 2008. [Online]. Available: http://dl.acm.org/citation.cfm?id=1482328.1482329
[16] E. H. Adelson and J. R. Bergen, “The Plenoptic Function and the Elements of Early Vision,” Processing, vol. 1, no. 1, pp. 3–20, 1991.
[17] M. N. Do, D. Marchand-Maillet, and M. Vetterli, “On the bandwidth of the plenoptic function.,” IEEE transactions on image processing : a publication of the IEEE Signal Processing Society, vol. 21, no. 2, pp. 708–17, Feb. 2012.
[18] M. Levoy, M. Levo, and P. Hanrahan, “Light field rendering,” in Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996.
[19] Taubman, D., "High performance scalable image compression with EBCOT," Image Processing, IEEE Transactions on , vol.9, no.7, pp.1158,1170, Jul 2000