

AVALIAÇÃO DA COMPRESSÃO DE LIGHT FIELDS COM O PADRÃO DE CODIFICAÇÃO DE VÍDEO HEVC

GIORDANO ESTÊVÃO BORGES LEAL; RUHAN CONCEIÇÃO; GUILHERME CORREA; BRUNO ZATT; LUCIANO AGOSTINI; MARCELO PORTO

*Universidade Federal de Pelotas – Video Technology Research Group (ViTech)
{gbleal, radconceicao, gcorrea, zatt, agostini, porto}@inf.ufpel.edu.br}*

1. INTRODUÇÃO

Imagens e vídeos digitais estão presentes na maioria dos dispositivos atuais. Dificilmente se encontra um indivíduo que não possua um *smartphone*, onde é possível visualizar as mais diversas formas de conteúdo distribuídas nas redes sociais e serviços de *streaming*. Considerando essa facilidade de acesso encontrada nos dias atuais, cada vez mais tecnologias de vídeo e imagem crescem em popularidade.

Muitas dessas tecnologias focam-se em melhorar a experiência do usuário, criando formas diferentes de conceber a interação aparelho-usuário. As televisões em 3D, por exemplo, chegaram ao mercado propondo uma maneira diferenciada de contato com o mundo tridimensional. Porém, com os cortes no desenvolvimento de televisões em 3D feitos pelas maiores empresas do ramo (KATZMAIER, 2016), nota-se como as restrições tecnológicas podem afetar o mercado. Cria-se então uma necessidade de novas tecnologias na área tridimensional, uma delas são os *light fields* (MATUSIK; PFISTER, 2004).

Os *light fields* são baseados na função plenóptica (ADELSON; BERGEN, 1991), que tenta listar todas as características do mundo visível e colocá-las numa função, que descreve tudo que é visualmente perceptível. Utilizando as informações contidas nessa função, é possível criar imagens tridimensionais que não perdem nenhum dado visual. A captação de algumas dessas informações discretizadas (removendo certos parâmetros de forma adequada) (IHRKE; RESTREPO; MIGNARD-DEBISE, 2016), gera um *light field*, que nada mais é do que a tecnologia da visão de diversos ângulos de uma imagem.

Porém, todas essas informações visuais importantes (ângulo, intensidade de cor, posição, etc.), sendo utilizadas para compor um *light field*, geram um enorme volume de dados para a sua representação. Enquanto vários métodos utilizam duas imagens bidimensionais (ou um conjunto igualmente restrito) para representar uma imagem em 3D, os *light fields* podem utilizar dezenas ou centenas, aumentando ainda mais o volume de dados e dificultando o seu processamento e armazenamento. Esse problema é consideravelmente minimizado através de codificação, permitindo atenuar informações redundantes nas imagens. É possível utilizar diversas formas de codificação e diferentes métodos. Uma possibilidade seria decompor o *light field* em uma sequência de imagens e utilizar o padrão JPEG (WALLACE, 1991). Outra opção seria agrupar esse conjunto de imagens em uma sequência de vídeo – utilizando ferramentas de codificação apropriadas. A segunda opção tem vantagem sobre a primeira, pois explora a redundância do movimento entre as imagens, não só dados visuais locais delas.

Este trabalho visa explorar duas formas de codificação de *light fields* transformados em vídeo e comparar a qualidade visual e compressão de dados entre elas, utilizando o padrão HEVC. Os resultados obtidos demonstram a eficiência dos métodos utilizados, apresentando formas de solução de problemas relacionados ao volume de dados por codificação de vídeo.

2. METODOLOGIA

Cada imagem em cada ângulo do *light field* é um *frame* no vídeo a ser codificado. Em um campo de visão de um *light field* as imagens variam da esquerda para direita na horizontal e de cima para baixo no ângulo vertical. Na Figura 1 (a) é possível ver esse campo e na Figura 1 (b) o temos aumentado, mostrando as diferentes vistas contidas nele.

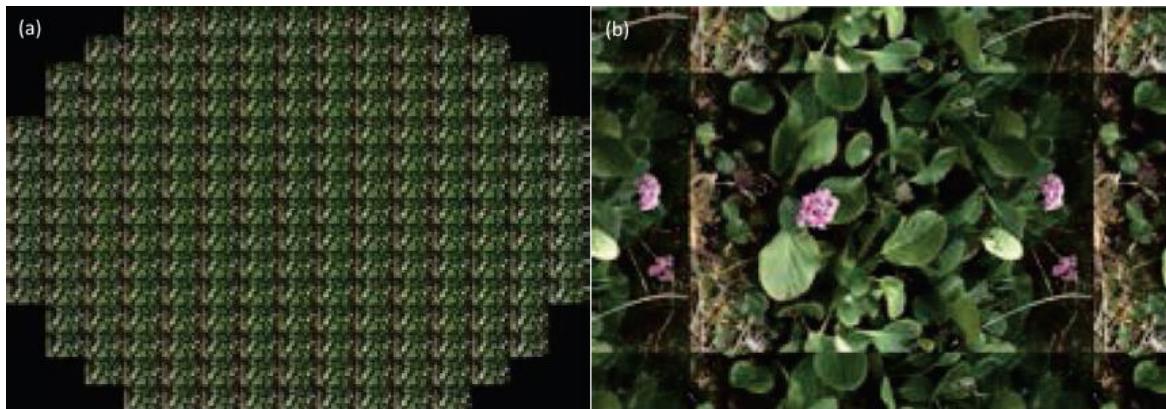


Figura 1. *Light Field* Completo (a) e zoom da imagem (b). Adaptado de: (VIEIRA et al., 2015).

Duas formas de transformação de *light fields* em vídeo foram utilizadas nesse trabalho. A primeira transformação (tipo “linha”) distribui as imagens no vídeo da seguinte forma: inicialmente, a primeira linha da esquerda para a direita, após, a segunda e assim por diante, como pode ser visto na Figura 2 (a). A segunda forma (tipo “circular”) começa no centro e vai distribuindo as imagens em espiral até chegar à última, como podemos ver na Figura 2 (b).

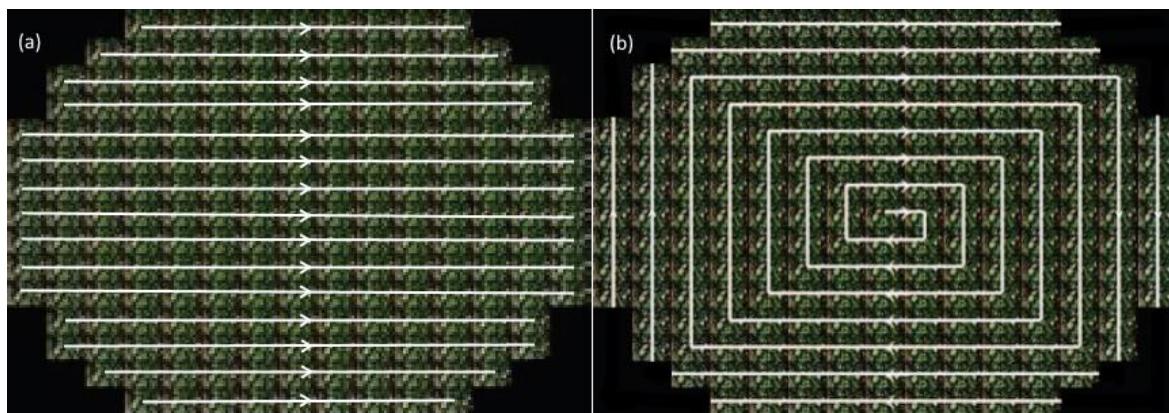


Figura 2. Transformações “linha” (a) e “circular” (b). Adaptado de: (VIEIRA et al., 2015).

Os *light fields* utilizados nos experimentos podem ser encontrados no dataset SMART (PAUDYAL et al., 2016). Treze *light fields* foram usados: *Building*, *Car*, *Chair*, *Flower*, *Grid*, *Ladder*, *River*, *Sky*, *Stone*, *Table*, *Tile*, *Whitesky* e *Window*. O programa no qual foram gerados os vídeos transformados foi o MatLab (THE MATHWORKS INC., 2014), com o Toolbox “Light Field Toolbox v0.4” (DANSEREAU, 2015). O padrão de compressão de vídeo utilizado foi o HEVC (*High Efficiency Video Coding*) (SULLIVAN et al., 2012), a partir do uso de seu software de referência, versão HM-16.9 (FRAUNHOFER HEINRICH HERTZ INSTITUTE, 2016), configurado em *Low Delay B*, com vídeos em *full chroma resolution* (4:4:4) e QPs (*quantization parameters*) em 22, 27, 32 e 37.

3. RESULTADOS E DISCUSSÃO

Os critérios de qualidade empregados foram a taxa de bits gerados após a codificação (*bitstream*), que define a eficiência de compressão dos dados, o PSNR (*Peak Signal-to-Noise Ratio*, ou “relação sinal-ruído de pico”), que define, de maneira objetiva, a qualidade visual do processo de codificação e a redução, que indica quantas vezes seria necessário se multiplicar o *bitstream* resultante para se obter o tamanho original do *light field*. A *bitstream* é uma sequência de dados em binário (bits) resultante do processo de codificação do vídeo original. O processo de decodificação recebe a *bitstream* e a transforma no vídeo final decodificado. PSNR é a medida que avalia a qualidade da imagem reconstruída (após o processo de descompressão) em relação à imagem original, a partir de cálculos matemáticos entre os pixels destas imagens. Os valores de PSNR e *bitstream* da Tabela 1 são equivalentes às médias dos resultados dos quatro QPs de cada codificação, os quais são parâmetros passados para o codificador de vídeo que, sumarizando, estabelecem quantas informações visuais prescindíveis podem ser eliminadas.

Avaliando os resultados, a média total do tamanho das *bitstreams* resultantes para o modo circular é de aproximadamente 393KB, enquanto que para o modo linha é de cerca de 353KB. Sendo que os *light fields* originais possuem 1,17 GB cada, a média dos *bitstreams* resultantes dos dois modos (373KB) mostra uma dedução de 99,97% (1,16GB) de tamanho do *light field* original, o que expressa a extrema eficiência desse método na redução do volume de dados.

Tabela 1. Médias de *Bitstream*, PSNR e Redução resultantes do processo de codificação dos *light fields* e suas médias gerais.

Light Field	PSNR	Bitstream	Redução	PSNR	Bitstream	Redução
	(dB)	(KB)	(x)	(dB)	(KB)	(x)
	Linha			Circular		
Building	38,09	871	1409	38,23	937	1309
Car	42,69	126	9719	43,07	130	9434
Chair	38,82	335	3663	39,03	350	3506
Flower	38,05	413	2970	38,41	473	2596
Grid	43,09	132	9273	43,43	133	9238
Ladder	37,18	620	1979	37,38	733	1674
River	38,09	693	1770	38,31	760	1615
Sky	40,13	463	2649	40,23	509	2410
Stone	36,38	571	2150	36,70	708	1733
Table	40,78	229	5368	41,13	223	5510
Tile	42,42	202	6063	42,69	214	5733
Whitesky	42,28	205	5985	42,70	206	5956
Window	41,62	271	4535	41,77	306	4015
Média	36,58	353	3476	36,85	393	3122

Podemos observar na Tabela 1 uma comparação entre as transformações utilizadas. A redução do *bitstream* do tipo linha é maior do que a do circular para a todos dos casos - com exceção de *Table*, que é altamente redundante na região central do campo - e nota-se o contrário para o PSNR.

Sendo assim, cada um é melhor para uma aplicação em específico. O tipo linha se mostra mais eficiente quando é necessária uma redução maior de

bitstream e o tipo circular quando menos perda de qualidade é o desejado – para valores similares, porém superiores, de *bitstream*.

4. CONCLUSÕES

Neste trabalho foram exploradas duas estratégias para a compressão de *light fields* utilizando o padrão de codificação de vídeo HEVC. A primeira fazendo uma transformação linha por linha, a segunda uma circular em espiral a partir do centro. Testes foram realizados com treze *light fields*, utilizando o HM 16.9 para gerar resultados de compressão e PSNR. Os resultados demonstraram que a taxa de compressão média foi superior a 99%, com PSNR médio entre 36,58 e 36,85dB. Cada técnica se mostrou eficiente para sua devida aplicação, pois uma delas apresentou valores menores de *bitstream* (tipo linha) e a outra valores maiores de PSNR (circular). Como trabalhos futuros temos a análise de outros métodos de transformação de *light fields* em vídeo, tais como um método de transformação em tipo *raster* (similar ao tipo linha, porém mudando de sentido a cada uma), trazendo mais dados de comparação de eficiência.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ADELSON, E. H.; BERGEN, J. R. The plenoptic function and the elements of early vision. **Computational Models of Visual Processing**, p. 3–20, 1991.
- DANSEREAU, D. **Light Field Toolbox v0.4MathWorks**, 2015. Disponível em: <<http://www.mathworks.com/matlabcentral/fileexchange/49683>>
- FRAUNHOFER HEINRICH HERTZ INSTITUTE. **High Efficiency Video Coding (HEVC)** | JCT-VC **Fraunhofer-Gesellschaft**, 2016. Disponível em: <<https://hevc.hhi.fraunhofer.de/>>
- HANHART, P.; EBRAHIMI, T. Calculation of average coding efficiency based on subjective quality scores. **Journal of Visual Communication and Image Representation**, v. 25, n. 3, p. 555–564, 2014.
- IHRKE, I.; RESTREPO, J.; MIGNARD-DEBISE, L. Principles of Light Field Imaging: Briefly revisiting 25 years of research. **IEEE Signal Processing Magazine**, v. 33, n. 5, p. 59–69, 2016.
- KATZMAIER, D. **With a bullet to the head from Samsung, 3D TV is now deader than ever**. Disponível em: <<https://www.cnet.com/news/3d-tv-is-now-more-dead-than-ever/>>. Acesso em: 4 set. 2017.
- MATUSIK, W.; PFISTER, H. 3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes. **ACM SIGGRAPH 2004 Papers on - SIGGRAPH '04**, v. 23, n. 3, p. 814, 2004.
- PAUDYAL, P. et al. SMART: a Light Field image quality dataset Pradip. **Proceedings of the International Conference on Multimedia Systems - MMSys '16**, p. 1–6, 2016.
- SULLIVAN, G. J. et al. Overview of the high efficiency video coding (HEVC) standard. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 22, n. 12, p. 1649–1668, 2012.
- THE MATHWORKS INC. **MATLAB** - **MathWorksWww.Mathworks.Com/Products/Matlab**, 2016. Disponível em: <<http://www.mathworks.com/products/matlab/>>
- VIEIRA, A. et al. Data formats for high efficiency coding of Lytro-Illum light fields. **5th International Conference on Image Processing, Theory, Tools and Applications 2015, IPTA 2015**, p. 494–497, 2015.
- WALLACE, G. K. The JPEG Still Picture Compression Standard. **Communications of the ACM**, v. 34, n. 4, p. 31–44, 1991.