

Codificação de mapas de profundidade usando casamento de padrões multiescalas

Danillo Graziosi^{1,2,3}, Carla L. Pagliari⁵, Nuno M. M. Rodrigues^{1,2}, Eduardo A. B. da Silva³, Sérgio M. M. de Faria^{1,2}, Murilo B. de Carvalho⁴

Resumo— Neste artigo vamos apresentar o desempenho taxa-distorção do algoritmo de casamento de padrões multiescala, o MMP (*Multidimensional Multiscale Parser*), para codificar mapas de profundidade para o novo formato de vídeo 3D. Também é proposta uma arquitetura híbrida de compressão de imagens multivistas mais profundidade. A eficiência de codificação do MMP contribui para a preservação dos elementos de alta frequência das imagens, que são representados pelas bordas dos objetos nos mapas de profundidade. Isso significa que as vistas reconstruídas com mapas codificados com o MMP apresentam objetos com bordas bem preservadas, que são característica importante para os métodos de síntese de vistas virtuais. Aplicamos o H.264/MVC na codificação das vistas e o MMP para a codificação das profundidades. Com esta nova proposta, as vistas reconstruídas apresentam bordas mais bem definidas e ganhos de qualidade de até 1 dB, se compararmos com uma arquitetura totalmente baseada no H.264/MVC.

Palavras-Chave— Imagens multivistas, mapas de profundidade, síntese de vistas, casamento aproximado de padrões multiescala

Abstract— In this article, the rate-distortion performance for coding depth maps of the new 3D video format using the multiscale pattern matching algorithm, MMP (*Multidimensional Multiscale Parser*), is presented. In addition, an hybrid architecture for multiview image plus depth compression is proposed. MMP's coding efficiency contributes to the preservation of high frequency details of an image, which in the case of depth maps represent the objects' edges in a scene. Hence, the reconstructed views present well defined edges, which are an important characteristic for view synthesis algorithms. We apply the H.264/MVC for coding the multiple views, and MMP for coding the respective depth maps. With this new proposal, the reconstructed views present better defined borders and objective quality gains of up to 1 dB, if we compare our results with an architecture based only on the H.264/MVC.

Keywords— Multiview images, depth maps, view synthesis, multiscale approximate pattern matching

I. INTRODUÇÃO

É cada dia mais evidente que chegamos na era do 3D, com o lançamento quase semanal de novos filmes em formato 3D, vendas de televisões *3D-ready*, etc. A difusão de meios de criação, transmissão e consumo de material 3D, isto é, de toda

a cadeia de operação do 3D, vem fomentar o recente interesse neste área.

No entanto, ainda existem diversos desafios a serem superados para a aceitação em massa do 3D, não só nos cinemas, mas também na casa do usuário. Atualmente, a visualização de conteúdo 3D é na maioria dos casos feita com auxílio de óculos 3D. As vistas estereoscópicas são multiplexadas temporalmente, e os óculos 3D permitem a visualização de cada imagem apenas por um único olho, proporcionando uma visão binocular e criando a impressão de profundidade. No entanto, o efeito prolongado do uso dos óculos é causa comum de dores de cabeça e cansaço. Novos monitores auto-estereoscópicos oferecem uma solução para a visualização de vídeo 3D sem a necessidade de óculos, com a multiplexação espacial de diversas vistas e o uso de lentes difusoras. O uso de múltiplas vistas permite que o telespectador possa se movimentar livremente e ainda assim continuar com a impressão de profundidade. No entanto ainda não está definido o número de vistas necessárias para o melhor efeito 3D. Outro agravante é o fato dos produtores de conteúdo não estarem propensos a aumentar ainda mais os custos de gravação, para capturarem mais vistas.

Motivados pela crescente demanda deste setor, o MPEG está finalizando a criação de um padrão para o novo formato de vídeos 3D, que é capaz de atender às necessidades tanto dos produtores de conteúdo quanto dos fabricantes de *display*. Com o auxílio de técnicas de visão computacional, o novo padrão será capaz de, com um número reduzido de vistas capturadas, sintetizar todas as vistas necessárias ao telespectador, sem a necessidade de se capturar um número elevado de vistas. Dessa forma a produção de conteúdo 3D fica independente do seu consumo e o novo padrão será eficaz, tanto para os produtores quanto para os consumidores [1]. Para que o novo padrão seja compatível com os padrões vigentes de multivistas [2] e vídeo mais profundidade [3], [4], o formato previsto deverá ser uma conjugação de múltiplas vistas com seus respectivos mapas de profundidade, também conhecido como multivistas mais profundidade.

O uso de múltiplas vistas acarreta em um aumento significativo na quantidade de dados para a codificação. O padrão de codificação de múltiplas vistas, o H.264/MVC (ITU-T Rec H.264 (3/2009) Advanced video coding for Generic Audio-visual Services, Annex H), explora a correlação entre vistas, porém é conhecido que a taxa resultante cresce linearmente com o número de câmeras [5]. Através de algoritmos de síntese de vídeo pode-se gerar vistas intermediárias, com o auxílio das vistas laterais e de seus respectivos mapas de

¹Instituto de Telecomunicações, Portugal;

²ESTG, Instituto Politécnico de Leiria, Portugal;

³PEE/COPPE/DEL/EE, Univ. Fed. do Rio de Janeiro, Brazil;

⁴TET, Univ. Fed. Fluminense, Brazil.

⁵IME, Instituto Militar de Engenharia, Brazil.

E-mails: danillo@lps.ufrj.br, carla@ime.eb.br, eduardo@lps.ufrj.br, nuno.rodrigues@co.it.pt, sergio.faria@co.it.pt, murilo@telecom.uff.br. Este trabalho foi apoiado pela Fundação Para a Ciência e Tecnologia (FCT), sob o projeto PTDC/EEA-TEL/66462/2006

profundidade. Os algoritmos de renderização de vistas virtuais usando a profundidade, também conhecidos como algoritmos DIBR (*Depth Image Based Rendering*), aplicam os conceitos de geometria projetiva para obter a textura de uma vista qualquer a partir das vistas adjacentes. Além disso, os mapas de profundidade requerem menos bits para serem codificados, sendo portanto uma solução eficiente para o problema crítico de aumento de taxa em sistemas multivistas.

Os mapas de profundidade são imagens que têm características bem distintas das imagens de textura. Eles apresentam geralmente superfícies suaves, sem textura, com bordas bem definidas [6]. Além da sua construção particular, os efeitos dos artefatos introduzidos na codificação são diversos, e estão particularmente relacionados com o software de síntese e de codificação usado no processo. Muitas propostas para codificação de mapas de profundidade foram feitas, desde adaptações de codificadores comuns de imagens, como o JMVC [7], [8], até a proposta de novos codificadores específicos para mapas de profundidade [9], [10]. No entanto a maioria delas têm dificuldade em preservar o elemento crítico dos mapas de profundidade: as bordas dos objetos de uma cena. Os artefatos introduzidos em áreas de borda, isto é, com conteúdo de alta frequência, causam a mistura da textura de elementos no plano frontal com a textura do plano de fundo da imagem. Isso acarreta uma diluição das bordas dos objetos e o aparecimento de pixels estranhos à textura de componentes da imagem, vindos de outras partes da imagem [11].

O algoritmo baseado em casamento aproximado de padrões multiescalas, o MMP (*Multiscale Multidimensional Parser*), é conhecido por apresentar um bom desempenho para imagens com conteúdo de alta frequências, como imagens compostas, que misturam texto e imagens. Suas recentes evoluções permitiram ao MMP atingir desempenho estado-da-arte também para imagens suaves [12], [13]. Neste artigo vamos avaliar o MMP como um codificador para mapas de disparidade. Algumas publicações recentes [14], [15] mostram a eficácia do método para a codificação de mapas de profundidade. Aqui vamos mostrar resultados mais detalhados, e apresentar uma proposta de codificador híbrido, onde a textura é codificada com um codificador baseado em transformadas, enquanto que a profundidade é codificada com o MMP.

Este artigo está organizado da seguinte maneira. Na seção II detalhamos o processo de síntese de imagens a partir da textura e do mapa de profundidade. Em seguida, apresentamos sucintamente o MMP na seção III. Vamos discutir os resultados do MMP codificando mapas de disparidade na seção IV. A seguir, propomos uma arquitetura híbrida de codificação de imagens multivistas e seus respectivos mapas de profundidade na seção V. A seção VI apresenta as conclusões deste trabalho.

II. GEOMETRIA PROJETIVA E ALGORITMO DE SÍNTESE DE VISTAS ATRAVÉS DE MAPAS DE PROFUNDIDADE

O processo de síntese de uma vista “virtual”, dado um número de vistas de referências e seus respectivos mapas de profundidade, é conhecido como DIBR (*Depth Image Based Rendering*) e é um elemento chave do novo padrão de vídeo 3D. Baseado em geometria projetiva, onde as coordenadas de

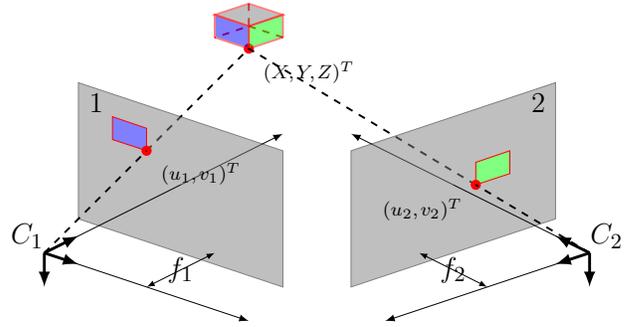


Fig. 1. Modelo de câmera *pinhole* e projeção em múltiplas vistas.

um elemento do espaço $(X, Y, Z)^T$ são descritas por quatro elementos $(X_1, X_2, X_3, X_4)^T$, também chamadas de coordenadas homogêneas, os objetos de uma cena são projetados em perspectiva no plano da imagem, usando o modelo de câmera *pinhole*. O centro da projeção em perspectiva é conhecido como centro ótico e a linha perpendicular ao plano da imagem que passa pelo centro ótico é o eixo ótico. O ponto onde cruza o eixo ótico com o plano da imagem é conhecido como o ponto principal. A Figura 1 mostra mais detalhes do modelo usado.

A projeção de um ponto no espaço $(X, Y, Z)^T$ para o plano da imagem, na posição $(u, v)^T$, pode ser determinada por um mapeamento projetivo (usando coordenadas homogêneas), dado por:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & \eta & O_x & 0 \\ 0 & \tau f & O_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = [K|0_3] P \quad (1)$$

onde f é a distância focal, $(O_x, O_y)^T$ corresponde a coordenada do ponto principal, η modela a razão de aspecto e τ o *skew* dos pixels. $P = (X, Y, Z, 1)^T$ corresponde ao ponto 3D definido pelas coordenadas homogêneas e o vetor com todos os elementos zerados corresponde ao vetor 0_3 . A matriz K também é conhecida como matriz intrínseca da câmera.

A Equação (1) assume que o posicionamento da câmera coincide com a origem do sistema de coordenadas, o que pode não necessariamente ocorrer. Portanto, dado a posição e orientação da câmera, operações de translação e rotação podem ser incorporadas à Equação (1). A rotação da câmera é modelada por uma matriz $R_{3 \times 3}$, e a sua posição é dada pelo vetor $C_{3 \times 1}$. A conjugação de ambos os elementos é conhecida como a matriz extrínseca da câmera. Logo, a Equação (1) fica da forma:

$$\begin{aligned} \lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= [K|0_3] \begin{bmatrix} R & 0 \\ 0_3^T & 1 \end{bmatrix} \begin{bmatrix} I_3 & -C \\ 0_3^T & 1 \end{bmatrix} P \\ &= KR \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - KRC \end{aligned} \quad (2)$$

Com a ajuda dos dados de profundidade, a imagem capturada por uma câmera de referência pode ser projetada para o mundo real 3D, e em seguida reprojetada para a vista de uma câmera alvo, um processo conhecido como *3D image warping* [16]. Ao princípio, o ponto 2D da vista de referência é projetado para o espaço 3D, que irá recair sobre um raio

$P(\lambda)$ conectando o ponto 2D ao centro da câmera $C_1 = (C_1^x, C_1^y, C_1^z)^T$, representado na Figura 1. A coordenada 3D será então dada por:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \underbrace{C + \lambda R^{-1} K^{-1} p}_{\text{raio } P(\lambda)} \quad (3)$$

Caso a profundidade Z seja conhecida, as coordenadas X e Y podem ser obtidas calculando o valor de λ através da relação $\lambda = \frac{Z - C_1^z}{\gamma}$, onde $(\alpha, \beta, \gamma)^T = R^{-1} K^{-1} (u_1, v_1, 1)^T$. Em seguida, a projeção do ponto para o plano da câmera alvo pode ser obtida usando a Equação (2) e as matrizes intrínscas e extrínscas da câmera em questão. Ao reprojeter uma imagem para uma câmera virtual, alguns pixels podem cair na mesma posição. Neste caso, um teste de visibilidade deve ser feito, e o pixel mais próximo da câmera virtual é que deverá ocupar o lugar. Note ainda que erros na operação de re-projeção podem ocorrer, tais como rachaduras, buracos na imagem e contornos fantasma. Tais erros ocorrem devido a problemas de arredondamento, amostragem ou bordas mal-definidas. No entanto, tais artefatos podem ser removidos através de técnicas de pós-processamento, como filtros de mediana para a profundidade ou filtros de remoção de ruído de borda, como o implementado no *software* de referência do MPEG, o VSRS (*View Synthesis Reference Software*, [17]).

III. CASAMENTO DE PADRÕES MULTIESCALA

O MMP é um algoritmo de codificação de imagem por blocos. Para cada bloco da imagem, uma predição é feita, e em seguida o resíduo é codificado usando padrões de um dicionário adaptativo. Similar ao padrão H.264/AVC, o MMP usa modos de predição direcionais com blocos de tamanho adaptativo. No caso do algoritmo apresentado em [13], os modos de predição eram similares aos do H.264/AVC, acrescido de um modo adaptativo, especialmente eficaz para codificar bordas com direção arbitrária. O casamento aproximado de palavras do dicionário com o resíduo da predição utiliza a transformação de escalas para fazer o casamento entre padrões de diferentes dimensões. Em seguida, os padrões codificados são concatenados para formar um novo padrão que será inserido no dicionário, adaptando o seu conteúdo às características da imagem que está sendo codificada. Maiores detalhes sobre o algoritmo podem ser encontrados em [12].

Como já mencionado anteriormente, as imagens de profundidade apresentam áreas sem textura alguma, com bordas em geral bem definidas, apresentando uma mistura de áreas de muito alta e muito baixa frequência espacial. Essa é uma característica vantajosa para o MMP, que consegue aprender estes padrões rapidamente e codificá-los eficientemente.

O MMP apresenta características semelhantes ao codificador de mapas de profundidade baseado em funções *platelet* apresentado em [9], conhecido como Platelet. Assim como este codificador, o MMP também utiliza uma segmentação flexível para codificar as bordas dos objetos por aproximações de blocos suaves. Porém ao aprender o novo padrão através da concatenação dos padrões suaves utilizados, o MMP é capaz de reutilizar esse padrão para codificar novas bordas que

venham a ocorrer durante a codificação. Desta forma o MMP evita a nova custosa segmentação da imagem, e portanto é mais eficiente que o Platelet, no sentido taxa-distorção. Outra vantagem do MMP em relação ao Platelet é o fato do MMP poder ser usado diretamente para a codificação das vistas, o que não parece ser o caso quando falamos do Platelet.

IV. AVALIAÇÃO DOS MAPAS DE DISPARIDADE CODIFICADOS COM O MMP

Uma análise da qualidade dos mapas de profundidade codificados deve ser feita. A Figura 2 apresenta diversas medidas objetivas para avaliar mapas de profundidade, desde valores comumente utilizados em análise de imagens, como o PSNR, assim como medidas propostas em outras publicações relacionadas à codificação de mapas de profundidade, como a distância de Hausdorff [9] (usada para medir distorção entre estruturas *mesh* 3D) ou a percentagem de erros da imagem [10]. Ainda sugerimos o uso do índice SSIM [18], que é uma medida qualitativa da imagem, baseada na preservação da estrutura dos objetos da imagem.

Em todos os gráficos da Figura 2 podemos ver que o MMP supera os outros concorrentes para todas as medidas em taxas médias e altas, até mesmo em compressão sem perdas. Já no caso das taxas baixas, o MMP não consegue adaptar o seu dicionário de maneira adequada, e apresenta problemas como um acentuado efeito de blocos. No entanto, as ferramentas usualmente utilizadas para medição de qualidade de imagens codificadas não são adequadas para avaliar os artefatos de codificação, uma vez que essas medidas não levam em conta os efeitos dos artefatos no processo de síntese. Avaliações de mapas de profundidade codificados devem ser feitas com base nas vistas reconstruídas a partir dos dados codificados. Em [15], mais resultados podem ser encontrados, além de vistas reconstruídas a partir da textura original e mapas codificados com o MMP, onde também mostramos a sua superioridade comparado com codificadores tradicionais como o H.264/AVC ou o JPEG2000 [19].

V. PROPOSTA DE UM CODIFICADOR HÍBRIDO E RESULTADOS EXPERIMENTAIS

Como já demonstrado, o MMP é particularmente eficaz para a codificação de mapas de profundidade, devido à preservação dos elementos de alta frequência. Vamos então propor uma solução de arquitetura híbrida, onde as texturas das vistas são codificadas com o padrão H.264/MVC para codificação de múltiplas vistas, enquanto que a profundidade é codificada com o MMP. Os resultados para o padrão H.264/MVC foram obtidas usando o *software* de referência JMVC, versão 8.4.

Para avaliar o desempenho da nossa proposta, usamos uma sequência com arranjo de câmera em arco (Breakdancers) e outra com arranjo paralelo (Book Arrival). Sintetizamos uma vista intermediária de acordo com o proposto pelo grupo MPEG, nos *Exploration Experiments* [20], utilizando a informação de textura e profundidade de duas vistas laterais e o *software* de síntese padrão do MPEG, o VSRS. Em seguida, sintetizamos a vista intermediária com a textura e o mapa original, e usamos essa imagem como referência para o

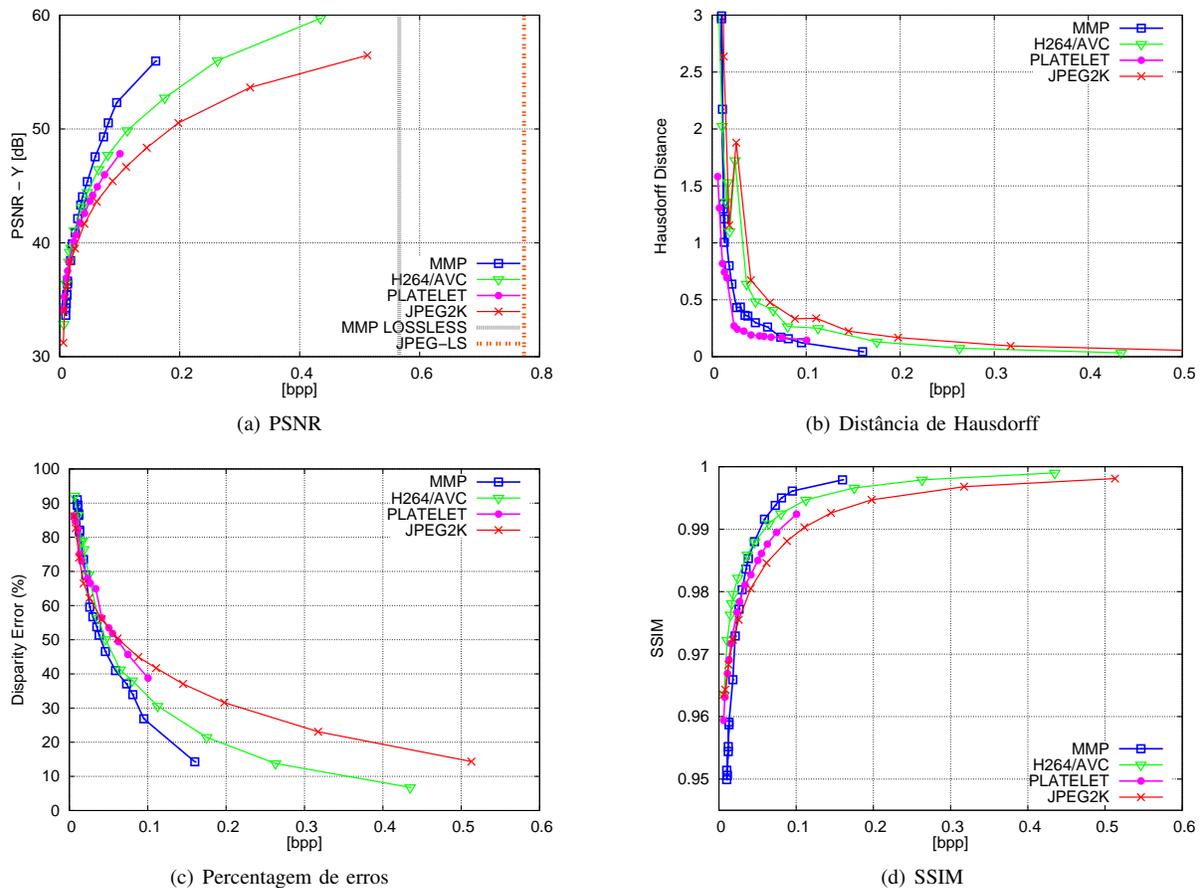


Fig. 2. Percentagem de erros e SSIM, medidas objetivas para avaliar a qualidade de mapas de profundidade da sequência Breakdancers (Frame 0, Camera 0).

cálculo do PSNR das imagens reconstruídas a partir de dados codificados. A configuração usada do software JMVC, para obter os resultados de codificação tanto da textura quanto da profundidade, foi a mesma indicada pelo MPEG [20].

Comparamos então o desempenho da nossa proposta de solução híbrida, com a solução inteiramente baseada no H.264/MVC. Note que os valores de taxa selecionados para avaliação também são semelhantes aos pontos ótimos taxa-distorção obtidos pelo MPEG, no âmbito dos experimentos para definição do padrão [20]. A Figura 3 mostra o desempenho taxa-distorção da vista intermediária, usando a textura e mapas codificados somente com o JMVC, e a nossa proposta. Ao utilizarmos o MMP no lugar do JMVC para a codificação dos mapas de profundidade, temos um ganho objetivo em todas as taxas avaliadas, para todas as sequências.

A Figura 4 mostra detalhes de uma imagem onde fica claro a vantagem do uso do MMP. As regiões de alta frequência, como o braço da cadeira, são mais bem preservados com os mapas codificados com o MMP, do que com mapas codificados com o JMVC.

VI. CONCLUSÕES

O novo formato de imagens 3D apresenta diversos novos desafios. O uso do algoritmo de síntese para a geração de vistas virtuais apresenta diversos problemas, e muito trabalho ainda deverá ser feito para aperfeiçoar esta técnica. O uso dos mapas de profundidade proporciona uma independência entre

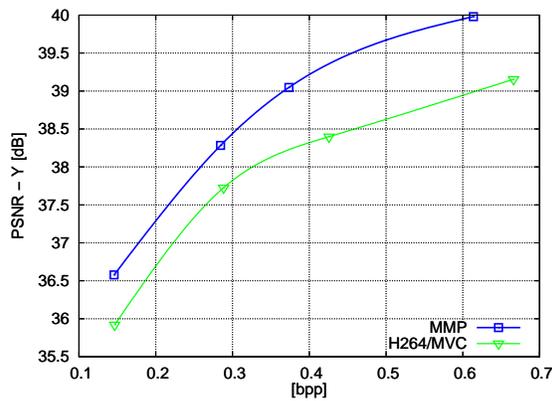
produção e visualização de conteúdo 3D, e é muito eficiente no que diz respeito à taxa necessária para a sua codificação. No entanto, a sua característica particular de mistura de áreas com alta e baixa frequência cria um desafio novo para os codificadores atuais.

Aqui apresentamos o MMP como uma alternativa para a codificação dos mapas de profundidade, e em seguida propomos uma solução híbrida usando o JMVC para a codificação das vistas e o MMP para a codificação dos mapas. Devido à sua propriedade de preservação das bordas, o MMP se mostrou particularmente eficaz nesse cenário, permitindo um aumento no desempenho taxa-distorção das vistas sintetizadas, notadamente por diminuir artefatos nas bordas dos objetos. Além de mostrar a eficácia do MMP, o resultado ressalta a importância da preservação das bordas nos mapas de profundidade.

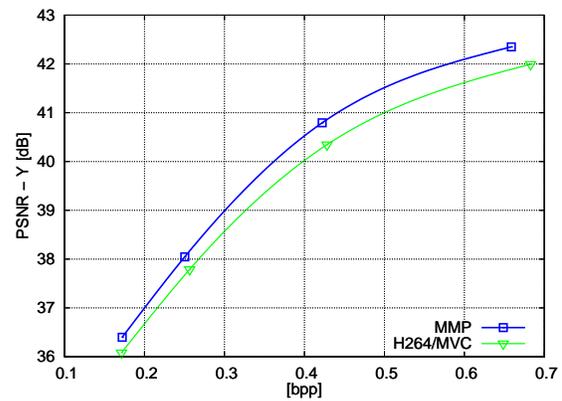
Como o MMP é flexível o suficiente para codificar textura e profundidade com o mesmo algoritmo, e em diversas taxas desejadas, uma proposta de trabalho futura é a realização de um codificador multivistas inteiramente baseado no MMP. Melhorias do método de síntese, baseadas em artefatos gerados pela codificação dos mapas de profundidade com o MMP, é outro tópico interessante de ser pesquisado.

REFERÊNCIAS

- [1], "ISO/IEC JTC1/SC29/WG11, Vision on 3D Coding, Doc. N10357," February 2009, Lausanne, Switzerland.



(a) Vista reconstruída 1 da sequência Breakdancers (frame 0)



(b) Vista reconstruída 9 da sequência Book Arrival (frame 0)

Fig. 3. Curvas taxa-distorção das vistas reconstruídas usando texturas e mapas de profundidade codificados.



(a) Detalhe da vista reconstruída usando o JMVC. A qualidade final é de 37.790 dB a 0.256 bits por pixel



(b) Detalhe da vista reconstruída usando o JMVC para codificar a textura e o MMP para codificar o mapa de profundidade. A qualidade final é de 38.046 dB a 0.250 bits por pixel.

Fig. 4. Análise subjetiva da reconstrução da vista da câmera 9 da sequência Book Arrival (frame 0), com textura e mapas de profundidade comprimidos.

[2] , "ITU-T, ISO/IEC JTC 1, Advanced video coding for generic audiovisual services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG4-AVC), Version 1: May 2003, Version 2: Jan. 2004, Version 3: Sep. 2004, Version 4: July 2005."

[3] , "ISO/IEC JTC1/SC29/WG11, Text of ISO/IEC DIS 23002-3 Representation of Auxiliary Video and Supplemental Information, Doc. N8768," January 2007, Marrakech, Morocco.

[4] , "ISO/IEC JTC1/SC29/WG11, Text of ISO/IEC 13818-1:2003/FDAM2 Carriage of Auxiliary Data, Doc. N8799," January 2007, Marrakech, Morocco.

[5] A. Smolic, K. Mueller, N. Stefanoski, J. Ostermann, A. Gotchev, G. Akar, G. Triantafyllidis, and A. Koz, "Coding algorithms for 3DTV: A survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1606–1621, November 2007.

[6] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47(1/2/3), pp. 7–42, April-June 2002.

[7] W.-S. Kim, A. Ortega, P. Lai, D. Tian, and C. Gomila, "Depth map distortion analysis for view rendering and depth coding," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '09*, November 2009, pp. 721–724.

[8] E. Ekmekcioglu, M. Mrak, S. Worrall, and A. Kondoz, "Utilisation of edge adaptive upsampling in compression of depth map videos for enhanced free-viewpoint rendering," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '09*, Cairo, Egypt, November 2009, pp. 733–736.

[9] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Mueller, P. H. N. de With, and T. Wiegand, "The effects of multiview depth video compression on multiview rendering," *Signal Processing: Image Communication*, vol. 24, no. 1-2, pp. 73–88, 2009.

[10] M. Sarkis and K. Diepold, "Depth map compression via compressed sensing," in *Proceedings of the IEEE International Conference on Image Processing, ICIP '09*, November 2009.

[11] K. Muller, P. Merkle, and T. Wiegand, "3-D video representation using depth maps," *Proceedings of the IEEE*, vol. PP, no. 99, pp. 1–14, 2010.

[12] N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria, and V. M. M. Silva, "On dictionary adaptation for recurrent pattern image coding," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1640–1653, September 2008.

[13] D. Graziosi, N. Rodrigues, E. A. B. da Silva, S. M. M. de Faria, and M. de Carvalho, "Codificação de imagens com predição adaptativa baseada no critério de mínimos quadrados," in *Proceedings of the Simpósio Brasileiro das Telecomunicações, SBrt '09*, Blumenau, SC, Brazil, September 2009.

[14] D. Graziosi, N. Rodrigues, C. Pagliari, S. de Faria, E. da Silva, and M. de Carvalho, "Compressing depth maps using multiscale recurrent pattern image coding," *Electronics Letters*, vol. 46, no. 5, pp. 340–341, 4 2010.

[15] D. Graziosi, N. Rodrigues, C. Pagliari, E. da Silva, S. de Faria, M. Perez, and M. de Carvalho, "Multiscale recurrent pattern matching approach for depth map coding," in *Proceedings of the Picture Coding Symposium, PCS '10*, Nagoya, Japan, December 2010, pp. 294–297.

[16] Y. Morvan, "Acquisition, compression and rendering of depth and texture for multi-view video," Ph.D. dissertation, Eindhoven University of Technology, Eindhoven, The Netherlands, 2009.

[17] , "ISO/IEC JTC1/SC29/WG11, Report on Experimental Framework for 3D Video Coding, Doc. N11631," October 2010, Guangzhou, China.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[19] D. S. Taubman and M. Marcelin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, 2nd ed. Norwell, Massachusetts: Kluwer Academic Publishers, 2001.

[20] , "ISO/IEC JTC1/SC29/WG11, Description of Exploration Experiments in 3D Video Coding, Doc. N11630," October 2010, Guangzhou, China.