

Research on Virtual View Rendering Based on DIBR

Sihan Wen*

Beihang University, Beijing, China

wsh_wency@163.com

doi:10.5729/lnes.2017.10.94

Keywords: Virtual view synthesis; DIBR; Holes filling; Image repairing.

Abstract. With the rapid development of video technology, interactive 3D video has been favored for its immersive effect, and virtual viewpoint rendering technology has played an increasingly important role. However, the traditional virtual viewpoint rendering technology still has many problems to be solved like holes, artifact and many other issues. Therefore, an optimized DIBR algorithm is proposed. First, we adopt the weighted sum method in the section of viewpoint fusion, and add the depth information in the process of determining the weight. Then, the depth image extrapolation method is used to repair the depth image, and the image restoration algorithm is used to fill and repair the color image. At last, we explore an reasonable evaluation to evaluate the quality of synthetic image, and compare our method with other existing methods.

基于深度图的虚拟视点绘制技术研究

温思寒*

北京航空航天大学, 北京, 中国

wsh_wency@163.com

doi:10.5729/lnes.2017.10.94

关键词: 虚拟视点绘制; DIBR算法; 空洞填充; 图像修复

中文摘要. 随着视频技术的快速发展, 交互式三维视频因其所营造的身临其境效果而受到人们的青睐, 其关键技术便依赖于虚拟视点绘制技术。然而, 传统的虚拟视点绘制技术存在空洞、伪影、无法实时绘制等诸多问题。因此本文提出一种改进的DIBR算法, 在图像融合部分, 对视点融合算法进行了改进, 采取加权求和法, 并在确定权重的过程中引入深度信息; 另外, 采用深度图外推法修复深度图像, 并采用改进的图像修复算法对彩色图像进行填充和修复; 最后, 通过合理的评价指标对合成图像进行评价, 并将本方法和其他现有方法进行对比分析。

1. 引言

近年来, 随着3D显示技术、视频处理技术[1-3]和计算机视觉算法的发展, 人们不再满足于追逐更高的清晰度和单一的观影模式, 而将目光投向了更加真实的可交互式3D视频。而虚

拟视点绘制技术作为交互式三维视频技术中的关键技术主要可以分为以下两类：一类是基于模型的虚拟视点绘制[4-5]，另一类是基于图像的虚拟视点绘制[6]。且目前随着深度学习的不断发展，其在图像分类[7-8]，视频处理[9-11]，机器视觉[12-13]及图像分割[14-16]等方向呈现出优异的表现。其中，由于卷积神经网络的应用使得立体匹配技术[17-19]获得的深度图精度不断提升。因此，研究基于深度图像的虚拟视点绘制技术(DIBR)[20-21]逐渐成为主流方向，不断有新的算法[22-23]被提出。本文在各算法基础上采用了一种双向DIBR算法，这种DIBR算法将虚拟视点左右两侧的参考视点同时向虚拟视点投影的图像进行视点融合，虽然需要利用更多已知参考信息，但能在相当程度上减少空洞数量。而且在视点融合算法中引入深度信息，对该算法进行了优化增加，提升了视点融合算法的可靠程度。

2. DIBR虚拟视点绘制技术

DIBR技术是输入一个或多个参考视点图像及其对应的深度图像，合成未知视点下图像的技术。具体实现过程可分为两步：第一步是反投影，即将参考视图中的每个像素点根据其深度值反投影到三维空间中；第二步则是其逆过程重投影，即将空间点位置再重投影到指定的虚拟视点平面上，从而建立起参考视点与虚拟视点像素位置之间的对应关系[24]，如图1所示。

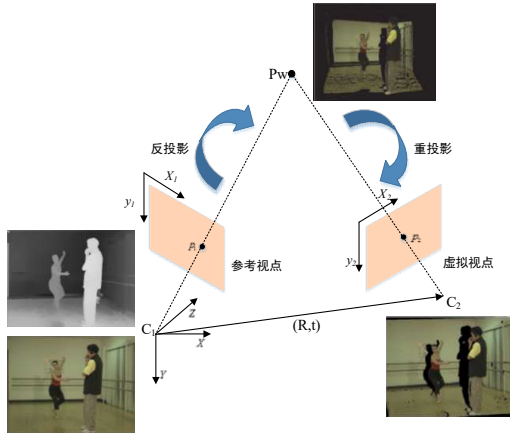


图1 虚拟视点绘制过程示意图

3. 改进的DIBR算法

3.1 3D warping算法

3D变换是DIBR技术的核心，我们首先由透视投影模型推出左摄像机和虚拟摄像机与三维空间点M之间的关系，将两式联立，并把其中的三维空间点M约去，就可以得到虚拟视点像素和参考视点像素的关系，如下：

$$s_V m_V = K_V R_V (K_L R_L)^{-1} (s_L m_L - K_L t_L) + K_V t_V \quad (1)$$

其中下标 L 和 V 分别表示参考视点位于左侧的摄像机和虚拟摄像机。并且我们假设左参考摄像机的坐标系与世界坐标系重合，就可以将上式简化为：

$$s_V m_V = K_V R_V K_L^{-1} Z_w m_L + K_V t_V \quad (2)$$

由于相机的内部参数 K 、平移矩阵 t 、旋转矩阵 R 和投影比例因子 s 都是已知的，所以我们可以根据参考视点图像 I_L 的像素位置信息 m_L 和它深度图所对应的景深信息 z_L 来获取虚拟视点图像中的像素位置信息，并将对应的颜色信息映射到与之对应的虚拟视点位置，合成虚拟

视点图像 I_v 。

但是，基于前向映射的3D warping算法有一个显著的缺点，参考视点中的某些像素首先会映射到虚拟视点的分数位置，再通过就近取整的方式将像素点的纹理映射到合成视点图像中，使得合成视点图像中有些位置没有纹理附着而产生细小的裂纹。

我们对其前向映射的方式进行了改进，采用反向3D warping算法[25]，从虚拟视点出发，首先对深度图进行正向3D变换，寻找到虚拟视点中的每个像素所对应的深度信息；其次根据深度信息将像素点反投影到三维空间内，计算得到其三维坐标；最后将三维坐标重投影到参考视点，此时利用周围的四个相邻区域的像素对它来进行插值，计算得到其纹理信息，避免裂纹的产生。

3.2 视点融合算法

如今虚拟绘制技术里最经典常用的融合策略是加权求和法。这种方法依据虚拟视点与参考视点之间的基线距离对左右虚拟视点图像设定相应的权重，输出是对二者进行加权求和的结果。但该算法只考虑了基线距离的单方面因素，依然可能会造成前景被背景像素遮挡，从而导致图像失真。

所以本文提出的改进算法将把距离和深度两方面因素都纳入到考虑范围内。我们知道，前景和背景深度值差异较大，且前景深度值远远大于背景，我们可以通过深度信息分辨出某一像素属于前景又或是背景，可以有效利用深度信息在很大程度上避免前景像素被背景像素遮挡，从而改善图像的质量。改进算法中，我们将基线距离和深度值的乘积一起作为视点融合的权重，这样当某像素位于背景时，它对最终融合结果的影响将大大减少。

经过我们的改进之后，公式如下所示：

$$I_v(u,v) = \frac{(1-\alpha) \cdot d_L(u,v) \cdot I_L(u,v) + \alpha \cdot d_R(u,v) \cdot I_R(u,v)}{(1-\alpha) \cdot d_L(u,v) + \alpha \cdot d_R(u,v)} \quad (3)$$

其中， $d_L(u,v)$ 和 $d_R(u,v)$ 分别表示左右两幅虚拟绘制图像中的像素坐标处 (u,v) 的深度值， I 和 t 的定义则与传统方法相同。

3.3 图像修复算法

3.3.1 深度图空洞填充

深度图中的空洞主要出现在横向区域，即水平方向。因而填充空洞时也主要从水平方向寻找背景信息填充空洞，即在深度值中寻找较小的像素值进行填充，公式如下：

$$D[p] = \min(D[m_l], D[m_r]) \quad (4)$$

式中， m_l 和 m_r 分别代表和空洞点处于同一条水平方向上的在空洞轮廓 $d(a,b)$ 上的前景和背景像素。

图2是对标准图像“Ballet”的深度图进行了处理，可以看到空洞区域被有效填充。

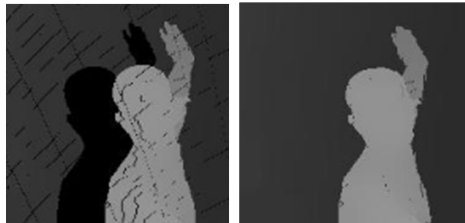


图2 空洞填充结果图

3.3.2 彩色图填充修复

彩色图修复部分，我们在Criminisi[27]的基础上对算法进行了改进。经典的彩色图填充算法包含以下两个方面：填充顺序估计和最佳匹配块准则。

在填充顺序估计部分，我们将优先级改进成如下表达式：

$$P(p) = e^{C(p)} \times D(p) + 3D(p) \quad (5)$$

引入了指数函数，在 $C(p)$ 减小速率不变的情况下，可以减慢优先级 $P(p)$ 的衰减速度，防止 $P(p)$ 趋于零。同时，加上数据量 $3D(p)$ 以后，可以更好的确保优先填充像素结构的正确性。

在最佳匹配块准则处，Criminisi算法寻找误差部分最小的最佳匹配块的定义如下所示：

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} d(\Psi_{\hat{p}}, \Psi_q) \quad (6)$$

$$d(\Psi_{\hat{p}}, \Psi_q) = SSD = \sum_{i=1}^m \sum_{j=1}^n (p_{ij} - q_{ij})^2 \quad (7)$$

在该算法的基础上，我们引入了深度信息，保证深度一致性，不但对纹理信息进行匹配，也对深度信息进行匹配，可以更加准确的完成图像修复的要求。我们的定义如下：

$$\Psi_{\hat{q}} = \arg \min_{\Psi_q \in \Phi} \{SSD_{RGB}(\Psi_{\hat{p}}, \Psi_q) + SSD_D(Z_{\hat{p}}, Z_q)\} \quad (8)$$

寻找到最佳匹配块并对目标匹配块进行填充之后，则将可信度项 $C(p)$ 更新，重复进行这一步骤，直到填充完毕。

4. 实验结果

图3给出了整个绘制过程中所得到的各阶段的虚拟视点图像，由图3(d)可看出，本文提出的算法对绘制过程中所产生的空洞进行了准确的填充，较好的绘制出了虚拟视点图像。

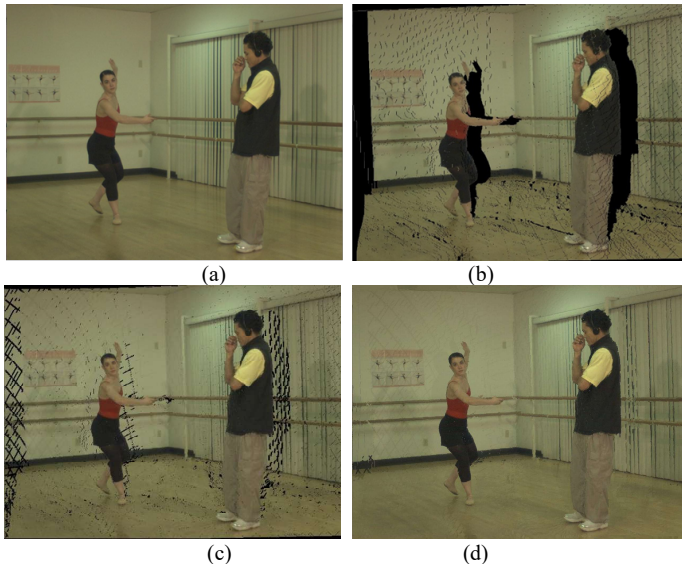


图3 (a) 左视初始图像 (b) 左视3Dwarping后图像 (c) 左右虚拟视点融合图像 (d) 图像修复后

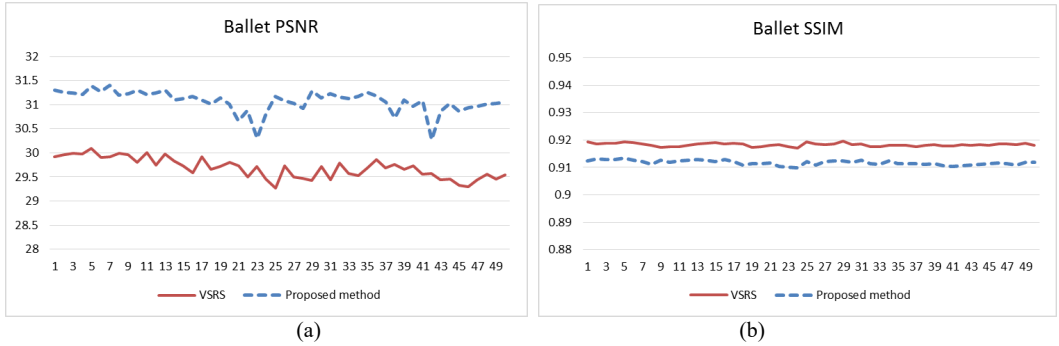


图4 本文提出算法与VSRS对比 (a) Ballet数据集PSNR计算结果对比 (b) Ballet数据集SSIM计算结果对比

此外，本文采用PSNR[27]以及SSIM[28]两个评价指标作为评价标准，对Ballet标准图像序列的前50帧分别进行了绘制，计算各帧的PSNR和SSIM指标，与VSRS[29]绘制图像的指标进行了对比，结果如图4所示，可以看出本文算法较VSRS算法SSIM提高了0.01~0.02，PSNR提高了0.5~2.37dB。

5. 结论

本文提出了一种新型的DIBR算法，采用双向DIBR的框架结构，在视点变换的过程中采用反向3D变换的算法，有效的减少细小裂纹的产生；处理大面积空洞时引入视点融合算法，并对该算法进行了改进，有效地填充和修复了空洞；在图像修复后处理算法中引入深度信息，对该算法进行了优化，对色彩进行了有效修复。最后，本文选取了SSIM、PSNR等指标进行质量评估，并且将结果与其它算法进行比较分析，验证了其有效性。

References

- [1] Shiping Zhu and Chunyan Zhang, A fast algorithm of intra prediction modes pruning for HEVC based on decision trees, Multimedia Tools and Applications, vol. 76, no. 20, pp. 21707-21728, 2017.
- [2] Shiping Zhu, Dongyu Zhao and Liyun Li, Adaptive fast intra prediction for high efficiency video coding, Multimedia tools and applications, vol. 75, no. 13, pp. 7559-7573, 2016.
- [3] Shiping Zhu, Yangshuan Hou, Zaikuo Wang and Kamel Belloulata, Fractal video sequences coding with region-based functionality, Applied Mathematical Modelling, vol. 36, no. 11, pp. 5633-5641, 2012.
- [4] Curless B, Levoy M, A volumetric method for building complex models from range images, Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. New Orleans, pp. 303-312, 1996.
- [5] Asai T, Kanbara M, Yokoya N, 3D modeling of outdoor environments by integrating omnidirectional range and color images, 3-D Digital Imaging and Modeling, Montreal, pp. 447-454, 2005.
- [6] Shum H, Kang S B, Review of image-based rendering techniques, Visual Communications and Image Processing, Perth, pp. 2-13, 2000.
- [7] He K, Zhang X, Ren S, et al, Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, pp. 770-778, 2016.

- [8] Kevin Jarrett, Koray Kavukcuoglu, Marc'Aurelio Ranzato, and Yann LeCun, What is the best multi-stage architecture for object recognition? *Computer Vision*, Kyoto, pp. 2146–2153, 2009.
- [9] Shiping Zhu and Ziyao Xu, Spatiotemporal visual saliency guided perceptual high efficiency video coding with neural network, *Neurocomputing*, 2017, DOI: 10.1016/j.neucom.2017.08.054
- [10] Shiping Zhu and Xianzi Zong, Fractal lossy hyperspectral image coding algorithm based on prediction, *IEEE Access*, vol. 5, pp. 21250-21257, 2017.
- [11] Shiping Zhu, Shupeizhang, and Chenhao Ran, An improved inter-frame prediction algorithm for video coding based on fractal and H.264, *IEEE Access*, vol. 5, pp. 18715-18724, 2017.
- [12] Raia Hadsell, Pierre Sermanet, Jan Ben, Ayse Erkan, Marco Scoffier, Koray Kavukcuoglu, Urs Muller, and Yann LeCun, Learning long-range vision for autonomous off-road driving. *Journal of Field Robotics*, vol. 26, no. 2, pp. 120–144, 2009.
- [13] Pierre Sermanet, Raia Hadsell, Marco Scoffier, Matt Grimes, Jan Ben, Ayse Erkan, Chris Crudele, Urs Miller, and Yann LeCun, A multirange architecture for collision-free off-road robot navigation. *Journal of Field Robotics*, vol. 26, no. 1, pp. 52–87, 2009.
- [14] Shiping Zhu, Xiaoyan Bu, and Qin Zhou, A novel edge preserving active contour model using guide filter and harmonic surface function for infrared image segmentation, *IEEE Access*, 2017, DOI: 10.1109/ACCESS.2017.2779278
- [15] Shiping Zhu, Qin Zhou and Ruidong Gao, A novel snake model using new multi-step decision model for complex image segmentation, *Computers and Electrical Engineering*, vol. 51, pp. 58-73, 2016.
- [16] Shiping Zhu and Ruidong Gao, A novel generalized gradient vector flow snake model using minimal surface and component-normalized method for medical image segmentation, *Biomedical Signal Processing and Control*, vol. 26, pp. 1-10, 2016.
- [17] Zbontar J and Lecun Y, Stereo matching by training a convolutional neural network to compare image patches, *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2287-2318, 2016.
- [18] Shiping Zhu and Lina Yan, Local stereo matching algorithm with efficient matching cost and adaptive guided image filter, *Visual Computer*, vol. 33, no. 9, pp. 1087-1102, 2017.
- [19] Shiping Zhu, Ruidong Gao and Zheng Li, Stereo matching algorithm with guided filter and modified dynamic programming, *Multimedia Tools and Applications*, vol. 76, no. 1, pp. 199-216, 2017.
- [20] Zhu C, Li S, Depth Image Based View Synthesis: New Insights and Perspectives on Hole Generation and Filling, *IEEE Transactions on Broadcasting*, vol. 62, no. 1, pp. 82-93, 2016.
- [21] Zarb T, Debono C J, Depth-based image processing for 3d video rendering applications, *Systems, Signals and Image Processing (IWSSIP)*, Piscataway, pp. 215-218, 2014.
- [22] Shiping Zhu and Yang Gao, Noncontact 3-D coordinates measurement of cross-cutting feature points on the surface of a large-scale workpiece based on the machine vision method, *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 7, pp. 1874-1887, 2010.
- [23] Shiping Zhu, Jiancheng Fang, Rui Zhou, Jianhui Zhao and Wenbo Yu, A new noncontact flatness measuring system of large 2-D flat workpiece, *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 12, pp. 2891-2904, 2008.

- [24] Wang L, Hou C, Lei J, et al, View generation with DIBR for 3D display system, *Multimedia Tools and Applications*, vol. 74, no. 21, pp. 9529-9545, 2015.
- [25] Cheng C, Liu J, Yuan H, et al, A DIBR method based on inverse mapping and depth-aided image inpainting, *Signal and Information Processing (ChinaSIP)*, Beijing, pp. 518-522, 2013.
- [26] Criminisi A, Perez P, Toyama K, Region filling and object removal by exemplar-based image inpainting, *IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society*, vol. 13, no. 9, pp. 1200-1212, 2004.
- [27] Huynh-Thu Q, Ghanbari M, Scope of validity of PSNR in image/video quality assessment, *Electronics letters*, vol. 44, no. 13, pp. 800-801, 2008.
- [28] Hore A, Ziou D, Image quality metrics: PSNR vs. SSIM, *International Conference on Pattern Recognition (ICPR)*, Cancun, pp. 2366-2369, 2010.
- [29] Gotfryd M, Wegner K, Doma'Nski M, et al. View synthesis software and assessment of its performance, *ISO/IEC JTC1/SC29/WG11, MPEG*, 2008: M15672.