

基于 DIBR 和图像修复的任意视点绘制

骆 凯^{1), 2)} 李东晓^{1), 2)} 冯雅美^{1), 2)} 张 明^{1), 2)}

¹⁾ (浙江大学信息与通信工程研究所, 杭州 310027)

²⁾ (浙江省综合信息网技术重点实验室, 杭州 310027)

摘 要 基于深度的图像绘制 (DIBR) 是高级视频应用的关键技术, 为提高视点变换的图像质量, 提出一种基于 DIBR 和图像修复的任意视点绘制方法。首先对深度图像进行形态学处理, 以减少视点变换产生的空洞, 平滑目标视点内部的物体轮廓; 利用视点变换方程生成目标视点; 对含有空洞的目标视点采用图像修复算法进行后处理, 设计了含有深度项的代价函数, 在深度的约束下进行纹理搜索, 将最佳匹配块填补到空洞; 在修复空洞的过程中采用亮度优先的策略以适应不同的色度采样格式。实验的主观效果对比和 PSNR 数据都显示本文算法比其他算法更为优越。

关键词 基于深度图像的绘制 视点变换 图像修复

中图分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2010)03-443-07

Arbitrary View Generation Based on DIBR and Depth-Aided Image Inpainting

LUO Kai^{1), 2)}, LI Dongxiao^{1), 2)}, FENG Yamei^{1), 2)}, ZHANG Ming^{1), 2)}

¹⁾ (Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027)

²⁾ (Zhejiang Provincial Key Laboratory of Information Network Technology, Hangzhou 310027)

Abstract Depth-image-based rendering (DIBR) is a key technology to the advanced video applications. In order to improve the image quality generated by view warping, this paper proposes an arbitrary view generation method based on DIBR and image inpainting. The depth image is processed by morphological operations to reduce the holes generated during image warping and to smooth the object contour in the target image. The view warping equation is utilized to generate an arbitrary view. The image inpainting method is employed to restore holes inside the warped view. Depth cue is incorporated into the cost function for texture estimation. The pixels in the most match block corresponds to the holes in the target image are copied to the holes. During the restoration process, luma-first strategy is applied to adapt to various color sampling format. Experimental results show the proposed algorithm has superiority in both subjective details and PSNR statistics.

Keywords depth image based rendering view warping image inpainting

0 引言

基于深度的图像绘制 (DIBR) 是各种高级视频

应用^[1-2], 例如 3 维电视、自由视点电视的关键技术之一。高级视频应用需要多个视点的图像, 基于 H. 264 的多视点视频编码标准 (MVC) 需要编码传输同接收终端相同数目的视点图像^[3]; 而利用单个视

基金项目: 国家自然科学基金项目 (60802013); 浙江省自然科学基金项目 (Y106574)

收稿日期: 2008-09-11; 改回日期: 2009-01-06

第一作者简介: 骆 凯 (1982—), 男。2009 年于浙江大学获信息与通信工程博士学位。主要研究方向为 3 维电视系统, 多视点编码技术。E-mail: lklion@126.com

点的图像及其对应的深度信息, DIBR 技术在理论上可以绘制出任意视点的图像。

DIBR 技术具有绘制速度快, 带宽需求小的优点。该技术的一个主要问题是绘制的目标视点图像内部存在暴露区域, 即空洞问题。空洞是由于在参考视点位置对场景的采样不充分造成的, 表现为其深度图像当中存在灰度值突变的边界。

为获得高质量的目标视点, DIBR 技术需要和数字图像处理等技术结合, 国内外许多学者对此进行了研究。已有的方法可以分为 3 类。针对参考视点采样点不充分, 第 1 类方法采用层次化的深度图技术^[4] (LDI)。LDI 图像的每个像素点包含了多层信息, 每层信息包含了色度和深度。这种特殊的信息存储方式使其能够有效地绘制目标视点的暴露区域, 但也带来了采样复杂, 信息量大的问题。针对深度图像当中的灰度值突变, 第 2 类方法结合采用预处理技术和后处理技术。一般来说, 通过高斯滤波^[5]来平滑深度图, 从而减少视点绘制产生的空洞; 通过均值滤波、水平方向像素内插或者外插来修复目标视点内遗留的空洞。当基线距离较大, 即目标视点和参考视点相距较远时, 需要采用大标准差、大窗口的高斯核函数, 由此会造成目标视点内物体的几何失真。Vázquez 在主观效果测试实验^[6]中发现简单后处理技术将不能够保持物体的纹理和结构。以上两类方法在视点绘制过程中采用单个参考视点, 第 3 类方法采用多个参考视点及其深度图^[7-8]。当两个参考视点分别绘制其相邻区域内部的同一目标视点时, 目标视点的暴露区域在理论上并不相交, 故而其纹理可以互补。该方法需要多个参考视点的深度图, 深度图的获取算法复杂, 且在融合两个目标的过程中, 由于深度图的不精确和对分像素点的重采样 (re-sampling), 目标视点内部的物体会产生虚假边界。

为获得高质量的目标视点, 提出一种基于 DIBR 和图像修复的视点绘制方法。其特征在于: 1) 利用了单个参考视点及其深度图像; 2) 对深度图像进行形态学处理以减少 DIBR 产生的空洞; 3) 采用图像修复 (inpainting) 技术来填补目标视点内部遗留的空洞; 4) 利用深度信息来优化图像修复算法中的纹理匹配函数, 优先考虑背景纹理; 5) 采用亮度优先的修复策略以适应不同的色度采样格式。算法的框

架如图 1 所示。

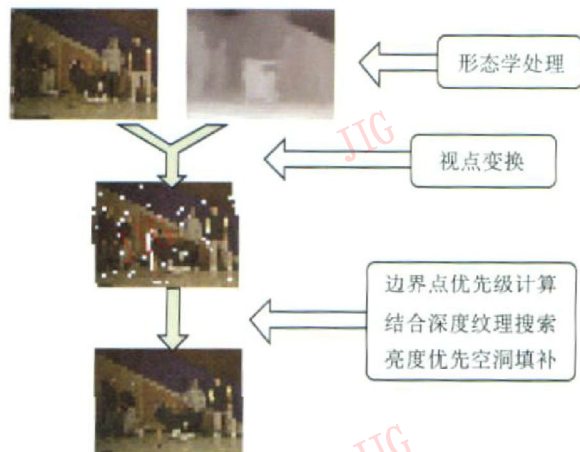


图 1 算法流程图

Fig 1 Algorithm flow

1 基于 DIBR 的目标视点绘制

DIBR 是利用参考视点及其对应的深度信息来绘制目标视点的过程。该过程可以看做从参考视点到目标视点的变换, 深度信息决定了该变换的尺度。本文对深度图像进行形态学上的预处理, 以减少目标视点内部空洞, 光滑物体轮廓。

1.1 深度图像的形态学处理

深度图像是一幅灰度图像, 其分辨率和参考图像一致, 其灰度值代表该点在世界坐标内的深度。灰度值在前景物体的轮廓处通常存在突变。

Zhang 采用非对称的高斯滤波^[5]、Wang 采用分段内插^[9]来平滑整个深度图, Chen^[10]、Darbo^[11]等采用边缘滤波来平滑深度突变区域。大范围地改变深度值尽管能减少甚至消除目标视点内部空洞, 但是会引入大面积的几何失真。

本文采用形态学上基本的腐蚀和膨胀算法, 采用 3×3 的单位掩模, 对深度图像进行一次开运算, 以断开狭窄的间断、消除细小的突出, 再进行一次闭运算, 来联结狭窄的间断、填补细小的空洞^[12]。

图 2(a) 显示了一条手臂的深度图, 限于当前的深度图生成技术, 该手臂的轮廓相当粗糙; 图 2(c) 显示了经过形态学处理之后的图像, 可见其轮廓已被平滑, 且背景和前景内部的深度值都未变; 图 2(b) 和图 2(d) 显示了分别由图 2(a) 和图 2(c) 绘制的目标图像, 可见图 2(d) 的空洞点较少, 且其手臂边缘更为光滑。

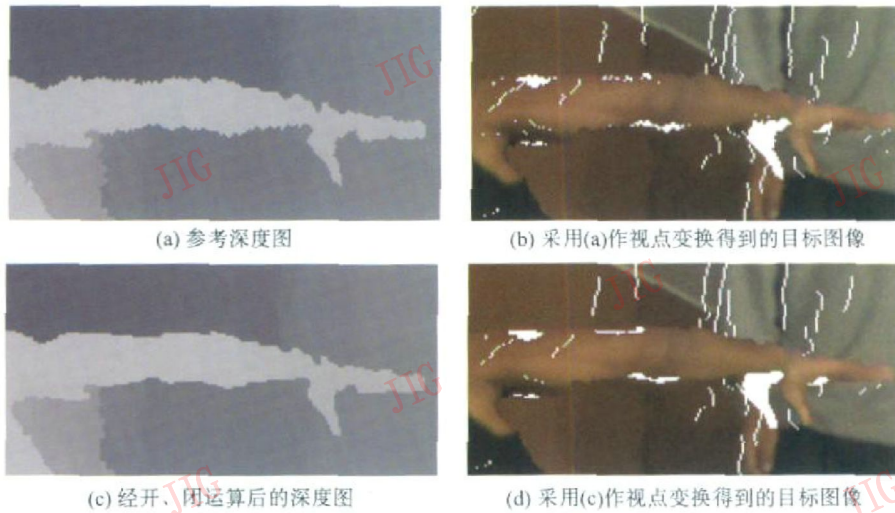


图 2 形态学处理效果

Fig 2 The impact of morphological operations

1.2 目标视点变换

视点变换可以分为 2 步: 对每一个参考像素点, 1) 根据其深度值, 将其从参考图像坐标空间投射到世界坐标空间; 2) 将空间点投射到目标视点相机坐标空间, 经归一化处理, 获得目标图像空间坐标。

Fehn 给出了一个适用于任意视点绘制的变换方程^[13]:

$$m = AP_nDM \quad (1)$$

式中, m 和 M 分别表示像素点的相机空间和世界空间坐标, A 和 D 分别表示相机的内参和外参矩阵, P_n 为归一化的透视投影矩阵。

具体操作中, 参考视点和目标视点的 A , P_n 和 D 已知, 令 $P = AP_nD$, 在步骤 1) 中, 已知参考点图像空间坐标 $m_{ref} = (u_{ref}, v_{ref})^T$ 和其在世界空间的深度 Z , 以及 P_{ref} , 通过解式 (1), 获得该点在世界空间的全部坐标 $M = (X, Y, Z)^T$; 在步骤 2) 中, 已知 M 和 P_{target} , 做矩阵乘法, 得到 $m_{target} = P_{target}M$ 。

图 3 给出了测试序列 breakdancers 的 4 号相机的第一帧图像和其深度图像, 图 3(c) 显示了经如上所述的变换过程, 投射到 5 号相机位置的目标图像, 白色表示暴露区域。

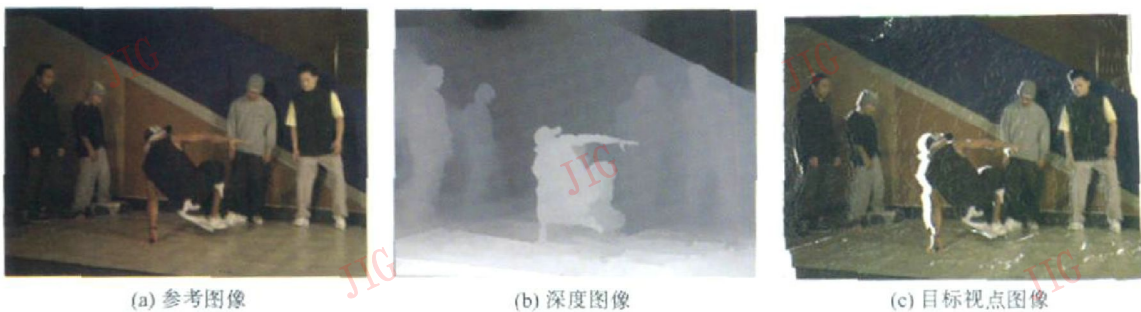


图 3 视点变换示例

Fig 3 Demonstration of view warping

2 基于图像修复的空洞填补

Tauber 认为 IBR 技术应当和图像修复^[14] (Inpainting) 技术结合, 以获得高质量的目标视点。图像修复是一类数字图像处理技术, 其目的在于填补图像中的缺损信息, 恢复图像的原本面貌, 使观察

者不会觉得图像有过损坏。

图像修复技术主要包括各向同(异)性扩散、整体变分、偏微分方程 (PDE)、光流场、纹理合成等方法。扩散方法对边缘和方向性纹理恢复效果较好, 但对大面积缺损区域修复效果不佳。PDE 方法对平滑区域的恢复效果较好, 但会模糊 (blur) 纹理细节。纹理合成技术则综合考虑了缺损区域边缘的结

构和纹理,可以获得一个相对较好的效果。

采用 Criminisi 提出的基于样本的修复方法^[15] (exemplar-based inpainting)来修复基于 DIBR 的视点变换所产生的空洞。根据 DIBR 技术和视频图像的特点,在纹理块匹配的代价方程中加入深度项,以区分空洞和其周围非空洞区域的深度层次;采用优先修复亮度分量的策略,使算法适应各种不同的色度采样格式;修改了原算法中的优先级计算公式,加强结构性纹理的传播。下面根据算法步骤的先后顺序,介绍以上 3 点。

2.1 优先级计算

Criminisi 算法^[15]的特点在于动态选择下一个待修复的空洞点,算法优先选择包含强等照度线 (isophote)的空洞块,使其向空洞内部区域延伸,因此可获得连结性较好的结构性纹理修复效果。

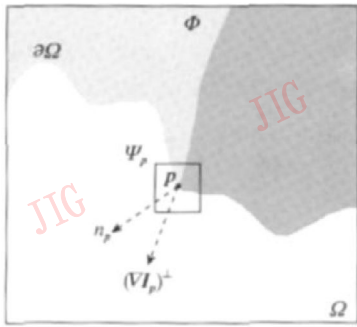


图 4 符号示意图
Fig 4 Notation diagram

图 4 给出了图像修复文献通常采用的符号示意图。用 I 表示图像整体, Ω 表示空洞区域, Φ 表示非空洞区域, $\partial\Omega$ 表示空洞边界。给出一个以 $\partial\Omega$ 上一点 p 为中心的块 Ψ_p , n_p 表示 p 点所在边界的法向, $(\nabla I_p)^\perp$ 表示其等照度线。Criminisi 设计了一个空洞边界点 p 的优先级计算公式:

$$P(p) = C(p)D(p) \quad (2)$$

式中, $C(p)$ 为该点的信任度,即非空洞点在 Ψ_p 中所占的百分比:

$$C(p) = \left(\sum_{q \in \Psi_p \cap \Phi} C(q) \right) / \text{area}(\Psi_p) \quad (3)$$

$D(p)$ 为该点的数值项,即法向量和等照度线的点积:

$$D(p) = |(\nabla I_p)^\perp \cdot n_p| / \alpha \quad (4)$$

$\alpha = 255$ (像素点用 8 位表示),表示对梯度 ∇I_p 的归一化运算。

设置初始图像非空洞区域点的信任度 $C(p)$ 为 1,式 (3) 中空洞边界点的 $C(p) \in (0, 1)$ 。每修复一个块,将改变空洞边界,需要更新受影响边界的优先

级 $P(p)$,其 $C(p)$ 更新为被修复块中心点的 $P(p)$ 。故随着修复的进行,在初始空洞内部, $C(p)$ 有变小的趋势。在式 (2) 的乘性条件下,新空洞边界点的 $C(p)$ 可能退化至零值,此时尽管点 p 的 $D(p)$ 值可能极大,但乘积项 $P(p)$ 为零,该点 p 仍然不被优先选择,延伸至该点的结构性纹理将不能继续保持。

将式 (2) 修改为加性:

$$P(p) = C(p) + D(p) \quad (5)$$

式 (5) 中, $C(p) \in [0, 1]$,将式 (4) 中对梯度的归一化修改为单位归一化,以使 $D(p)$ 的动态范围和 $C(p)$ 一致。

2.2 考虑深度信息的纹理搜索

逐点计算完空洞边界的优先级后,算法选择包含最高优先级的块 $\Psi_{p_{max}}$,搜索同该块最为匹配的纹理块 Ψ_q

$$\Psi_q = \arg \min_{\Psi_q \in \Phi} d(\Psi_{p_{max}}, \Psi_q) \quad (6)$$

以填补 $\Psi_{p_{max}}$ 中的空洞点。

同原算法^[15]相比,在式 (6) 的匹配函数 $d(\cdot)$ 中增加了深度项 $\lambda | \cdot |$,将 $d(\cdot)$ 设计为

$$d(\Psi_p, \Psi_q) = SAD(\Psi_p, \Psi_q) + \lambda | \text{ave}(\sum \text{depth}(p_1)) - \text{ave}(\sum \text{depth}(q_0)) | \quad (7)$$

式中, SAD 表示块匹配时的绝对差值和, Ψ_q 和 Ψ_p 中非空洞点的纹理愈相似, $SAD(\cdot)$ 愈小。

式 (7) 中 p_1 为 Ψ_p 中的非空洞点:

$$p_1 = p \in \Psi_p \cap \Phi, \forall p \in I \quad (8)$$

记 p_0 为 Ψ_p 中的空洞点:

$$p_0 = p \in \Psi_p \cap \Omega, \forall p \in I \quad (9)$$

式 (7) 中 q_0 为搜索块 Ψ_q 中的点,其位置对应于 Ψ_p 中的空洞点 p_0 的位置。

式 (7) 中 $\text{depth}(p)$ 表示该点的深度值, $\text{ave}(\cdot)$ 表示取均值, λ 为拉格朗日乘子:

$$\lambda = \begin{cases} \lambda_0 & \text{ave}(\sum \text{depth}(p_1)) \leq \text{ave}(\sum \text{depth}(q_0)) \\ \lambda_1 & \text{ave}(\sum \text{depth}(p_1)) > \text{ave}(\sum \text{depth}(q_0)) \end{cases} \quad (10)$$

其中 $\lambda_1 > \lambda_0$ 。

式 (7) 中的 $SAD(\cdot)$ 倾向于选择同 Ψ_p 中非空洞点纹理相近的点来填补其空洞,深度项 λ 倾向于选择比 Ψ_p 中非空洞点深度更大的纹理来填补其空洞。深度图不准确可能造成视点变换中,参考视点中的纹理被投射到目标视点中的错误位置。仅考虑 $SAD(\cdot)$ 项,在修复中可能造成错误纹理的生长,当错误

纹理属于前景区域时,对主观效果的影响较大,该现象可以在图 6 中观察到; λ_1 优先采用背景(深度较大)来修复空洞,可以减弱上述错误纹理的生长。 $SAD(\cdot)$ 和 λ_1 的共同作用将使匹配函数达到动态平衡。

将纹理搜索放在参考视点中进行,搜索窗口的中心设置在 Ψ_p 中梯度值最大点的投影点。参考视点图像具有目的视点图像的所有信息,且没有空洞,因此在搜索窗口大小一定时,可以增加有效的候选纹理块数目。

2.3 亮度优先策略

针对视频图像存在多种色度采样格式,而通常为 $Y:U:V=4:2:0$ 的情况,在上述的纹理搜索过程中,式(7)中的 SAD 项仅考虑亮度(Y)分量。

从搜索窗口内选出最佳匹配块 $\Psi\hat{q}$ 后,将 q_0 点填补到 p_0 点,同时记录 q_0 点的位置,以便在目的视点的亮度空间空洞点修复完成之后,进行对色度空间空洞点的修复。

修复完一个块,将改变空洞边界,因此重复第 2.1~2.3 节的上述操作,直至亮度空间空洞点全部修复。根据色度点的采样位置,查询对应的亮度点的投影点 q_0 ,将参考视点在 q_0 位置的色度空间的样点值填补到待修复的色度点位置,完成目的视点图像色度空间空洞点的修复。

上述亮度优先的修复策略适用于各种色度采样格式,对于 $Y:U:V=4:4:4$ 的情况,可简单修改式(7)的 SAD 项,加入色度分量,以获得更相似的匹配块。

3 实验与讨论

为验证本文算法效果,采用微软研究院^[16]提供的视频测试序列 *breakdancers* 进行实验。该序列包含 8 个两两水平相距约 20 cm 的视点,其基于点(point based)的深度图像同彩色图像的分辨率一致,为 1024×768 。在视点变换中,以 4 号相机(该相机处于中心位置)为参考视点,为方便 PSNR 比较,以 5 号相机为目标视点。第 1.2 节图 3 给出了参考视点的第一帧彩色图像和深度图像。实验的环境为 P4 3 0 G, 1 024 M 内存。

在本文算法实现中,采用 3×3 的单位掩模对深度图作形态学处理,图像修复的块大小设置为 9×9 ,搜索窗口设置为 15×15 ,拉格朗日乘子 λ_1 和 λ_2 分别设置为 10 和 0。上述参数对实验序列可以达到

比较好的效果。

对 *Crim in is* 算法^[15]实现,采用和本文算法同样的参数设置,其没有的参数除外。

选择较为典型的 Zhang 的非对称高斯滤波^[15]作比较。因测试序列的分辨率和基线距离都较大,故采用强参数设置。2 维高斯核的水平和垂直标准差分别设置为 20 和 60 对应的窗口分别为 61 和 193。

图 5 给出了目标视点的真实图像和上述 3 种算法的视点绘制结果。

将图 5(b)~图 5(d)同图 5(a)对比,可见图 5(b)在物体边界处存在伪像,图 5(c)存在物体轮廓、位置和线条变形。图 6 给出了局部图像对比。

图 6 中,图 6(d)中左手臂的边缘处存在类似前景的伪像;图 6(e)中左手臂变粗,小腿变窄,右手臂在红圈处有向左的位移。由前文所述,本文算法没有大面积改变深度值,故绘制的视点几乎不存在失真;由于在纹理搜索时综合考虑了匹配块的纹理和深度,式(7)中的深度项 λ_1 优先选择背景来修补空洞,故在物体边缘抑制了伪像的生长。注意图 6(a)中的红圈处存在较明显的深度错误,认为图 6(f)中仍然存在少量伪像,部分是由于深度图的不精确造成,该现象在多参考视点的算法^[8]中也存在。

本文算法能够克服 *Crim in is* 算法^[15]的不足和 Zhang 算法^[15]的弱点,目标视点的主观效果更好。但实验中也发现本算法对于包含自然纹理的空洞修复效果较好,对于包含结构性纹理较多的空洞,例如多个垂直性的线条,将不能保持其结构性。

图 7 给出了非洞区域的 PSNR 比较。非洞区域定义为根据参考视点和未经处理的深度图像,经视点变换得到的目标视点内的非洞点的坐标集合。根据上述定义,分别将图 5(c)和图 5(d)的非洞区域和图 3(c)做 PSNR 计算,计算结果绘于图 7。*Crim in is* 算法没有预处理深度图,故其非洞区域像素值和图 3(c)完全一致。图 7 显示本文算法中对深度图进行形态学处理对目标视点的影响甚微。

为衡量目标视点和真实图像的整体差距,分别将图 6(d)、(e)、(f)和图 6(c)做 PSNR 计算,得到图 8。图 8 显示本文算法绘制的目标视点在整体质量上也较其他算法更好。注意到其 PSNR 的绝对值不高,这是因为在理论上,视点变换得到的目标图像应该和真实图像一致,实际上由于深度图的准确性、重采样等因素的影响,目标图像和实际图像存在差别,且恢复的空洞信息并不是其真实信息。



图 5 真实图像和不同算法的视点绘制结果

Fig. 5 Real image and view warping results of different algorithms



图 6 局部图像对比

Fig. 6 The comparison of segments of images

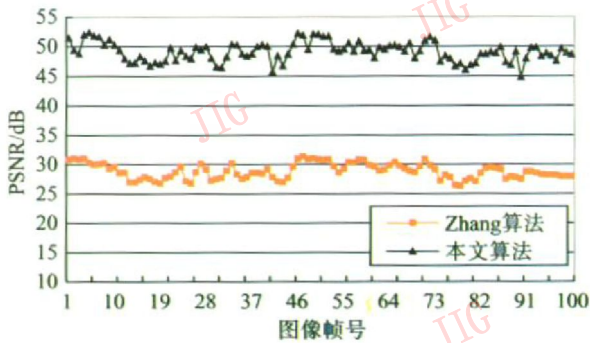


图 7 非洞区域 PSNR 对比

Fig. 7 Non-hole region PSNR comparisons

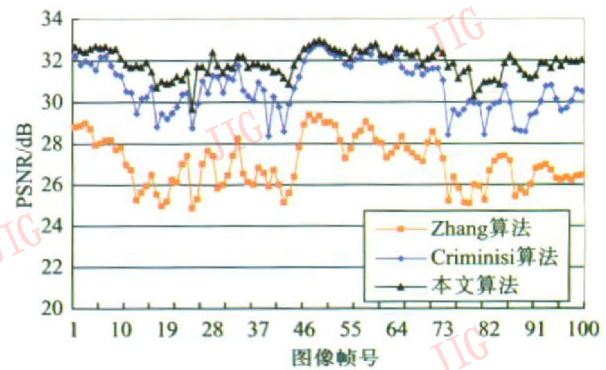


图 8 整体图像 PSNR 对比

Fig. 8 Whole image PSNR comparisons

4 结 论

利用 DIBR 技术, 由单个参考视点经视点变换可得到任意位置的目标视点; 对深度图像进行形态学处理以减少视点变换产生的空洞, 平滑目标视点内的物体轮廓; 对含有空洞的目标视点采用结合深度信息的图像修复算法进行修补。与其他算法相比, 绘制的目标视点不变形, 空洞的纹理不模糊, 空洞边缘的结构性能加以延伸, 由于深度不匹配而产生的伪像大大减少。下一步的工作方向在于如何进一步减少直至消除目标视点内前景物体边缘的伪像, 增强算法处理包含规整性结构的空洞的能力。

参考文献 (References)

- [1] Fehn C, de la Barre R, Pastoor S, et al. Interactive 3-dtv-concepts and key technologies [J]. *Proceedings of the IEEE*, 2006, 94(3): 524-538
- [2] Kauff P, Atzpadin N, Fehn C, et al. Depth map creation and image-based rendering for advanced 3dtv services providing interoperability and scalability [J]. *Signal Processing-Image Communication*, 2007, 22(2): 217-234
- [3] Meikle P, Smolic A, Muller K, et al. Efficient prediction structures for multiview video coding [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2007, 17(11): 1461-1473
- [4] Shade J, Gortler S, He L W, et al. Layered depth images [C] // *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: ACM Press, 1998, 231-242
- [5] Zhang L, Tan W J. Stereoscopic image generation based on depth images for 3DTV [J]. *IEEE Transactions on Broadcasting*, 2005, 51(2): 191-199
- [6] Vazquez C, Tan W J, Speranza F. Stereoscopic imaging: Filling disoccluded areas in depth image-based rendering [C] // *Proceedings of Three-Dimensional TV, Video and Display V*. Boston, MA, USA: SPIE, 2006, 63920D-12
- [7] Narayanan P J, Sashi Kumar P, Sireesh Reddy K. Depth+ texture representation for image based rendering [C] // *Proceedings of the 4th Indian Conference on Computer Vision, Graphics and Image Processing*. Kolkata, India: Allied Publisher, 2004, 113-118
- [8] Liu Zhanwei, An Ping, Liu Suxing, et al. Arbitrary view generation based on dibr and image merge [J]. *Journal of Image and Graphics*, 2007, 12(10): 1696-1700 [刘占伟, 安平, 刘苏醒, 等. 基于 dibr 和图像融合的任意视点绘制 [J]. *中国图象图形学报*, 2007, 12(10): 1696-1700.]
- [9] Wang W, Huo L, Zeng W, et al. Depth image segmentation for improved virtual view image quality in 3-DTV [C] // *Proceedings of Intelligent Signal Processing and Communication Systems (ISPACS 2007)*. Xiamen, China: IEEE Press, 2007, 300-303
- [10] Chen W Y, Chang Y L, Lin S F, et al. Efficient depth image based rendering with edge dependent depth filter and interpolation [C] // *Proceedings of IEEE International Conference on Multimedia and Expo (ICME 2005)*. Amsterdam, Netherlands: IEEE, 2005, 1314-1317
- [11] Darbo I, Tillier C, Pesquet-Popescu B. Distance dependent depth filtering in 3d warping for 3DTV [C] // *Proceedings of Multimedia Signal Processing (MMSP 2007)*. Chania, Crete, Greece: IEEE Press, 2007, 312-315
- [12] Gonzalez R C, Woods R E. *Digital Image Processing*. Second Edition [M]. Upper Saddle River, New Jersey, USA: Prentice Hall Inc, 2002
- [13] Fehn C. Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV [J]. *Proceedings of SPIE*, 2004, 5291: 93-104
- [14] Tauber Z, Li Z N, Drew M S. Review and preview: Disocclusion by inpainting for image-based rendering [J]. *IEEE Transactions on Systems, Man, and Cybernetics Part C-Applications and Reviews*, 2007, 37(4): 527-540
- [15] Criminisi A, Perez P, Toyama K. Region filling and object removal by exemplar-based image inpainting [J]. *IEEE Transactions on Image Processing*, 2004, 13(9): 1200-1212
- [16] Zinick C L, Sing Bing K, Matthew U, et al. High-quality video view interpolation using a layered representation [J]. *ACM Transactions on Graphics*, 2004, 23(3): 600-608