

基于深度图像分割的虚拟视点绘制

娄达平, 王晓东, 富显祖, 章联军

(宁波大学 信息科学与工程学院, 浙江 宁波 315000)

摘要: 针对虚拟视点绘制中出现的伪轮廓和空洞问题, 提出一种改进的虚拟视点绘制算法。根据深度信息标记出深度图像的不可信区域, 通过多阈值图像分割技术对图像进行分割处理, 利用三维图像变换生成虚拟视点图像, 并采用局部中值滤波消除目标图像中存在的细小空洞。对绘制出的两幅图像进行融合处理, 并应用图像修复法填充剩余的空洞。实验结果表明, 该算法能获得的虚拟视点图像质量较高, 同时与 MPEG 提供的虚拟视点绘制标准参考软件相比, 生成的虚拟视点 PSNR 值提高了 0.74 dB。

关键词: 虚拟视点; 深度图像; 图像分割; 三维图像变换; 局部中值滤波; 图像修复

中文引用格式: 娄达平, 王晓东, 富显祖, 等. 基于深度图像分割的虚拟视点绘制[J]. 计算机工程, 2016, 42(10): 12-19.

英文引用格式: Lou Daping, Wang Xiaodong, Fu Xianzu, et al. Virtual View Point Rendering Based on Range Image Segmentation[J]. Computer Engineering, 2016, 42(10): 12-19.

Virtual View Point Rendering Based on Range Image Segmentation

LOU Daping, WANG Xiaodong, FU Xianzu, ZHANG Lianjun

(School of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315000, China)

[Abstract] This paper proposes an improved virtual view point rendering algorithm to solve the problems of ghosting and holes in the virtual view point rendering. This method marks the unreliable regions of depth map based on the depth information. Through multi-threshold image segmentation, the image is divided into different parts. A virtual view point is generated by making use of 3D image warping. In order to deal with the small cracks, local median filtering is used to obtain a single-view point image. Two generated virtual images are blended together to eliminate the disoccluded holes in the desired view. An inpainting method is utilized to eliminate the remaining holes. Experimental results show that high quality virtual view points are generated, the Peak Signal to Noise Ratio (PSNR) of virtual view point can be 0.74 dB higher than that standard reference software of virtual view point rendering.

[Key words] virtual view point; range image; image segmentation; 3D image transformation; local median filtering; image inpainting

DOI: 10.3969/j.issn.1000-3428.2016.10.003

1 概述

基于深度图像的绘制 (Depth Image-based Rendering, DIBR) 技术在三维视频和自由视点电视中是一项关键技术^[1-4]。在理论上, 使用单个视点的彩色图像及其对应的深度图像, 应用 DIBR 技术可以绘制出任意给定视点的图像并且具有绘制速度快、带宽小等优点。然而, 在实际采用该技术的自由视点视频框架中存在的主要问题

是, 绘制后的虚拟视点图像通常存在空洞, 即空洞问题。根据空洞产生的原因, 大致可以分为两类: (1) 由参考视点采样点不充分引起的空洞, 这类空洞通常比较大。(2) 由深度图像中同一物体表面的灰度值不连续引起的空洞, 这类空洞通常表现为细小的裂纹。

为提高目标视点视图的质量, 国内外许多学者对此进行了研究。已有方法可以分为 3 类^[5]: (1) 采用深度图像分层 (Layered Depth Image, LDI) 技术^[6], 在构造 LDI 时, 使得每个像素点既包含可见物体的彩色

基金项目: 国家自然科学基金资助重点项目 (U1301257); 国家科技支撑计划基金资助项目 (2012BAH67F01); 浙江省教育厅科研计划基金资助项目 (Y201327703); 浙江省科技厅创新团队自主设计基金资助项目 (2012R10009-08); 宁波市科技创新团队研究计划基金资助项目 (2011B81002)。

作者简介: 娄达平 (1991—), 男, 硕士研究生, 主研方向为虚拟现实、多媒体信号处理; 王晓东, 副教授; 富显祖, 硕士研究生; 章联军, 工程师、硕士。

收稿日期: 2015-12-21 **修回日期:** 2016-01-27 **E-mail:** wangxiaodong@nbu.edu.cn

信息和深度信息,又保存了遮挡物体的信息。通过这种方式能够有效地避免目标视图中大块空洞现象,但同时造成了采样过程繁琐和信息存储量大的问题。因此,该方法的计算复杂度较高。(2)将深度图像预处理技术与彩色图像后处理技术相结合达到平滑深度图像的目的。通常做法是,通过高斯滤波器来平滑深度图像,从而减少绘制视图中的空洞。通过均值滤波和像素插值的方法来修复目标视图中残留的空洞。文献[7]针对深度图像边缘存在突变的问题,提出采用高斯滤波器对深度图像进行平滑滤波的方案,在进行虚拟视点绘制前先对深度图像进行滤波处理,使得整体图像变得平滑,从而在一定程度上减少了空洞出现的数量。由于文献[7]是对整幅深度图像进行滤波处理,比较容易造成几何失真和时间复杂度高的问题,因此文献[8]选用高斯双边非对称滤波器对深度图像进行处理,根据彩色图像信息自适应地对深度图像边缘进行滤波处理,既节省时间又能够减轻几何失真的问题。类似还有文献[9],利用 A-Trous 小波滤波器有效保留对象边缘深度信息。此类方法虽然能在一定程度上减少空洞甚至消除空洞,但只适用于目标视点和参考视点距离较小时,即小基线距离的情况;否则需要使用强参数设置的高斯核函数,但会引起目标视图中对象的几何失真。(3)将图像修复方法引入到空洞填补中。文献[10]将基于样本图像修复的方法运用于空洞填补问题,利用彩色图像梯度信息确定当前空洞点的修复优先级,对优先级最高的点进行整幅图像搜索匹配及空洞填补,以达到修复空洞的目的。为提高修复的准确性,文献[11-12]在此基础上结合深度图像信息,文献[11]以先修复深度图像空洞再修复彩色图像空洞的顺序对图像进行修复,文献[12]利用深度图像的梯度信息,对修复优先级的判定公式进行改进,更加精确地对目标视图中的空洞进行修复。由于视频信息存在时域冗余,而在相邻图像中可能存在更加准确的匹配信息,据此文献[13]指出利用时域上的图像信息对空洞进行填补,将最佳匹配的搜索范围进行扩大,显著提高空洞修复的质量。此类方法能够保持目标视图的纹理和结构信息,图像质量也较高,但是对象边界处会发生模糊和出现伪轮廓,并且算法复杂度较高。

针对上述方法存在的伪轮廓和空洞问题,本文提出一种基于深度图像分割的虚拟视点绘制算法。该算法主要包含三部分内容:(1)对深度图像进行不可信区域标记来消除 DIBR 产生的伪轮廓。(2)利用图像分割技术对深度图像进行分割。(3)利用局部中值滤波技术来修补目标视图内的空洞。

2 虚拟视点绘制

在本文提出的虚拟视点绘制算法中,选择 2 个参考视点的彩色图像及其对应的深度图像,利用 DIBR 技术绘制虚拟视点图像。深度信息不仅用于

三维图像变换,而且用于标记不可信区域和分割图像。算法流程如图 1 所示,深度图像的不可信区域标记和多阈值分割是本文算法的重点。

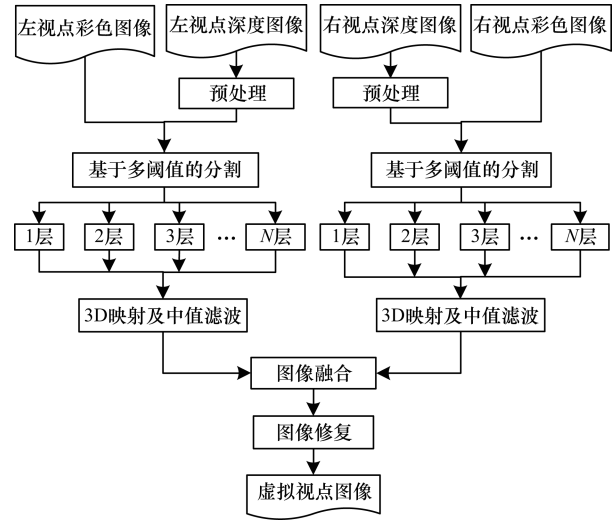


图1 本文算法流程

2.1 三维图像变换

三维图像变换是虚拟视点绘制技术中的重要步骤,变换方式可分为 2 种:3D 映射和 1D 映射^[14]。3D 映射是根据深度图像、参考图像和相机参数,将参考图像中的像素点变换到世界坐标系下,将世界坐标系的点变换到目标视图上。这种方式可以绘制出任意视点下的虚拟视点视图。1D 映射被称为视点插值,不同之处在于直接将参考图像的像素点水平变换到目标视图中同一行的对应位置,使用平行相机采集的图像,适用于目标视点视角位于基线上的虚拟视点视图。本文主要研究的是通过汇聚相机采集的视频图像,因此,通过 3D 映射进行虚拟视点绘制,下面对 3D 映射方式进行具体介绍。

(1)将 2D 平面上的像素点变换到世界坐标系上。令 (x, y) 为图像上某一个像素点的坐标位置, (X, Y, Z) 是世界坐标系上的位置, Z 是该像素点的实际深度值,变换公式如下:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R}_{\text{ref}}^{-1} \mathbf{K}_{\text{ref}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} z - \mathbf{R}_{\text{ref}}^{-1} \mathbf{t}_{\text{ref}} \quad (1)$$

在式(1)中, $\mathbf{K}_{\text{ref}} = \begin{bmatrix} f & 0 & u_0 \\ dx & 0 & \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ 是相机内矩阵;

$\mathbf{R}_{\text{ref}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ 是相机旋转矩阵; $\mathbf{t}_{\text{ref}} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$ 是相

机平移向量; $z = z_{\text{far}} + D(x, y) \times \frac{z_{\text{near}} - z_{\text{far}}}{255}$ 为深度图像

的灰度值; z_{far} 表示物体实际的最远距离; z_{near} 表示物体实际的最近距离。

(2)将世界坐标系上的像素点变换到目标视图的坐标上,变换公式如下:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \mathbf{K}_{\text{vir}} \mathbf{R}_{\text{vir}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \mathbf{K}_{\text{vir}} \mathbf{t}_{\text{vir}} \quad (2)$$

其中, \mathbf{K}_{vir} , \mathbf{R}_{vir} , \mathbf{t}_{vir} 分别为目标相机内矩阵、旋转矩阵和平移向量。坐标 $[u, v, w]^T$ 通过式(3)和式(4)的计算可以得到原像素点在目标视图上的坐标 (U, V) 。

$$U = \text{round}(u/w) \quad (3)$$

$$V = \text{round}(v/w) \quad (4)$$

图2(a)和图2(b)分别给出了 Ballet 序列的3号相机的第1帧图像及其深度图像,图2(c)是经过3D映射,投影到4号相机的虚拟图像,其中,黑色为出现空洞的部分。



(a)参考图像



(b)深度图像



(c)虚拟视点图像

图2 虚拟视点绘制示意图

2.2 深度图像预处理

在对深度图像的预处理中,通常使用各类滤波器进行处理,从而达到减少空洞和消除伪轮廓的目的。然而,这类方法只适用于于基线距离较小的情况,对于绘制目标距离较远的虚拟图像,会造成伪轮廓。因此,本文从引起伪轮廓的根本原因出发来更好地解决该问题。

2.2.1 深度图像可信区域分析

深度图像是一幅灰度图像,一般通过深度相机采集或者利用彩色视频序列进行软件估计来获得深度图像^[15],其分辨率和彩色图像一致,其灰度值表示物体离相机的距离(深度值)。深度图像的灰度值在前景和背景的边界处通常存在突变,然而彩色图像的亮度分量在对象边界处是缓慢变化的,因此,边界过渡区域包含前景和背景的混合像素点,如图3所示,其中可疑像素点在通过 DIBR 技术绘制虚拟视点图像时,可能被错误地映射到背景区域上,从而导致伪轮廓现象,如图4所示。

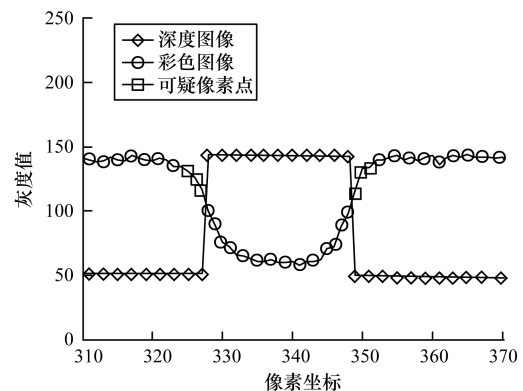


图3 边缘像素点灰度值和水平位置的关系

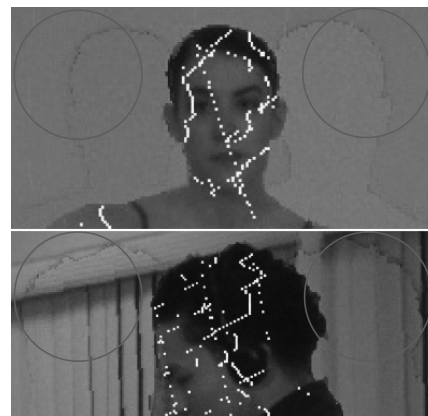


图4 伪轮廓现象

2.2.2 伪轮廓去除技术

经过上述分析可知,通过对深度图进行预处理标记出可疑像素点,能够达到去除伪轮廓的目的。受文献[16]出现空洞区域的启发,利用左参考视点绘制得到的目标视点图像,伪轮廓出现在前景的左侧;利用右参考视点绘制得到的目标视点图像,伪轮廓出现在前景的右侧。采用 Canny 算子对深度图像中物体边缘进行检测并结合文献[9]标记出不可信区域像素点。边缘检测结果如图5(a)所示。不可信区域标记如下:

$$Mask(i, j) = \begin{cases} 1, D(i-1, j) - D(i, j) > T \text{ 左参考图像} \\ 2, D(i, j) - D(i-1, j) > T \text{ 右参考图像} \\ 0, \text{其他} \end{cases} \quad (5)$$

其中, $D(i, j)$ 是深度图像中坐标为 (i, j) 像素点的灰度值; T 是一个阈值, 本文算法取 T 的范围为 $[3, 5]$ 。经过大量统计分析发现前景、背景之间的混合像素点一般为二三个, 为了正确地标记出这些不可信区域, 这里对于标记为 1 的像素点, 将其左侧的 3 个像素点标记为不可信区域; 对于标记为 2 的像素点, 将其右侧的 3 个像素点标记为不可信区域。这些不可信区域在深度图像中设置成黑色, 如图 5(b) 所示。

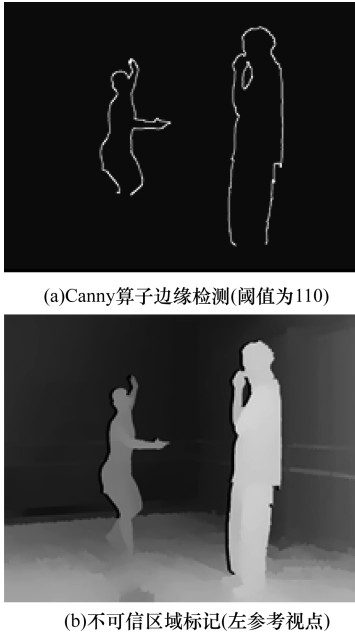


图 5 深度图像预处理

在图像三维变换时, 只对深度图像中的可信区域进行变换, 在变换之后得到的目标视点图像中不会存在伪轮廓现象。通过该处理方式不但能够有效去除背景中残留的前景像素点, 而且部分背景像素点不会被设置为不可信区域。因为由这些被错误标记的背景像素点缺失而导致的空洞点, 可以利用另一个参考视点的虚拟图像进行有效填补。采用经过不可信区域标记后的深度图像进行三维图像变换的结果如图 6 所示, 可以发现几乎不存在伪轮廓现象。

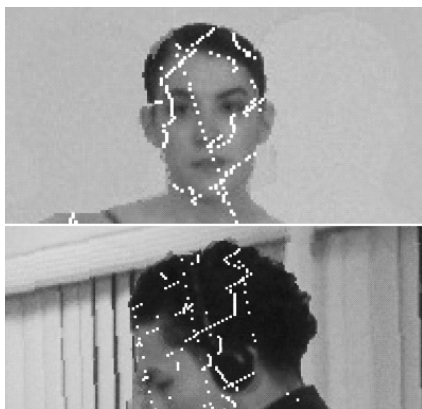


图 6 伪轮廓去除后的效果

2.3 图像分割及视点绘制

图像分割是图像处理中的一项重要处理技术, 现有方法基本仅适用于彩色灰度图像分割, 通常采用大津阈值提取灰度图像中的前景, 但该方法不准确, 当背景区域被错误地判断为前景处理时会造成虚拟视点图像质量的下降。为了有效地区分深度图像中的各类深度值, 本文将对多阈值灰度图像分割技术^[17]进行改进并将其运用到深度图像分割中, 通过分割后的深度图像来指导三维图像变换。该方法的主要优点有: (1) 在三维图像变换中, 能够解决在前景区域中出现背景区域像素点的问题, 如图 7 所示, 前景人物中出现了许多背景像素点, 直接影响了人的主观质量感受; (2) 避免了不同区域像素点的错误映射情况。

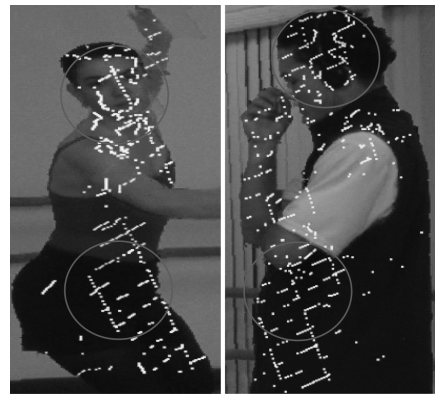


图 7 前景区域中出现背景像素点的现象

2.3.1 多阈值深度图像分割

本文采用多阈值 (Multi-level Thresholding, MTL) 图像分割技术对深度图像进行分割。

(1) 获取深度图像的统计直方图。

通过山峰聚类算法找出直方图的峰值, 具体步骤如下:

1) 深度图像的统计直方图定义如下:

$$H(k) = \sum_{\forall f(x,y)=k} f(x,y), k \in R \quad (6)$$

其中, k 表示深度图像的灰度值; $f(x, y)$ 表示图像上像素坐标为 (x, y) 处的灰度值; R 表示整幅图像。

2) 给定所需的峰值个数 M_0 , 使用迭代聚类过程计算峰值个数。当迭代获取的峰值个数收敛于 M_0 时, 终止迭代过程, 得到最后的峰值点。初始化迭代值 $ITER = 0$ 。

3) 开始迭代计算峰值个数, 设置 $ITER = ITER + 1$, $N_{C, ITER} = N_{CO} / 2^{ITER-1}$, 其中, N_{CO} 为直方图初始灰度级个数, 本文为 256。每个灰度级 f_i 出现的频率定义为:

$$g_{i, ITER} = 0.5 (g_{i, ITER-1} + g_{i-1, ITER-1})$$

$$i = 1, 2, \dots, N_{C, ITER}$$

4) 判定变化方向, d_i 用来标记灰度级为 i 处的变化方向, 定义如下:

$$d_i = \begin{cases} +1, (g_{i-1} > g_{i+1}) \cap (g_{i-1} \geq g_i) \neq 0 \\ -1, (g_{i+1} > g_{i-1}) \cap (g_{i+1} \geq g_i) \neq 0 \\ 0, \text{其他} \end{cases} \quad (7)$$

5) 峰值的判定:

①如果 $(d_i = 0) \cap (d_{i-1} = -1) \cap (d_{i+1} = +1)$, 那么灰度级 i 被判定为峰值点位置。

②如果 $(d_i = -1) \cap (d_{i+1} = +1)$, 那么峰值被判定为在灰度级 i 和 $i+1$ 之间。重复步骤 4), 直到在直方图上找到所有峰值点。

6) 如果在步骤 5) 中找到的峰值点个数收敛于给定的 M_0 值, 那么跳转到步骤 7); 否则, 跳转到步骤 3)。

7) 终止迭代过程, 确定所寻找到的峰值点坐标。

(2) 通过线性估计算法对 2 个峰值之间的点进行拟合, 得到一个有理函数 $R(w)$ 。

(3) 利用一维黄金搜索算法找到每个有理函数的谷值, 选取为阈值点, 具体步骤如下:

1) 搜索出 $R(w)$ 中的最小值, 定义为 R_{\min} 。

2) 定义函数 $Y(w)$ 如下:

$$Y(w) = \begin{cases} R(w), R(w) \leq R_{\min} \\ R_{\min}, \text{其他} \end{cases} \quad (8)$$

3) 搜索出 $Y(w)$ 中的最小值, 定义为 Y_{\min} 。

4) 如果 $Y_{\min} = R_{\min}$, 则进行步骤 5); 否则, 令 $R_{\min} = Y_{\min}$, 跳转到步骤 2)。

5) 得到函数的全局最小值 R_{\min} , 终止搜索过程。

通过使用 MTL 技术, 将深度图像分割为不同部分。下面利用该技术对 Ballet 序列第 3 个相机的第 1 帧深度图像进行分割, 灰度直方图分割结果如图 8 所示, 直方图分割的灰度值范围和阈值选取如表 1 所示。根据阈值对彩色图像进行分割得到如图 9 所示的图像。

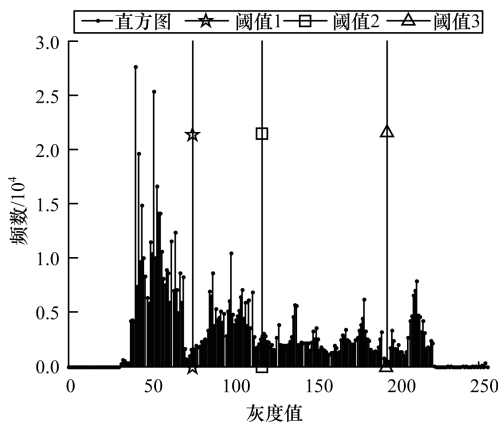
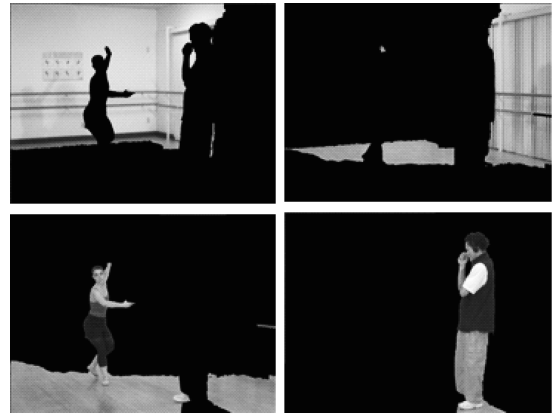


图 8 第 1 帧深度图像的直方图分割

表 1 直方图分割范围及阈值选取

层数	灰度值范围	阈值
第 1 层	(0, 76)	76
第 2 层	(76, 118)	118
第 3 层	(117, 194)	194
第 4 层	(194, 255)	-



(a) 相机 3 彩色参考图像分割



(b) 相机 5 彩色参考图像分割

图 9 根据深度图像阈值的彩色图像分割结果

2.3.2 分割图像的虚拟视点绘制

利用分割后的图像及相应的深度信息进行三维图像变换, 在绘制过程中, 由于采样的不充分, 如图 10(a) 所示, 本文采用局部中值滤波算法对对象表面出现的细小空洞进行填补, 若空洞周围的非空像素点个数大于给定阈值, 则进行修复; 反之不进行处理。具体公式如下:

$$I(U, V) = \begin{cases} M\left(I\left(U+k-\frac{w+1}{2}, V+l-\frac{w+1}{2}\right)\right), \\ I(U, V) = 0 \\ I(U, V), \text{其他} \end{cases} \quad (9)$$

其中, $I(U, V)$ 表示坐标为 (U, V) 像素点的值; 0 表示该像素点为空洞; $k=1, 2, \dots, w$; $l=1, 2, \dots, w$, w 表示中值滤波器窗口的大小, 本文取 3, 给定的阈值取 4。由图 10(b) 可看出, 经过局部中值滤波后, 目标视点中的空洞显著减少。



(a) 未经滤波处理结果 (b) 经滤波处理结果

图 10 局部中值滤波处理前后的效果对比

在分割绘制的过程中, 对每个分割对象绘制后的结果进行局部中值滤波处理, 最终将所有分割对象融合成一幅完整的虚拟视点图像, 在前景中不会残留背景像素点, 各像素点之间也不会相互错误映射, 如图 11 所示。



(a) 未分割绘制结果

(b) 分割绘制结果

图 11 单视点融合后的虚拟视点图像

3 图像融合

本文采用 2 个参考视点进行绘制之后, 得到 2 幅目标视点的图像, 每幅图像都含有较大的空洞区域。这些空洞出现的原因是, 在参考视点中不可见的背景区域经过绘制后, 在目标视点成为可见区域, 由于缺乏相应的纹理参考信息造成了大块空洞。对于这类空洞, 可以由不同视点的像素进行填补, 最终融合得到新的虚拟视点图像。这里采用双向 DIBR 绘制技术中常用的融合方法^[18], 具体计算公式如下:

$$I_V(U_V, V_V) = \begin{cases} 0, I_L(U_L, V_L) = 0; I_R(U_R, V_R) = 0 \\ I_L(U_L, V_L), I_L(U_L, V_L) \neq 0; \\ I_R(U_R, V_R) = 0 \\ I_R(U_R, V_R), I_L(U_L, V_L) = 0; \\ I_R(U_R, V_R) \neq 0 \\ \alpha I_L(U_L, V_L) + (1 - \alpha) I_R(U_R, V_R), \\ I_L(U_L, V_L) \neq 0; I_R(U_R, V_R) \neq 0 \end{cases} \quad (10)$$

其中, $I_V(U_V, V_V)$ 表示虚拟视点图像坐标为 (U_V, V_V) 处的像素值; $I_L(U_L, V_L)$ 和 $I_R(U_R, V_R)$ 表示左右参考视点经过三维图像变换后得到左右参考图像的像素值; (U_L, V_L) 和 (U_R, V_R) 表示虚拟视点图像坐标为 (U_V, V_V) 在由左右参考图像生成的虚拟图像上的对应点, 0 表示空洞, α 是一个权重系数, 由虚拟视点与参考视点之间的距离决定, 如式 (11) 所示。

$$\alpha = \frac{|t - t_L|}{|t - t_L| + |t - t_R|} \quad (11)$$

其中, t 表示目标相机外部参数中的平移向量; t_L 表示左侧相机外部参数中的平移向量; t_R 表示右侧相机外部参数中的平移向量。

通过上述图像融合处理消除了由视点遮挡引起的大空洞, 并采用快速图像修复法^[19]消除剩余的少部分空洞。

4 实验结果与分析

为验证本文方法的有效性, 采用微软研究院的交互视觉媒体组提供的 3D 测试序列^[20] Ballet 和 Breakdancers 进行实验。这 2 组测试序列由 8 个相机采集 100 帧彩色图像和由软件对应计算得到的深度图像组成, 图像分辨率为 1 024 像素 \times 768 像素, 相邻相机之间的水平距离是 20 cm, 以 1 维弧形摆放。在三维图像变换中, 以 3 号相机和 5 号相机为参考视点, 4 号相机为目标视点, 将目标视点绘制的图像与 4 号相机采集到的图像进行比较, 利用两者之间的峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似度 (Structural Similarity, SSIM) 分析虚拟视点绘制效果, PSNR 和 SSIM 是对图像质量的客观度量^[21]。本文将图像 RGB 彩色空间转换到 YUV 彩色空间, 利用亮度分量 Y 值计算最终结果, 转换公式如式 (12) 所示, PSNR 和 SSIM 的计算公式如式 (13)、式 (14) 所示。

$$Y(U, V) = 0.299R(U, V) + 0.587G(U, V) + 0.114B(U, V) \quad (12)$$

$$P_{\text{PSNR}} = 10 \lg \frac{255^2}{(1/W \times H) \sum_{s,t=0}^{W-1, H-1} (Y(s, t) - \hat{Y}(s, t))^2} \quad (13)$$

$$S_{\text{SSIM}(x, y)} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (14)$$

其中, W 和 H 是图像像素矩阵的行和列; $Y(s, t)$ 表示 4 号相机拍摄的真实图像; $\hat{Y}(s, t)$ 表示绘制的虚拟视点图像; μ_x, μ_y 和 σ_x, σ_y 分别表示各自图像块的均值和方差 (下标 x 表示参考相机拍摄的真实图像块, 下标 y 表示绘制的虚拟视点图像块); σ_{xy} 表示协方差; C_1 和 C_2 均为常数。PSNR

值越大,SSIM 值越接近于 1,表示绘制的虚拟视点图像质量越高。

本文选择典型的 Zhang 非对称高斯滤波、Criminisi 算法以及 MPEG 提供的虚拟视点绘制参考软件(VSRS3.5)^[22]与本文算法进行性能对比。在 Zhang 算法中,由于实验序列的分辨率和基线距离较大,因此采用强参数设置,二维高斯核的水平与垂直标准差分别设置为 20 和 60,对应窗口大小分别为 61 和 193。在 Criminisi 算法中,图像修复的匹配块大小设置为 9×9,对整幅图像搜索最佳匹配块。在 VSRS3.5 中,绘制文件参数设置为整像素精度,使用边界去噪和 3D 映射模式。

图 12 和图 13 分别给出了上述 4 种算法的视点绘制结果,可以看到,Zhang 算法存在人物几何畸变和像素点错误映射问题,Criminisi 算法会在对象侧面出现模糊情况,VSRS3.5 存在错误空洞填补问题。本文算法绘制的图像没有出现几何失真且在物体边缘处几乎不存在伪轮廓现象。

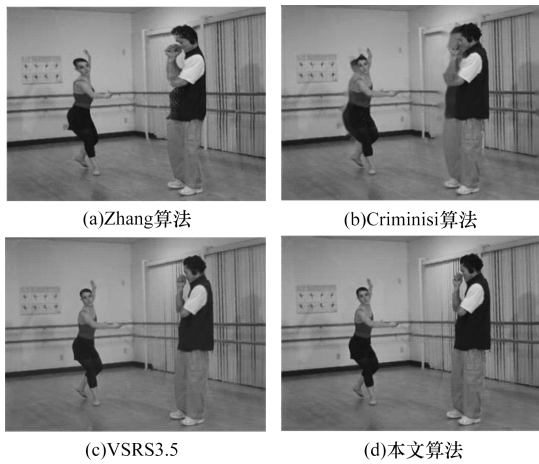


图 12 Ballet 序列测试结果

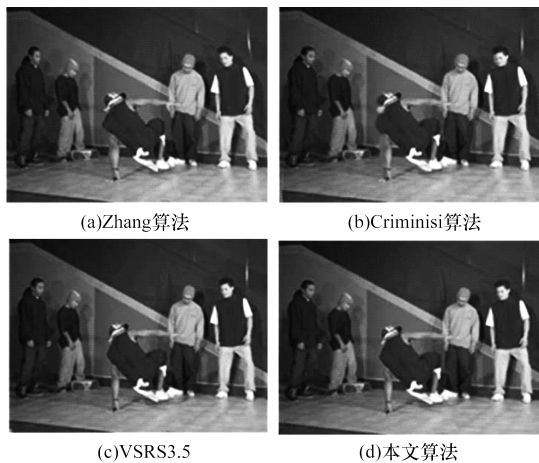
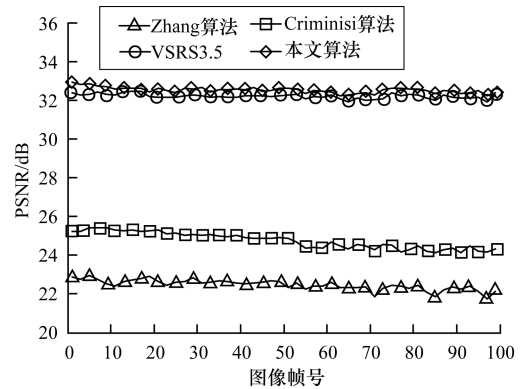
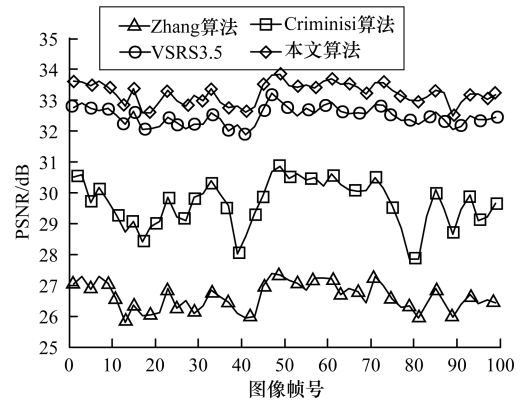


图 13 Breakdancers 序列测试结果

图 14 分别给出了上述 4 种算法的 100 帧虚拟视点图像与真实图像的 PSNR 比较。表 2、表 3 给出了相应算法的平均 PSNR 值、SSIM 值。



(a) Ballet 序列的 PSNR



(b) Breakdancers 序列的 PSNR

图 14 Ballet 和 Breakdancers 视频序列的 PSNR 对比

表 2 绘制虚拟视点的平均 PSNR 比较 dB

算法	Ballet	Δ PSNR1	Breakdancers	Δ PSNR2
Zhang 算法	22.46	10.06	26.65	6.58
Criminisi 算法	24.75	7.77	29.67	3.56
VSRS3.5	32.20	0.32	32.49	0.74
本文算法	32.52		33.23	

表 3 绘制虚拟视点的平均 SSIM 比较

算法	Ballet	Δ SSIM1	Breakdancers	Δ SSIM2
Zhang 算法	0.692 8	0.212 3	0.753 0	0.112 2
Criminisi 算法	0.811 6	0.093 5	0.816 7	0.048 5
VSRS3.5	0.896 6	0.008 5	0.860 2	0.005 0
本文算法	0.905 1		0.865 2	

其中, Δ PSNR1 表示 Ballet 视频序列中对比方法中的 PSNR 值与本文算法 PSNR 值的差值; Δ PSNR2 表示 Breakdancers 视频序列中对比方法中的 PSNR 值与本文算法 PSNR 值的差值; Δ SSIM1 表示 Ballet 视频序列中对比方法中的 SSIM 值与本文算法 SSIM 值的差值; Δ SSIM2 表示 Breakdancers 视频序列中对比方法中的 SSIM 值与本文算法 SSIM 值的差值。主观视觉效果和客观 PSNR,SSIM 数据都表明,本文算法比其他算法更优越。本文算法没有对深度图像进行任何滤波处理,没有改变物体实际的深度值,因此,不会造成图像的几何失真。在对深度图像不可信区域

标记后进行三维图像变换, 从而抑制可疑像素点的映射, 避免对象边缘出现伪轮廓现象, 并且通过图像分割技术, 消除像素点存在错误映射的情况。

为进一步验证本文算法的有效性, 对分割阈值进行修改。表 4 给出了深度图像均匀分割(阈值取为 64, 128, 192)与本文分割方案的对比结果。如果仅简单地对深度图像进行均匀地分割处理, 而没有结合深度图像本身的特性, 客观质量结果会比本文方案略低。本文方案根据深度图像的特征信息, 更加合理地确定出分割阈值, 从而绘制出质量更高的虚拟图像。

表 4 2 种分割方案的性能比较

方案	Ballet		Breakdancers	
	PSNR/dB	SSIM	PSNR/dB	SSIM
均匀分割方案	32.45	0.899 3	33.15	0.860 0
本文分割方案	32.52	0.905 1	33.23	0.865 2
差值	0.07	0.005 8	0.08	0.005 2

5 结束语

本文提出一种基于深度图像分割的虚拟视点绘制算法。首先对深度图像进行预处理, 通过边缘检测技术对可能造成伪轮廓的像素点进行标记, 并且对可信像素点进行三维图像变换。利用深度信息的图像分割方法能够较好地去除虚拟视图中的错误像素。与 Zhang 算法、Criminisi 算法和 VSRS3.5 相比, 本文算法绘制出的虚拟视图不存在几何失真, 在对象边缘处几乎不存在伪轮廓现象, 图像质量较高。随着深度图像分割数目的增加, 本文算法能在一定程度上提高绘制质量, 但是所需时间也会递增, 为了能够实时地绘制出虚拟视点, 如何权衡绘制时间和图像质量两者之间的关系, 是下一步工作需要研究的内容。

参考文献

[1] 李朋, 韦穗, 孙庆, 等. 基于图像金字塔修复的虚拟视点绘制[J]. 计算机工程, 2013, 39(7): 31-34.

[2] Liu Jinxia, Liu Yanwei, Qi Honggang. 3D Video Rendering Adaptation: A Survey [J]. 3D Research, 2015, 6(1): 1-13.

[3] Fehn C, Rene D L B, Pastoor S. Interactive 3-DTV-Concepts and Key Technologies [J]. Proceedings of the IEEE, 2006, 94(3): 524-538.

[4] 鞠芹, 安平, 张倩, 等. 基于多目立体匹配的深度获取方法[J]. 计算机工程, 2010, 36(14): 174-176.

[5] 骆凯, 李东晓, 冯雅美, 等. 基于 DIBR 和图像修复的任意视点绘制[J]. 中国图象图形学报, 2010, 15(3): 443-449.

[6] Shade J, Gortler S, He L, et al. Layered Depth Images [C] // Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA: ACM Press, 1998: 231-242.

[7] Zhang L, Tam W J. Stereoscopic Image Generation Based on Depth Images for 3D TV [J]. IEEE Transactions on Broadcasting, 2005, 51(2): 191-199.

[8] Lee P J. Nongeometric Distortion Smoothing Approach for Depth Map Preprocessing [J]. IEEE Transactions on Multimedia, 2011, 13(2): 246-254.

[9] 高凯, 陈贺新, 赵岩, 等. 面向虚拟视点绘制的深度图滤波及上采样方法[J]. 中国图象图形学报, 2013, 18(9): 1085-1092.

[10] 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.

[11] Xu X, Po L M, Cheung C H, et al. Depth-aided Exemplar-based Hole Filling for DIBR View Synthesis [C] // Proceedings of 2013 IEEE International Symposium on Circuits and Systems. Washington D. C., USA: IEEE Press, 2013: 2840-2843.

[12] Wang Dan, Zhao Yan, Wang Zheng, et al. Hole-filling for DIBR Based on Depth and Gradient Information [J]. International Journal of Advanced Robotic Systems, 2015, 12(1): 39-42.

[13] Schmeing M, Jiang Xiaoyi. Superpixel-based Occlusion Filling in Depth Image Based Rendering [C] // Proceedings of the 22nd International Conference on Pattern Recognition. Washington D. C., USA: IEEE Press, 2014: 1073-1078.

[14] 赵猛, 金一丞, 尹勇. 立体显示中双目视差模型和深度感知研究[J]. 计算机工程, 2011, 37(17): 271-273.

[15] 田寨兴, 彭宗举, 蒋刚毅, 等. 面向 HEVC 的深度视频预处理[J]. 光电工程, 2014, 41(1): 87-94.

[16] Yao C, Tillo T, Zhao Y, et al. Depth Map Driven Hole Filling Algorithm Exploiting Temporal Correlation Information [J]. IEEE Transactions on Broadcasting, 2014, 60(2): 394-404.

[17] Papamarkos N, Gatos B. A New Approach for Multilevel Threshold Selection [J]. CVGIP Graphical Models & Image Processing, 1994, 56(5): 357-370.

[18] Mueller K, Smolic A, Dix K, et al. View Synthesis for Advanced 3D Video Systems [J]. EURASIP Journal on Image and Video Processing, 2008(4): 1-11.

[19] Telea A. An Image Inpainting Technique Based on the Fast Marching Method [J]. Journal of Graphics Tools, 2004, 9(1): 23-34.

[20] Zitnick C L, Kang S B, Uyttendaele M, et al. High-quality Video View Interpolation Using a Layered Representation [J]. ACM Transactions on Graphics, 2004, 23(3): 600-608.

[21] Wang Z, Bovik A C, Sheikh H R, et al. Image Quality Assessment: From Error Visibility to Structural Similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

[22] MPEG. 3DV Depth Estimation and View Synthesis Software Package; ISO/IEC JTC1/SC29/WG11 N12188—2011 [S]. Geneva, Switzerland: International Organization for Standardization, 2011: 9.