

Kinect 深度图像快速修复算法

王奎¹, 安平^{1,2}, 张兆杨^{1,2}, 程浩¹, 李贺建¹

(1. 上海大学 通信与信息工程学院, 上海 200072; 2. 新型显示技术及应用集成教育部重点实验室, 上海 200072)

摘要: 深度提取是基于“纹理+深度”自由立体视频系统的关键技术, 而立体视频实际应用系统需要高效快速的深度图提取. 提出一种针对 Kinect 提取深度图的快速修复算法. 首先, 对 Kinect 提取的彩色纹理图和深度图进行对齐裁剪, 并采用背景填充算法对裁剪后的深度图进行初步修复; 然后, 对初步修复后的深度图进行基于颜色匹配的快速修复, 得到质量较好的可用深度图. 实验结果表明, 本算法能有效修复原始深度图中由于遮挡而引起的空洞, 获取的深度图整体平滑度好、边缘清晰; 在普通 PC 机上达到 25~30 帧/s 的处理帧率, 实现了深度图的实时提取.

关键词: 深度图修复; 实时获取; 颜色匹配; Kinect; 立体视频

中图分类号: TP 391.41

文献标志码: A

文章编号: 1007-2861(2012)05-0454-05

Fast Inpainting Algorithm for Kinect Depth Map

WANG Kui¹, AN Ping^{1,2}, ZHANG Zhao-yang^{1,2}, CHENG Hao¹, LI He-jian¹

(1. School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China;

2. Key Laboratory of Advanced Display and System Applications, Ministry of Education, Shanghai 200072, China)

Abstract: Depth extraction is a key step in 3D video system based on texture plus depth. Real application systems require high efficiency and fast depth extraction. This paper presents a fast inpainting algorithm for Kinect depth map. This paper clips and aligns the origin color and depth images captured by Kinect, and partially fill holes in the clipped depth image using a background based method. Then this paper uses a fast inpainting algorithm based on color match to fill the remaining holes in the depth image to obtain a better depth image. Experimental results show that the proposed algorithm can efficiently repair errors in the original depth map such as holes caused by occlusion. Finally we obtain a smooth depth map with clear edges. The processing speed can reach 25~30 frame/s by using an ordinary desktop computer to realize real-time depth map extraction.

Key words: depth map inpainting; real-time capture; color match; Kinect; stereo video

随着立体显示技术和视频处理技术的发展, 3D 视频技术成为近年来的研究热点^[1]. 3D 视频技术研究首先要解决 3D 信息的表示问题, 即采用何种方式来描述 3D 场景. 目前主要有两种方案: 一种是采用

多视方案, 用摄像机阵列拍摄 3D 场景, 在立体显示器上显示出来; 另一种是“纹理+深度”的方案^[1], 用彩色纹理图描述 3D 场景的纹理信息, 用深度图描述 3D 场景的深度信息, 在终端采用基于深度图的绘

收稿日期: 2011-11-18

基金项目: 国家自然科学基金资助项目(60832003, 61172096, 61171084); 上海市科委重点资助项目(10510500500); 上海市教委重点资助项目(09ZZ90)

通信作者: 安平(1968~), 女, 教授, 博士生导师, 博士, 研究方向为数字视频处理及立体视觉. E-mail: anping@shu.edu.cn

制(depth image based rendering, DIBR)技术绘制虚拟视点,将虚拟视点在立体显示器上显示出来.第一种方案由于拍摄的视点较多,数据的编码、存储以及传输成为技术瓶颈.第二种方案由于有了场景的深度信息,在拍摄过程中只需要捕获很少的视点,因此数据量相对第一种方案少了很多,便于数据的编码存储和传输;并且在显示端可以利用DIBR技术进行一定范围内的任意视点绘制^[2],从而产生更好的视觉效果.基于“纹理+深度”3D信息描述的多视点+深度(multi-video+depth, MVD)系统以其传输带宽小、易于虚拟视点绘制等优点而被认为是一种有应用前途的3D视频方案^[2],已成为当前的研究热点.

但是基于“纹理+深度”的立体视频系统中,3D场景深度信息的获取是十分困难的.目前对深度信息的获取主要有两种方法^[3]:①采用立体匹配算法被动获取深度信息,一般利用双目相机拍摄3D场景信息,用立体匹配算法提取场景中物体的视差,再根据相机参数计算出深度信息.这种方案中的算法实现复杂度高,很难做到深度实时提取,同时这种方法受输入场景以及遮挡问题影响非常大;②采用测距设备主动获取场景的深度信息,即使用深度相机直接获取深度信息.深度相机目前分为两类,一类是TOF(time of flight)相机,另一类是Kinect,它们都是通过发射和接收反射光来计算深度信息以实现深度信息的实时提取.但是这种深度提取设备价格昂贵,且获取的深度图分辨率低,给实际应用带来困难.

TOF相机的优点是其发射和接收的传感器较密集,输出深度图质量较好.但深度图中的空洞或者深度错误会对虚拟视点绘制质量带来较大的影响,这些错误通常发生在遮挡区域,即深度图不连续的地方.文献[4]提出了使用高斯滤波来平滑深度图,随后一些文献也给出了对高斯滤波的改进算法,如文献[5]提出了一种非对称高斯滤波算法.这些算法都只利用了空域信息(即单帧图像局部区域),对深度图的修复效果有限.文献[6]针对TOF相机深度图空域跳动的问题,提出了一种利用时域和空域信息来减少其跳变的方法.文献[7]则提出了一种结合立体相机纹理数据的深度图分辨率提升算法,通过结合深度图对应的纹理图来提升深度图的质量.然而,TOF相机存在获取深度图分辨率低、价格昂贵等缺点.

Kinect是微软研发的一种可以获取场景深度的

设备,主要应用在人机交互(如Xbox游戏机)、3D场景重建、机器视觉等领域^[8].Kinect可以同时采集场景纹理和深度,相对于TOF深度提取设备,Kinect价格便宜,可以提取较高分辨率的深度图.但是Kinect提取的深度图质量较差,在遮挡区域、光滑物体表面存在较大的深度信息缺失空洞,必须进行填充修复处理.

针对Kinect深度图修复,文献[9]提出了时域滤波深度图修复增强算法,但是这种方法复杂度高,算法实现的实时性差,帧率不到2帧/s.因此本研究提出一种Kinect深度图实时修复算法,结合时域背景拟合信息以及空域颜色信息对Kinect输出的原始深度信息进行修复,取得了较好的效果,同时在普通PC机上可以实现20~30帧/s的帧率.

1 算法框架

在三维场景中,由于背景被前景遮挡或者前景的个别部分距Kinect过近使该部分不在红外光场之内等因素,部分Kinect发射的红外光无法反射回来而产生背景或者前景空洞,显然,当前景距相机比较近时产生的空洞问题更严重.这些空洞可以采用图像修复算法进行修复,但是对于大空洞,如果单纯依靠空域图像修复,不仅修复非常耗时,而且修复效果也不好.由于Kinect出现的空洞大部分是场景中的背景部分,因此可以利用场景中的背景对空洞进行初步修复,然后再结合彩色纹理信息对局部小空洞进行二次修复.

基于上述考虑,本研究的算法的主要步骤如下:

- (1) 对Kinect输出的深度图和纹理图进行对齐裁剪;
- (2) 利用帧差法结合原始深度图估计背景,并对背景深度图采用结合纹理颜色信息的方法进行修复;
- (3) 用背景填充算法对原始深度图进行初步修复;
- (4) 结合纹理颜色信息对深度图进行二次修复;
- (5) 对修复的空洞区域采用中值滤波去除噪声.

2 算法描述

2.1 纹理图和深度图的对齐裁剪

Kinect最大可输出640×480分辨率的深度图和纹理图.如前所述,Kinect提取的深度图质量较差,在遮挡区域、光滑物体表面存在较大的深度信息缺失空洞;而且Kinect采集的原始深度图与纹理图

虽然已作了匹配和对齐处理,但并不是完全对齐的,纹理图相对于深度图在大小上存在差异(见图1(a)和图1(b)).因此,本研究首先通过实验对采集的深度图和纹理图进行裁剪处理,裁剪结果如图1(c)和图1(d)所示.

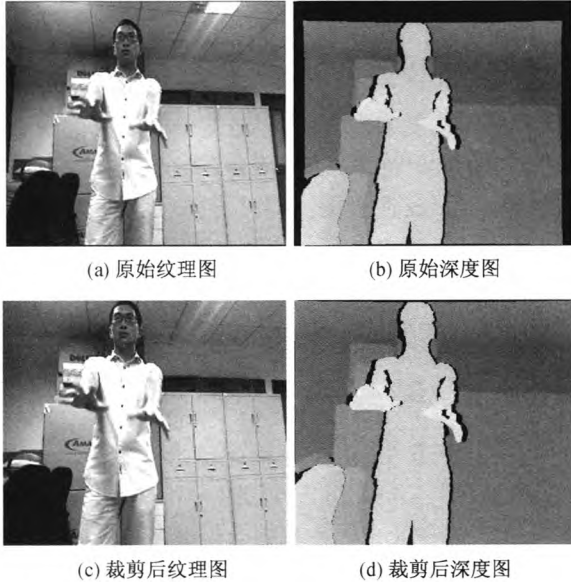


图1 对齐裁剪处理结果
Fig.1 Result after align and cut

2.2 背景纹理图和深度图的估计

将采集的第一帧深度图和纹理图作为初始的背景深度图和纹理图.对于深度图背景的更新,可以直接使用深度图帧差法进行背景更新检测,也可以用纹理图帧差法进行背景更新检测.由于 Kinect 深度图的稳定性差、噪声大,而帧差法^[10]本身对噪声非常敏感,因此本研究对纹理图采用帧差法进行背景更新检测.

帧差法以视频序列中相邻两帧之间基于像素的帧差来区分图像中的运动部分和背景部分,即

$$D_k(x, y) = |f_k(x, y) - f_{(k-1)}(x, y)|, \quad (1)$$

式中, $D_k(x, y)$ 表示第 k 帧的帧差图, $f_k(x, y)$ 和 $f_{(k-1)}(x, y)$ 分别表示第 k 和第 $k-1$ 帧中坐标 (x, y) 处的像素值.

帧差图中的噪声和运动部分有明显的差别,可以根据下式来滤除噪声的干扰:

$$\begin{cases} d_k(x, y) = 0, & D_k(x, y) < C_{\text{threshold}}; \\ d_k(x, y) = 1, & D_k(x, y) \geq C_{\text{threshold}}, \end{cases} \quad (2)$$

式中, $C_{\text{threshold}}$ 是针对噪声滤除设定的阈值(多次实验得出的经验值为45), $d_k(x, y) = 1$ 表示 (x, y) 处背景

需要更新, $d_k(x, y) = 0$ 表示 (x, y) 处背景不需要更新.

用帧差法得到变化区域($d_k(x, y) = 1$)后,比较当前帧深度与背景深度,如果当前帧深度值小于背景深度值,就以当前帧深度值更新背景深度值,否则不进行更新.在更新深度值时,同时更新对应位置的纹理图,这样就完成了深度图和纹理图背景的实时更新.

经背景估计的初始帧会出现小范围的空洞,此时采用下文所述的基于颜色匹配算法进行修复.估计的背景纹理图和深度图结果如图2所示.

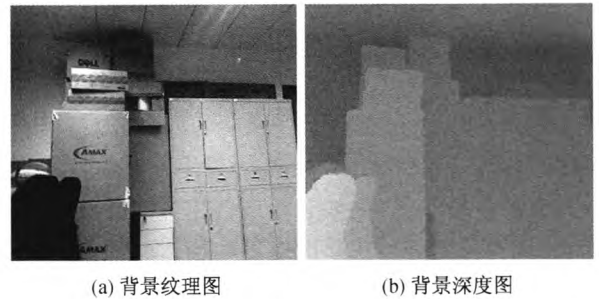


图2 更新后的背景图
Fig.2 Background image after update

2.3 基于背景填充的深度图初步修复

得到背景纹理图和背景深度图后,下面对深度图进行初步修复,利用背景深度值去填充当前帧中处于背景位置的空洞.首先,利用当前帧的彩色纹理图与背景纹理图作帧差,检测出背景与前景;然后,使用背景深度图填充当前帧深度图中处于背景区域的空洞.

背景深度填充的表达式为

$$\begin{cases} d_k(x, y) = b_k(x, y), & h_k(x, y) = 1, u_k(x, y) \neq 1; \\ d_k(x, y) = 0, & h_k(x, y) = 1, u_k(x, y) = 1, \end{cases} \quad (3)$$

式中, $d_k(x, y)$ 表示 Kinect 提取的第 k 帧深度图, $b_k(x, y)$ 表示第 k 帧估计的背景深度, $u_k(x, y)$ 表示根据帧差法检测出背景更新区域和非更新区域, $u_k(x, y) = 1$ 表示前景,否则表示背景. $h_k(x, y) = 1$ 表示第 k 帧深度图的空洞区域,当且仅当 (x, y) 处深度为空洞,并且该处是背景区域时,才进行深度背景填充.

图3给出了初步修复的结果,图3(a)是经过裁剪对齐后的原始深度图,图3(b)是经过初步修复后的深度图,可见,原始深度图中的大部分空洞已被修复.

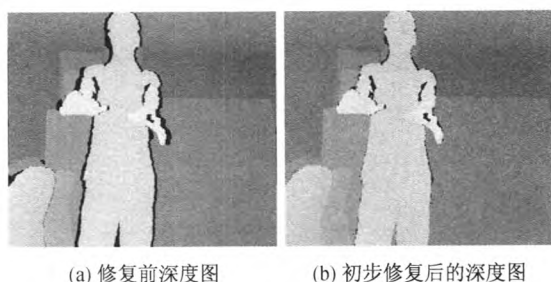


图3 初步修复结果

Fig.3 Result after inpainting initially

2.4 基于纹理颜色匹配的深度图二次修复及消噪处理

经过背景深度填充算法修复后的深度图空洞已经大大减小,但是仍然存在前景空洞(如图3(b)所示).针对这种空洞,本研究采用基于搜索提前终止的颜色匹配搜索算法进行修复.

基于颜色匹配的搜索算法根据空洞处纹理图像的颜色,在纹理图像中找出最佳颜色匹配点,然后将该处的深度值作为空洞处的深度值.为提高处理速度,本算法没有采用基于块的匹配,而是直接采用基于像素的颜色匹配,并且设定匹配阈值为6(即当前待修复深度处对应的彩色像素与非空洞区域的颜色像素进行比较匹配),以控制搜索长度.这样虽然匹配点不一定是最佳的,但实验表明,这样做可大幅提高空洞修复速度,同时兼顾了修复质量.

基于像素的快速修复算法在修复区域会出现修复噪声,本研究采用中值滤波器对修复后的图像进行中值滤波以去除噪声.但仅对修复区域作中值滤波处理,对于非修复区域不进行中值滤波处理,这样一方面提高了中值滤波速度,同时也最大限度地保持了原始正确的深度值.

图4给出了二次修复后的深度图以及进一步经过中值滤波后的深度图.

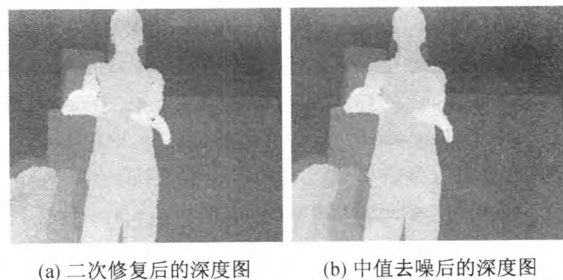


图4 二次修复及消噪处理结果

Fig.4 Result after the secondary inpainting and denoising

3 实验结果

本研究对所提出的深度修复算法用C语言实现,并在普通PC机(Intel E4800 3.0 GHz 双核CPU, 2 G内存)上进行了效果和修复速度测试,图5给出了几组经过裁剪后的场景的初始纹理图(见图5(a))和初始深度图(见图5(b))以及最终修复的深度图(见图5(c)),由图5(c)可见,空洞几乎已全部修复.

图5中场景从上到下分别记为场景1、场景2、场景3、场景4.这4个场景的处理耗时分别为24, 32, 26, 34 ms,可见已达到25~30帧/s的实时修复要求.当前景物体距离Kinect相对较远时,可以实现30帧/s左右的修复速度;当前景物体距离Kinect较近时,Kinect输出的原始深度图存在非常大的空洞,可以实现25帧/s左右的修复速度.



图5 多个场景的修复结果

Fig.5 Inpainted result of a number of different scenarios

为了进一步检验本算法的修复效果,下面对修复后的深度图以及对应的纹理图进行虚拟视点绘制.绘制出左边8路虚拟视点和右边8路虚拟视点(相邻视点间的视差为1 cm),并进行8路虚拟视点交织,在裸眼立体显示器上观看,可以观看到较好的立体效果.

图6给出了利用修复后的深度图和纹理图绘制出来的左边4路和右边4路虚拟视点.图7给出了左边8路虚拟视点合成的立体视.

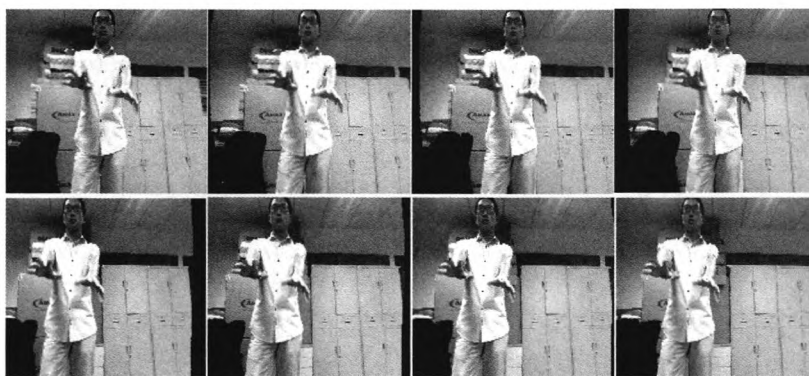


图6 通过 DIBR 得到 8 路虚拟视点

Fig.6 8 virtual views by DIBR



图7 虚拟视点交织得到的立体视

Fig.7 Stereoscopic view by interlace

4 结束语

本研究提出一种针对 Kinect 提取的深度图的快速修复算法,该算法主要利用了背景估计以及纹理图的颜色等信息对 Kinect 输出的带有空洞的深度图进行修复.首先,根据 Kinect 输出深度图中空洞的特点,利用背景估计信息对深度图中背景空洞进行初步修复;然后,结合纹理图对深度图进行二次修复.实验结果表明,本算法有效修复了原始深度图的空洞等问题,同时在普通 PC 机上实现了 25~30 帧/s 的修复速度.在以后的工作中,将致力于继续提高深度图的修复质量并进行多视绘制.

参考文献:

- [1] 张兆杨,安平,张之江,等.二维和三维视频处理及立体显示技术[M].北京:科学出版社,2010:131-133.
- [2] DARIBO I, SAITO H. A novel inpainting-based layered depth video for 3DTV [J]. IEEE Transactions on Broadcasting, 2011:533-541.
- [3] UM G M, KIM T, BANG G. Multi-view 3D video acquisition using hybrid camera with beam splitter [C]// 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON). 2011:1-4.
- [4] DARIBO I, TILLIER C, PESQUET-POPESCU B. Distance dependent depth filtering in 3D warping for 3DTV [C]// Multimedia Signal Processing. 2007:312-315.
- [5] CHEN W Y, CHANG Y L, LIN S F, et al. Efficient depth image based rendering with edge dependent depth filter and interpolation [C]// IEEE International Conference on Multimedia and Expo. 2005:1314-1317.
- [6] KIM S Y, CHO J H, KOSCHAN A, et al. Spatial and temporal enhancement of depth images captured by a time-of-flight depth sensor [C]// Pattern Recognition (ICPR). 2010:2358-2361.
- [7] ZHU J J, WANG L, YANG R G, et al. Fusion of time-of-flight depth and stereo for high accuracy depth maps [C]// Computer Vision and Pattern Recognition (CVPR2008). 2008:1-8.
- [8] WILSON A. Using a depth camera as a touch sensor [C]// ITS'10 ACM International Conference on Interactive Tabletops and Surfaces. 2010:69-72.
- [9] MATYUNIN S, VATOLIN D, BERDNIKOV Y. Temporal filtering for depth maps generated by kinect depth camera [C]// The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON). 2001:1-4.
- [10] JI X P, WEI Z Q, FENG Y W. Effective vehicle detection technique for traffic surveillance systems [J]. Visual Communication & Image Representation, 2005, 17(3):647-658.

作者: [王奎](#), [安平](#), [张兆杨](#), [程浩](#), [李贺建](#), [WANG Kui](#), [AN Ping](#), [ZHANG Zhao-yang](#), [CHENG Hao](#),
[LI He-jian](#)

作者单位: [王奎, 程浩, 李贺建, WANG Kui, CHENG Hao, LI He-jian \(上海大学通信与信息工程学院, 上海, 200072\)](#), [安平, 张兆杨, AN Ping, ZHANG Zhao-yang \(上海大学通信与信息工程学院, 上海200072; 新型显示技术及应用集成教育部重点实验室, 上海200072\)](#)

刊名: [上海大学学报 \(自然科学版\)](#) 

英文刊名: [Journal of Shanghai University \(Natural Science Edition\)](#)

年, 卷(期): 2012, 18 (5)

被引用次数: 2次

参考文献(10条)

1. [张兆杨; 安平; 张之江](#) [二维和三维视频处理及立体显示技术](#) 2010
2. [DARIBO I; SAITO H](#) [A novel inpainting-based layered depth video for 3DTV](#) 2011
3. [UM G M; KIM T; BANG G](#) [Multi-view 3D video acquisition using hybrid camera with beam splitter](#) 2011
4. [DARIBO I; TILLIER C; PESQUET-POPESCU B](#) [Distance dependent depth filtering in 3D warping for 3DTV](#) 2007
5. [CHEN W Y; CHANG Y L; LIN S F](#) [Efficient depth image based rendering with edge dependent depth filter and interpolation](#) 2005
6. [KIM S Y; CHO J H; KOSCHAN A](#) [Spatial and temporal enhancement of depth images captured by a time-of-flight depth sensor](#) 2010
7. [ZHU J J; WANG L; YANG R G](#) [Fusion of time-of-flight depth and stereo for high accuracy depth maps](#) 2008
8. [WILSON A](#) [Using a depth camera as a touch sensor](#) 2010
9. [MATYUNIN S; VATOLIN D; BERDNIKOV Y](#) [Temporal filtering for depth maps generated by kinect depth camera](#) 2001
10. [JI X P; WEI Z Q; FENG Y W](#) [Effective vehicle detection technique for traffic surveillance systems](#) 2005 (03)

引证文献(2条)

1. [徐欢, 郑江滨](#) [Kinect深度图像修复算法](#) [期刊论文] - [微电子学与计算机](#) 2013 (11)
2. [李应彬, 冯杰, 张华熊, 李晨勤](#) [基于改进双边滤波的Kinect深度图像空洞修复算法研究](#) [期刊论文] - [工业控制计算机](#) 2013 (11)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_shdxxb201205003.aspx