

Kinect 深度图像修复算法

徐 欢, 郑江滨

(西北工业大学 陕西省语音与图像信息处理重点实验室, 陕西 西安 710129)

摘 要: 针对 Kinect 深度数据不稳定, 深度图像与彩色图像存在偏差, 图像噪声较大和空洞等问题, 提出了一种 Kinect 深度图像修复算法。首先, 设计了一种 Kinect 标定方法, 完成深度图像与彩色图像对准; 第二, 利用深度数据的时间相关性, 实现了深度图中不稳定数据的重计算; 第三, 借助于深度数据空间相关性, 并结合颜色一致性约束, 填补了深度图中(缺失的数据)空洞; 最后, 设计了一种改进的双边滤波器, 在平滑深度图像数据的同时保留了边缘信息。实验结果表明, 修复后的深度图像整体平滑、边缘清晰, 与原始深度图像相比, 空洞明显减少, 深度图像质量大大提高。

关键词: 深度图像修复; Kinect 标定; 时间相关性; 空间相关性; 颜色一致性约束; 双边滤波器

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2013)11-0098-06

A Novel Inpainting Algorithm for Kinect Depth Image

XU Huan, ZHENG Jiang-bin

(College of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: The quality of a Kinect depth image is easy to be decreased due to the influence of the instability of depth data, image noise, holes with unknown depth and bias between depth and color image, this paper proposed a novel algorithm to inpaint Kinect depth image. Firstly, this paper designed a Kinect calibration scheme and used the calibration results to align the depth image and the color image. Secondly, according to the temporal correlation of depth data, an optimization was applied to improve the data stability. Thirdly, this paper proposed a hole-filling algorithm by combing the spatial correlation of depth data and the color consistency constraint. Finally, an improved bilateral filter was used to smooth the image as well as reserve image edges. Experimental results show that holes are reduced and a smooth with significant depth-edge image could be obtained.

Key words: depth image inpainting; Kinect calibration; temporal correlation; spatial correlation; color consistency constraint; bilateral filter

1 引言

Kinect 可以实时获得场景深度信息, 近几年来被广泛地应用于机器视觉、三维重建、目标跟踪等研究领域。但是 Kinect 深度图像还存在以下问题: ①测量时场景的光照变化, 会导致得到的深度数据不稳定; ②由于物体表面材质的红外反射特性和区域遮挡等原因, 产生空洞问题; ③与彩色图像之间存在偏差。

针对 Kinect 获取的深度图像数据存在的上述

问题, 文献[1-5]提出一些解决方法, 但这些算法往往没有充分利用彩色图像信息来解决空洞问题, 同时没有考虑 Kinect 深度图像和彩色图像之间的偏差。本文提出 Kinect 深度图像修复算法。

2 算法描述

针对 Kinect 的深度图像特性, 同时结合彩色图像信息, 本文提出 Kinect 深度图像修复算法如图 1 所示。算法主要步骤如下:

收稿日期: 2012-12-03; 修回日期: 2013-02-21

基金项目: 国家自然科学基金项目(60970069); 国家“八六三”计划项目(2012AA011803); 航天创新基金项目(2011XW0001)

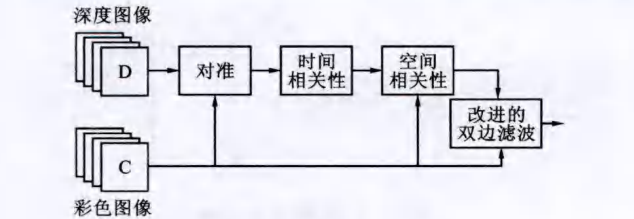


图1 本文算法流程图

- (1)对 Kinect 进行标定,结合标定结果通过坐标变换,对深度图像与彩色图像进行对准;
- (2)利用深度数据之间的时间相关性,分析多帧深度图像,采用直方图统计的方法,对深度图像进行优化;
- (3)利用深度数据之间的空间相关性,结合彩色图像颜色一致性约束,填补深度图像中的空洞;
- (4)通过改进的双边滤波器,对深度图像进行整体去噪和平滑.

2.1 深度图像与彩色图像对准

Kinect 由一个红外发射器,红外摄像机及 RGB 彩色摄像机组成. RGB 相机和红外相机的参数不一致,使得深度图像和彩色图像之间存在偏差,会对后期的处理产生不可预料的结果.

2.1.1 Kinect 标定

普通相机的标定一般通过拍摄棋盘格图案的标定板,结合张氏标定等经典标定方法.如图 2(b),在 Kinect 深度图像中,无法显示标定板上的棋盘图案,从而无法对红外相机直接进行标定.使用 Kinect SDK 可以获取 Kinect 原始红外图像,从图 2(c)中可以看出在红外原始图像中可以观测到棋盘图案,但是由于红外发射器的功率较小以及红外发射器发出点阵图案干扰,拍摄出的棋盘图案非常不清晰,很难用于标定.为此,设计了一种 Kinect 标定方法:将 Kinect 的红外发射器挡住,用红外补光灯照射标定板,这样一来就可以得到清晰的标定板红外原始图像,如图 2(d)所示.结合经典的标定方法和标定工具包,可以对 Kinect 完成标定,标定的结果见表 1.

表 1 Kinect 标定结果

彩色相机		深度相机	
$fx_c = 1\ 060.22 \pm 0.89$		$fx_d = 575.23 \pm 0.50$	
$fy_c = 1\ 053.87 \pm 0.88$		$fy_d = 544.72 \pm 0.51$	
$cx_c = 624.45 \pm 0.70$		$cx_d = 326.76 \pm 0.42$	
$cy_c = 506.22 \pm 0.62$		$cy_d = 261.11 \pm 0.42$	
$R = \begin{bmatrix} 0.999\ 8 & -0.001\ 51 & 0.006\ 48 \\ 0.001\ 58 & 0.999\ 93 & -0.011\ 72 \\ -0.006\ 46 & 0.011\ 73 & 0.999\ 1 \end{bmatrix}$		$c_0 = -0.002\ 883$	
		$c_1 = 3.15$	
		$T = [-0.025\ 80 \quad -0.027\ 93 \quad 0.001\ 68]$	

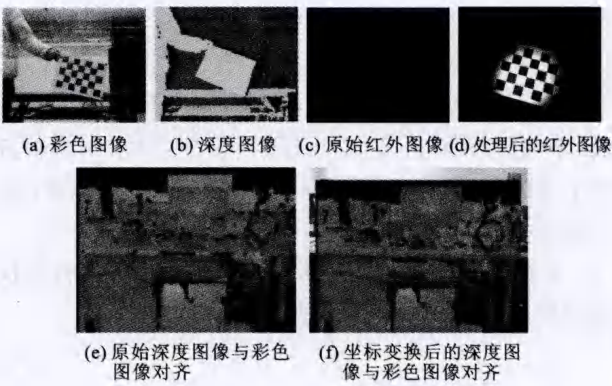


图2 标定图像与对准结果

2.1.2 对准

定义 Kinect 原始深度图像中的一点为 P ,其坐标为 (u_d, v_d) ,对应的深度测量值为 z_{raw} ,根据式(1)可以得到点 P 真实的深度值 z_d :

$$z_d = \frac{1}{c_0 z_{raw} + c_1} \tag{1}$$

$$\begin{cases} x_d = (u_d - cx_d) * z_d / fx_d \\ y_d = (v_d - cy_d) * z_d / fy_d \end{cases} \tag{2}$$

通过式(2)将深度图像坐标系中的点 $P(u_d, v_d)$ 转换到世界坐标系,得到在世界坐标系中的对应点 P_d ,记其坐标为 (x_d, y_d, z_d) .

其中, fx_d 和 fy_d 为深度相机在 x 和 y 方向的焦距, cx_d, cy_d 为主点偏移.

通过式(3)利用彩色相机的外参,包括旋转矩阵 R 和平移矩阵 T ,将世界坐标中系的点 $P_d(x_d, y_d, z_d)$ 转换到彩色相机坐标系中,得到坐标为 (x_c, y_c, z_c) 的点 P_c .

$$P_c = RP_d + T \tag{3}$$

最后利用式(4)将彩色相机坐标系中的点 P_c 转换到彩色图像坐标系,得到坐标为 (u_c, v_c) 的点 Q ,点 Q 即为深度图像中的点 $P(u_d, v_d)$ 在彩色图像中的对应点.

$$\begin{cases} u_c = x_c * fx_c / z_c + cx_c \\ v_c = y_c * fy_c / z_c + cy_c \end{cases} \tag{4}$$

通过坐标变换的方法,可以将深度图像和彩色图像进行对准.图 2(e) 是将原始图像与对应的彩色图像对其的效果,可以看出深度图像和彩色图像无法完全对齐,图像之间存在水平和垂直位移以及缩放.图 2(f) 是将经过坐标变换的深度图像与彩色图像对齐的效果,从图中可以看出两幅图像基本上是对齐的.

2.2 时间相关性

Kinect 的深度数据是不稳定的.对于深度图中

的任何一个像素点,它的深度往往也是随时间变化的,在某一些帧中,该点的深度信息可能会缺失,大量的深度信息缺失的像素点会在深度图像中形成空洞.根据多帧深度数据的时间相关性,可以由当前深度图像的前若干帧和后若干帧深度图像去预测空洞区域某些像素点的深度.

首先对 Kinect 原始深度数据进行量化,将深度值转换到 0 ~ 255 区间,便于显示和计算.

$$D = \frac{Z_d}{Z_{dmax}} * 255 \quad (5)$$

其中 Z_d 为 Kinect 输出的深度值, Z_{dmax} 为 Kinect 输出的最大深度值, D 为量化后的深度值.

对于量化后深度图像中的任意一个像素点,记坐标为 (u, v) , 获得其在 N 帧深度图像每一帧的深度值 $D_i(u, v)$, 用直方图统计的方法, 分析该点的深度分布情况. 在直方图中, 选取峰值对应的深度, 即出现次数最多的深度值作为该像素点的最终深度 $D(u, v)$, 如式(6):

$$D(u, v) = \text{peak}\{\text{hist}\{\sum_{i=1}^N D_i(u, v)\}\} \quad (6)$$

2.3 空间相关性

Kinect 的深度图像中存在大量空洞, 产生这些空洞的原因很多, 包括 Kinect 本身设备原因, 场景光照情况, 物体表面材质的红外发射特性以及区域遮挡等. 本文中主要把空洞分为两类, 如图 3(a) 所示, 红色区域为第一类空洞, 主要是由于物体部分区域对红外的反射特性造成的空洞, 这一类空洞主要出现在物体中间, 如图 3(b) 深度图像中的红色标记的区域; 如图 3(a) 所示的蓝色区域为第二类空洞, 第二类空洞主要是由于区域遮挡原因造成的, 这类空洞往往出现在物体的边缘, 如图 3(b) 深度图像中的蓝色标记的区域.

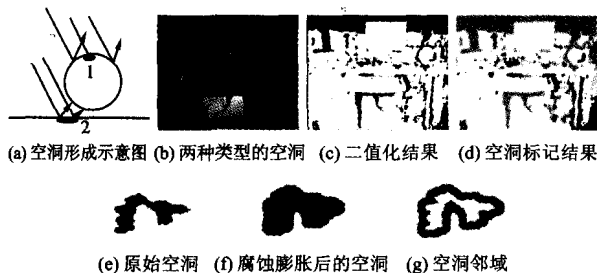


图 3 空间相关性分析

通过对空洞产生原因的分析, 本文提出以下算法进行深度图像补洞, 算法流程如图 4 所示.

首先对采集到的深度图像进行二值化处理. 在 Kinect 深度图像 D 中, 会将无法测量的点的深度

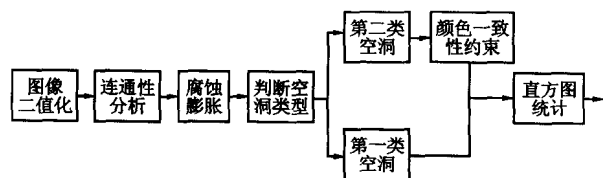


图 4 补洞算法流程图

$D(u, v)$ 输出为 0 对于深度图像 D 中的每一个像素点, 把深度 $D(u, v)$ 为 0 的像素点的灰度值 $I(u, v)$ 记为 0, 见式(7). 二值化的结果如图 3(c) 所示.

$$I(u, v) = \begin{cases} 0 & D(u, v) > 0 \\ 255 & D(u, v) = 0 \end{cases} \quad (7)$$

对于图 3(c) 中的空洞, 使用区域连通性分析, 对每一个空洞进行标记, 如图 3(d) 为标记完的结果, 每一种颜色代表一个连通的空洞区域.

虽然无法使用 Kinect 直接获得空洞区域的深度信息, 但是这些空洞区域周围的像素点是有深度信息的, 可以充分利用周围像素的深度信息, 来预测和估算空洞区域的深度信息. 对于每一个空洞, 这里以图 3(e) 所示的空洞区域 H_O 为例, 首先对其进行形态学腐蚀膨胀处理, 扩大空洞区域的范围, 使得该空洞周围的一些像素点被包含进来, 得到如图 3(f) 所示的区域 H_{ED} , 通过式(8) 将腐蚀膨胀之后的区域与原始的空洞区域做差, 就可以得到空洞区域的邻域 H_{NP} , 如图 3(g) 的所示.

$$H_{NP} = H_{ED} - H_O \quad (8)$$

计算 H_{NP} 中所有像素点的深度值的方差, 然后根据计算出的深度方差的大小判断空洞类型, 见式(9):

$$\text{std}_{\text{depth}} \begin{cases} > \text{Threshold}_{\text{depth}}, \text{第一类空洞} \\ \leq \text{Threshold}_{\text{depth}}, \text{第二类空洞} \end{cases} \quad (9)$$

①如果深度值方差 $\text{std}_{\text{depth}}$ 小于等于阈值 $\text{Threshold}_{\text{depth}}$, 则该空洞属于第一类空洞, 即在物体中间的空洞. 这里假设物体表面是分片光滑的, 因此物体表面的一个小区域中的深度值也应该是相近的.

②如果深度值方差 $\text{std}_{\text{depth}}$ 大于阈值 $\text{Threshold}_{\text{depth}}$, 说明该空洞周围的像素点的深度差别较大, 空洞应该出现在边缘区域, 该空洞属于第二类空洞.

当空洞出现在物体边缘时, 其邻域往往会同时包含前景和背景的像素点, 如果直接使用所有像素点进行预测, 得到的结果往往是不准确的. 本文提出了颜色一致性约束, 找出邻域中与空洞颜色较为接近的像素点, 使用符合颜色一致性约束的像素点来进行预测, 从而得到较为准确的结果.

首先在彩色图像中找到空洞 H_O 所对应的区域 C_O , 统计 C_O 中所有彩色信息, 根据式(10) 计算 C_O 平均 RGB 颜色值 $\tilde{V}$:

$$\tilde{V} = \text{mean}_{i=R,G,B} \{V_i(u,v), (u,v) \in C_O\} \quad (10)$$

对于 H_{NP} 中的每一个的像素点 (u,v) , 在彩色图像中找到对应的坐标 (u_c, v_c) , 验证其是否满足颜色一致性约束:

$$\Delta = \sum_{i=R,G,B} (\tilde{V}_i - V_i(u_c, v_c))^2 \quad (11)$$

$$\Delta \begin{cases} \leq \text{Threshold}_{\text{color}}, \text{满足颜色一致性约束} \\ > \text{Threshold}_{\text{color}}, \text{不满足颜色一致性约束} \end{cases} \quad (12)$$

根据式(11) 计算点 (u_c, v_c) 与 C_O 的平均颜色值的差异 Δ , 如果 Δ 小于等于颜色一致性阈值 $\text{Threshold}_{\text{color}}$, 说明该像素点与空洞区域的颜色较为接近, 可以用来预测空洞区域的深度; 反之, 如果 Δ 大于 $\text{Threshold}_{\text{color}}$, 说明该像素点与空洞区域的颜色差异较大, 很有可能是属于其他区域的点, 不能用来预测空洞区域的深度. 通过颜色一致性约束, 可以提高深度预测的准确度和可信度.

在进行颜色一致性判断时, 需要对深度图像中的像素点, 找到其在彩色图像中对应的像素点. 在 2.1 节中, 通过坐标变化对深度图和彩色图进行了对准, 这样一来就可以在彩色图像中较为准确的找到对应点, 最终预测出的深度也更加可信.

H_{NP} 满足颜色一致性约束的像素点的集合记为 H_c , 对 H_c 中的所有像素点的深度值进行直方图统计, 然后对统计结果进行加权平均, 所得的结果作为 H_O 的预测深度值 D_0 :

$$D_0 = \sum \omega * \{ \text{hist} \{ D_{H_c}(u,v) \} \} \quad (13)$$

式中, 权重 ω 是该深度值的像素个数除以 H_c 中的像素总数.

2.4 改进的双边滤波

经过补洞处理之后的 Kinect 深度图像, 往往仍存在少量的深度信息缺失空洞, 同时图像整体噪声仍然较大, 需要使用滤波算法进行滤波对图像整理进行. 普通的滤波算法, 例如高斯滤波、均值滤波等, 虽然能够很好的对图像进行平滑, 但是往往是物体的边缘变得模糊, 丢失大量的高频信息. 双边滤波可以在去除系统噪声平滑图像的同时很好的保留图像的边缘信息, 被广泛的使用的深度图像处理中. 双边滤波公式如下:

$$S(p) = \frac{1}{w} \sum_{q \in N_p} I(q) G_s(\|p-q\|) G_I(\|I(p) - I(q)\|) \quad (14)$$

$$w = \sum_{q \in N_p} G_s(\|p-q\|) G_I(\|I(p) - I(q)\|) \quad (15)$$

对于深度图像中的一个像素点 p , q 是 p 的邻域 N_p 中的点, S 表示滤波之后的图像, G_s 基于像素距离的高斯函数, G_I 是基于像素颜色差异的高斯函数, I 为深度图像对应的彩色图像, w 为归一化常数.

但是双边滤波往往会带来伪边缘效应, 即彩色图像中的物体上的文字和图案, 会被错误地当成边缘, 出现在最后的深度图像当中, 同时也不能够有效地去除外点, 并且计算量相对较大. 为了解决上述问题, 本文提出该并双边滤波算法:

$$S(p) = \frac{1}{w} \sum_{q \in N_p} D(q) G_s(\|p-q\|) G_d(\|D(p) - D(q)\|) \tilde{G}_I(\|I(p) - I(q)\|) \quad (16)$$

其中:

$$G_d(\|D(p) - D(q)\|) = \begin{cases} 1, & \|D(p) - D(q)\| < \delta_d \\ 0, & \|D(p) - D(q)\| \geq \delta_d \end{cases} \quad (17)$$

$$\tilde{G}_I(\|I(p) - I(q)\|) = \begin{cases} 1, & \text{std}_{D(N_p)} \leq \text{Threshold}_{D(N_p)} \\ \exp\left\{-\frac{(I(p) - I(q))^2}{2\sigma_I^2}\right\}, & \text{std}_{D(N_p)} > \text{Threshold}_{D(N_p)} \end{cases} \quad (18)$$

G_d 用来去除外点, 当邻域内的像素点与 p 点的深度差大于阈值 δ_d 时就被认为是外点, 并在计算过程中忽略. $\text{std}_{D(N_p)}$ 表示 p 的邻域中深度方差, 如果 $\text{std}_{D(N_p)}$ 小于等于阈值 $\text{Threshold}_{D(N_p)}$ 说明邻域中所有像素点的深度相差不大, 邻域处于平坦区域, 可以直接用高斯函数进行平滑; 如果 $\text{std}_{D(N_p)}$ 大于 $\text{Threshold}_{D(N_p)}$, 说明邻域处于边缘区域, 需要使用基于颜色差异的高斯函数来保留边缘高频信息. 即使在彩色图像的物体中间出现了文字或者图案的区域, 但是在深度图像中对应区域仍然为平坦区域, G_I 的值为 1, 双边滤波器变为普通的高斯滤波器, 这样就可以抑制伪边缘效应, 同时也能一定程度上减少计算量.

3 实验结果与分析

在 PC 机上基于 Matlab 语言对提出的 Kinect 深度图像修复算法进行了实现, 程序运行环境为

Windows7 64 位, Matlab2012a, 计算机配置为 Intel(R) Core (TM) i5 750@2.67 GHz 四核 CPU, 4 GB 内存. Kinect 驱动为 Kinect SDK for Windows v1.6. 本实验使用 Kinect 采集了不同场景的深度图像和彩色图像, 图像大小均为 640×480 .

为了进行时间相关性分析, 对同一场景进行连续拍摄, 然后对近 1 000 帧深度图像进行时间相关性分析. 在图 5(a) 中随机选取一点 p , 图 5(c) 为 p 在所有帧深度图像中的深度分布, 可以看出深度是非常不稳定的, 图 5(a) 中红色方框标记的区域是由于 Kinect 数据不稳定性产生的系统随机噪声. 图 5(c) 是对图 5(a) 中的点 p 的深度直方图统计的结果, 选取直方图中峰值作为 p 点的深度值 $D(p)$. 在图 5(b) 为根据时间相关性优化的结果, 对比图 5(b) 和图 5(a) 中对应红色方框区域可以发现, 对于在单帧图像中深度信息缺失的像素点, 利用多帧深度数据的时间相关性进行预测, 可以获得较为可信的深度值, 同时也提高了深度数据的稳定性, 为后续的处理提供更加可靠的深度数据.

如图 5(e), 选取原始深度图像中物体的部分进行放大显示. 利用空间相关性对原始深度图像中的黑色空洞区域进行填补, 在无颜色一致性约束的结果中可以发现空洞区域的最终于物体的深度较为一致, 而实际上该空洞区域应该出现在背景当中, 而在结合颜色一致性约束进行补洞的结果中, 可以看出空洞区域的深度与背景较为一致, 对深度的预测更加准确, 更符合实际情况. 在本文的实验中, $\text{Threshold}_{\text{color}}$ 的值为 10, $\text{Threshold}_{\text{depth}}$ 的值为 10, 腐蚀膨胀操作选取的结构元为半径为 6 的 disk.

图 5(f) 是使用不同滤波算法对深度图像进行整体滤波的结果, 选取物体边缘区域进行放大显示, 通过对比边缘细节可以看出, 均值滤波和高斯滤波均会使边缘变得模糊, 而本文提出的改进双边滤波可以得到清晰边缘. 实验中, 均值滤波的模板大小为 3×3 , 高斯滤波的模板大小为 3×3 , 高斯函数的方差为 0.5, G_s 中的 σ_r 为 0.5, σ_d 为 10, σ_l 为 0.5, $\text{Threshold}_{D(Np)}$ 为 10, 模板大小为 3×3 .

另外, 实验对不同的场景进行了拍摄, 从图 6 中可以看到深度图像修复的结果. 通过对比上下两行可以发现, 单修复后的深度图像中基本不存在空洞, 图像整体平滑边缘清晰, 图像质量大大提高. 但是对于图 6 中红色区域标记的这种较大的空洞, 由于面积较大邻域信息较复杂, 修复效果不理想.

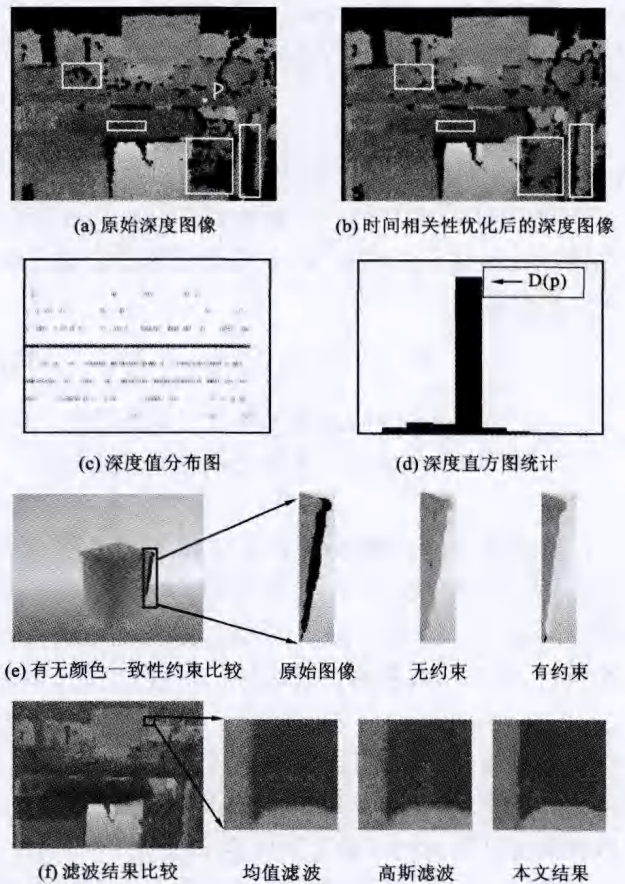


图 5 实验结果

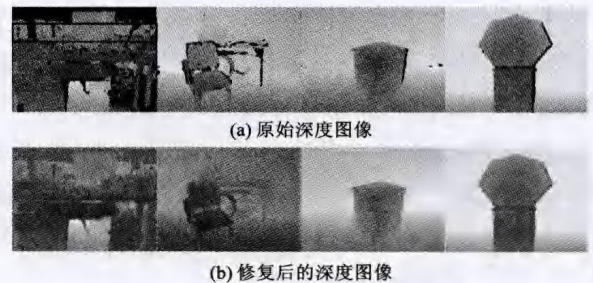


图 6 深度图像修复前后对比图

4 结束语

Kinect 深度图像存在大量空洞和图像噪声, 本文提出的算法充分利用深度数据之间的时间相关性和空间相关性, 同时结合彩色图像的颜色一致性约束, 能够有效地填补空洞, 提高数据稳定性. 同时通过对原始双边滤波器的改进, 解决了伪边缘效应问题并减少了计算量, 最终得到整体平滑边缘清晰的深度图像. 但本文算法也存在一定的不足, 对于较大面积的空洞修复效果不理想, 希望通过对空洞的进一步分割来解决这个问题.

参考文献:

- [1] Fu Jingjing, Wang Shiqi, Lu Yan, et al. Kinect-like depth denoising[C]//Circuits and Systems (ISCAS). Seoul: IEEE, 2012:512-515.
- [2] Massimo Camplani, Luis Salgado. Efficient spatio-temporal hole filling strategy for kinect depth maps[C]//Three-Dimensional Image Processing (3DIP) and Applications II. California: SPIE, 2012:82-90.
- [3] Milani S, Calvagno G. Joint denoising and interpolation of depth maps for MS Kinect sensors[C]//Acoustics, Speech and Signal Processing (ICASSP). Koyoto: IEEE, 2012:797-800.
- [4] Xu Kang, Zhou Jun, Wang Zhen. A method of hole-filling for the depth map generated by kinect with moving objects detection[C]//Broadband Multimedia Systems and Broadcasting (BMSB). Seoul: IEEE, 2012: 1-5.
- [5] 王奎,安平,张兆杨,等. Kinect 深度图像快速修复算法[J]. 上海大学学报:自然科学版,2012,18(5):454-458.
- [6] Matyunin S, Vatolin D, Berdnikov Y, et al. Temporal filtering for depth maps generated by Kinect depth camera[C]//3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON). Antalya: IEEE, 2012:1-4.
- [7] Johannes Kopf, Michael F Cohen, DANI Lischinski, et al. Joint bilateral upsampling[J]. Transaction on Graphics, 2007, 26(3):96-102.

作者简介:

徐欢 男,(1988-),硕士研究生.研究方向为多媒体信息处理与计算机视觉.

郑江滨 男,(1971-),教授,博士生导师.研究方向为多媒体信息处理与计算机视觉.

(上接第 97 页)

- [2] 张琳,王汝传,张永平.一种基于模糊集合的可用于网络环境的信任评估模型[J].电子学报,2008,36(5):862-868.
- [3] Shunan Ma, Jingsha He, Shuai Xunbo. Application of fuzzy comprehensive evaluation method in trust quantification [J]. International Journal of Computational Intelligence Systems, 2011,4(5):768-776
- [4] 王守信,张莉,李鹤松.一种基于云模型的主观信任评价方法[J].软件学报,2010(6):1341-1352.
- [5] 路峰,吴慧中.网格环境下基于云模型的信任评估与决策方法研究[J].系统仿真学报,2009,21(2):421-426.
- [6] Wang Shangguang, Zheng Zibin, Sun Qibo, et al. Cloud model of service selection[C]//Proc of IEEE INFOCOM 2011 Workshop on Cloud Computing. Shanghai:IEEE,2011:677-682.
- [7] 蔡红云,杜瑞忠,田俊峰.基于多维信任云的信任模型研究[J].计算机应用,2012,32(1):5-7.

作者简介:

蔡红云 女,(1980-),硕士,副教授.研究方向为云计算、可信计算.

田俊峰 男,(1965-),博士,教授.研究方向为网络安全、可信计算.