

软分割约束边缘保持插值的半自动 2D 转 3D

吴少群, 袁红星, 安 鹏, 程培红

(宁波工程学院电子与信息工程学院, 浙江宁波 315016)

摘 要: 半自动 2D 转 3D 将用户标注的稀疏深度转换成稠密深度, 是解决 3D 片源不足的主要手段之一. 针对现有方法利用硬分割增强深度边缘引入误差的问题, 提出像素点与超像素深度一致性约束的边缘保持插值方法. 首先, 建立像素点深度和超像素深度传播的能量模型, 通过像素点与所属超像素间深度差异的约束项将二者关联起来; 其次, 利用矩阵表示形式将两个能量模型的最优化转换成稀疏线性方程组的求解问题. 通过超像素提供的约束项, 可避免深度传播穿过低对比度边缘区域, 从而能保持对象边缘. 实验结果表明, 本文方法对象边缘处深度恢复的准确性优于融合图割的随机游走方法, PSNR 改善了 1.5dB 以上.

关键词: 二维转三维; 图割; 随机游走; 软分割

中图分类号: TN911.73

文献标识码: A

文章编号: 0372-2112 (2015)11-2218-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.11.012

Semi-Automatic 2D-to-3D Conversion Using Soft Segmentation Constrained Edge-Aware Interpolation

WU Shao-qun, YUAN Hong-xing, AN Peng, CHENG Pei-hong

(School of Electronics and Information Engineering, Ningbo University of Technology, Ningbo, Zhejiang 315016, China)

Abstract: Semi-automatic 2D-to-3D conversion is a promising solution to 3D stereoscopic content creation. Its main process is to estimate the dense depth map from user-defined strokes on the image. Existing methods preserve depth boundaries by incorporating hard segmentation. However, the inexact segmentation around object boundaries will decrease depth accuracy around these regions. To help solve this problem, an edge-aware interpolation method is developed which is constrained by depth consistency between pixels and superpixels. First, we formulate depth propagation in terms of two energy functions of pixels and superpixels, which are influenced by each other through the constraint of soft segmentation. Second, the energy functions are reformulated in matrix forms and they are solved jointly in a sparse linear equation. We recover depth boundaries with help of the superpixels constraint which prevents depth propagation across low contrast edge regions. Experimental comparisons with existing algorithms show that our method demonstrates significant advantages over object boundaries. The PSNR is improved by more than 1.5 dB compared with hybrid graph-cuts and random-walks approach.

Key words: 2D-to-3D conversion; graph-cuts; random-walks; soft segmentation

1 引言

随着影视业步入 3D 时代, 市场对 3D 片源的需求急剧增加. 2D 转 3D 是解决片源不足的重要突破口, 其核心技术在于通过 2D 画面产生准确的深度图^[1]. 基于追求良好 3D 品质与转换效率方面的考量, 结合人工交互的半自动转换是该领域的未来发展趋势^[2].

现有方法^[3-5]大多采用基于边缘保持的插值^[6-8]进行深度传播, 即假设相似的颜色或亮度具有相似的深

度, 根据相邻点颜色或亮度差异构造权重, 建立近邻传播系数矩阵, 通过求解线性方程组实现用户标注深度向整个图像域的传播. 然而, 这种传播方法没有充分考虑相邻像素点深度的全局一致性, 即深度跳变仅发生在对象边缘处, 对象内部深度渐进变化^[9]. 因而, 上述方法很难阻止深度传播穿过低对比度边缘区域, 导致估计深度的对象边缘较模糊. 图 1(d) 是利用文献^[6]中近邻传播方法获得的深度图. 从放大的白色方框中已很难分辨出女孩脖子与背景交界处的深度边缘.

Phan 等^[10]和 Xu 等^[11]试图通过图像硬分割提供的初始深度抑制该现象的发生.然而,硬分割产生的错误对象边缘会影响估计深度的准确性.针对该问题,本文提出软分割约束边缘保持的深度传播方法.与 Phan 等^[10]和 Xu 等^[11]不同的是,我们要求像素点的深度与所属超像素的深度逼近,从而可避免硬分割边缘错误带来的误差,并获得边缘保持的平滑深度图.本文将超像素和像素点之间的这种深度一致性关系称为软分割约束.

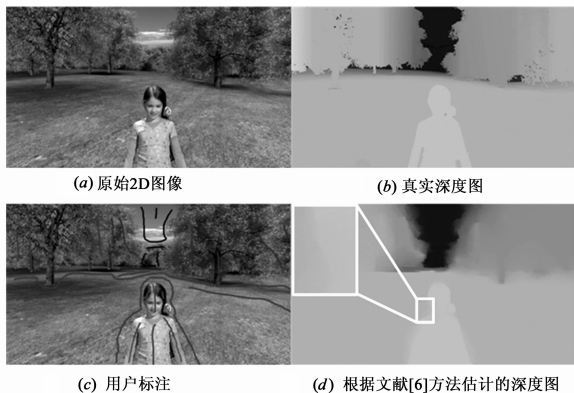


图1 一个半自动2D转3D深度估计实例

2 相关工作

目前,2D 转 3D 技术主要分为三类:人工转换,全自动转换和半自动转换.人工转换能够提供最佳的深度图,但也是最耗时和成本最高的方法,视频转换费用高达 10 万美金每分钟^[3].全自动方法试图通过场景内各种深度线索进行深度估计^[12].然而,这些线索和深度间的关系不是绝对的,使得估计深度难以达到用户期望的品质要求.半自动方法利用图像的局部相似性,将用户分配的稀疏深度传播到整个图像域.作为品质和效率最佳均衡的转换方式,近年来半自动方法迅速得到学术界和产业界的追捧.

Guttman 等^[4]首先在视频的首末帧上为部分像素分配深度值;然后采用边缘保持的插值方法^[6]将用户指定的深度传播到首末帧;其次提取视频帧的 SIFT 特征点,通过 SVM 学习获得中间帧的控制点,将用户在首末帧上的深度扩散到控制点;最后在时空一致性和控制点约束条件下建立线性方程组,通过最小二乘法得到所有帧的估计深度.该方法可得到较理想的深度,但其计算过程复杂,涉及 SIFT 特征点提取, SVM 训练、分类和两次线性方程组的求解.为此, Rzesutek 等^[5]提出基于随机游走的方法,将用户标注的深度映射为像素点对应前景的概率,然后计算未标注像素点到达标注点的概率,最后将概率重新映射成深度值.由于随机游走的平滑性,该方法估计的深度图丢失了对象边缘.而深度图的对象边缘是影响用户 3D 感官体验的重要因

素^[3]. Wang 等^[3]试图通过图像的边缘信息保持估计深度图的边缘,在用户标注深度的扩散过程中,将处于图像边缘位置的邻域像素点分配较小的权重,以阻止深度传播穿过对象边缘. Phan 等^[10]将用户标注点视为种子像素,用户分配的深度作为标注值;然后利用图割算法对图像进行多标注分类,与种子相连的区域被分配与种子相同的深度;其次根据图割算法得到的初始深度,对随机游走算法的边权重进行调节,初始深度越相似,邻域像素点间边的权重越大;最后通过随机游走算法得到最终的估计深度.图割算法的优势是可以产生较尖锐的对象边缘.结合图割算法的硬分割特性和随机游走的平滑性, Phan 等^[10]可获得对象边缘保持的平滑深度图.然而,对于复杂场景图割算法难以准确分离出对象边缘.图割算法这一缺陷影响了 Phan 等^[10]方法的性能. Xu 等^[11]针对图割算法计算复杂度高的问题,提出将分水岭分割替代图割,但该方法同样存在分割对象边缘的准确提取问题.

3 算法流程

本文算法流程如图 2 所示,主要分为用户标注、标注提取与深度转换、超像素分割和软分割约束边缘保持的插值四个部分:(1)用户标注,利用画笔在输入图像上为部分像素点分配深度,画笔颜色越亮表示距离摄像机越近;(2)标注提取与深度转换,提取用户输入的画笔标注,并根据颜色深浅转换成相对深度,得到人工设定的稀疏深度图;(3)超像素分割,根据纹理、颜色和亮度等相似性,通过过分割算法^[13]将图像划分成若干个图像块;(4)软分割约束边缘保持的插值,根据像素点深度与所属超像素深度一致性的约束条件,建立边缘保持的深度插值能量模型,通过求解该模型得到最终估计的深度图.

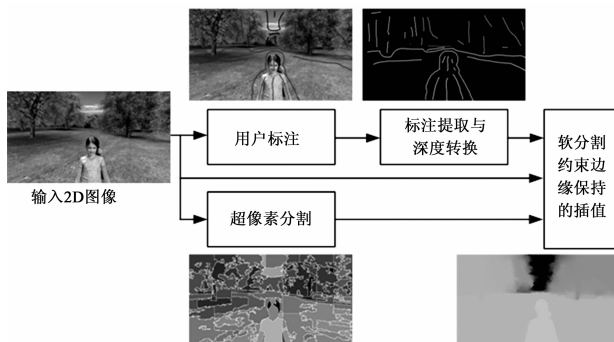


图2 本文算法流程

4 软分割约束边缘保持插值的深度传播模型

假设输入 2D 图像的像素点个数为 N ;过分割得到的超像素个数为 M ;用户输入稀疏深度图的列向量表示为 \mathbf{u}^p .则待估计深度 \mathbf{d} 的能量函数定义如下:

$$E(\mathbf{d}) = \sum_{i,j} w_{ij}^p (d_i - d_j)^2 + \lambda \sum_{i=1}^N m_i^p (d_i - u_i^p)^2 + \eta \sum_{i=1}^N (d_i - \sum_{k=1}^M w_{ik}^s d_k^s)^2 \quad (1)$$

$$w_{ij}^p = \begin{cases} \exp(-\beta \| \mathbf{I}_i - \mathbf{I}_j \|^2), & \text{若像素点 } i \text{ 与像素点 } j \text{ 相邻} \\ 0, & \text{否则} \end{cases} \quad (2)$$

式(1)中 d_i 表示第 i 个像素点的深度; w_{ij}^p 为像素点 i, j 之间的相似度, 其定义如式(2)所示, 其中 β 为空间尺度系数, \mathbf{I}_i 表示第 i 个像素点的 RGB 颜色值; m_i^p 为用户标注的像素点掩膜, 若像素点 i 位于用户标注区域则 $m_i^p = 1$, 否则 $m_i^p = 0$; d_k^s 表示第 k 个超像素的深度; w_{ik}^s 为像素点 i 与超像素 k 间的关联系数, 若像素点 i 属于超像素 k 则 $w_{ik}^s = 1$, 否则 $w_{ik}^s = 0$; λ 用于控制标注区域估计像素点深度和标注像素点深度之间的逼近程度; η 用于控制像素点与所属超像素之间的深度差异。

通过最小化式(1)得到待估计的深度图. 然而式(1)中包含了未知量 \mathbf{d}^s , 为此定义 \mathbf{d}^s 的能量函数, 如式(3)所示.

$$E(\mathbf{d}^s) = \sum_{k,l} w_{kl}^s (d_k^s - d_l^s)^2 + \lambda \sum_{k=1}^M m_k^s (d_k^s - u_k^s)^2 + \theta \sum_{k=1}^M (d_k^s - \sum_{i=1}^N w_{ki}^p d_i)^2 \quad (3)$$

$$w_{kl}^s = \sum_{(m,n) \in B_{kl}} \exp(-\beta \| \mathbf{I}_m - \mathbf{I}_n \|^2) \quad (4)$$

式(3)中 w_{kl}^s 为超像素 k, l 之间的相似度, 定义如式(4)所示, 其中 B_{kl} 为超像素 k, l 之间相邻像素点构成的边集合; m_k^s 表示用户标注的超像素掩膜, 若超像素 k 包含了用户标注的像素点则 $m_k^s = 1$, 否则 $m_k^s = 0$; \mathbf{u}^s 为用户标注的超像素深度列向量表示, 若超像素 k 含有用户标注的像素点则 u_k^s 为其内所有标注点深度的平均值, 否则 $u_k^s = 0$; λ 用于控制标注区域估计的超像素深度和标注超像素深度之间的逼近程度; θ 用于控制超像素深度与其内所有像素点平均深度间的差异; w_{ki}^p 表示超像素 k 与像素点 i 间的关联系数, 定义为 $w_{ki}^p = \frac{w_{ik}^s}{\sum_j w_{jk}^s}$.

5 深度传播模型的求解

为得到估计的深度图, 需最小化式(1)和式(3)表示的能量函数. 为此, 将其表示成矩阵形式, 分别如式(5)和式(6)所示.

$$E(\mathbf{d}) = \mathbf{d}^T \mathbf{L}^p \mathbf{d} + (\mathbf{d} - \mathbf{u}^p)^T \mathbf{A}^p (\mathbf{d} - \mathbf{u}^p) + \eta (\mathbf{d} - \mathbf{W}^{ps} \mathbf{d}^s)^T (\mathbf{d} - \mathbf{W}^{ps} \mathbf{d}^s) \quad (5)$$

$$E(\mathbf{d}^s) = (\mathbf{d}^s)^T \mathbf{L}^s \mathbf{d}^s + (\mathbf{d}^s - \mathbf{u}^s)^T \mathbf{A}^s (\mathbf{d}^s - \mathbf{u}^s) + \theta (\mathbf{d}^s - \mathbf{W}^{sp} \mathbf{d})^T (\mathbf{d}^s - \mathbf{W}^{sp} \mathbf{d}) \quad (6)$$

式(5)中 \mathbf{L}^p 为 $N \times N$ 矩阵, 定义为 $\mathbf{L}^p = \mathbf{D}^p - \mathbf{W}^p$. 其中 \mathbf{D}^p 为 $N \times N$ 对角矩阵, $D_{ii}^p = \sum_j w_{ij}^p$. $N \times N$ 矩阵 \mathbf{W}^p

的元素定义如式(2)所示. \mathbf{A}^p 表示 $N \times N$ 对角矩阵, 若第 i 个像素点属于用户标注区域则 $A_{ii}^p = \lambda$, 否则 $A_{ii}^p = 0$. $N \times M$ 矩阵 $\mathbf{W}^{ps} = (w_{ik}^s)$ 由像素点与超像素的关联系数构成. 由于一个像素点只能属于一个超像素, 因而 \mathbf{W}^{ps} 的每一行仅有一个非零元素.

式(6)中 \mathbf{L}^s 为 $M \times M$ 矩阵, 定义为 $\mathbf{L}^s = \mathbf{D}^s - \mathbf{W}^s$. 其中 \mathbf{D}^s 为 $M \times M$ 对角矩阵, $D_{kk}^s = \sum_l w_{kl}^s$. $M \times M$ 矩阵 $\mathbf{W}^s = (w_{kl}^s)$ 的元素定义如式(4)所示. \mathbf{A}^s 表示 $M \times M$ 对角矩阵, 若第 k 个超像素包含了用户标注的像素点则 $A_{kk}^s = \lambda$, 否则 $A_{kk}^s = 0$. $M \times N$ 矩阵 $\mathbf{W}^{sp} = (w_{ki}^p)$ 由超像素与像素点的关联系数构成. 由 w_{ki}^p 的定义可知 \mathbf{W}^{sp} 是 \mathbf{W}^{ps} 转置后, 每一行再除以该行非零元素的个数.

为获得最小化式(5)的解, 将式(5)和式(6)分别对 \mathbf{d} 和 \mathbf{d}^s 求导, 得到如式(7)所示的稀疏线性方程组, 其中 \mathbf{I}_N 和 \mathbf{I}_M 分别表示 $N \times N$ 和 $M \times M$ 的单位矩阵.

$$\begin{bmatrix} \mathbf{L}^p + \mathbf{A}^p + \eta \mathbf{I}_N & -\eta \mathbf{W}^{ps} \\ -\theta \mathbf{W}^{sp} & \mathbf{L}^s + \mathbf{A}^s + \theta \mathbf{I}_M \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \mathbf{d}^s \end{bmatrix} = \begin{bmatrix} \mathbf{A}^p \mathbf{u}^p \\ \mathbf{A}^s \mathbf{u}^s \end{bmatrix} \quad (7)$$

本文采用预条件共轭梯度算法求解式(7)^[14], 然后将列向量 \mathbf{d} 重新表示成 $2D$ 矩阵形式, 即得到最终估计的深度图.

6 实验与分析

深度图重构实验采用清华大学宽带网数字媒体技术实验室提供的测试数据^[15], 视点合成实验采用微软的视频序列^[16]. 深度图重构重点比较了其中四个有代表性的测试场景 Dice-2 (颜色歧义性、尖锐边缘), Philips-the-3D-experience (无纹理区域), Philips-the-3D-experience-2 (无纹理区域、尖锐边缘、遮挡) 和 HeadRotete (颜色歧义性、遮挡).

实验中定量比较的指标有: (1) 峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR), 具体定义为 $\text{PSNR} = 10 \log(255^2 N / \sum_{i=1}^N (d_i - \tilde{d}_i)^2)$, 其中 $\tilde{\mathbf{d}}$ 为真实深度图的列向量表示; (2) 真实深度图和估计深度图之间的归一化互协方差 (Normalized Cross-Covariance, NCC)^[17], 具体定义为 $\text{NCC} = \frac{1}{N\sigma\bar{\sigma}} \sum_i (d_i - \mu)(\tilde{d}_i - \bar{\mu})$, 其中 σ 和 $\bar{\sigma}$ 分别是估计深度图和真实深度图的标准差, μ 和 $\bar{\mu}$ 分别是二者的均值. NCC 的取值区间为 $[-1, 1]$, 其值越接近 1 说明估计的深度图越接近真实深度图.

本文方法的参数设置在所有实验中均相同, 具体如表 1 所示. 图 2 中超像素分割采用了文献[13]中的 SLIC 算法, 实现时利用了 MATLAB 工具箱 VLFeat 提供的 API 函数 `vl_slic`^[18], 其中区域大小设为 10, 空间正则化参数设为 0.5.

表 1 本文方法参数设置

β	λ	η	θ
30	1000	0.1	1

实验中,将本文方法和图割(GC)、随机游走(RW)、融合图割与随机游走的方法(RW + GC)^[10],边缘保持的最优化插值(OPT)^[6]进行比较.为进行性能比较,首先按照真实深度图的远近变化关系,在近景和远景处的部分区域分别用亮色和暗色画笔予以标注;其次将这些标注转换成稀疏深度图;然后利用上述算法将稀疏深度图转换成稠密深度图;最后将估计的深度图与真实深度图进行比较.

6.1 深度图重构定量比较

表 2 和表 3 分别给出了 GC、RW、RW + GC 和本文方法对四个测试数据估计深度图的 PSNR 和 NCC.从表 2 和表 3 可看出,本文引入软分割约束后,明显改善了深度图的估计质量,本文方法的 PSNR 和 NCC 指标在四个测试数据上均是最好的.与 RW + GC 方法相比,本文方法 PSNR 提高了 1.5dB 以上.除了 Dice-2 外,在其他三个测试数据上 RW + GC 算法的性能均不如 RW 算法.这是因为,对于复杂场景 GC 很难得到准确的对象边缘,导致估计深度图在对象边缘处误差较大.此时利用 GC 得到的深度图对 RW 的权重进行调节,虽然可以增强估计深度图的边缘,但也会引入误差.与 RW + GC 不同的是,本文并不是根据分割结果调节相邻像素间的相似度,而是根据过分割得到的超像素增加一个约束项,通过控制超像素和其内像素点深度的差异达到增强深度边缘的目的.因而本文方法既可增强估计深度的边缘,又能避免 RW + GC 中因分割错误带来的边缘误差.

表 2 深度图 PSNR 比较(单位: dB)

数据 方法	Philips-the-3D-experience	Philips-the-3D-experience-2	Dice-2	HeadRotete
GC	20.28	17.87	24.81	22.28
RW	22.91	19.81	24.34	24.49
RW + GC	20.36	18.49	25.29	22.27
OPT	22.87	19.59	25.13	24.85
本文方法	24.13	20.05	27.62	25.40

表 3 深度图 NCC 比较

数据 方法	Philips-the-3D-experience	Philips-the-3D-experience-2	Dice-2	HeadRotete
GC	0.6458	0.8202	0.9897	0.9481
RW	0.7384	0.8655	0.9888	0.9704
RW + GC	0.6480	0.8368	0.9908	0.9485
OPT	0.7610	0.8612	0.9907	0.9731
本文方法	0.8066	0.8760	0.9946	0.9755

6.2 深度图重构定性比较

GC、RW、RW + GC 和本文方法对四个测试数据估

计深度图视觉比较分别如图 3 ~ 6 所示.图 3 中 RW 和 OPT 方法获得的深度图边缘效果较差,从白色方框的放大区域看,已很难分辨出女孩头部和背景的深度边界. GC 虽然能够得到尖锐的对象边缘,但在天空与树木的边界分割出现较大偏差,使得这一区域出现较大的估计误差.这正是表 2 和表 3 中,融合 GC 后 RW 性能反而下降的原因.对比这些方法估计深度的白色方框区域,可发现本文方法结果与真实深度最接近.

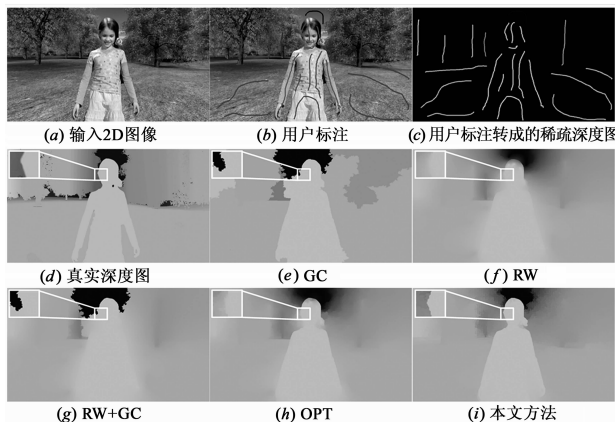


图 3 Philips-the-3D-experience 实验结果

图 4 中 GC 同样在树木与天空分界处出现了分割偏差.另外,由于 GC 很难精确分离出对象边缘,导致图 4(e),(g)中白色方框内前景对象的边缘不够完整.而 RW 和 OPT 由于没有考虑深度的全局性约束,使得对象边缘处出现晕斑现象.本文方法引入软分割约束后,同时考虑了深度的局部和全局一致性,因而在获得平滑深度图的同时能够增强深度边缘.由图 4(i)可看出,本文估计深度的对象边缘与真实深度最相似.图 4 中蒲公英叶子处深度,本文方法没有明显改善.这是因为这些地方主要是细长结构,需要超像素足够小才能保持其边缘.

图 5 中背景的深度没有变化,前景的深度内部变化

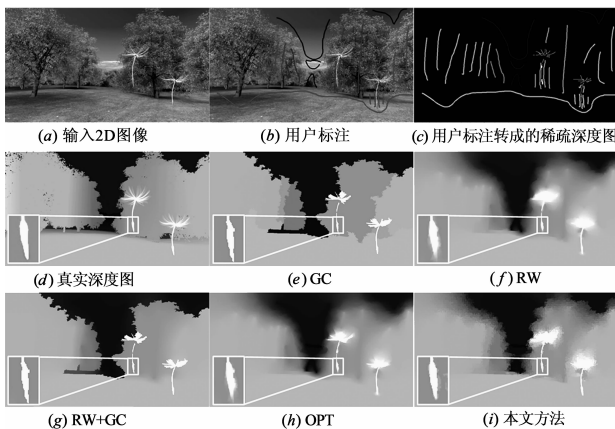


图 4 Philips-the-3D-experience-2 实验结果

较小. 这是 GC 方法在此情况下 PSNR 和 NCC 优于 RW 的原因. 但由于边缘分割不准确, 图 5(e), (g) 中对象边缘的完整性较差, 从放大区域看, 估计结果没有恢复对象边缘的弧线过渡. 本文方法既增强了对象边缘, 又较好地保持了边缘处的细节信息. 图 6(e), (g) 再次说明硬分割错误会引入较大的深度估计偏差, 如鼻子下方出现了明显错误的估计深度. 而本文方法在边缘恢复和估计深度准确性方面取得较好均衡.

6.3 视点合成比较

为验证不同方法估计深度图对视点合成的影响, 使用微软的测试序列 Ballet 和 Breakdancer^[16] 进行实验比较. 其中, 视点合成方法采用本文的前期工作^[19]. 实验时, 以第 4 路和第 5 路摄像机作为参考视点, 合成第 4 路摄像机画面. 以第 1 帧画面为例, 合成视点的 PSNR 如表

4 所示. 图 7 和图 8 是合成图像的视觉对比. 由表 4 可知, 本文方法估计深度生成的视点图像质量明显优于现有方法, 与 RW 方法相比, 本文方法 PSNR 提高了 0.3dB 以上. 从图 7 和图 8 的白色方框内放大区域可看出, 所提方法重构深度合成的视点图像更接近真实图像.

表 4 视点合成的 PSNR 比较 (单位: dB)

方法 \ 数据	Ballet	Breakdancer
GC	23.45	23.92
RW	23.03	23.63
RW + GC	23.45	23.66
OPT	23.40	23.76
本文方法	23.46	23.99

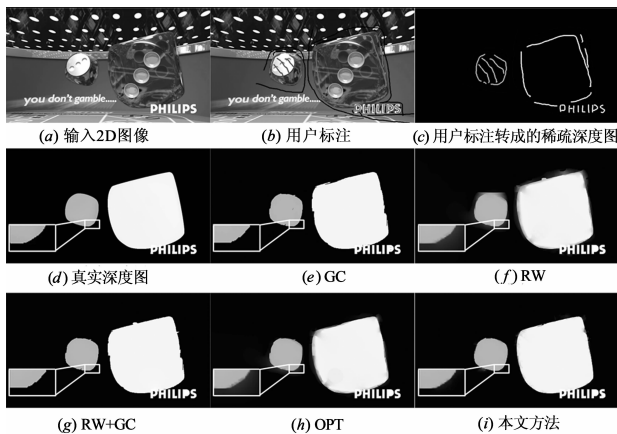


图 5 Dice-2 实验结果

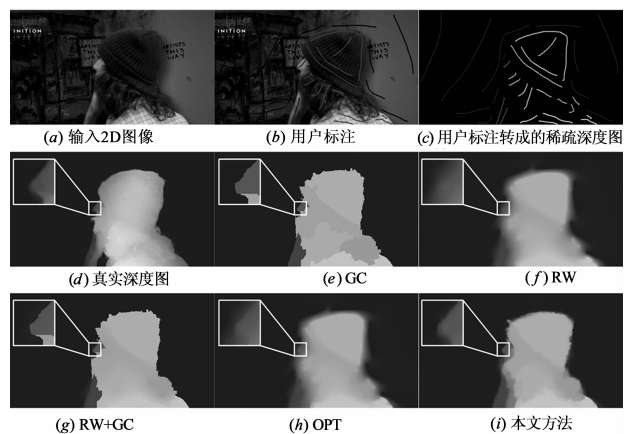


图 6 HeadRotete 实验结果

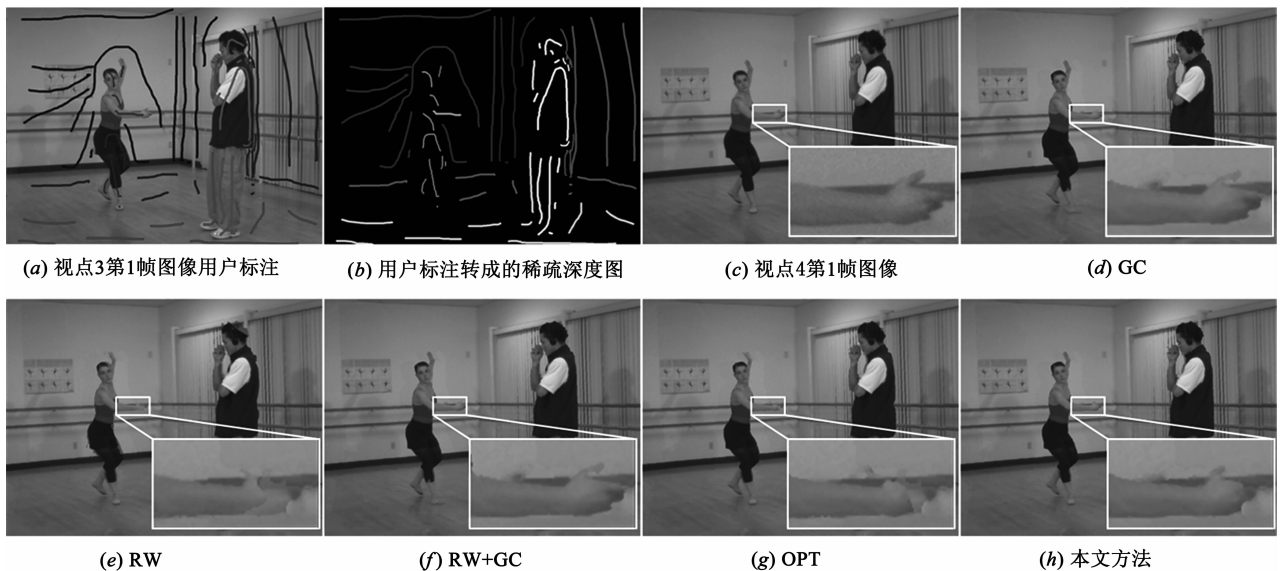


图 7 Ballet 实验结果

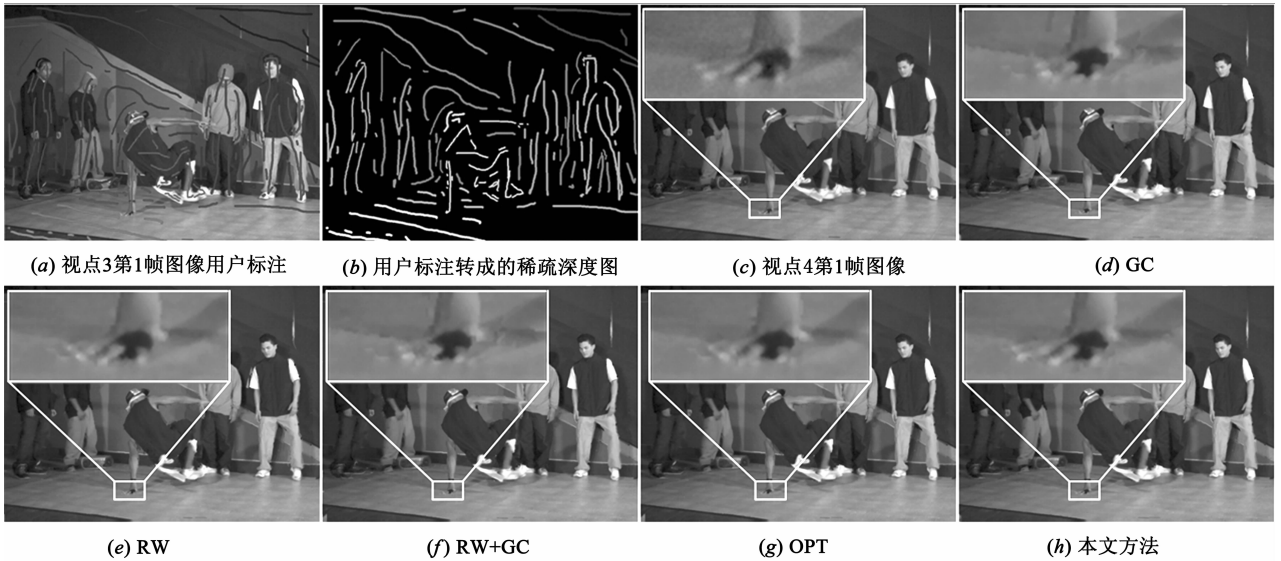


图8 Breakdancer实验结果

7 总结

半自动 2D 转 3D 面临的主要问题是如何将用户标注的稀疏深度图转换成稠密深度图,并保持对象边缘.融合图割的随机游走方法试图利用图割提供的硬分割结果增强对象边缘.然而,硬分割难以准确分离对象边缘,使得估计深度在对象边缘处误差较大.针对该问题,本文在边缘保持的插值方法基础上,引入软分割约束,具体做法是:(1)利用过分割算法将图像划分成若干个超像素,由于超像素准确保留了对象边缘信息,因而可用于抑制深度传播穿过边缘区域;(2)建立软分割约束的深度传播能量模型,包含了度量估计深度与标注深度差异的数据保真项,邻域像素点深度一致性的平滑项和像素点与超像素深度一致性的软分割约束项;(3)为求解深度传播能量模型,用类似方法建立超像素深度估计的能量模型,通过矩阵表示形式将两个能量模型转换成一个稀疏线性方程组,通过求解该方程组得到最终估计的深度图.

深度图重构实验结果表明,本文方法比融合图割的随机游走方法 PSNR 提高 1.5dB 以上.下一步,我们拟采用光流提供的运动信息,将这种方法扩展到视频上,形成时空一致性边缘保持的深度传播方法.此外,本文没有考虑用户标注误差对深度传播的影响,我们将探讨消除标注噪声的问题.

参考文献

- [1] 刘伟,吴毅红.基于图层优化与融合的 2D-3D 视频转换方法[J].计算机辅助设计与图形学学报,2012,24(11):1426-1439.
Liu Wei, Wu Yi-hong. A 2D-3D video conversion method

- based on layer optimization and intergration [J]. Journal of Computer-Aided Design & Computer Graphics, 2012, 24(11): 1426-1439. (in Chinese)
- [2] 赖文能,陈韦志.浅谈 2D 至 3D 视讯转换技术[J].影像与识别,2010,16(2):61-75.
Lai Wen-eng, Chen Wei-zhi. 2D-to-3D video conversion technologies overview [J]. Images & Recognition, 2010, 16(2): 61-75. (in Chinese)
- [3] Wang O, Lang M, Frei M, et al. StereoBrush: interactive 2D to 3D conversion using discontinuous warps [A]. Proceedings of Eurographics Symposium on Sketch-Based Interfaces and Modeling [C]. New York: ACM Press, 2011. 47-54.
- [4] Guttman M, Wolf L, Cohen-Or D. Semi-automatic stereo extraction from video footage [A]. Proceedings of IEEE International Conference on Computer Vision [C]. Los Alamitos: IEEE Computer Society Press, 2009. 136-142.
- [5] Rzeszutek R, Phan R, Androustos D. Semi-automatic synthetic depth map generation for video using random walks [A]. Proceedings of IEEE International Conference on Multimedia and Expo [C]. Los Alamitos: IEEE Computer Society Press, 2011. 1-6.
- [6] Levin A, Lischinski D, Weiss Y. Colorization using optimization [A]. Proceedings of ACM SIGGRAPH [C]. New York: ACM Press, 2004. 689-694.
- [7] 褚宏莉,李元祥,周则明,等.基于黑色通道的图像快速去雾优化算法[J].电子学报,2013,41(4):791-797.
Chu Hong-li, Li Yuan-xiang, Zhou Ze-ming, et al. Optimized fast dehazing method based on dark channel prior [J]. Acta Electronica Sinica, 2013, 41(4): 791-797. (in Chinese)
- [8] Hu Wei, Dong Zhao, Yuan Guo-dong. Edit propagation via edge-aware filtering [J]. Journal of Computer Science and

- Technology, 2012, 27(4): 830 – 840.
- [9] 袁红星, 吴少群, 朱仁祥, 等. 融合对象性和视觉显著度的单目图像 2D 转 3D [J]. 中国图象图形学报, 2013, 18(10): 1478 – 1485.
Yuan Hong-xing, Wu Shao-qun, Zhu Ren-xiang, et al. Single-view image 2D-to-3D conversion based on objectness and visual saliency[J]. Journal of Image and Graphics, 2013, 18(10): 1478 – 1485. (in Chinese)
- [10] Phan R, Androutsos D. Robust semi-automatic depth map generation in unconstrained images and video sequences for 2D to stereoscopic 3D conversion[J]. IEEE Transactions on Multimedia, 2014, 16(1): 122 – 136.
- [11] Xu X, Po L M, Ng K H, et al. Watershed and random walks based depth estimation for semi-automatic 2D to 3D image conversion [A]. Proceeding of International Conference on Signal Processing, Communications and Computing [C]. Los Alamitos: IEEE Computer Society Press, 2012. 84 – 87.
- [12] Zhang L, Vazquez C, Knorr S. 3D-TV content creation automatic 2D-to-3D video conversion[J]. IEEE Transactions on Broadcasting, 2011, 57(2): 372 – 383.
- [13] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274 – 2282.
- [14] Barrett R, Berry M, Chan T F, et al. Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods [M]. Philadelphia: SIAM Press, 1994. 35 – 48.
- [15] 曹汛. 2D to 3D conversion test sequences[EB/OL]. <http://media.au.tsinghua.edu.cn/2Dto3D/testsequence.html>, 2014-05-12/2014-05-12.
- [16] 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.
- [17] Ranftl R, Gehrig S, Pock T, et al. Pushing the limits of stereo using variational stereo estimation[A]. Proceedings of IEEE Intelligent Vehicles Symposium [C]. Los Alamitos: IEEE Computer Society Press, 2012. 401 – 407.
- [18] Vedaldi A, Fulkerson B. VLFeat: an open and portable library of computer vision algorithms[EB/OL]. <http://www.vlfeat.org>, 2008-01-01/2014-05-12.
- [19] 郁理, 郭立, 袁红星. 基于深度图像的视点绘制新方法[J]. 中国科学院研究生院学报, 2010, 27(5): 638 – 644.
Yu Li, Guo Li, Yuan Hong-xing. A novel method of depth-image-based view synthesis [J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2010, 27(5): 638 – 644. (in Chinese)

作者简介



吴少群 女, 1981 年 12 月出生, 安徽安庆人. 2011 年毕业于华东师范大学职教教研究所, 获教育学硕士学位. 现为宁波工程学院实验师, 主要从事图像与信号处理、3D 视频处理等方面的研究工作.

E-mail: vison101@163.com



袁红星 (通信作者) 男, 1980 年 10 月出生, 安徽安庆人. 2010 年毕业于中国科学技术大学电子科学与技术系, 其后在美满电子科技有限公司从事信号处理理论与技术研究工作, 2011 进入宁波工程学院电子与信息工程学院, 主要从事图像与信号处理、3D 视频处理、2D 转 3D 等方面的研究工作.

E-mail: yuanhx@mail.ustc.edu.cn