

文章编号:1007-2780(2016)03-0301-09

一种基于 MRF 的单幅图像数据的 三维重构方法研究

李 蓉^{1*}, 邓春健^{1,2}, 邹 昆^{1,2}

(1. 电子科技大学 中山学院, 广东 中山 528402;

2. 电子科技大学 计算机科学与工程学院, 四川 成都 611731)

摘要:综合分析了常见的基于图像的三维重构方法的优缺点,提出了一种基于单张图像,采用马尔科夫随机场(MRF)推断 3D 位置和方向的 3D 重构方法。该算法首先将图片分割成多个小的区域(超像素块),并假定空间场景由许多很小的平面组成,超像素块与平面相互对应,对图像中每个超像素块求取出一组特征向量(纹理、颜色等),使用 MRF 模型化平面参数之间、超像素特征向量与平面参数之间的关系,采用监督学习的方式求取相关参数,求解 MRF 模型,并根据平面参数进行场景重建。这种算法不需对场景结构做明确的假定,因此较之前的方法可以获得更多 3D 结构细节信息。用该方法对 200 张图片样本进行验算,样本中有 60% 实现了较为准确的 3D 重构。

关 键 词:三维重构;超像素块;最大后验概率;监督学习

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/YJYXS20163103.0301

3D reconstruction method based on single image data by MRF

LI Rong^{1*}, DENG Chun-jian^{1,2}, ZOU Kun^{1,2}

(1. *University of Electronic Science and Technology of China, Zhong Shan Institute, Zhong Shan 528400, China;*

2. *School of Computer Science & Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)*

Abstract: The mainstream methods of image-based 3D reconstruction are overviewed. The advantages and disadvantages of these methods are analyzed. A new 3D-reconstruction method is put forward; it

收稿日期:2015-04-13; **修订日期:**2015-06-15.

基金项目:国家自然科学基金项目(No. 61302115);广东省自然科学基金(No. S2013010015764);广东省高等学校优秀青年教师培养计划项目(No. Yq2013204, No. Yq2013206);中山市科技攻关项目(No. 2013A3PC0337);电子科技大学中山学院青年基金(No. 413YJ05);电子薄膜与集成器件国家重点实验室中山分实验室开放基金(No. 412S0605)

Supported by National Natural Science Foundation of China(No. 61302115); Science Foundation of Guangdong Province(No. S2013010015764); Cultivation Plan of Outstanding Young Teachers in Guangdong Province(No. Yq2013204, No. Yq2013206); Key Program for Science and Technology Development in Zhongshan City(No. 2013A3PC0337); Young Teacher Project of University of Electronic Science and Technology, Zhongshan Institute(No. 413YJ05); Open Funding of Nation Key Laboratory of Electronic Film and Integrated Devices, Zhongshan Branch(No. 412S0605)

* 通信联系人, E-mail: windylee007@126.com

is based upon a single image of an unstructured environment. For each small homogeneous patch in the image, a Markov Random Field (MRF) is used to infer a set of ‘parameters’ that contain both the 3D location and 3D orientation of the patch. The space is assumed being made up of many tiny planes, and each plane corresponds with a super-pixel. For each super-pixel in the image, a series of characteristic vectors such as color and texture are calculated. The relations between different planar parameters, as well as between super-pixel vectors and planar parameters are probed by using the MRF model. The MRF is trained via supervised learning. The scene is reconstructed based on the planar parameters. This algorithm does not require explicit assumptions about the structure of the scene, and compared with previous methods it enable us to gain more detailed information about the 3D structure. Using this approach, we have conducted checking computations for 200 pictures, among which 60% have been correctly 3D reconstructed.

Key words: 3D reconstruction; superpixel; MAP; supervised learning

1 引言

物体三维重构是计算机辅助几何设计(CAGD)、计算机视觉、医学图像处理和虚拟现实等领域的核心问题。在计算机内完成物体的三维构成主要有两种方式,第一种使用三维建模软件通过技术人员建模生成物体的几何三维模型,第二种通过一定的技术途径获取物体的真实几何形状。第一种技术实现已经十分成熟,现在已经有许多软件支持,如:3DS MAX、UG、Maya等,这些软件一般用数学函数的曲线曲面来表示几何形状。第二种方式一般采用高精度设备获取场景几何信息,该方法一般应用在对精度有较高的要求的场合,但其数据采集过程比较复杂,设备也比较昂贵,并且该方法使用时会受到被测量的物体属性以及环境的限制,限制了它的应用范围^[1]。

随着各种带有拍摄功能的智能电子产品的普及,获取图像的途径越来越多,基于图像数据进行3D重构成为真实三维场景建模的一个新的研究热点。在计算机视觉中,三维重构的定义是利用2D投影重建物体三维信息的数学过程和技术,一般步骤包括数据获取、预处理、点云拼接和特征分析等。目前三维重构理论涉及到的概念和理论有双目融合,灰度相关,视差,深度,基于纹理的重建,基于颜色值的重建,基于运动模型的重建,大规模场景重建等^[2-3]。

三维重构过程可以基于单视图,也可以基于多视图。一般情况下,单张图片信息是不完整的,所以在重构的过程中要利用先验知识,而基于多

视图的三维重构原理与人的双目定位类似,实现起来相对较容易,对于基于单视图和基于多视图进行重构的研究都已有不少文献。

2 相关研究

基于多张图片的三维重构算法有很多,一般步骤是在多幅图中分别找出相应的对应特征点,然后匹配对应的特征点,然后根据对应的特征点的几何约束,颜色值约束,运动模型约束等其他约束来估算出空间中对应点的深度。其中最难的部分在与怎样将每幅图像中的点对应起来,即多目融合问题。基于多张图片的三维重构方法的文献已有不少,有的是基于序列图,如文献[4],该方法能够通过对几幅序列图分析实现交互式重构,但该方法对输入的图有较高的要求,且需要使用序列图作为输入对象。有的是基于相机校准的,如文献[5]中,相机校准过程采用了基于平行六面体中的正交性约束来进行,进而实现3D重构,该方法不用进行人机交互,但建模时间较长。

基于单张图像的三维重构技术也有不少,其中成熟的算法有阴影恢复形状方法(shape from shading, SFS),该方法是基于单张图像的灰度图中的阴影线索去获取相关的三维信息,以完成3D重构。该方法有很多优点,如成本低、操作简单、效率高,因此备受关注。SFS是由Horn等^[6]最先提出的,近年来,国内外的很多学者利用阴影恢复形状法实现了基于单幅图像的三维重构,并在工业检测与测量、逆向工程及自然景物模拟等领域进行了应用。然而,SFS方法也存在下述缺陷:①图像的背

景信息被包含在重构模型中;②对大尺寸的物体重构时,由于受图像分辨率的限制,其分辨力通常无法满足工程实际要求。文献[7]提出了一种首先提取物体的二维轮廓信息,然后把二维轮廓信息与三维形貌信息与融合建立物体的数字化模型,通过B样条方法实现数字化模型的重构。该方法能够有效地去除图像的背景信息。文献[8]以3种基元为基础,通过人机交互、基元构建以及图像处理,快速得到逼真的三维场景。文献[9]通过把现有模型依据照片的剪影轮廓来行变形,把模型与剪影轮廓进行匹配,该方法能够构建复杂模型,但需要提供如基元模型之类的现有模型作为基础,且只能实现单个物体的建模。

3 理论基础

人看到一张静态图像去理解它的3D结构是一件很容易的事,因为人有图像里面这些物体的先验知识(当然得包含生理和心理效应),例如水是在地下的,天空和白云在上面,大部分的墙壁是直立的平面,桌子是方的,树是长在地面的,而且上面有叶子,还有山水、动物等等,人看到它的一面,可以联想到它的立体结构。这也是为什么人用一个眼睛也能“看到”三维世界的原因。

如上一章所述,以前为了实现3D重构,一般都是通过两个(模拟人的双眼)或者多个摄像头来获取图像,再配准的。当然,也有研究单张图像的,但是都会对场景做很多的假设,例如由明暗恢复形状 shape from shading(SFS),运用光学辅助等,使用有很大的局限性。随着技术的发展,很多带有深度的摄像头已经广泛使用,例如 kinect,但还是两个摄像头(一个彩色,一个红外用于测图像深度,也就是测场景中物体到相机的距离);以及三星开发的在传统的CMOS图像传感器里面增加可以测距的红外感光元,这样一个摄像头就可以获取彩色和深度图像了,但是目前技术还不是很成熟,应用还不广。

本文讨论的算法旨在从单张静态图像去重构其3D场景且不需要对场景做假设。图像为3D空间在2D平面的投影,那么同一个投影可能是由无限种可能的3D场景投影产生的,这就是重构的困难。但是,不是每一种可能都是合理的,我们要做的就是找到最有可能投影为这张2D图像

的3D场景,因为我们生活的环境是结构性的,也就是说物体大部分是规则和有联系的,所以我们才可以通过一些描述这类规则和联系的先验知识去预测哪一种3D场景的可能性最大,最合理。

透视投影是多对一的关系,投影线上的任何一点对应同一个像点。如果用两个摄像机,则可以消除这种多对一,从而能够确定第三维坐标Z的值,即深度信息。其原理如图1所示。

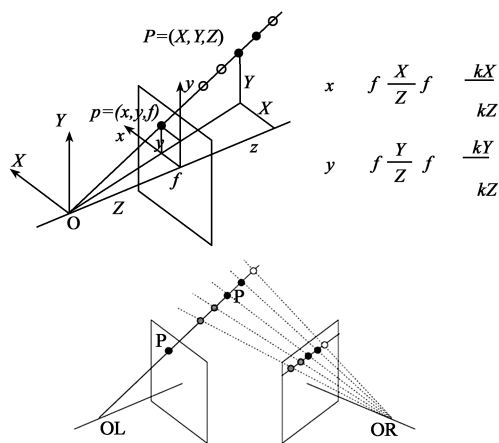


图1 透视投影

Fig.1 Perspective projection

我们可以借鉴人的理解原理,将人对场景理解的先验知识赋予计算机,让计算机通过分析图像的特征去获得额外的3D结构的信息,通过这些信息去进行3D重构。所以首先要确定图像中有哪些信息对场景的理解有用。人类单眼确立深度的信息包括:遮挡关系、线条透视、纹理梯度、明暗线索、调节聚焦等。在本论文中,基本思路如下:

用文献[10]的聚类算法将图像划分为2000个小的区域,每个区域的像素的属性相似,例如纹理、颜色等,这个区域称为超像素块,如图2。一个超像素块一般是结构的一小部分,例如墙壁或者平面的一部分等,假定空间由许多小的平面组成,超像素块与平面相互对应。

求取每个超像素块的特征向量及边缘特征向量(纹理、颜色等)。

MRF模型化平面参数之间、超像素块特征向量与平面参数之间的关系。

采用监督学习的方式求取相关参数。

求解MRF模型,并根据平面参数进行场景重构。

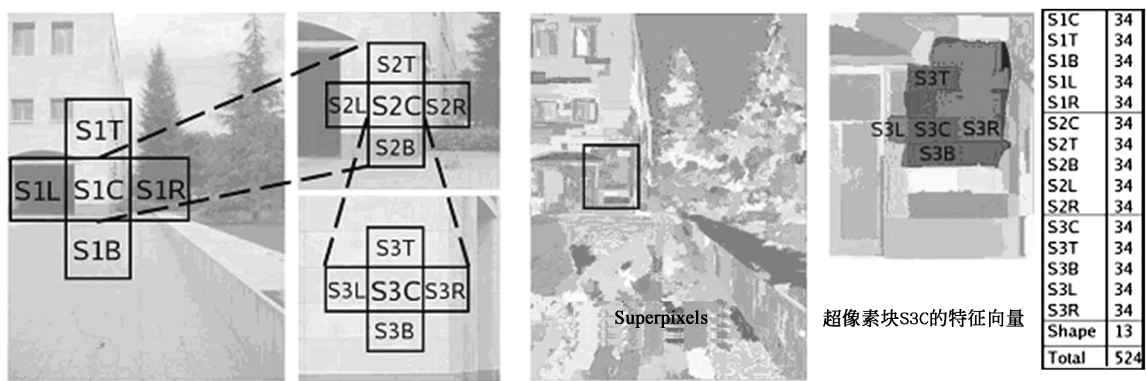


图 2 超像素块的特征向量

Fig. 2 The feature vector for a superpixel

4 MRF 模型

根据以上讨论,下面将通过提供图像和相应的深度图像去训练得到 MRF 模型,用这个模型去描述超像素块的 3D 位置和方向与图像特征的关系,即确定一个模型,我们输入一张图像的特征,然后这个模型就输出对应的 3D 位置和方向,也就实现了 3D 重构。在 MRF 模型里力争求取图像的以下属性:

- (1)图像特征与深度:超像素块的图像特征与深度(和方向)存在着一定的联系^[11];
- (2)连接结构:除了遮挡的情况,一般相邻的超像素块是最有可能互相连接的;
- (3)共面结构:相邻超像素块很有可能属于同一个平面,如果他们有相似的特征,而且两者之间没有边缘;
- (4)共线结构:图像中的长直线代表了 3D 场景中的直线,例如建筑物的边缘、人行道、窗等。

4.1 空间平面参数 MRF

在这个 MRF 中,每个节点代表一个超像素块,我们假定每个超像素块位于一个平面,我们将用参数 $\alpha \in R^3$ 推断平面的方向和位置。平面参数 α 特征:对空间平面上任意点 q ,满足 $\alpha^T q = 1$; $1/|\alpha|$ 为摄像机中心到空间平面距离;任意射线 R_i 的空间点 i 的深度 $d_i = 1/R_{iT} \alpha$ 。如图 3 所示。

$$P(a \mid X, v, y, R, \theta) = \frac{1}{Z} \prod_i f_1(a_i \mid X_i, v_i, R_i, \theta) \prod_{i,j} f_2(a_i, a_j \mid y_{i,j}, R_i, R_j), \quad (1)$$

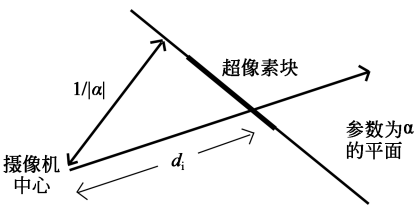


图 3 平面参数 α

Fig. 3 Plane parameter α and rays R from the camera

$f_1(\cdot)$:表征超像素块(以下简称 SP_i)的特征向量与空间平面参数之间的统计相关性。

$f_2(\cdot)$:表征 SP 对应空间平面参数彼此间的统计相关性。具体由 SP 之间的互连性、共面性、共线性、遮挡关系等几部分共同决定,而这些关系的决定通过衡量在相邻 SP 中所选择的不同像素点之间的关系来实现,即:

$$f_2(\cdot) = \prod_{\{s_i, s_j\} \in N} h_{S_i, S_j}(\cdot), \quad (2)$$

其中 s_i, s_j 为选择的像素, h 为像素间的关系。
 α_i : SP_i 对应的空间平面参数。
 X_i : SP_i 特征向量,即 $X_i = \{x_{i,s_i} \in IR^{524} : s_i = 1, \dots, S_i\}$ 。

R_i :摄像机中心到 SP_i 所包含的像素点射线集合, $R_i = \{R_{i,s_i} : s_i = 1, \dots, S_i\}$ 。

v_i : SP_i 估计深度的置信度。

$y_{i,j}$: SP_i, SP_j 对应空间平面之间是否为遮挡关系的标记,如图 4。

4.2 特征与深度关系

图像某区域特征(纹理、颜色、边缘等)在一定程度上与该区域相应的空间平面的深度、方向等信息相关。设 d' 为 SP_i 中像素 s_i 对应的估计深

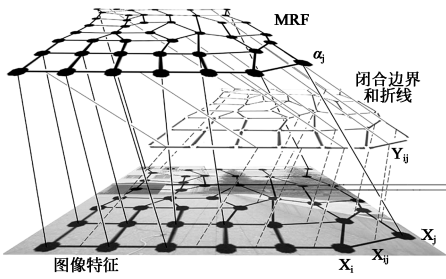


图4 MRF示意图

Fig. 4 Illustration of the Markov Random Field (MRF)

度,其与 s_i 的特征向量之间存在如下关系:

$d' = x_{i,s_i}^T \theta_r$, 其中 x_{i,s_i}^T 是 SP_i 中第 s_i 个像素的特征向量, θ_r 为模型参数。 d 为像素点 s_i 的真实深度,则相对深度误差为:

$$\Delta d = (d' - d)/d = d'/d - 1 = \frac{R_{i,s_i}^T a_i (x_{i,s_i}^T i s_i \theta_r) - 1}{1}, \quad (3)$$

其中: R_{i,s_i}^T : SP_i 中第 s_i 个像素对应的射线, α_i : SP_i 对应平面参数,则图像SP特征向量与其对应的平面参数之间统计关系为:

$$f_1(a_i | x_i, v_i, R_i, \theta) =$$

$$\exp\left(-\sum_{s_i=1}^N V_{i,s_i} | R_{i,s_i}^T a_i (x_{i,s_i}^T \theta_r) - 1 \right), \quad (4)$$

V_{i,s_i} : SP_i 中第 s_i 个像素估计深度的可信度。

4.3 遮挡关系及边界

用 $y_{ij} \in \{0, 1\}$ 表示超像素块 i 和 j 之间的边界特征向量, $y_{ij} = 0$ 表示遮挡,而 $y_{ij} = 1$ 表示边没有遮挡。计算公式为: $P(y_{ij} = 1 | x_{ij}; \psi) = 1/(1 + \exp(-\psi^T x_{ij}))$, x_{ij} 为超像素 i 和 j 的特征(5.2节)。

y_{ij} 通过逻辑回归分析与机器学习的方法获取。

4.4 连接关系

求取连接关系的基本思路:选择摄像机中心到 SP_i 与 SP_j 边界上点相对距离来实现连接性的约束。

$$h_{s_i,s_j}(a_i, a_j, y_{ij}, R_i, R_j) =$$

$$\exp(-y_{ij} | (R_{i,s_i}^T a_i - R_{j,s_j}^T a_j) d_0 |), \quad (5)$$

其中 $R_{i,s_i}^T a_i = 1/d_{i,s_i}$, $R_{j,s_j}^T a_j = 1/d_{j,s_j}$,

$$d_0 = \sqrt{d_{i,s_i} d_{j,s_j}}, (R_{i,s_i}^T a_i - R_{j,s_j}^T a_j) d_0 = (d_{i,s_i} - d_{j,s_j}) / \sqrt{d_{i,s_i} d_{j,s_j}}$$

4.5 共面关系

选择摄像机中心到 SP_i 与 SP_j 中心点的相对

距离来实现共面性的约束,如图5。

$$\begin{aligned} h_{s_i}^r(a_i, a_j, y_{ij}, R_{i,s_i}) &= \exp(-y_{ij} | (R_{i,s_i}^T a_i - R_{j,s_j}^T a_j) d_{0,s_i} |) \\ h_{s_j}^r(a_i, a_j, y_{ij}, R_{j,s_j}) &= \exp(-y_{ij} | (R_{j,s_j}^T a_j - R_{i,s_i}^T a_i) d_{0,s_j} |) \\ h_{s_i,s_j}^r(\cdot) &= h_{s_i}^r(\cdot) h_{s_j}^r(\cdot), \end{aligned} \quad (6)$$

其中: $R_{i,s_i}^T a_i = 1/d_{i,s_i}$, $R_{j,s_j}^T a_j = 1/d_{j,s_j}$,

$$d_{0,s_i} = \sqrt{d_{i,s_i} d'_{i,s_i}}$$

$$R_{j,s_j}^T a_j = 1/d_{j,s_j}, R_{i,s_i}^T a_i = 1/d'_{i,s_i},$$

$$d_{0,s_j} = \sqrt{d_{j,s_j} d'_{j,s_j}} (R_{j,s_j}^T a_i - R_{j,s_j}^T a_j) d_{0,s_j} = (d'_{j,s_j} - d_{j,s_j}) / \sqrt{d_{j,s_j} d'_{j,s_j}}.$$

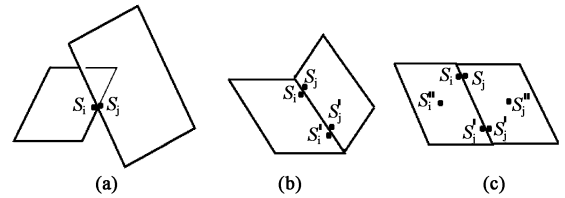


图5 相交(a)、共线(b)、共面(c)结构示意图

Fig. 5 (a) connected, (b) Co-linearity, (c) Co-planar

4.6 共线关系

求取共线关系的基本思路:选择摄像机中心到 SP_i 与 SP_j 共线点的相对距离来实现共线性的约束,如图6。

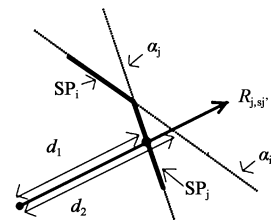


图6 共线计算示意图

Fig. 6 2-d illustration to explain the co-linearity term

4.7 参数学习及MAP-MRF求解

MAP是MRF建模中最常用的最优化准则,以此来确定目标函数。

$$\alpha^* = \arg \max_{\alpha} \log P(\alpha | X, V, y, R, \theta_r) =$$

$$\arg \max_{\alpha} \log \frac{1}{Z} \prod_i f_1(a_i | X_i, v_i, R_i, \theta)$$

$$\prod_{i,j} f_2(a_i, a_j | y_{i,j}, R_i, R_j) =$$

$$\arg \min_{\alpha} \sum_i^K (a \sum_{s_i=1}^{M_i} V_{i,s_i} | (R_{i,s_i}^T a_i) d_{i,s_i} - 1 | +$$

$$\sum_{j \in N(i), s_j \in B_{ij}} y_{ij} | (R_{i,s_i}^T a_i - R_{j,s_j}^T a_j) d_{s_i, s_j} | + \sum_{j \in N(i), s_j \in C_j} y_{ij} | (R_{j,s_j}^T a_i - R_{j,s_j}^T a_j) d_{s_j} |, \quad (7)$$

其中, K 为图像中所有 SP 数量, $N(i)$ 为 SP_i 的邻域 SP 集, B_{ij} 是 SP_i 与 SP_j 边界点集(互连性), C_j 是 SP_j (共面性与共线性)中心。

每一项皆为 α 的线性函数, 最终的目标函数为:

$$\arg \min_x \|Ax - b\|_1 + \|Bx\|_1 + \|Cx\|_1, \quad (8)$$

$$Cx = \begin{bmatrix} a_1 \\ \vdots \\ a_K \end{bmatrix},$$

参数学习采用已知真实深度与图像特征向量进行训练。

(1) 遮挡关系参数的估算如下:

$$P(y_{ij} = 1 | e_{ij}; \varphi) = 1 / (1 + \exp(-\varphi^T e_{ij})), \quad (9)$$

其中: e_{ij} 为 SP_i 、 SP_j 之间边界特征向量, φ 为模型参数(机器学习)。

(2) 深度估计的可信度的估计:

$$P(V_{i,s_i} = 1 | X_{i,s_i}, \Phi) = 1 / (1 + \exp(-\Phi^T X_{i,s_i})), \quad (10)$$

其中: V_{i,s_i} 为 SP_i 中第 s_i 个像素估计深度的可信度。 Φ 为模型参数(机器学习)。

(3) 特征向量与深度关系参数 θ_r 的估计:

$$\theta_r^* = \arg \max_{\theta_r} \log P(\alpha | X, v, y, R, \theta) = \arg \max_{\theta_r} \sum_i \log r f_1(a_i | X_i, V_i, R_i, \theta_r) + \sum_{i,j} f_2(a_i, a_j | y_{ij}, R_i, R_j), \quad (11)$$

第二项与 θ_r 无关, 上式可简化为:

$$\theta_r^* = \arg \min_{\theta_r} \sum_i \sum_{s_i=1}^{M_i} V_{i,s_i} | \frac{1}{d_{i,s_i}} (X_{i,s_i}^T \theta_r - 1) |$$

最后采用 Multi-Conditional Learning^[12] 方法求取。

5 图像特征提取

5.1 立体线索

对于处于某一位置的每一个超像素块, 计算出基于纹理的统计学特征, 以及基于位置和方向的特征。使用 Laws 模板滤波, 图像模式采用

YCbCr, 平滑、边缘及角点检测 (3×3) 模板全部作用于图像 Y 通道, 另外, 第一滤波模板单独作用于图像的 Cb 与 Cr 通道。用于方向检测 (5×5) 模板作用于图像的 Y 通道。如图 7 所示。

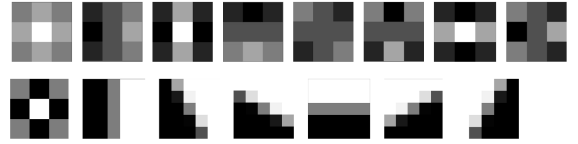


图 7 所用的卷积滤波器

Fig. 7 The convolutional filters used

设 17 个滤波器输出为 $F_n(x, y)$ ($n=1, \dots, 17$), 则

$$E_i(n) = \sum_{(x,y) \in SP_i} |I(x, y) * F_n(x, y)|^k, \quad (12)$$

令 $K=2, 4$, 则分别求出能量值及峰态值, 共生成 34 维特征向量。为了充分利用上下文信息, 对指定 SP, 从 3 个尺度上对其进行特征提取, 并将每个尺度中其四邻域的 SP 特征也作为其特征的一部分, 如图 2 所示。加入 14 个 SP 形状(如 SP 边数、像素密度等)、位置(如水平与垂直方向平均坐标等)及离心率特征。最终特征向量维数: $34 \times (4+1) \times 3 + 14 = 524$ 。

5.2 边界特征向量

边界特征向量用于 SP 对应空间平面之间遮挡关系的判断。分别根据图像纹理、颜色、边缘等 7 个不同的属性, 从两个尺度上对图像进行分割, 共生成 14 个分割, 对应 14 维特征向量, 每个分量为任意两个 SP 是否属于同一分割区域的标记。如果 SP_i 、 SP_j 在 14 个分割中都属于同一个分割区域, 则 SP_i 、 SP_j 对应的空间平面更可能为共面或互连关系, 否则, 则可能为遮挡关系。

6 实验

为了验证以上结论, 收集了 518 幅带有深度的分辨率为 $2\,272 \times 1\,704$ 图片, 其深度分辨率为 55×305 , 400 幅用于训练学习, 118 用于测试, 另外用 200 幅分辨率为 800×600 普通图片进行测试, 测试结果为基本正确地完成了 60% 的图片的三维重建。

图 8 为模型计算出来的深度图, 图 9 是部分

3D 重建结果,其中第 1~5 张是基本正确的重构,第 6 张是不正确的重构。重构失败的原因是多方

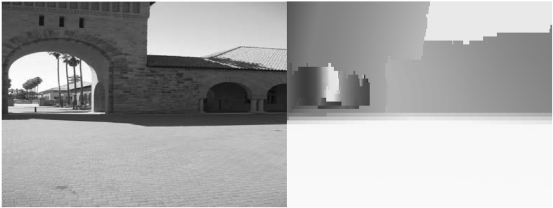


图 8 原始图片与计算出的深度图

Fig. 8 Original image and depth from MRF

面的,一是由于系统本身的缺陷,每张图划分的超像素块有限(2 000 个),训练的样本也不算多;二是样本本身的特点,如:遮挡、缺乏纹理、重复性纹理、基线过大等因素都会导致重建出错。为提高重建精度,可以融入对象检测与识别方法,大场景重建时可与常规 SFM 算法相结合。

与目前已有的单目视觉 3D 重建方法相比,本文算法有表现突出的方面,也有不足的方面,详情如表 1 所示。



图 9 部分实验结果(1-5 基本正确,第 6 张不正确)

Fig. 9 Typical results from our algorithm

表 1 各种单目视觉 3D 重建方法定性比较

Tab. 1 Comparison of 3D reconstruction methods by monocular vision

方法	复杂度	自动化程度	重建效果	应用情况
SFS	使用单张图片重构,空间和时间复杂度都较低,重建时间短,偏微分求解的病态问题难以解决	能实现完全自动化	重构效果较差,易受光源影响,鲁棒性较差	对光源要求严格,难以应用于室外环境等光线条件复杂情况下的重建
光度立体视觉法	使用少量图片进行重构,空间和时间复杂度都较低,重建时间较短	能实现一定程度的自动化	重建效果较好,受光源影响,鲁棒性较差	对光源要求严格,难以应用于室外环境等光线条件复杂情况下的重建
纹理法	使用单幅图像重建,空间时间复杂度都较低,重建时间短	能实现完全自动化	重建效果较差,但对光照,噪声等不敏感,鲁棒性较好	只适用于有规则纹理的物体,应用范围较小
轮廓法	重建空间复杂度较高,时间复杂度较低,重建时间可以满足实时要求	可实现完全自动化	取决于轮廓像数量,轮廓像越多重建越精确	凹陷,空洞的物体的重建无法应用于表面有
调焦法	使用少量图像进行重建,空间和时间复杂度都较低	难以实现自动化重建	能计算每个点的深度,重建效果比较精细	对纹理信息少的物体重建效果较差
本文方法	重建空间复杂度较高,重建时间较长,不能保证满足实时要求	可以自动化,也可以加入人工约束	能计算超像素块之间的关系及其的深度,重建效果比较精细	对于结构复杂的物体及场景重建效果差

7 结 论

试验证明,我们的算法实现重构的正确率及质量上都很可观,这种用 MRF 对图像中微小的同一区域(超像素块)进行建模,用先验知识进行训练,用图像特征估算平面参数,推测图像中相邻区域的关系,不需要对场景做预先的假设,但它也有很多可以进一步思考的地方,例如:每个超像素块的特

征向量维数比较高(524 维),如果进行降维处理(如 KL 变换),是否能对学习、预测的速度与精度有所提高;用 14 维超像素块间的边缘特征向量来判断对应空间平面遮挡情况存在,可信度值得进一步验证。总之,三维重建的难点不在于三维重建的过程,更在于对二维图像的理解。通过从原始图像中提取与理解尽量多的特征信息,将特征信息用于对三维重建过程的指导或对三维重建结果的优化的思路是可行的。

参 考 文 献:

[1] 佟帅,徐晓刚,易成涛,等. 基于视觉的三维重建技术综述[J]. 计算机应用研究,2011,28(7):2411-2417.
TONG S, XU X G, YI C T, *et al.* Overview on vision-based 3D reconstruction [J]. *Application Research of Computers*, 2011, 28(7): 2411-2417. (in Chinese)

[2] 王欣,袁坤,于晓,等. 基于运动恢复的双目视觉三维重建系统设计[J]. 光学精密工程,2014,22(5):1379-1387.
WANG X, YUAN K, YU X, *et al.* Design of binocular vision 3D reconstruction system based on motion recovery [J]. *Optics and Precision Engineering*, 2014, 22(5): 1379-1387. (in Chinese)

[3] 邹虹,曾鑫,温鑫. 基于两视点视频融合技术的裸眼 3D 显示的研究[J]. 液晶与显示,2014,29(5):824-829.
ZOU H, ZENG X, WEN X. Glasses-free 3D display based on two-view video synthesize technology [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(5): 824-829. (in Chinese)

- [4] 姜翰青,章国锋,董子龙,等. 基于图像序列的交互式快速建模系统[J]. 计算机辅助设计图形与图形学学报, 2008, 20(9): 1196-1203.
JIANG H Q, ZHANG G F, DONG Z L, *et al.* Fast interactive scene modeling from an image sequence [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2008, 20(9): 1196-1203. (in Chinese)
- [5] 段福庆,吴福朝,胡占义. 基于平行性约束的摄像机标定与 3D 重构[J]. 软件学报, 2007, 18(6): 1350-1360.
DUAN F Q, WU F C, HU Z Y. Camera calibration and 3D reconstruction using parallelism constraint [J]. *Journal of Software*, 2007, 18(6): 1350-1360. (in Chinese)
- [6] HORN B K P. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view [D]. Massachusetts: Massachusetts Institute of Technology, 1970.
- [7] 吴凤和,张晓峰,施法中. 基于单幅图像数据的三维重构方法研究[J]. 中国机械工程, 2007, 18(17): 2071-2075.
WU F H, ZHANG X F, SHI F Z. Research on 3D reconstruction method based on single image data [J]. *China Mechanical Engineering*, 2007, 18(17): 2071-2075. (in Chinese)
- [8] 胡春雨,刘弘,陆佃杰. 一种基于单幅图像的 3D 场景建模方法[J]. 计算机应用研究, 2014, 31(6): 1909-1913.
HU C Y, LIU H, LU D J. Monocular image based 3D-scene reconstruction method [J]. *Application Research of Computers*, 2014, 31(6): 1909-1913. (in Chinese)
- [9] XU K, ZHENG H L, ZHANG H, *et al.* Photo-inspired model-driven 3D object modeling [J]. *ACM Trans on Graphics*, 2011, 30(4): 80-89.
- [10] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. *International Journal of Computer Vision*, 2004, 59(2): 167-181.
- [11] SAXENA A, SUN M, NG A Y. Make3D: Learning 3D scene structure from a single still image [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2008, 31(5): 824-840.

作者简介:李蓉(1982—),女,四川绵阳人,硕士,讲师,主要从事机器视觉、多媒体技术等方面的研究。E-mail: windylee007@126.com

邓春健(1980—),男,广东乐昌人,博士,教授,硕士生导师,主要从事嵌入式系统、智能设备等方面的研究。

邹昆(1981—),男,湖北郧西人,博士,副教授,主要从事图形图像处理等方面的研究。