

中图法分类号: TP242.6+2 文献标志码: A 文章编号: 1006-8961(2011)06-1057-07

论文索引信息: 韩坤芳, 沈会良, 程岳. 双目立体重建中的去噪及补洞方法 [J]. 中国图象图形学报, 2011, 16(6): 1057-1063

双目立体重建中的去噪及补洞方法

韩坤芳, 沈会良, 程岳

(浙江大学信息与电子工程学系, 杭州 310027)

摘要: 通过双目立体视觉重建得到的真实物体 3 维模型一般具有较大的噪声, 并可能存在孔洞。针对这一问题, 提出在规则 3 维点集下, 通过非局部 (NL) 方法来滤除 3 维模型的噪声并填补孔洞。为提高算法效率, 对含孔区域和非孔区域的点集, 分别做基本 NL 滤波和基于主成分分析的 NL 滤波, 然后进一步通过 NL 方法对孔洞做填补。实验结果表明, 对于多种真实物体, 所提出的方法均能有效地重建 3 维模型, 且与其他方法相比具有更高的精度。

关键词: 双目立体; 非局部; 去噪; 补洞; 主成分分析

Denoising and hole-filling for binocular stereo reconstruction

Han Kunfang, Shen Huiliang, Cheng Yue

(Department of Information and Electronic Engineering, Zhejiang University, Hangzhou 310027 China)

Abstract: The 3D models reconstructed from binocular stereo always have noticeable noise, and even contain holes. To deal with this issue, this paper presents a non-local (NL) method to denoise and hole-fill the original model on a regular vertex set. To improve computational efficiency, the proposed method uses a basic NL method and principal component analysis (PCA) based NL method to remove model noise in hole and non-hole regions, respectively. The filled holes are then processed by NL denoising to improve the consistency with the surrounding surfaces. The experimental results show that, when evaluated on different real objects with various sizes, the proposed method can reconstruct 3D models effectively and performs much better than previous methods.

Keywords: binocular stereo; non-local; denoising; hole filling; principal component analysis

0 引言

双目立体重建利用两个摄像机模仿人类视觉, 在同一时间从不同角度拍摄场景中某一物体, 并在计算视差基础上, 通过三角测量原理, 获得两个摄像机公共视场内空间物体的 3 维坐标, 最终重建物体的形状和位置^[1]。双目立体重建在工业测量、机器人视觉导航、地图生成、航空勘测和近距照相测量等领域发挥越来越重要的作用, 有极高的应用价值。

本文主要讨论基于双目立体视觉的真实物体 3 维模型重建。当测量中的物体表面只有很少甚至没有纹理时, 如何建立致密立体匹配成为一个难点。Davis 等人^[2]和 Zhang 等人^[3]分别提出了时空光条方法, 通过对空间物体投射一系列的光条模式, 增加空间物体的纹理信息, 从而降低了匹配难度。为提高匹配精度, 采用左右摄像机对称性测试, 即只有在左右相机完全匹配的情况下才确定该像素点的视差^[4]。在测量精度要求较高的应用中, 进一步需要亚像素级别的立体匹配^[5]。

收稿日期: 2010-05-04; 修回日期: 2010-06-13

基金项目: 国家重点基础研究发展计划 (973) 项目 (2009CB320801); 国家自然科学基金项目 (60778050)。

第一作者简介: 韩坤芳 (1986—), 女, 浙江大学信息与电子工程学系硕士研究生, 主要研究方向为立体视觉、3 维重建。

E-mail: hankunfang@ sina. com。

通讯作者: 沈会良, E-mail: shenhl@ zju. edu. cn。

基于视差重建得到的 3 维模型通常具有较多噪声,究其原因有以下几点:

1) 两个摄像机从不同角度拍摄物体,导致物体表面点在两个摄像机中的透视成像点并不完全一一对应,必然存在某些像素点不能真正匹配。

2) 计算物体表面点位置信息的三角测量原理建立在理想成像模型基础之上,而实际中的两个摄像机光心与投影点形成的光线通常不相交于一点^[6]。

3) 摄像机标定的精度和外界光线的变化等因素也会造成 3 维模型重建中的噪声。

Nehab 等人^[7]结合双目立体视觉中的位置信息和光度立体视觉中的法向信息,通过优化融合来滤除噪声,得到较好的 3 维模型,但其必需额外的法向信息。Fleishman 等人^[8]用双边滤波器(bilateral filter)对重建模型去噪,根据两个顶点的距离远近以及深度的相似性做加权平均。由于局部滤波的局限性,该方法在边缘和非平滑区域的去噪效果不够理想。

除噪声外,两个摄像机的不同视场和物体表面高光可能会造成位置信息缺失,导致 3 维重建模型存在孔洞。Wang 和 Oliverira^[9]利用 MLS(moving least squares)方法对孔洞周边顶点拟合小平面,用于计算孔洞的位置信息。该方法能够较好地恢复平滑曲面,但对于纹理丰富的曲面填补效果不佳。

提出采用非局部(NL)噪声滤波方法^[10-11]对双目立体重建模型进行去噪,并在非孔区域利用主成分分析(PCA)^[12]来提高运算效率。在补洞处理中,根据待补顶点的周围邻接信息,在全局范围内寻找最相似顶点,填补孔洞处的位置信息。在实验中,测试了本文所提出方法在多种真实物体上的性能,并与其他方法做比较。

1 双目立体重建

采用的成像系统由两个摄像机以及一台投影仪组成,投影仪向物体表面投射光条图案以增加纹理信息,摄像机同步拍摄一系列图像。摄像机标定和图像校正采用 Camera Calibration Toolbox for Matlab^[13]。

采用 Davis^[2]和 Zhang^[3]的光条时空立体技术,通过投影仪向物体投射一系列用 10 位格雷码^[14]编码的光条模式,将像素的邻域窗从空间扩展到时空,从而降低匹配难度。像素空间邻域窗是指以该像素为中心,大小为 $m \times n$ 的窗口。时空邻域窗增加了时间维度,其窗口大小为 $m \times n \times t$ 。本文中采用的

匹配窗口大小为 $1 \times 1 \times 12$ 。

将邻域窗口中像素集合记为向量 \mathbf{p} , 像素 i 和 j 的相似性用归一化互相关系数(normalized cross correlation coefficient) ρ 来评价:

$$\rho = \frac{\sum_{k \in \Omega} (\mathbf{p}_i(k) - \mu_{p_i})(\mathbf{p}_j(k) - \mu_{p_j})}{\sigma_{p_i} \sigma_{p_j}} \quad (1)$$

式中, Ω 表示邻域窗, μ_p 表示向量 \mathbf{p} 的平均值。 $\rho \in [-1, 1]$, ρ 为正时,其值越大表示像素匹配程度越高。为提高匹配率及匹配精度,通过左右对称性测试确定视差值 d , 并采用高斯曲线拟合计算亚像素视差:

$$\tilde{d} = d + \delta \quad (2)$$

其中

$$\delta = \frac{\ln \rho_{i-1} - \ln \rho_{i+1}}{2 \ln \rho_{i-1} - 4 \ln \rho_i + 2 \ln \rho_{i+1}} \quad (3)$$

得到图像中每点的亚像素视差后,根据针孔成像模型和摄像机内外参数容易求出空间点在摄像机坐标系下的 3 维坐标 (x, y, z) ^[4]。

2 基于 NL 去噪及补洞方法

2.1 去噪

噪声滤波具有多种方法,都以平均化为出发点,可在局部区域内平均化,也可在非局部区域平均化。局部区域平均化有 Gaussian 滤波^[15]和双边滤波^[16]等。Buades 等人^[10]提出的 NL 噪声滤波方法去噪效果明显。这些噪声滤波方法在 2 维图像上的应用比较成熟,但目前还很少应用于 3 维重建模型。

3 维重建模型是规则的点集,每个顶点的邻域信息明确,可以将重建模型中的深度值 z 类比于 2 维图像的灰度值,即只考虑顶点深度信息 z , (x, y) 则认为是规则点阵,故可以采用 2 维图像的去噪方法对 3 维重建模型进行去噪。

NL 噪声滤波方法耗时较长,为提高运算速度,本文对含孔区域的点集和非孔区域分别做基本 NL 滤波和基于主成分分析的 NL 滤波。

2.1.1 基本 NL 滤波

NL 去噪方法的基本思想是,每一个顶点的新深度值是全局范围内所有类似顶点的加权平均。由于搜索整个区域将消耗很长的运算时间,选取一个合适的窗口 S 作为搜索域:

$$z_{NL}(i) = \sum_{j \in S} w(i, j) z(j) \quad (4)$$

式中, $z(j)$ 表示顶点 j 的深度信息,权重 $w(i, j)$ 由两

个顶点邻域窗向量的欧氏距离 (Euclidean distance) 决定:

$$w(i, j) = C \cdot e^{-\|v_i - v_j\|^2/h^2} \quad (5)$$

式中: v_i 和 v_j 分别为顶点 i 和 j 的邻域窗向量; 窗口中像素数量为 K ; $\|v_i - v_j\|^2$ 为邻域窗向量欧氏距离; h 为衰减因子, 一般取值为 $0.3 \leq h \leq 0.5$; C 为权重归一化因子。

考虑到孔洞 (本文对其深度信息标记为零值) 的存在, 在计算权重时, 两顶点邻域窗向量中零值位置必须为当前参考点邻域窗向量和比较点邻域窗向量非零位置的交集。

考虑到权重较小的顶点对去噪并没有正面影响, 不选取 S 域内所有顶点, 而只选取权重不小于搜索域内最大权重 w_{\max} 一定比例的顶点, 即要求顶点权重 $w(i, j) \geq \kappa_1 w_{\max}$, κ_1 取值范围为 $[0, 1]$ 。

实际模型中, 顶点之间的绝对深度相似较少, 而相对深度相似较多, 为此先对每个顶点的邻域窗向量减去其平均值, 再计算权重。由于孔洞的存在, 邻域窗向量平均值 μ_v 定义为向量中每个元素之和除以向量的非零元素个数。由此, 式(4)变为

$$z_{\text{NL}}(i) = \sum_{j \in S} w(i, j) [z(j) - \mu_{v_j} + \mu_{v_i}] \quad (6)$$

式中

$$w(i, j) = C \cdot e^{-\|v'_i - v'_j\|^2/h^2} \quad (7)$$

$$v'_i = (v_i(1) - \mu_{v_i}, \dots, v_i(K) - \mu_{v_i}) \quad (8)$$

向量 v'_i 与式(8) v'_i 的计算类似。

2.1.2 基于主成分分析的 NL 滤波

基本 NL 去噪方法耗时较长, 为提高运算速度, Mahmoudi 等人^[17]先利用顶点邻域窗平均值或梯度值对邻域窗进行筛选, 然后再进行加权平均。但用平均值进行筛选会误剔相对相似的顶点, 而由于噪声的存在用梯度进行筛选误差会比较大。采用 Tasdizen^[12]提出的主成分分析的方法, 先对顶点邻域窗向量进行降维, 继而使用 NL 去噪。但由于孔洞的存在, 对存在孔洞的邻域窗向量进行主成分分析是没有意义的, 故对于邻域窗向量中存在孔洞或搜索域中存在孔洞的顶点依然采用基本 NL 去噪。

设重建模型顶点有效总个数 (含孔洞) 为 M , 邻域窗内像素数量为 K 。将重建模型存储为 $M \times K$ 的矩阵 Q , 每一行存储该行序号顶点邻域窗向量。然后求解该矩阵的协方差矩阵, 计算其特征向量 (b_1, \dots, b_K) 以及特征值 ($\lambda_1, \dots, \lambda_K$)。根据特征值的大小, 本文选取 d 个特征值较大的特征向量作为主成分分析后的新空间坐标。将 Q 矩阵投射到 d 维向量空间, 得到大小为

$M \times d$ 的矩阵, 即每个邻域窗向量由原来的 K 维降低到 d 维, 新邻域窗向量记为 $u_i, u_i = \sum_{p=1}^d f_p(i) b_p$, 其中, $f_p(i) = v_i^T b_p$, 从而大幅度减少运算时间。

PCA 在提高运算速度的同时也丢失了重建模型的一些信息, 因此需要谨慎选择 d 。根据 $\sum_{i=1}^K \lambda_i \geq \kappa_2$ 来确定 d , 一般 $0.4 \leq \kappa_2 \leq 0.9$ 。

考虑到特征向量的规范正交化性, 新邻域窗向量之间的欧几里得距离为

$$\|u_i - u_j\|^2 = \sum_{p=1}^d (f_p(i) - f_p(j))^2 \quad (9)$$

故式(5)为

$$w(i, j) = C \cdot e^{-\sum_{p=1}^d (f_p(i) - f_p(j))^2/h^2} \quad (10)$$

同样, 为增加相似顶点的个数, 采用相对相似的方法, 式(10)变为

$$w(i, j) = C \cdot e^{-\sum_{p=1}^d (f'_p(i) - f'_p(j))^2/h^2} \quad (11)$$

其中

$$f'_p(i) = f_p(i) - \sum_{k=1}^d f_k(i)/d \quad (12)$$

然后用式(6)求顶点新值。

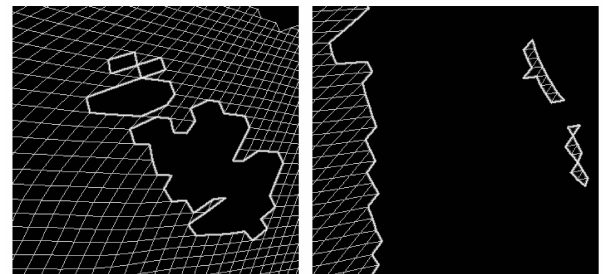
2.2 补洞

正确判定孔洞是填补孔洞的必要前提, 孔洞的判定可分为 3 步:

1) 寻找网格边界边 (boundary edge), 任何一条边界边只邻接一个三角面, 而非边界边被两个三角面所共享。

2) 检索全部边界边, 记录可构成环的边界边及其顶点, 判定为洞。

3) 判定 2) 中洞的真伪。如图 1 所示, 既有需要填补的真实孔洞, 也有类似岛屿 (island) 无须填补的孔洞。通过检查孔洞边界外顶点深度值是否为零, 判决其真伪。



(a) 真正的洞

(b) 岛屿

图 1 洞的不同类型

Fig. 1 Different types of holes

确定孔洞之后,采用 NL 方法对孔洞进行填补:对于需要填补的某个顶点,利用其周边邻接信息,在搜索域内寻找与其最相似的顶点,并将其值填补给该顶点,并记该顶点的深度信息为 z_H 。搜索相似点时,根据式(7)计算该补顶点的邻接窗向量与搜索域内顶点的邻接窗向量相似度,选取相似度最大顶点的深度值作为该补顶点的填补值,即

$$z_H(i) = z(j)_{\max(u(i,j))} - \mu_{v_j} + \mu_{v_i} \quad (13)$$

为保持较好的边界连续性,降低累积误差,本文采用沿孔洞的边界线逐圈向内的填补方法:首先,遍历洞内顶点,通过判断顶点是否为边界邻接点,确定该顶点是否为填补对象,对需要填补的顶点在搜索域内寻找与其最相似的顶点,用式(13)计算其深度信息;然后,在下次遍历孔洞的时候,把前步已填补的顶点作为已知顶点,重复填补操作;如此循环,直至洞内所有顶点均完成填补。图 2 为补洞的步骤示意图,网格中灰色区域为孔洞,第 1 次遍历时黑色点为填补对象,第 2 次遍历时黑色点作为已知顶点,白色点为填补对象。

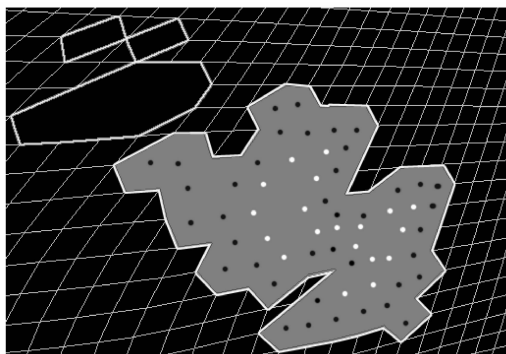


图 2 补洞步骤示意图

Fig. 2 Procedure of hole-filling

由于孔洞内顶点是逐个进行填补,初步填补完后,深度之间可能没有达到最佳连接,故对每个填补完的孔洞使用 NL 去噪方法对其进行平滑。实验表明,沿孔洞边界线逐圈向内填补,并对其进行 NL 去噪,能够较平滑地将孔洞填补完整。

3 实验结果与分析

实验系统中采用分辨率为 1024×768 的液晶投影仪对物体投射光条模式,镜头焦距为 8 mm。分别对浮雕、石膏头像、茶壶、贝壳和猴子 5 种物体(如图 3)进行 3 维重建,物体的高度范围从 6 ~

25 cm 不等。由于投影仪分辨率限制,最细光条投射到物体表面不能达到 1 个像素宽度,故重建出的物体表面模型会产生一些条纹。其他学者的研究也出现相似结果^[2-3],如何消除此类条纹已超出本文的讨论范围。

在实验中,将本文提出的 NL 去噪方法和双边噪声滤波方法^[8]做比较,NL 补洞方法与 Wang 等人^[9]的 MLS 方法进行对比。双边滤波是以两个顶点的距离远近以及深度相似性这两个因子影响权重,对原顶点进行加权平均。MLS 补洞方法则是利用该补顶点周边邻接信息,拟合二次曲面,根据拟合平面计算该补顶点的深度信息。

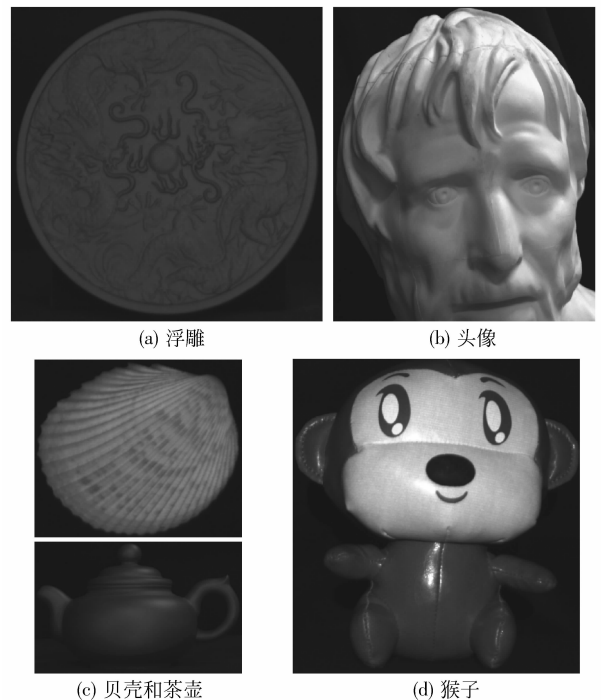


图 3 各物体图像

Fig. 3 Images of five objects

3.1 去噪结果

NL 去噪方法参数设置如下:搜索窗口 S 为 21×21 ,邻域窗口为 5×5 (即 $K = 25$), $h = 0.3$, $\kappa_1 = 0.5$, $\kappa_2 = 0.8$ 。为定量评价去噪效果,本文采集形状已知球体的光条编码图像,重建其 3 维形状并用 NL 方法和双边滤波去噪,并以均方根(RMS)误差 $\varepsilon = \left(\sum_{i=1}^n (z_{\text{denoised}}(i) - z_{\text{real}}(i))^2 / n \right)^{1/2}$ (其中 n 为顶点数量, $z_{\text{real}}(i)$ 为真实深度)来衡量去噪效果。表 1 给出了两种去噪方法的均方根误差,可以看出 NL 去噪效果优于双边去噪。

表 1 球体模型去噪结果的均方根误差
 Tab. 1 RMS error of denoising results for the sphere model

方法	双边去噪	NL 去噪
RMS	0.407	0.386

进一步地,对浮雕、茶壶、石膏头像 3 维重建模型做去噪滤波。从图 4 可看到,浮雕重建模型的深

度变化不显著,经双边滤波后,模型的很多细节深度变化都被严重平滑,而 NL 去噪后的浮雕模型仍然保存了细致的细节结构。图 5 进一步表明,由于双边去噪本质为局部滤波,不能在整体上达到统一效果,表现在对茶壶的边缘去噪不够,而 NL 方法的去噪效果明显优于双边去噪。图 6 中石膏头像的眼睛经过双边滤波后被严重平滑,但经 NL 方法去噪,眼睛细节部分得到很好的保留。

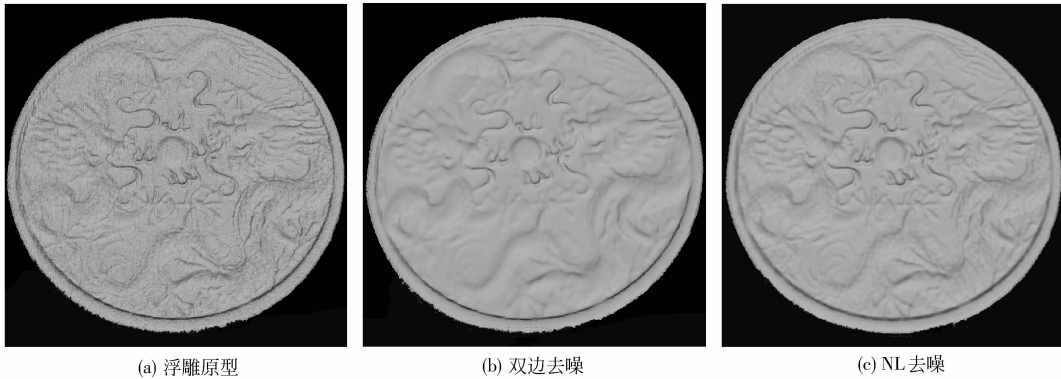


图 4 浮雕 3 维重建模型去噪结果

Fig. 4 Denoising results of the basso-relievo model

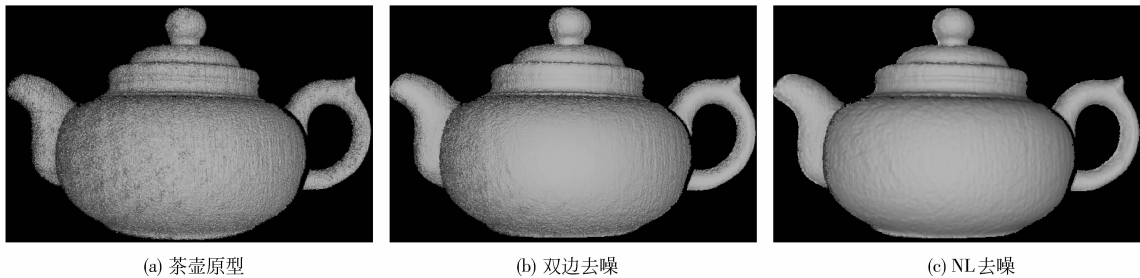


图 5 茶壶 3 维重建模型去噪结果

Fig. 5 Denoising results of the teapot model

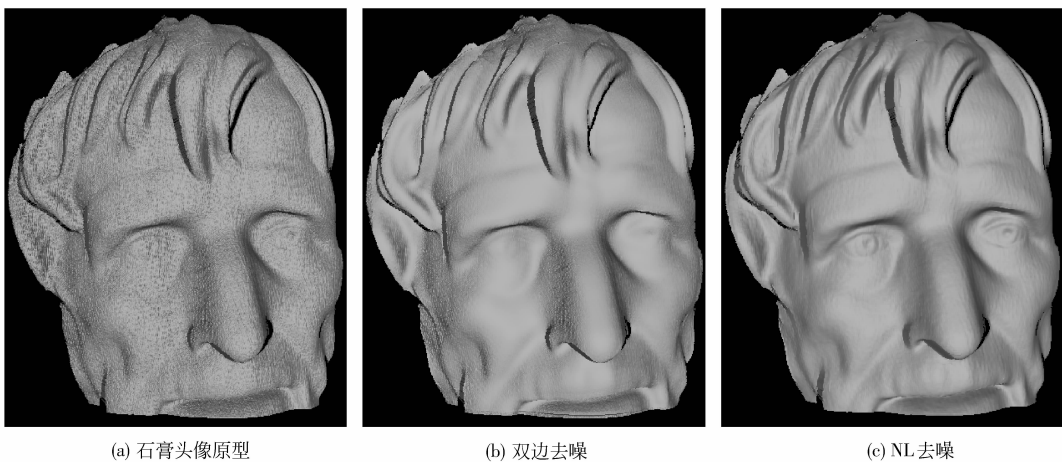


图 6 石膏头像 3 维重建模型去噪结果

Fig. 6 Denoising results of the plaster head model

3.2 补洞结果

为定量评价补洞效果,对贝壳重建模型挖去不同规格的孔洞,将填补结果与实际值进行比较,用 RMS 误差来衡量补洞效果。MLS 与 NL 方法的孔洞填补效果如图 7 所示,MLS 方法填补后的贝壳孔洞区域失去纹理特征。这是由于 MLS 补洞方法采用了局部连续曲面的思想,不适合处理纹理丰富的表面。表 2 进一步给出了 MLS 补洞与 NL 补洞方法的均方根误差,可以看到,NL 方法的误差在多种孔洞大小下都明显低于 MLS 方法。

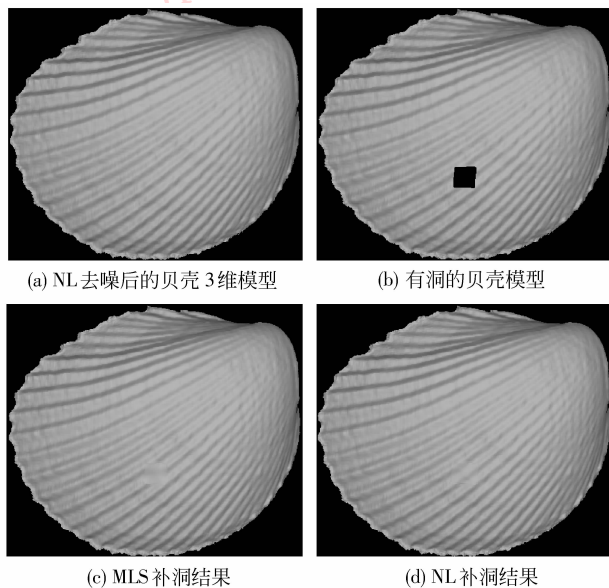


图 7 贝壳重建模型补洞结果(孔洞规格为 20×20 像素)

Fig. 7 Hole-filling results of the shell model
(Hole size: 20×20 pixels)

表 2 贝壳模型填补结果的均方根误差

Tab. 2 RMS error of hole-filling results for the shell model

孔洞大小/像素	/mm			
	10×10	10×20	20×20	20×40
MLS 补洞	0.174	0.254	0.284	0.349
NL 补洞	0.077	0.113	0.140	0.176

此外,本文还对实际中的真实孔洞进行填补。图 8 所示的猴子模型由于高光原因在脸、耳朵、身体等部位产生孔洞,由于两个摄像机视场不同在手臂、脚等部位产生孔洞,由于光线视角以及物体本身质地的原因在鼻子右上方无反射而产生孔洞。本文采用 NL 补洞方法对孔洞进行填补,填补结果如图 8 (c) 所示。由图可知,本文基于 NL 的补洞方法有较好的填补效果。

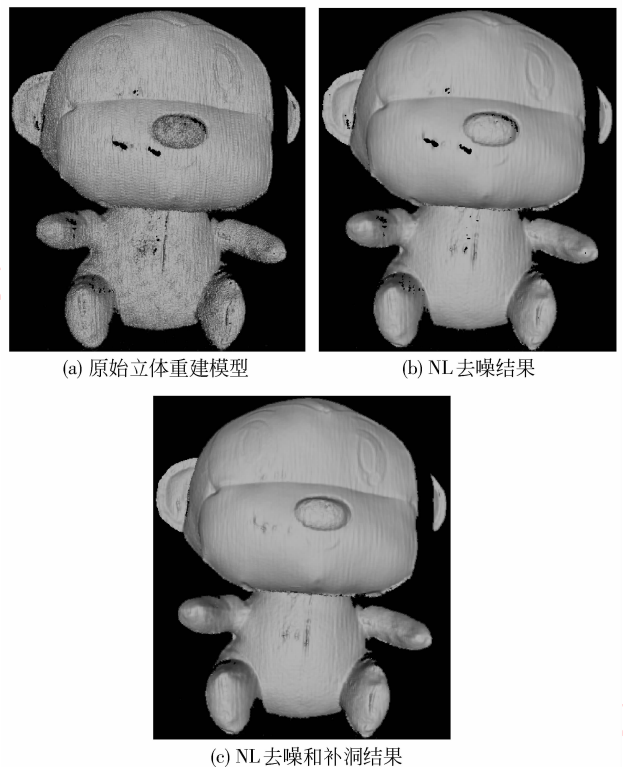


图 8 猴子重建模型的去噪和补洞结果

Fig. 8 Results of NL denoising and hole-filling
for the monkey model

4 结 论

基于双目立体视觉的 3 维重建模型往往具有较大的噪声,并由于两个摄像机视角不同以及物体表面高光等原因可能会造成位置信息丢失,导致重建模型存在孔洞。

针对这些问题,提出一种非局部(NL)的去噪和补洞方法,将 3 维重建模型的规则顶点深度信息类比于 2 维图像中像素灰度信息,对双目立体 3 维重建模型进行去噪和补洞。为提高运算速度,本文对含孔洞区域的点集和非孔洞区域分别做基本 NL 滤波和基于主成分分析的 NL 滤波。实验结果表明,相对于局部方法,如双边滤波去噪和 MLS 补洞,本文提出的 NL 去噪与补洞方法能够较好地保留物体表面细节,在定量和定性比较上都优于已有方法。

研究过程中发现,随着孔洞尺寸的增大,补洞误差会逐步增加。对于较大的孔洞,不管是用局部还是非局部的的方法,精度都会下降。对此可以考虑增加约束条件或将 MLS 与基于 NL 补洞的方法结合起

来填补更大的孔洞。此外,本文提出的补洞方法无法对没有形成闭环的孔洞进行填补,如图8中猴子的左耳。这些都是将来需要进一步研究解决的问题。

参考文献 (References)

- [1] Forsyth D A, Ponce J. *Computer Vision: A Modern Approach* [M]. New Jersey: Prentice Hall Professional Technical Reference, 2002: 199-212.
- [2] Davis J, Nehab D, Ramamoorthi R, et al. Spacetime stereo: a unifying framework for depth from triangulation [J]. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 2005, 27(2): 296-302.
- [3] Zhang L, Curless B, Seitz S M. Spacetime stereo: shape recovery for dynamic scenes [C]// *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Madison, Wisconsin, USA: IEEE Computer Society, 2003: 367-374.
- [4] Zhang Guangjun. *Vision Measure* [M]. Beijing: Science Press, 2008. [张广军. 视觉测量[M]. 北京: 科学出版社, 2008.]
- [5] Yu W R, Xu B G. A sub-pixel stereo matching algorithm and its applications in fabric imaging [J]. *Machine Vision and Applications*, 2009, 20(4): 261-270.
- [6] Hartley R, Zisserman A. *Multiple View Geometry in Computer Vision* [M]. 2nd ed. Cambridge, England: The Press Syndicate of the University of Cambridge, 2003: 311-324.
- [7] Nehab D, Rusinkiewicz S, Davis J, et al. Efficiently combining positions and normals for precise 3D geometry [J]. *ACM Transactions on Graphics*, 2005, 24(3): 536-543.
- [8] Fleishman S, Drori I, Cohen D. Bilateral mesh denoising [J]. *ACM Transactions on Graphics*, 2003, 22(3): 950-953.
- [9] Wang J N, Oliverira M M. A hole-filling strategy for reconstruction of smooth surface in range images [C]// *XVI Brazilian Symposium on Computer Graphics and Image Processing*. Sao Carlos, Brazil: Computer Graphics and Image Processing, 2003: 11-18.
- [10] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising [C] // *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA: IEEE Computer Society, 2005: 60-65.
- [11] Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one [J]. *Multiscale Modeling & Simulation*, 2005, 4(2): 490-530.
- [12] Tasdizen T. Principal neighborhood dictionaries for nonlocal means image denoising [J]. *IEEE Transactions on Image Processing*, 2009, 18(12): 2649-2660.
- [13] Bouguet J Y. *Camera Calibration Toolbox for Matlab* [CP/OL]. (2008-06-02) [2009-12-01]. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.
- [14] Salvi J, Pages J, Battle J. Pattern codification strategies in structured light systems [J]. *Pattern Recognition*, 2004, 37(4): 827-849.
- [15] Lindenbaum M, Fischer M, Bruckstein A. On gabor contribution to image enhancement [J]. *Pattern Recognition*, 1994, 27(1): 1-8.
- [16] Tomasi C, Manduchi R. Bilateral filtering for gray and color images [C]// *Sixth International Conference on Computer Vision*. Bombay, India: Narosa Publishing House, 1998: 839-846.
- [17] Mahmoudi M, Sapiro G. Fast image and video denoising via nonlocal means of similar neighborhoods [J]. *IEEE Signal Processing Letters*, 2005, 12(12): 869-842