

文章编号:1001-4888(2018)05-0751-06

基于运动散斑的物体三维形貌测量方法^{*}

孙俊锋¹, 徐玉华¹, 赵晓枫², 谢良¹, 张小虎¹

(1. 国防科学技术大学 航天科学与工程学院, 湖南长沙 410073; 2. 火箭军工程大学 发射工程系, 陕西西安 710025)

摘要:设计了一种基于近红外运动散斑的低成本双目立体视觉系统。该系统将一个低成本散斑投影器安装在一个摇摆运动机构上,向目标表面投射运动的散斑,利用空间-时间立体视觉原理,在立体匹配的时候,不同于传统的立体视觉方法(采用二维的滑动窗口进行对应点查找),它不仅利用图像像素之间的空间信息,还引入了图像的时间信息,可以采用一个三维的滑动窗口进行立体匹配。与传统的立体视觉方法相比,该方法具有更高的重建精度。此外,本文采用金字塔搜索策略、积分图像和 GPU 对立体匹配算法进行加速,使得立体匹配的计算速度提高了 800 多倍。

关键词:立体视觉; 立体匹配; 空间-时间立体视觉; 运动散斑

中图分类号: O39 **文献标识码:** A **DOI:** 10.7520/1001-4888-17-131

0 引言

基于视觉的三维重建可以描述为从相机拍摄的二维图像恢复目标表面的三维结构,是当前计算机视觉中的热点问题。它在三维建模、工业检测、地形重建、逆向工程、人机交互、人体测量、3D 打印、电影和动画制作、虚拟现实等众多领域中具有广泛的应用^[1]。

常见的三维重建方法包括双目立体视觉^[2]、多视图立体重建^[3]、结构光三维扫描^[4]、光度立体视觉^[5]、从阴影恢复形状(Shape-from-Shading, SFS)^[6]等。本文方法在本质上属于双目立体视觉。双目立体视觉利用两个同步的相机采集的一对立体图像,通过立体匹配算法建立左右视图像素之间的对应关系,从而得到视差图,再根据预先标定好的相机参数,就可以得到目标的三维点云。对于双目立体视觉来说,立体匹配是其中的核心。经过几十年的研究,立体匹配仍然是一个活跃的研究课题。基于该方法的系统的优势是结构简单,成本相对较低;缺点是重建结果受被测目标表面纹理状况影响很大。Zhang^[7]等人提出的空间-时间立体视觉(space-time stereo)方法,使用数字投影机向目标表面投射随时间变化的条纹,利用多帧立体图像进行匹配,可有效提高立体匹配的可靠性。

近年来,随着 Kinect、Xtion 等深度传感器的出现,大大降低了三维数据获取的成本。这些传感器通过向目标表面投射人眼不可见的近红外散斑,有效增加了目标表面的纹理特征,使得立体匹配更加可靠。这些深度传感器虽然在价格和实时性上具有优势,但测量精度比较低。文[8,9]通过对从不同视角多次采集的三维数据进行加权平均,从而降低噪声水平,提高三维点云的精度。

本文的技术贡献是:

(1)提出一种基于运动散斑的低成本三维形貌测量方法。与以往的基于散斑的方法^[10]不同,本文将一个廉价的散斑投影器安装在一个运动机构上,投影器向目标表面投射出运动的散斑。重建时,不仅可以利用图像像素点的空间信息,还可以利用像素点的时间信息,从而可以采用一个三维的匹配窗口进

^{*} 收稿日期: 2017-06-20; 修回日期: 2017-08-28

通讯作者: 徐玉华(1985—),男,博士后,工程师。主要研究领域:计算机视觉,三维重建。Email: xyh_nudt@163.com

行立体对应计算。实验结果表明,该方法可有效提高三维测量的精度。与 Zhang 等人^[7]的方法相比,他们采用一个数字投影机向目标表面投射随时间变化的图案,而本文方法只需要将一个价格几十元人民币的近红外散斑投影器安装在运动的机构上,即可获得随时间变化的图案。因此,本文设计的系统在成本上具有显著的优势。此外,本文采用肉眼不可见的红外光散斑进行三维重建,由于没有任何刺眼的感觉,在人体重建方面,具有更好舒适度。

(2)多分辨率技术、积分图像方法和 GPU 对归一化互相关算法(Normalized Cross Correlation, NCC)进行加速,与单核 CPU 相比,可以加速 875 倍。

在接下来的第 1 节中,将阐述运动散斑三维重建的基本原理;第 2 节详细描述本文设计的三维重建系统;第 3 节是匹配算法加速方法描述;第 4 节给出实验结果;最后一节是本文总结。

1 运动散斑三维重建方法原理

立体匹配问题是计算机视觉领域中的一个热点问题。传统的立体匹配问题是建立一对立体图像对之间的点对应关系,通过计算像素之间的对应关系,可以获得对应点的三维坐标。根据事先标定的相机参数对立体图像做立体校正,可以使得对应点搜索只需要在该像素点所在的水平线上进行^[2],如图 1(a)所示。当目标表面没有足够丰富的纹理特征时,就难以建立正确的点对应关系,要得到精确的立体匹配会变得非常困难。为了解决这一问题,可以通过在立体视觉系统中增加一个可以主动投射预定图案的投影器,在目标表面人为地增加纹理特征,使得立体匹配更加可靠。这种方法可以显著提高表面重建的质量,但仍需要采用比较复杂的立体匹配算法才能获得致密的重建^[11],或者设置很大的用于匹配的图像块的尺寸^[10],而增大匹配块的尺寸会导致重建结果丢失目标表面的细节。

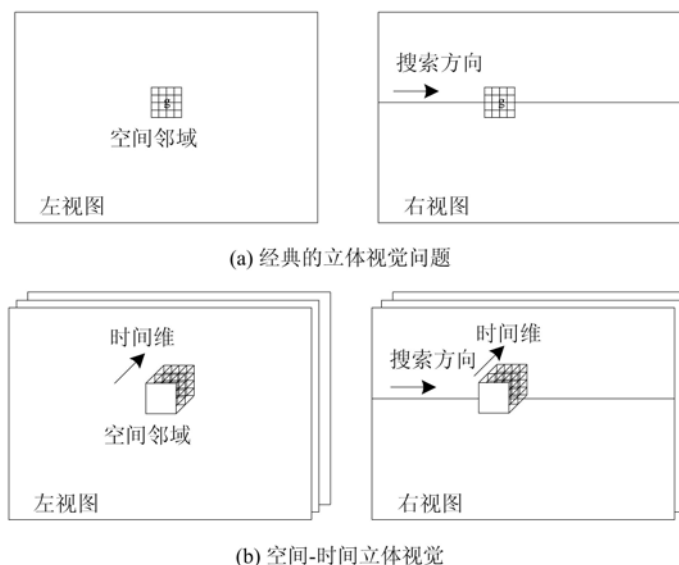


图 1 经典的立体视觉问题和空间-时间立体视觉比较

Fig. 1 Traditional stereo matching and space-time stereo matching

尺寸的空间窗口,在细节重建方面具有显著优势。

本文设计的系统将一个近红外散斑投影器安装在一个做往复运动的机构上,在表面形成运动的散斑阵列,从而可以采用空间-时间立体视觉方法进行三维重建。

2 基于近红外运动散斑的低成本双目立体视觉系统设计

2.1 系统硬件组成

本系统硬件主要有两个同步的 CMOS 相机和一个安装在运动机构上的近红外散斑投影器组成。镜头上安装窄带滤光片,用于过滤掉其他波长的环境光。运动机构由电机驱动的一个偏心轮构成,它带动投影器做周期性往复运动。系统示意图如图 2 所示。

空间-时间立体视觉方法在传统的立体视觉方法基础上,增加了时序信息。仍然是通过建立左右相机图像之间的点关系,利用三角测量原理恢复目标表面的三维信息。所不同的是,前者在图像采集的过程中,投射的主动光是不断变化的。匹配时就可以增加主动光在时间上的变化信息,采用一个三维的滑动窗口进行对应点搜索,如图 1(b)所示。采用三维窗口进行对应点搜索的好处主要在于:(1)由于引入了时间信息,在空间窗口的尺寸设的很小(如 5×5 或 7×7)的情况下,采用很简单的立体匹配算法,就可以得到致密的重建;(2)由于可以采用小

Kinect 深度传感器中的投影器(图 2(b))可以投射出密集的人眼不可见的近红外散斑点(波长为 830nm), 对人眼没有任何刺激作用, 且价格便宜。如果把该投影器安装在一个运动机构上, 如图 2 所示, 就可以在目标表面形成随时间变化的运动光斑。利用空间-时间立体视觉原理, 可以得到可靠的密集重建。与 Zhang 等人采用数字投影机的系统^[7]相比, 该系统在硬件成本控制方面具有明显优势。

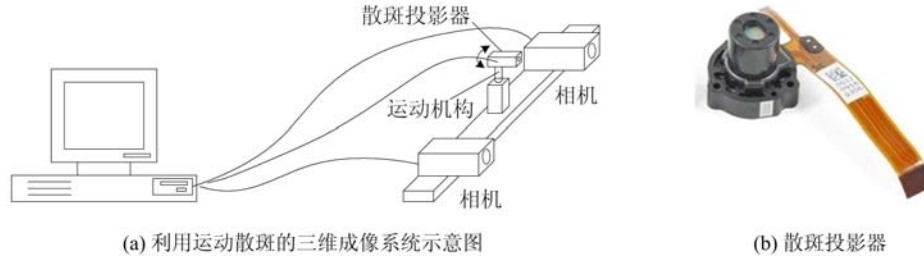


图 2 基于空间-时间立体视觉原理的三维成像系统

Fig. 2 The 3D reconstruction system based on space-time stereo matching

2.2 重建流程

假设两个相机的内外参数均已标定。重建的流程如下:

(1) 运动散斑立体图像对采集

投影器在计算机的控制下做往复运动, 在目标表面产生随时间变化的运动光斑。双目相机连续采集 K 对图像对。

(2) 立体校正

利用标定好的相机参数, 对立体图像对进行立体校正, 使得左、右相机图像的极线处于同一条水平线上。从而在立体匹配时, 在图像上只需要沿着水平方向进行一维搜索。

(3) 立体匹配

在图 2 所示的系统中, 由于向目标表面投射了密集的斑点, 且斑点是随时间变化的, 根据空间-时间立体视觉原理, 采用三维的窗口进行立体匹配(如图 2(b)所示)。由于三维的窗口中包含丰富的图像信息, 采用归一化互相关匹配方法(Normalized Cross Correlation, NCC)^[2]即可得到致密的视差图。三维窗口 NCC 匹配度计算方法如式(1)所示。

$$ncc(x, y, d) = \frac{\sum_{i=1}^K \sum_{u, v \in \Omega(x, y)} (I_{i,L}(u, v) - \bar{I}_L)(I_{i,R}(u + d, v) - \bar{I}_R)}{\sqrt{\sum_{i=1}^K \sum_{u, v \in \Omega(x, y)} (I_{i,L}(u, v) - \bar{I}_L)^2 \sum_{i=1}^K \sum_{u, v \in \Omega(x, y)} (I_{i,R}(u + d, v) - \bar{I}_R)^2}} \quad (1)$$

其中, d 为视差值; $ncc(x, y, d)$ 表示以左图 I_L 中的像素点 (x, y) 为中心的 K 个图像块与右图 I_R 中以像素点 $(x - d, y)$ 为中心的 K 个图像块之间的归一化相关值; $\Omega(x, y)$ 为以 (x, y) 为中心的邻域; $\bar{I}_L$ 和 $\bar{I}_R$ 分别为左图和右图三维窗口中的像素灰度平均值。

为了加快匹配速度, 采用金字塔搜索策略, 实现由粗到精的匹配。采用三层金字塔, 当前层的图像宽度或高度为它下一层的 $1/2$ 。采用双向匹配剔除误匹配点, 即假设左图中的一个像素点 P_L 在右图中找到的对应点为 P_R , 再对右图中的点 P_R 在左图中进行反向匹配, 得到对应点 P'_L , 必须要满足 $|P_L - P'_L| \leq 1$, 否则认为是误匹配。

NCC 得到的视差图是整像素级的。为了得到亚像素级的匹配精度, 对以 NCC 得到的匹配位置为中心的 2 个像素进行细分, 细分间隔为 0.1 像素。再用 NCC 相似度对这 21 个位置进行搜索, 找到 NCC 得分最高的位置, 即为最终的亚像素匹配结果。采用这个方法, 理论上可以得到 $1/10$ 像素的匹配精度。

立体匹配得到视差图以后, 根据事先标定好的相机参数, 可将视差图转为三维点云。假设 B 为立体相机的基线长度, F 为等效焦距, 左相机的主点坐标为 (c_{xL}, c_{yL}) , 右相机的主点坐标为 (c_{xR}, c_{yR}) 。由于左右视图已经经过立体校正, 所以 $c_{yL} = c_{yR}$ 。假设左相机图像上的某个像素点 (x, y) 处的视差值

为 $d(x, y)$, 则该像素点在三维空间中的位置 $(X, Y, Z)^{[2]}$ 为:

$$X = \frac{(x - cx_L)Z}{F}, Y = \frac{(y - cy_L)Z}{F}, Z = \frac{BF}{d(x, y) - cx_L + cx_R} \quad (2)$$

3 匹配算法加速

3.1 积分图像优化

除了采用 1.2 节中提到的金字塔搜索测量之外, 为了对匹配算法进行进一步的加速, 本文采用积分图像方法计算式(1)分子中的 $\bar{I}_L, \bar{I}_R$ 和分母。

图像 $f(u, v)$ 的积分 $s(u, v)$ 的计算方法^[2]为:

$$s(u, v) = f(u, v) + s(u-1, v) + s(u, v-1) - s(u-1, v-1) \quad (3)$$

其中, 当 $u, v < 0$ 时, $s(u, v) = 0$ 。

同理, $f^2(u, v)$ 的积分 $s^2(u, v)$ 计算方法为:

$$s^2(u, v) = f^2(u, v) + s^2(u-1, v) + s^2(u, v-1) - s^2(u-1, v-1) \quad (4)$$

用积分图像计算以 (u, v) 为左上起点、宽度为 N 的图像块的像素灰度值累加和的方法为:

$$\text{sum}(u, v) = s(u+N-1, v+N-1) - s(u-1, v+N-1) - s(u+N-1, v-1) + s(u-1, v-1) \quad (5)$$

采用积分图像方法, $\text{sum}(u, v)$ 的计算与图像块的尺寸无关。利用式(5), 可快速计算出式(1)中的 $\bar{I}_L, \bar{I}_R$ 。以类似的方法, 将式(1)分母展开, 可以利用积分图像进行快速计算。

3.2 GPU 并行化加速

式(1)除了分子中的相关运算 $\sum_{i=1}^K \sum_{u,v \in \Omega(x,y)} I_{i,L}(u, v) I_{i,R}(u+d, v)$ 不能采用积分图像计算之外, 其余的部分都可以采用积分图像方法进行快速计算。本文对这部分采用 GPU 进行并行计算。由于本文采用局部的 NCC 计算图像块之间的相似度, 因此每个像素块的计算是独立的, 因此, 非常适合采用 GPU 做一步的加速。本文采用 CUDA 在 NVIDIA GPU 上实现相关运算的并行加速。

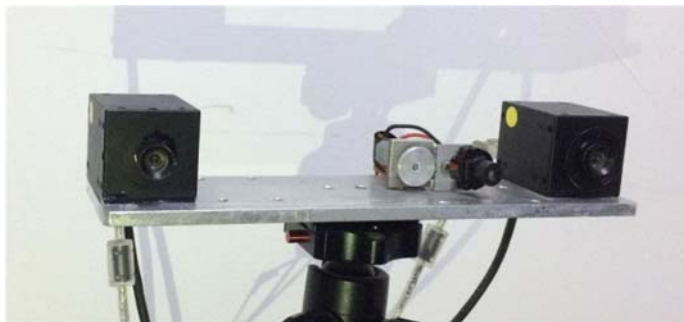


图 3 本文设计的三维测量系统样机

Fig. 3 The prototype of the proposed 3D measurement system

4 实验结果

图 3 为本文设计的三维测量系统样机。相机的分辨率为 1280×960 , 帧率为 36fps, 散斑投影器的波长为 830nm, 相机镜头的焦距为 6mm, 双目相机的基线长度为 25mm 左右, 散斑投射器和相机通过硬触发电路板进行同步。图 4(a)为测量系统采集的一对散斑图像, (b)~(d)分别为空间-时间立体视觉方法在 $K=1$ 、 $K=5$ 和 $K=10$ 时的重建结果, 可以看出, 随

着 K 的增大, 重建精度得到显著提升。图 4(e)为 Kinect 得到的同一个人的点云数据, 显然, 空间-时间立体视觉方法的结果具有更好的重建精度。

在配置为 I7-6700HQ CPU@2.6GHz、NVIDIA Geforce GTX 960M 显卡的笔记本电脑上, 用 10 对立体图像 ($K=10$) 计算一片点云的时间为 1.2 秒左右。与仅利用 CPU 原图分辨率上直接计算式(1)相比, 加速了 875 倍。

为了能够对人体进行快速三维扫描, 可以采用多套如图 3 所示的三维成像单元, 从多个视角对人体进行扫描。图 5 为由 13 套这样的三维成像单元得到的三维人体表面模型, 其中 1 套安装在人的头顶上方, 另外 12 套安装在 4 根立柱上 (每根立柱上安装 3 套)。最终的三角网格表面模型由泊松重建算法^[9]得到。

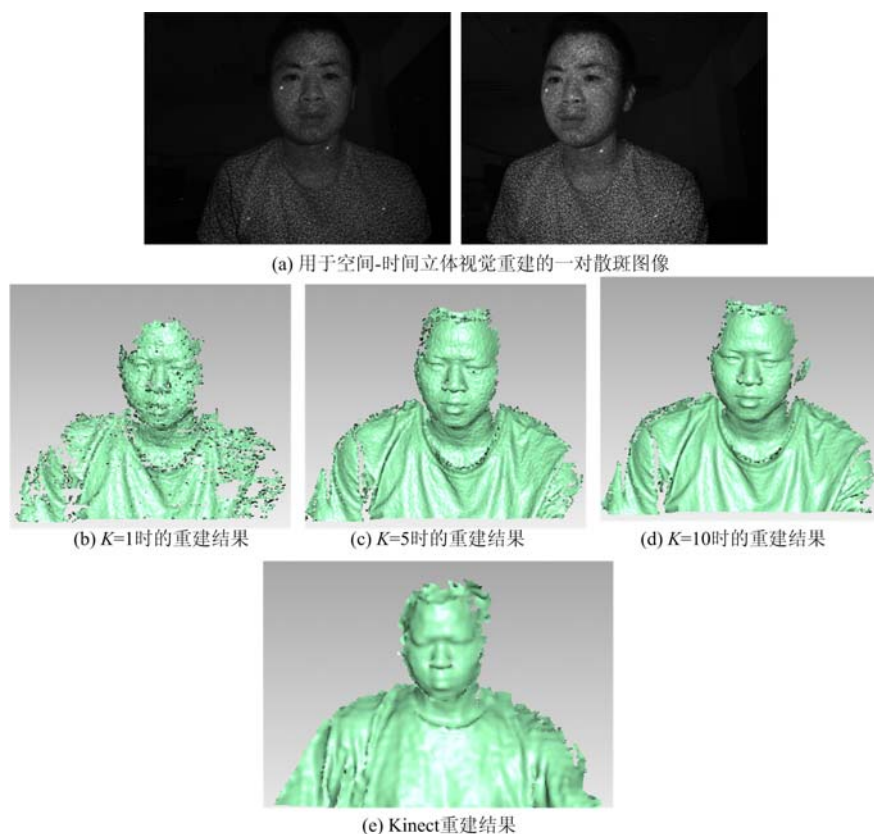


图 4 空间-时间立体视觉重建结果与 Kinect 比较

Fig. 4 The comparison results between the space-time stereo matching and Kinect



图 5 多视系统得到的人体全身重建结果

Fig. 5 The whole body reconstruction results based on multi-view system

5 结论

本文设计了一种基于近红外运动散斑的低成本双目立体视觉系统。将一个售价只需要几十元的散斑投影器安装在一个摇摆机构上,在目标表面形成运动的散斑。利用空间-时间立体视觉原理,采用一个三维的滑动窗口进行立体匹配。与传统的立体视觉方法相比,该方法具有更高的重建精度。此外,本文采用金字塔搜索策略、积分图像和 GPU 对立体匹配算法进行加速,使得立体匹配的计算速度提高了 800 多倍。在未来的工作中,将基于本文方法,开发手持式三维测量系统。

参考文献:

- [1] Blais F. Review of 20 years of range sensor development[J]. Journal of Electronic Imaging, 2004, 13(1):228—240.
- [2] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. International Journal of Computer Vision, 2002, 47(1-3):7—42.
- [3] Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(8):1362—1376.
- [4] Zhang S, Huang P S. High-resolution, real-time three-dimensional shape measurement[J]. Optical Engineering, 2014, 45(12):1269—1278.
- [5] Basri R, Jacobs D, Kemelmacher I. Photometric stereo with general, unknown lighting[J]. International Journal of Computer Vision, 2007, 72(3):239—257.
- [6] Zhang R, Tsai P S, Cryer J E, et al. Shape-from-shading: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(8):690—706.
- [7] Zhang L, Curless B, Seitz S M. Spacetime stereo: shape recovery for dynamic scenes[C]//Computer Vision and Pattern Recognition, 2003. Proceedings 2003 IEEE Computer Society Conference on. IEEE, 2003, 2:II-367-74 vol. 2.
- [8] Newcombe R A, Izadi S, Hilliges O, et al. KinectFusion: real-time dense surface mapping and tracking[C]//Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on. IEEE, 2011:127—136.
- [9] Newcombe R A, Fox D, Seitz S M. Dynamicfusion: reconstruction and tracking of non-rigid scenes in real-time [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015:343—352.
- [10] Dekiff M, Berssenbrügge P, Kemper B, et al. Three-dimensional data acquisition by digital correlation of projected speckle patterns[J]. Applied Physics B, 2010, 99(3):449—456.
- [11] Sagawa R, Sakashita K, Kasuya N, et al. Grid-based active stereo with single-colored wave pattern for dense One-shot 3D scan[C]. Second Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization & Transmission, 2012.
- [12] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction[C]//Proceedings of the fourth Eurographics symposium on Geometry processing. 2006, 7. Chinese Journal of Computers, 2007, 30(3):361—367.

On the measurement method of three-dimensional topography of object based on moving speckle

SUN Jun-feng¹, XU Yu-hua¹, ZHAO Xiao-feng², XIE Liang¹, ZHANG Xiao-hu¹

(1. College of Aero Science and Engineering, National University of Defense Technology, Changsha 410073, Hunan, China;

2. Xi'an Research Institute of Hi-Tech, Xi'an 710025, Shannxi, China)

Abstract: A low-cost binocular stereoscopic vision system based on near-infrared (NIR) moving speckle is designed and presented in this paper. In this system, a low cost NIR projector is mounted on a wobbler mechanism to project moving speckles on the surface of the target. The space-time stereo method is used to compute the disparity map by combing the spatial and temporal information. The advantage is that a 3D sliding matching window can be used to establish the pixel correspondences between the images of the left and right cameras. Experimental results show that the method presented outperforms the traditional stereo methods in accuracy. Additionally, the stereo matching algorithm is speed up by more than 800 times by combing the pyramid searching strategy, integral image technique, and the graphics processing unit (GPU).

Keywords: stereo vision; stereo match; spacetime stereo; moving speckles