http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2017.0653

基于置信度的 TOF 与双目系统深度数据融合

孙哲，张勇*，常衢通
（北京航空航天大学 仪器科学与光电工程学院，北京 100083）

摘 要：为了解决铁路火车等结构单一环境的三维重建问题，提出了置信度的概念，将 TOF 系统与双目系统的互补性特点有效结合。通过联合标定，建立起 TOF 系统与双目系统的坐标关系，将 TOF 中的点映射到左相机视角下，得出双目系统左相机视角下的 TOF 测量视差图，再利用图像分割以及曲面拟合对其上采样处理至双目图像的分辨率大小。根据各系统特点定义置信度，确定数据融合的不同系统权重。利用 Middlebury 的数据集处理结果，融合后的匹配精度较双目系统精度提高一倍以上，且视差图的分辨率提升至与双目系统相同大小。

关 键 词：双目立体匹配；TOF；数据上采样；数据融合；置信度
中图分类号：V219；TP391.4
文献标识码：A **文章编号**：1001-5965(2018)08-1764-08

场景的三维重建技术一直是计算机视觉领域研究的热点和重点。摄像机模型将空间中三维的点映射到二维的平面上，使得深度信息缺失，给三维重建带来了挑战。目前，主流获取深度信息的方法有雷达和声呐等非光学测量方法，以及以激光扫描和红外线测量为代表的主动光测量方法，以多视角立体视觉为代表的被动光测量方法[1]。其中非光学方法在检测透明物体上有很大优势，尤其在室内环境检测中应用较多，但是其劣势在于信号衰弱严重、精度较低。

近年来，3D 视觉的研究与应用正呈现迅猛发展的态势，这是由于低成本的 3D 感知硬件的可用性增长所驱动的。如 Microsoft 公司于 2010 年发布了 Kinect，利用了结构光技术快速重建 3D 场景。Occipital 于 2015 年使用相同的原理，提供可连接的 3D 传感器。然而这类传感器的使用会有明显限制，主动光由于发散等物理条件约束，干扰后续算法如中心位置提取等处理[1]。最重要的是，这样会丢失图像的精细化结构，比如高纹理区、物体边缘和明光照射环境等。此外，对于高速运行的物体进行重建，面阵芯片的分辨率较低且采样速率较低，无法满足场景需求。TOF(Time of Flight)相机同样采用主动光探测方式，但其目的不是照明，而是利用入射光信号与反射光信号的变化来进行距离测量[2]。因此在室外等环境中，TOF 比结构光技术应用更加广泛。同时，TOF 的深度测量受物体表面灰度特征影响较小，测量准确度相对较高。但是，TOF 相机分辨率较低，无法重建精细物体结构。

与 TOF 相机获得深度数据的方法不同，双目立体视觉三维重建可以恢复高纹理区更精细化的结构，且在高速检测中易获得较高分辨率图像[3]。但是双目立体匹配算法对场景要求较高，弱纹理区会产生很多误匹配，非前向平行平面易重建为阶梯状，边缘发生膨胀。因此，在有些应用场景中，双目立体系统获取的深度信息较不可靠。

目前，铁路等交通运输行业对零件等缺损检测，愈加依赖于三维信息。而市场上的 3D 产品

收稿日期：2017-10-23；**录用日期**：2018-01-05；**网络出版时间**：2018-02-10 17:11
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20180210.1540.002.html
* **通信作者**. E-mail：06952@buaa.edu.cn

引用格式：*孙哲，张勇，常衢通. 基于置信度的 TOF 与双目系统深度数据融合[J]. 北京航空航天大学学报，2018，44(8)：1764-1771. SUN Z，ZHANG Y，CHANG Q T. In-depth data fusion of TOF and stereo vision system based on confidence level[J]. Journal of Beijing University of Aeronautics and Astronautics，2018，44(8)：1764-1771 (in Chinese).*

很少能满足铁路货车车底三维重建精度和速度的要求。以运行速度为 120 km/h 的车底应用场景为例，基于 TOF 的深度相机拍摄速度目前可以达到 5 000 帧/s，即深度信息的精度仅可以达到 5 mm。对于车底大型部件的细节和螺纹等小型部件的轮廓，会造成深度信息缺失。而对双目系统而言，车底的图像环境简单，缺少纹理，极易造成立体算法的误匹配，导致错误的深度信息。

综合上述 TOF 相机和双目立体匹配算法等技术的优缺点，可以建立一个多相机融合系统。包括 TOF 相机（低分辨率）、双目立体相机（高分辨率），实现 3D 与 2D 协同成像，在拍摄高质量火车 2D 图像的同时，输出与 2D 图像高度匹配的高精度 3D 图像。利用数据融合策略，既提升了 3D 图像的分辨率，也改善了重建精度。

基于双目与 TOF 系统具有互补的特性，科研人员开展了许多将两系统结合的研究工作。一些研究的重点放在 TOF 系统与单相机系统的结合上。文献[4]提出了基于双边滤波器的方法，利用彩色图像信息对于 TOF 上采样进行引导。文献[5]利用边缘保护的方法对 TOF 数据进行插值处理。但是 TOF 与双目系统的结合会得到更准确的视差结果，因为 2 个系统都可以得到各自的深度图。文献[6]根据贝叶斯理论提出了基于后验概率的马尔可夫随机场模型，并利用置信传播的方法进行优化求解，然而此类全局性算法计算复杂度较大。文献[7]将 TOF 深度数据转化为代价约束双目立体匹配算法，并未考虑 TOF 数据无效情况对双目立体匹配算法的干扰，且融合效果很大程度依赖于双目立体匹配算法。文献[8]采用局部算法将 TOF 系统与双目系统置信度聚合，但是默认两系统的权重值相等。本文根据两系统的优缺点，针对不同区域对其测量结果进行权重赋值，以得到最优效果。

1　双目立体匹配算法

1.1　算法原理

双目系统由 2 个相机从不同的位置对同一目标进行拍摄，通过内外参标定、畸变校正[9]和立体校正[10]，输出标准前向平行图像，保证匹配点对的像素纵坐标 Y 相同，如图 1 所示。然后通经过立体匹配算法实现图像的稠密匹配，即可计算得到视差值。在得到视差值后根据三角测量原理可以计算得出物体的深度信息，进而得到 3D 点云数据。

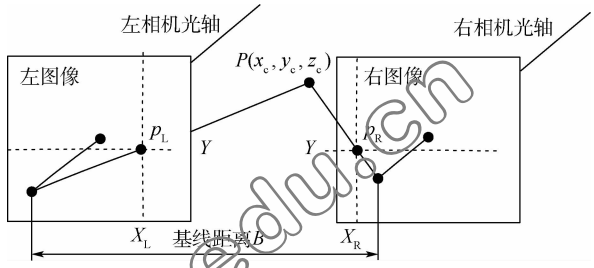


图 1　双目系统示意图

Fig. 1　Schematic diagram of stereo vision system

根据图 1，对空间点 P 在左右图像上的对应像点 p_L 和 p_R 具有相同的 Y 坐标。设 C_L 和 C_R 分别为左、右相机的光心，光心之间的距离为 B，相机焦距为 f。

设 P 点在左相机坐标系下的坐标为（x_L，y_L，z_L），在左、右图像上像点对应横坐标为 X_L 和 X_R，则视差 d 可表示为

$$d = X_L - X_R \tag{1}$$

根据三角原理，点 P 在相机坐标系下的三维坐标为

$$\begin{cases} x_L = \dfrac{BX_L}{d} \\ y_L = \dfrac{BY}{d} \\ z_L = \dfrac{Bf}{d} \end{cases} \tag{2}$$

因此，当基线距离 B、相机焦距 f 通过标定得到后，若左相机图像上的一点能够在右相机图像上找到对应的匹配点，即得到视差 d，继而确定出该点在做相机坐标系下的三维坐标[9]。

1.2　局部匹配算法

一套完整的双目立体匹配系统主要包括：图像采集、相机标定与畸变矫正、立体匹配和三维重建。其中立体匹配是整个双目立体视觉中的核心部分，一般分为全局和局部 2 种算法。全局匹配算法通过构建匹配代价函数，不断迭代优化求解全局最优解，其精度比局部匹配算法高，但计算复杂度高、耗时长。而局部匹配算法计算复杂度低、实时性好，精度也可达到一定要求，因此得到广泛的应用。为保证系统的实时性能，本文采用局部匹配算法。

局部匹配算法一般包括代价计算、代价聚合、视差计算和视差优化 4 个部分[11]。基于生物视觉上的观察，以人眼立体匹配为参考给出“由粗到细”（coarse-to-fine）的策略进行立体匹配。

1.2.1　代价计算

常用的匹配代价主要有灰度绝对差和、平方

http://bhxb.buaa.edu.cn

差和、Census 转换等。

绝对差和：

$$C(p,d)=\sum_{p\in I}|I_{\mathrm{L}}(p)-I_{\mathrm{R}}(p-d)| \tag{3}$$

平方差和：

$$C(p,d)=\sum_{p\in I}\|I_{\mathrm{L}}(p)-I_{\mathrm{R}}(p-d)\| \tag{4}$$

Census 转换：

$$\begin{aligned}&\mathrm{cen}(p)=\otimes_{q\in w_p}c(p,q)\\&c(p,q)=\begin{cases}1 & I(p)>I(q)\\0 & \text{其他}\end{cases}\\&C(p,d)=\mathrm{cen}^{\mathrm{L}}(p)\oplus\mathrm{cen}^{\mathrm{R}}(p-d)\end{aligned} \tag{5}$$

式中：$C(p,d)$为左图像 p 处像素在视差 d 下的初始匹配代价；I_{L} 和 I_{R} 分别为左、右图像上对应位置的像素大小；w_p 为像素 p 的邻域窗口；⊗表示串联操作；⊕表示求汉明距离，即逐位进行异或运算。Census 变换是利用当前点与周围像素点之间的分布关系构建像素点特征向量，并采用汉明距离度量 2 个特征子串的匹配程度，如式(5)所示。Census 保留了窗口中像素的位置特征，并且对亮度变化有较强稳健性，即能够减少光照误差引起的误匹配[11]。

1.2.2 代价聚合

通过对比基于 ST(Segment-Tree)、EF(Edge-aware Filter)、BF(Bilateral Filter)、GF(Guided Filter)、NL(Non-Local)和 Box 等聚合算法的视差结果，以 Middlebury 数据库中图像 Teddy 为例，如图 2所示。对比图中 Teddy 熊与周围背景的视差细节可以发现，ST、EF、BF 和 Box 聚合算法的结果图中存在大量噪点，而 GF 聚合算法在前景与背景的深度不连续区域存在很多误匹配。综合考虑算法复杂度和算法效率等因素，最终确定双目聚合算法采取 NL。

1.2.3 视差计算

局部算法在视差计算中一般采用胜者为王(Winner-Takes-All,WTA)的计算策略。在匹配代价聚集之后能够得到一系列待定的视差值与对应的匹配代价，WTA 是在这些视差值中，选择匹配代价最小的视差值作为视差计算的结果[3]。

1.2.4 视差优化

视差计算所得的初步视差图存在许多误匹配与无效视差。由于双目图像中存在许多由遮挡、视场角不同等原因所带来的匹配难点，并非所有像素点都在双目图像中可以得到匹配像素[4]。对于只在单幅图像中可见的区域，并不能由双目立体匹配算法进行重建，并且之前处理中会存在一些误匹配的区域。因此，在视差计算之后还需要一个优化求精的过程。



图 2 不同代价聚合算法的视差图对比

Fig. 2 Comparison of disparity maps generated by different cost aggregation algorithms

1）左右视差不一致性检测

在得到初始视差图像后，首先对左、右图像的视差图像进行不一致性检验[3]。

$$|D_{\mathrm{L}}(p)-D_{\mathrm{R}}(p-D_{\mathrm{L}}(p))|\leqslant\gamma \tag{6}$$

式中：D_{L} 和 D_{R} 分别为左图像和右图像的视差；γ 为检测阈值，通常取 $\gamma=2$。当左图像上一点的视差与右图像对应点视差相差较大时，认为该点视差为异常值。

2）空洞填充

经过上述步骤后，标记出不可信的点。这些点在视差图中产生大量的无效空洞，需要用周围的有效值进行填充。

定位所有的无效点，对每个无效点遍历如下操作：①分别向左、向右搜索，直到发现左边或者右边第 1 个有效视差的点，停止搜索；②若①中同时找到左、右第 1 个有效视差，则选取其中最小的一个赋值给当前点；若只发现其中的一个，则直接将其赋值给当前点。

由于离有效点近的无效点赋值的可信度远高于远离有效点的无效点，因此，执行完以上过程，并不能输出准确的视差图。尤其是一些空洞比较大的区域，赋值后视差的不可信程度急剧增加。

http://bhxb.buaa.edu.cn

故还需要把 2D 图像作为引导图像,参与滤波修补。

2　TOF 系统置信度

TOF 相机的测量原理:首先,通过连续发射经过调制的特定频率的光脉冲到被观测物体上;然后,接收从物体反射回去的光脉冲,通过探测光脉冲飞行的往返时间来计算被测物体离相机的距离,同时记录光线强度信息[12]。

如文献[1]中所述,TOF 的置信度测量受很多因素的影响,如被测物体表面反射率,测量距离,以及对深度不连续处周围的像素测量,因此很难实现对每个像素位置的置信度测量。但是在后续的融合算法中,需要对每个像素位置给出准确的置信度值。因此,本文利用了 TOF 相机的辐射特性来定义每个像素处的置信度,包括接收信号的振幅和强度。

TOF 接收到返回信号的振幅和强度取决于很多因素,其中最重要的是表面反射特性,以及与相机的距离 2 个方面。虽然也可以直接使用距离定义置信度,但是使用振幅和强度定义置信度考虑了除距离外的其他信息。接收到的振幅对测量的精度影响很大,振幅越大,信噪比越高,因此,测量精度越高。根据文献[1,13],TOF 相机测得深度值的噪声分布符合高斯标准差:

$$\sigma_z = \frac{c}{4\pi f_{\mathrm{mod}}} \cdot \frac{1}{\mathrm{SNR}} = \frac{c}{4\pi f_{\mathrm{mod}}} \cdot \frac{\sqrt{I/2}}{A} \tag{7}$$

式中:f_{mod} 为 TOF 信号的频率;SNR 为信噪比;A 为振幅;I 为强度;c 为光速。双目系统获得的初始信息是视差值,因此,需要将 TOF 的深度值转换为视差值。同理,TOF 系统的置信度也需要转化为与视差值的关系。对于测得的深度值 z,设其标准差为 σ_z,根据式(2),对应的视差标准差 σ_d 计算式为

$$2\sigma_d = |d_1 - d_2| = \frac{Bf}{z - \sigma_z} - \frac{Bf}{z + \sigma_z} = Bf\frac{2\sigma_z}{z^2 - \sigma_z^2} \tag{8}$$

化简为

$$\sigma_d = Bf\frac{\sigma_z}{z^2 - \sigma_z^2} \tag{9}$$

为了利用视差值的标准差构造 TOF 系统置信度 r_{T},经过多次试验后,定义 2 个阈值 $\sigma_{\min}$ 和 $\sigma_{\max}$。其中 $\sigma_{\min}$ 为接收到的景深最近处最亮像素的标准差;$\sigma_{\max}$ 为接收到的景深最远处最暗像素的标准差。当测得该处视差值标准差小于 $\sigma_{\min}$ 时,认为该点完全稳定,即置信度为 1;当标准差大于 $\sigma_{\max}$ 时,认为该点不可靠,即置信度为 0;标准差处于 2 个阈值之间时,设置信度范围为(0,1),具体如下:

$$r_{\mathrm{T}} = \begin{cases} 1 & \sigma_d \leqslant \sigma_{\min} \\ \dfrac{\sigma_{\max} - \sigma_d}{\sigma_{\max} - \sigma_{\min}} & \sigma_{\min} < \sigma_d < \sigma_{\max} \\ 0 & \sigma_d \geqslant \sigma_{\max} \end{cases} \tag{10}$$

3　融合算法

根据双目立体数据和 TOF 相机 3D 数据的特性,采用如下方案生成最终的高分辨率和高精度 3D 点云数据。分别获取双目系统和 TOF 系统的视差数据,根据 2 个系统的优缺点分配不同的权重,融合视差数据,输出点云文件。方案的整体流程如图 3 所示。

多系统间数据融合的核心技术主要有以下 3 个方面:

1)数据角度的统一,即从不同相机视角获得的数据通过联合标定,转换到左相机视角下。

2)TOF 深度图上采样,生成高分辨率、稠密视差[14]。根据 2 个系统的数据,逐个像素点进行融合,因此,必须通过一定的插值算法,将低分辨率图像转化为高分辨率图。

3)融合策略[15]。假设 d_{T} 和 d_{S} 分别为 TOF 系统和双目系统的视差,最终生成的视差为 d,则有如下关系:

$$d = w_{\mathrm{T}}d_{\mathrm{T}} + w_{\mathrm{S}}d_{\mathrm{S}} \tag{11}$$

式中:w_{T} 为 TOF 系统求得的视差值权重;w_{S} 为双目系统求得的视差值权重。

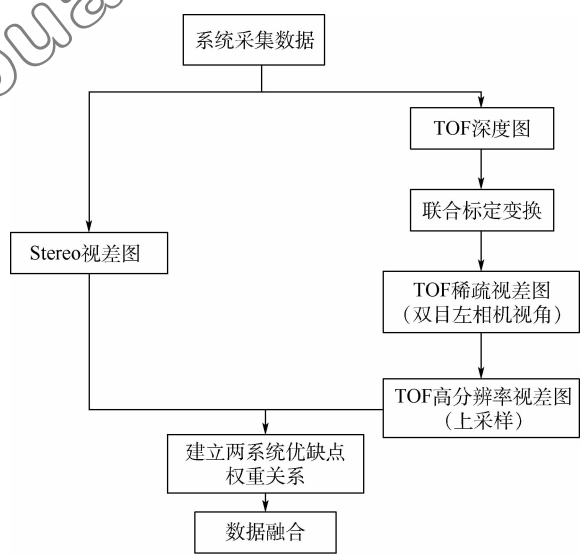


图 3　TOF 与双目系统数据融合示意图

Fig. 3　Schematic diagram of data fusion of TOF and stereo vision system

http://bhxb.buaa.edu.cn

3.1 TOF系统高分辨率视差图

TOF相机获取的深度图分辨率较低,因此在TOF深度图和双目匹配得到的视差图融合前,首先对TOF深度图进行稀疏式上采样,步骤如下:

步骤1 将TOF相机产生的点云通过联合标定的变换矩阵[$\boldsymbol{R},\boldsymbol{T}$]映射到左相机的相机坐标系[5],$\boldsymbol{R}$为旋转矩阵,$\boldsymbol{T}$为平移矢量。即将TOF相机的中点的三维坐标$X_{\mathrm{T}i}(i=1,2,\cdots,N)$变换到以左相机为参考的双目系统的三维点$X_{\mathrm{L}i}(i=1,2,\cdots,N)$。

步骤2 将双目系统的三维点$X_{\mathrm{L}i}$,通过左相机的内参矩阵投影到左相机图像坐标系中,从而形成点阵$x_{\mathrm{L}j}(j=1,2,\cdots,n)$。这里相机将空间3D点投影变换为平面2D点,因此,必然存在部分点被遮挡,导致成像点数目与空间点数目不同。

步骤3 根据双目系统中左相机坐标系下的深度z与视差d的关系,可以将TOF得到的深度图转换为与该双目系统对应的视差图。

这样实现将TOF相机视差图的分辨率提升到与左相机一样的水平,如图4所示。然而此时TOF的视差图仍为稀疏,还需要对缺失的数据进行填充[16-18]。

根据稀疏的已知视差,在视差空间拟合出平滑的视差曲面,即可根据此视差曲面进行上采样对未知视差进行填充。显然二次曲面在一些复杂较扭曲场景下不能满足需求,而曲面方程的幂次越高,计算复杂度越大。因此,为了保证拟合的视差曲面最大程度上接近真实场景,此处采用二元三次方程对视差曲面进行拟合[19]:

$$d(x,y)=a_1+a_2x+a_3y+a_4x^2+a_5xy+a_6y^2+a_7x^3+a_8x^2y+a_9xy^2+a_{10}y^3 \tag{12}$$

式中:$d(x,y)$表示一个三维视差曲面;$a_1,a_2,\cdots,a_{10}$为系数;x和y为像素坐标。

经过以上步骤,便可得到左相机视角下的高分辨率TOF相机视差图。

3.2 双目系统与TOF系统数据融合

数据融合的目的是通过不同权重将2个系统对应的视差值融合在一起,以得到高分辨率、高稠密的视差图。具体的融合权重确定方法如下:

1)弱纹理区域

对双目图像进行梯度提取,可以确定弱纹理区域,在此部分的融合视差全部取TOF相机获得的视差值,即$w_{\mathrm{S}}=0,w_{\mathrm{T}}=1$。

2)非弱纹理区域

首先计算双目系统视差图的置信度r_{S}:



(a) 左相机视角2D图像(4 096像素×2 000像素)



(b) TOF深度图(1 536像素×667像素)



(c) 升分辨率后左相机视角下的视差图(4 096像素×2 000像素)

图4 TOF系统高分辨率视差图

Fig. 4 Disparity map of TOF system with high resolution

$$r_{\mathrm{S}}=\begin{cases}1-\dfrac{C_p^{1\mathrm{st}}}{C_p^{2\mathrm{nd}}} & C_p^{2\mathrm{nd}}>T_{\mathrm{c}}\\ 0 & \text{其他}\end{cases} \tag{13}$$

式中:$C_p^{1\mathrm{st}}$和$C_p^{2\mathrm{nd}}$分别为双目中求得的最小和次小代价。此外,为避免$C_p^{2\mathrm{nd}}$为0,设常参数$T_{\mathrm{c}}>0$。

此外,对TOF数据上采样后得到的视差图不可避免地会出现无效视差,针对上采样后的TOF视差图,将像素分为2类TOF数据有效区域T_{valid}以及无效区域T_{invalid},p为图像上的像素。双目系统视差权重分配如下:

$$w_{\mathrm{S}}=\begin{cases}\dfrac{r_{\mathrm{S}}}{r_{\mathrm{S}}+r_{\mathrm{T}}} & p\in T_{\mathrm{valid}}\\ 1 & p\in T_{\mathrm{invalid}}\end{cases} \tag{14}$$

式中:r_{T}为式(10)定义的TOF系统置信度;r_{S}为式(13)定义的双目系统置信度。即当TOF像素视差有效时,采取两系统视差加权结果;否则采用双目视差结果。

确定双目权重后,相应的TOF系统权重也可确定:

http://bhxb.buaa.edu.cn

$$w_T = 1 - w_S \tag{15}$$

最后，即可根据式(11)的融合策略进行两系统视差结果的融合。

4　实验结果与分析

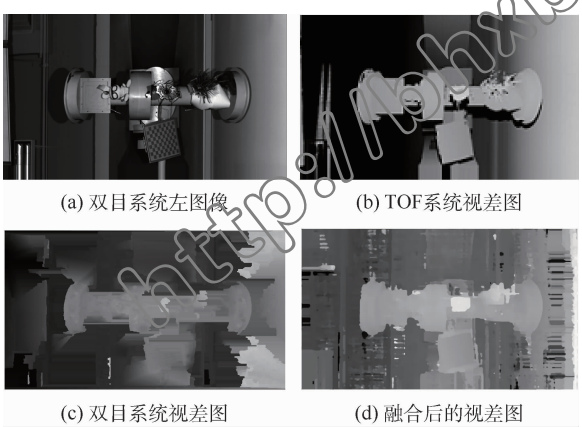
首先，搭造出模拟火车车底的环境拍摄图像，进行联合标定，建立 TOF 系统与双目系统之间的坐标关系。然后，对 TOF 图像进行深度转视差以及上采样处理，对双目图像进行畸变矫正，立体匹配。图 5 展示了样图处理结果。

根据图 5 可以直观地看出，TOF 系统直接得到的视差图比双目系统得到的视差图结果更精确，但是分辨率较低。经过数据融合之后，得到的视差图在分辨率方面比 TOF 结果较好，精度比双目系统较好。由图 5(b)和(c)可知，TOF 系统测量结果基本没有将左侧背景深度有效测量出，而双目系统测量结果对于物体细节的深度测量表现不好。经过融合后，对两系统各自不同测量缺陷都有效减小。

但图 5 示例只能定性地分析系统性能，为了可以定量分析系统的精度，本文利用 Middlebury 数据集进行定量评测。

TOF 系统得到视差图精度高，但分辨率低。首先，将 Middlebury 数据集的 4 组数据的视差图同时沿横、纵轴方放大 4 倍[20]，对其进行高斯滤波；再利用本文方法采样得到高分辨率视差图，与双目立体匹配算法的视差图结果进行融合。以 Teddy 为例，具体结果如图 6 所示。

从图 6 直观上看，Teddy 图像的数据融合后的视差图误匹配区域较双目系统减少。此外，数据集处理结果的定量分析如表 1 所示。

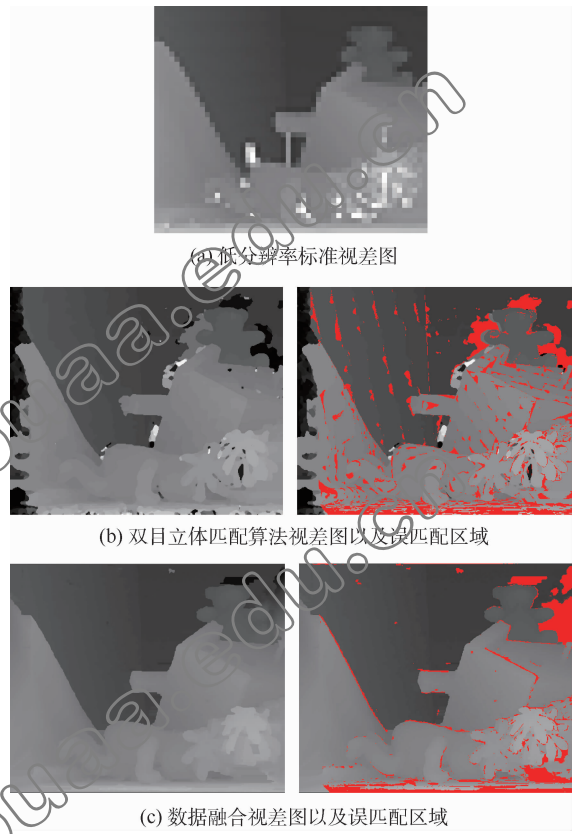


(a) 双目系统左图像　(b) TOF系统视差图

(c) 双目系统视差图　(d) 融合后的视差图

图 5　样图处理结果示例

Fig. 5　Examples of processing results



(a) 低分辨率标准视差图

(b) 双目立体匹配算法视差图以及误匹配区域

(c) 数据融合视差图以及误匹配区域

图 6　Teddy 处理结果示例

Fig. 6　Examples of Teddy processing results

表 1　双目立体匹配和融合算法结果分析

Table 1　Analysis of stereo vision and fusion algorithm results　%

图　像	双目立体匹配算法误匹配率	融合算法误匹配率
Teddy	13.34	6.34
Cones	8.56	4.63
Tsukuba	6.52	2.39
Venus	4.79	1.43

5　结　论

1）针对 TOF 系统，考虑到系统精度与接收到信号振幅的关系，确定了 TOF 系统的置信度与振幅的关系式；针对双目系统，分为弱纹理区与非弱纹理区，并根据聚合代价的全局最小值与次小值，确定其置信度；即本文提出的融合算法将两系统的优点有效结合。

2）根据标准数据集的实验结果，不同系统的视差融合之后的精度，比双目立体匹配算法的视差图精度提高了一倍以上，尤其平面区域更加平滑，且分辨率也较 TOF 系统高。

参考文献（References）

[1] ZANUTTIGH P, MARIN G, DAL MUTTO C D, et al. Time-of-flight and structured light depth cameras: Technology and appli-

http://nhxb.buaa.edu.cn

cations[M]. Berlin:Springer,2016:23-25.

[2] HANSARD M,LEE S,CHOI O,et al. Time-of-flight cameras: Principles, methods and applications[M]. Berlin: Springer, 2013:50-53.

[3] TIPPETTS B,LEE D J,LILLYWHITE K,et al. Review of stereo vision algorithms and their suitability for resource-limited systems[J]. Journal of Real-Time Image Processing, 2016, 11(1):5-25.

[4] YANG Q,AHUJA N,YANG R,et al. Fusion of median and bilateral filtering for range image up-sampling[J]. IEEE Transactions on Image Processing,2013,22(12):4841-4852.

[5] GARRO V,MUTTO C D,ZANUTTIGH P,et al. A novel interpolation scheme for range data with side information[C]// Proceedings of the 2009 Conference on Visual Media Production. Piscataway,NJ:IEEE Press,2009:52-60.

[6] ZHU J,WANG L,GAO J,et al. Spatial-temporal fusion for high accuracy depth maps using dynamic mrfs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2010,32(5):899-909.

[7] VINEET G, JAN C, RADU H. High-resolution depth maps based on ToF-stereo fusion[C]// Proceedings of IEEE Conference on Robotics and Automation. Piscataway,NJ:IEEE Press, 2012:4742-4749.

[8] MUTTO C D,ZANUTTIGH P,MATTOCCIA S,et al. Locally consistent ToF and stereo data fusion[C]// Proceedings of the 12th International Conference on Computer Vision-Volume Part I. Berlin:Springer,2012:598-607.

[9] FUSIELLO A,TRUCCO E,VERRI A. A compact algorithm for rectification of stereo pairs[J]. Machine Vision and Applications,2000,12(1):16-22.

[10] KAHLMANN T,INGENSAND H. Calibration and development for increased accuracy of 3D range imaging cameras[J]. Journal of Applied Geodesy,2008,2(1):1-11.

[11] 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.

[12] PIATTI D,RINAUDO F. SR-4000 and Camcube 3.0 time of flight (ToF) cameras:Tests and comparison[J]. Remote Sensing,2012,4(4):1069-1089.

[13] MUTTO C D,ZANUTTIGH P,CORTELAZZO G. Probabilistic ToF and stereo data fusion based on mixed pixels measurement models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(11):2260-2272.

[14] DOLSON J,BAEK J,PLAGEMANN C,et al. Up-sampling range data in dynamic environments[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway,NJ:IEEE Press,2010,1141-1148.

[15] HU X,MORDOHAI P. A quantitative evaluation of confidence measures for stereo vision[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(11) 2121-2133.

[16] GUDMUNDSSON S,AANAES H,LARSEN R. Fusion of stereo vision and time of flight imaging for improved 3D estimation [J]. International Journal Intelligent System Technologies and Applications,2008,5(3):425-433.

[17] EVANGELIDIS G,HANSARD M,HORAUD R. Fusion of range and stereo data for high-resolution scene-modeling[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,37(11):2178-2192.

[18] NAIR R,LENZEN F,MEISTER S,et al. High accuracy ToF and stereo sensor fusion at interactive rates[C]// Proceedings of European Conference on Computer Vision Workshops. Berlin:Springer,2012:7584.

[19] 刘娇丽,李素梅,李永达,等. 基于 TOF 与立体匹配相融合的高分辨率深度获取[J]. 信息技术,2016(12):190-193.
LIU J L,LI S M,LI Y D,et al. High-resolution depth maps based on TOF-stereo fusion[J]. Information Technology,2016 (12):190-193(in Chinese).

[20] 张康. 基于图像深度感知中的立体匹配和深度增强算法研究[D]. 北京:清华大学,2015:84-87.
ZHANG K. Stereo matching and depth enhancement in image-based depth perception[D]. Beijing: Tsinghua University, 2015:84-87(in Chinese).

作者简介:

孙哲 女,硕士研究生。主要研究方向:计算机视觉、三维重建。

张勇 男,博士,副教授。主要研究方向:模式识别、图像采集与处理。

常衢通 男,硕士研究生。主要研究方向:计算机视觉、三维重建。

http://bhxb.buaa.edu.cn

In-depth data fusion of TOF and stereo vision system based on confidence level

SUN Zhe, ZHANG Yong*, CHANG Qutong

(School of Instrumentation Science and Opto-electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: To solve the problem of 3D reconstruction in textureless environment like railway, the confidence coefficient is proposed to combine TOF and stereo vision system effectively. Through the joint calibration, the coordinate relationship between TOF and stereo vision systems is established. Then by projecting the points in TOF to left camera in stereo vision system, the disparity map of TOF is obtained. After image segment and surface fitting the disparity map is up-sampled and its resolution is equal to that of stereo images. According to the confidence coefficients of different systems, the system weight values of data fusion are defined. Finally, the proposed method is evaluated with Middlebury dataset, and the results show that the accuracy has been raised twofold or more, and the resolution of disparity map is equal to that of stereo images as well.

Keywords: stereo vision; TOF; data upsampling; data fusion; confidence coefficient

http://bhxb.buaa.edu.cn

Received: 2017-10-23; **Accepted**: 2018-01-05; **Published online**: 2018-02-10 17:11

URL: kns.cnki.net/kcms/detail/11.2625.V.20180210.1540.002.html

* **Corresponding author**. E-mail: 06952@buaa.edu.cn