

引用格式: ZHUANG Sufeng, JI Yong, TU Dawei, et al. Underwater RGB-D Camera Based on Binocular Stereo Vision[J]. Acta Photonica Sinica, 2022, 51(4):0404003

庄苏锋,吉勇,屠大维,等. 基于双目立体视觉的水下 RGB-D 相机[J]. 光子学报, 2022, 51(4):0404003

基于双目立体视觉的水下 RGB-D 相机

庄苏锋,吉勇,屠大维,张旭

(上海大学 机电工程与自动化学院, 上海 200444)

摘 要:为满足水下机器人水下目标识别、精细作业的需要,研制了一种融合 3D 点云和二维 RGB 图像为一体的 RGB-D 水下相机工程样机。针对水下环境的特殊性,考虑水下多层折射效应,建立水下相机成像及双目立体成像模型。将水下图像转换为空气中的图像来消除折射对重建精度的影响,进行对应点像素匹配,重建 3D 点云。采用基于颜色校正和暗原色先验的水下图像增强算法对左、右水下图像进行图像增强处理。最后,构建对准叠加模型将 3D 点云和二维彩色数据叠加融合,获得水下 RGB-D 图像数据。实验表明:研制的水下 RGB-D 工程样机具有较好的三维测量精度,3 m 远处的系统重建误差为 2.6 mm,颜色再现真实,3D 点云数据与 RGB 二维图像对准精准。这对于水下目标识别和水下机器人精细作业具有重要意义。

关键词:双目立体视觉;水下相机成像模型;水下图像增强;水下三维重建;点云色彩对准叠加模型;水下 RGB-D 相机

中图分类号:TH74;TP3

文献标识码:A

doi:10.3788/gzxb20225104.0404003

0 引言

正如人类利用两只眼睛获取场景二维、三维视觉信息才能在客观世界中行动和作业自如一样,机器人在环境中行动、开展目标识别和精细作业,同样需要场景及目标的 RGB 二维彩色图像及 3D 点云信息。目前市面上出现的微软 Kinect 及 Intel 英特尔 R200、R461 等产品具有同时获取 RGB 及 3D 点云的能力,是一种 RGB-D 相机,可以满足空气中机器人应用需要。水下机器人同样需要 RGB-D 视觉传感器,以完成水下目标识别及精细作业任务,但至今为止鲜有水下 RGB-D 相机商用产品报道。

目前,空气中的基于双目立体视觉技术的三维测量技术已经趋于成熟^[1-2]。但是,水下环境的特殊性给水下 RGB-D 相机研制带来了困难。首先,由于水介质与水中悬浮颗粒的吸收和散射特性,使得水下图像比空气中的图像存在更严重的颜色失真、对比度低、细节特征模糊等问题^[3-4]。国内外研究者为此开展相关研究,并取得了一定的效果。CHIANG J Y 等^[3]提出了一种波长补偿和去模糊的水下图像增强方法,但该方法无法处理降质明显的图像。CARLEVARIS B N 等^[5]根据不同波长光线在水下衰减速率不同的特性,利用水下图像三个颜色通道的衰减差异进行场景深度估计,进而对图像进行散射光消除提高图像质量。HITAM M S 等^[6]提出一种混合对比度受限的自适应直方图均衡化颜色模型对水下图像的可视性差问题进行改善,该方法虽然提高了水下图像的对比度,减少了噪声和伪影,但对图像颜色失真问题的改善不明显。王永鑫等^[7]提出一种基于迭代直方图均衡化水下图像增强算法,来提升水下图像的视觉对比度。各种图像增强技术的研究,为水下立体视觉提供了高质量的图像保障。此外,工业相机应用于水下场景时,相机须封装封在防水壳中,存在的折射现象使得光路弯曲,传统的相机模型失效。为此,KWON Y H 等^[8]深入研究了折射现象,并提出构建水下折射模型才是解决水下三维重建问题的根本解决方法。AGRAWAL A 等^[9]提出

基金项目:国家自然科学基金(Nos.62176149, 61673252)

第一作者:庄苏锋(1991—),男,博士研究生,主要研究方向为图像处理、机器视觉及三维测量。Email: sufengzhuang@163.com

导师(通讯作者):屠大维(1965—),男,教授,博士,主要研究方向为光电检测、精密机械及仪器、机器视觉及伺服控制。Email: tdw@shu.edu.cn

收稿日期:2021-11-11;录用日期:2021-12-24

<http://www.photon.ac.cn>

一种多层平面折射几何理论,用数学公式对光线在多种介质中的折射传播进行准确的描述。张文明等^[10]提出一种基于非平行双目视觉系统的转化条件和转化模型,将水下图像转化为摄像机与目标间无水存在的一般空气图像后再利用空气中的方法进行处理。WU X等^[11]提出了一种基于几何关系将水下图像转化为空中图像的图像校正方法以消除折射的影响。

虽然水下RGB-D技术的研究还不多见,但以上研究都为水下RGB-D相机作了技术铺垫。本文设计了一款基于双目立体视觉的水下RGB-D相机,并对相关技术开展研究。主要包括:考虑水下多层折射效应,建立水下相机成像及双目立体成像模型;将水下图像转换为空气中的图像来消除折射对重建精度的影响,进行对应点像素匹配,重建3D点云;采用基于颜色校正和暗原色先验的水下图像增强算法对左、右水下图像进行图像增强处理;最后,构建对准叠加模型将3D点云和二维彩色数据叠加融合,获得水下RGB-D图像数据。

1 水下相机成像及双目立体视觉模型

研究的水下RGB-D相机以双目立体视觉为基础,本节对水下相机成像模型和水下双目立体相机的成像过程进行分析。

1.1 水下相机成像模型

当相机应用于水下时,需要将它封装于防水密封舱体中,舱体前端具有玻璃窗口,光线的传播依次经过“水-玻璃-空气”三种介质,并最终在相机像平面成像。然而,由于三者折射率不同,光线在不同介质的分界面处发生折射,使得水下相机成像的物像关系不满足空气中小孔成像原理。考虑到窗口玻璃厚度较小且两面平行,同时相对于物距而言其折射影响可忽略不计,故简化的模型如图1所示。

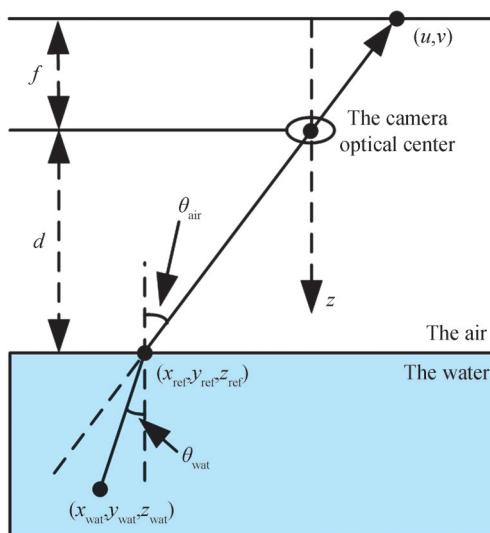


图1 水下相机成像模型

Fig.1 Underwater camera imaging model

水下目标点 $(x_{\text{wat}}, y_{\text{wat}}, z_{\text{wat}})$ 的成像光线经过折射平面上一点 $(x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}})$, 发生折射后在像平面的坐标为 (u, v) , 其中入射光线和折射光线的方向向量分别表示为 $(\alpha_{\text{wat}}, \beta_{\text{wat}}, \gamma_{\text{wat}})$ 与 $(\alpha_{\text{air}}, \beta_{\text{air}}, \gamma_{\text{air}})$, 与分界面法向量之间的夹角分别为 θ_{wat} 和 θ_{air} , f 为相机焦距, d 为折射平面到相机镜头中心的距离。由折射定律知

$$n_{\text{wat}} \sin \theta_{\text{wat}} = n_{\text{air}} \sin \theta_{\text{air}} \quad (1)$$

式中, n_{air} 与 n_{wat} 分别为光线在空气和水介质中的折射率, 水和空气分界面法向量的方向向量在相机坐标系下可表示为 $(0, 0, 1)^T$, 根据入射光线、折射光线及分界面法向量之间的关系有

$$\begin{bmatrix} \alpha_{\text{wat}} \\ \beta_{\text{wat}} \\ \gamma_{\text{wat}} \end{bmatrix} = k_1 \begin{bmatrix} \alpha_{\text{air}} \\ \beta_{\text{air}} \\ \gamma \end{bmatrix} + k_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (2)$$

式中, k_1 和 k_2 为

$$\begin{cases} k_1 = \frac{n_{\text{wat}}}{n_{\text{air}}} \\ k_2 = \sqrt{1 - \frac{n_{\text{wat}}^2}{n_{\text{air}}^2} \sin^2 \theta_{\text{air}}} - \frac{n_{\text{wat}}}{n_{\text{air}}} \cos \theta_{\text{air}} = \sqrt{1 - \frac{n_{\text{wat}}^2}{n_{\text{air}}^2} (1 - \gamma_{\text{air}}^2)} - \frac{n_{\text{wat}}}{n_{\text{air}}} \gamma_{\text{air}} \end{cases} \quad (3)$$

设 (x_u, y_u) 为相机坐标系下像平面上某像素点, 与该点对应的折射光线的方向向量为

$$\begin{bmatrix} \alpha_{\text{air}} \\ \beta_{\text{air}} \\ \gamma_{\text{air}} \end{bmatrix} = \frac{1}{\sqrt{x_u^2 + y_u^2 + f^2}} \begin{bmatrix} x_u \\ y_u \\ f \end{bmatrix} \quad (4)$$

设 $r = \sqrt{x_u^2 + y_u^2 + f^2}$, 表示折射光线和分界面的交点与像平面像点之间的直线距离, 结合式(2)和(4)可得入射光线方向向量和像点坐标 (x_u, y_u) 的关系为

$$\begin{bmatrix} \alpha_{\text{wat}} \\ \beta_{\text{wat}} \\ \gamma_{\text{wat}} \end{bmatrix} = \frac{n_{\text{wat}}}{n_{\text{air}}} \begin{bmatrix} \frac{x_u}{r} \\ \frac{y_u}{r} \\ -\sqrt{\frac{n_{\text{wat}}^2}{n_{\text{air}}^2} + \left(\frac{f}{r}\right)^2} - 1 \end{bmatrix} = \frac{n_{\text{wat}}}{n_{\text{air}}} \begin{bmatrix} \frac{1}{r} & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & -\sqrt{\frac{n_{\text{wat}}^2}{n_{\text{air}}^2} + \left(\frac{f}{r}\right)^2} - 1 \end{bmatrix} \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} \quad (5)$$

设在相机坐标系下像点对应的空间目标点的坐标为 $(x_{\text{cam}}, y_{\text{cam}}, z_{\text{cam}})$, 其对应光线在传播过程中与水 and 空气分界面的交点坐标为 $(x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}})$, 则有

$$k \begin{bmatrix} \alpha_{\text{wat}} \\ \beta_{\text{wat}} \\ \gamma_{\text{wat}} \end{bmatrix} = \begin{bmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ z_{\text{cam}} \end{bmatrix} - \begin{bmatrix} x_{\text{ref}} \\ y_{\text{ref}} \\ z_{\text{ref}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -dx_u/f \\ 0 & 1 & 0 & -dy_u/f \\ 0 & 0 & 1 & d \end{bmatrix} \begin{bmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ z_{\text{cam}} \\ 1 \end{bmatrix} \quad (6)$$

结合式(5)与(6)可求得像点与空间目标点之间的关系为

$$k \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix} = \frac{n_{\text{air}}}{n_{\text{wat}}} \begin{bmatrix} \frac{1}{r} & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & -\sqrt{\frac{n_{\text{wat}}^2}{n_{\text{air}}^2} + \left(\frac{f}{r}\right)^2} - 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & -dx_u/f \\ 0 & 1 & 0 & -dy_u/f \\ 0 & 0 & 1 & d \end{bmatrix} \begin{bmatrix} x_{\text{cam}} \\ y_{\text{cam}} \\ z_{\text{cam}} \\ 1 \end{bmatrix} \quad (7)$$

设相机的外参矩阵为 P , 内参矩阵为 B , 有

$$\begin{cases} P = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \\ B = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \end{cases} \quad (8)$$

令 $n_0 = n_{\text{air}}/n_{\text{wat}}$, 结合式(7)和(8)可得水下相机成像模型

$$k \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & u_0 \\ 0 & 1/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} n_0 r & 0 & 0 & \frac{-n_0 dx_a r}{f} \\ 0 & n_0 r & 0 & \frac{-n_0 dy_a r}{f} \\ 0 & 0 & \frac{-n_0}{\sqrt{n_0^2 + \left(\frac{f}{r}\right)^2 - 1}} & \frac{n_0 d}{\sqrt{n_0^2 + \left(\frac{f}{r}\right)^2 - 1}} \\ 0 & 0 & \sqrt{n_0^2 + \left(\frac{f}{r}\right)^2 - 1} & \sqrt{n_0^2 + \left(\frac{f}{r}\right)^2 - 1} \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\text{world}} \\ y_{\text{world}} \\ z_{\text{world}} \\ 1 \end{bmatrix} \quad (9)$$

1.2 水下双目立体视觉模型

基于双目立体视觉的水下 RGB-D 相机的核心硬件由左、右相机组成(图 2)。不失一般性,设左、右相机的焦距为 f , 两者的基线距离为 B , 两者的光心连线平行于玻璃封装窗, 同时也平行于设定的坐标系 x 轴。光心与封装窗口的距离为 d , 空气和水的折射率分别为 n_0, n_1 。设水中目标点在左相机坐标系下的坐标为 $P(x_{\text{cam}}, y_{\text{cam}}, z_{\text{cam}})$ 。它在左图上的图像坐标为 $p(u_{\text{lef}}, v_{\text{lef}})$, 主点坐标为 $(u_{0\text{lef}}, v_{0\text{lef}})$ 。在右图上的图像坐标为 $p(u_{\text{rig}}, v_{\text{rig}})$, 主点坐标为 $(u_{0\text{rig}}, v_{0\text{rig}})$ 。像元尺寸为 $d_X \times d_Y$ 。基于上述水下相机模型以及标定参数中的两个相机坐标系之间的关系, 可将左、右水下图像进行立体校正, 使其符合行对齐原则^[12]。水下双目立体视觉模型如图 3 所示。

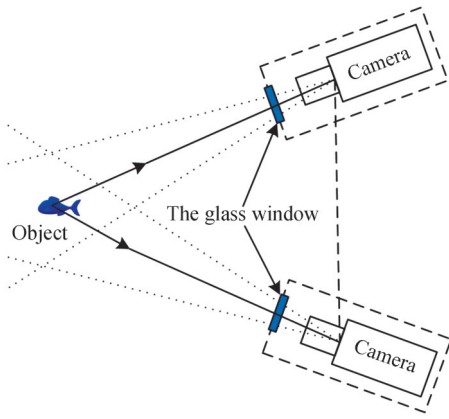


图 2 系统示意图
Fig.2 System Schematic

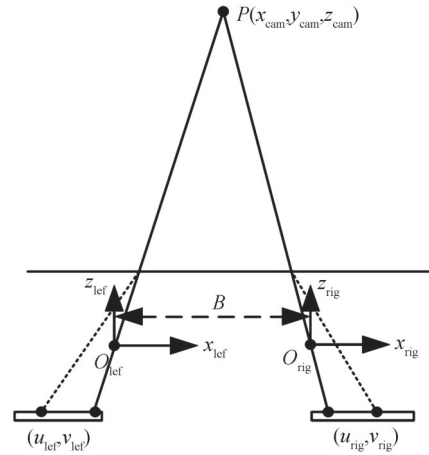


图 3 水下双目视觉模型
Fig.3 Underwater binocular vision model

此时, 相机坐标系下的 P 点三维坐标为

$$\begin{cases} x_{\text{cam}} = \frac{B \cdot (u_{\text{lef}} - u_{0\text{lef}})}{u_{\text{lef}} - u_{0\text{lef}} - u_{\text{rig}} + u_{0\text{rig}}} \\ y_{\text{cam}} = \frac{B \cdot (v_{\text{lef}} - v_{0\text{lef}}) \cdot d_Y}{(u_{\text{lef}} - u_{0\text{lef}} - u_{\text{rig}} + u_{0\text{rig}}) \cdot d_X} \\ z_{\text{cam}} = \frac{B \cdot f}{(u_{\text{lef}} - u_{0\text{lef}} - u_{\text{rig}} + u_{0\text{rig}}) \cdot d_X} \end{cases} \quad (10)$$

式中, $(u_{\text{lef}} - u_{0\text{lef}} - u_{\text{rig}} + u_{0\text{rig}}) \cdot d_X$ 为像素点的视差值, 获取该值需进行对应点匹配。

1.3 对应点匹配策略

基于像素的立体匹配过程即搜寻左、右图像上的所有对应点, 根据对应点之间的位置偏差得到视差值。简言之, 针对左图像中的每个像素点, 沿着它在右图像中的水平极线方向寻找与它互相匹配的像素点。但在实际进行像素点匹配时会存在几个问题: 首先, 要保证两个相机完全共面且参数一致是非常困难的, 且计算过程中也会产生误差累积, 右图的匹配点未必恰好在极线上(但肯定在极线附近); 其次, 单个像素点进行比较鲁棒性很差, 很容易受到光照变化和视角不同的影响, 所以搜索范围需要适当放宽。

因此,实际的详细像素匹配过程为:在高精度的标定数据基础上,获取校正后的标准图像对,此时对应点应在一条水平线附近。在左图像中建立一个以待匹配点 $a(x,y)$ 为中心位置的 $m \times n$ 窗口(图中的红色虚线框),在视差 d 范围($d \in [d_{\min}, d_{\max}]$,且为整数,大小由实际场景决定)内以1个像素的步长水平移动这个模板窗口穿过右侧目标图像,在每个位置都进行相似度检测。当该相似度达到极值时,该位置可被认为是匹配点,此时的 d 值即为该点的视差值。匹配过程示意图如图4所示。

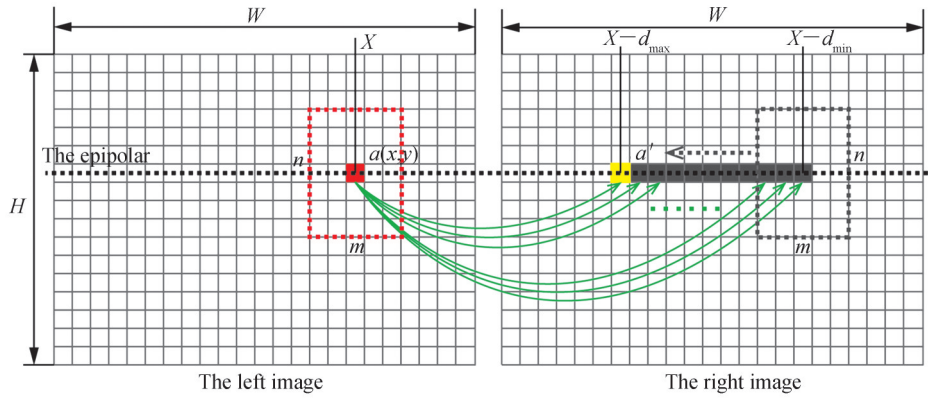


图4 基于像素的匹配过程示意图
Fig.4 Pixel-based matching process diagram

2 基于颜色校正和暗原色先验的水下图像增强算法

水分子、水下微粒和杂质对光线的散射作用使得水下图像存在模糊和对比度下降等问题,同时水介质对光线的吸收作用使得水下图像的颜色失真。由于水下特殊的成像环境,采集到的图像清晰度下降、颜色失真,影响到立体匹配过程中对应点的匹配精度及RGB图像颜色的真实性。因此对于水下RGB-D相机而言,水下图像的增强处理尤为重要。本文根据水下图像成像原理,结合Retinex理论提出一种基于水下图像衰减校正系数的颜色校正算法用于对水下图像进行颜色校正,再基于暗原色先验理论改进了水下均匀光成分估计方法对水下图像进行清晰度增强处理。并选取几种常用图像增强算法作为对比,利用信息熵、Sift特征点与Canny特征点数量作为评价指标,以衡量不同算法的效果。

2.1 基于Retinex的水下图像颜色校正

Retinex理论解释了传统颜色视觉理论解释不了的颜色的恒常性,且物体在不同光线的照射下其颜色保持恒定不变^[13-14]。然而,光线在水下进行传播时受到水介质吸收散射的影响会被衰减,衰减程度也随着波长发生变化,其中红色光相比蓝绿色光衰减的更为严重。为此,本文结合Retinex理论提出了一种基于水下图像衰减校正系数的颜色校正算法用于对水下图像进行颜色校正。根据水下图像RGB三通道均值求解衰减校正系数,利用Retinex算法结合衰减校正系数分别对RGB三个颜色通道图像的反射分量进行求解,合成三通道反射分量输出颜色校正后的水下图像。算法流程如图5所示。

JAFFE J S在MCGLAMERY B L建立的光学模型的基础上进行了拓展,将水下图像成像时相机平面接受的光线表示为前向散射光、后向散射光以及目标反射衰减光的相加的形式^[15-16]。根据水下图像成像模型,相机接收到的总光强 $E_{LF}^c(u)$ 可简化为

$$E_{LF}^c(u) = E^c(u) \cdot \exp[-a(c) \cdot d(u)] + A \cdot \{1 - \exp[-a(c) \cdot d(u)]\} \quad (11)$$

式中, u 表示图像中的像素点; $E^c(u)$ 表示未经衰减的水下目标的反射光; c 表示图像的RGB颜色通道; $d(u)$ 为相机到水下目标的距离; $a(c)$ 为光线水下衰减系数(由于不同波长光线在水下的衰减程度不同,衰减系数 $a(c)$ 的值会随着颜色通道的改变而变动)。

由于光线在水中的衰减程度因波长而异,这使得水下成像图像中三通道分量的强度大小存在很大差异。本文提取图像R、G、B三通道值,并计算均值,即

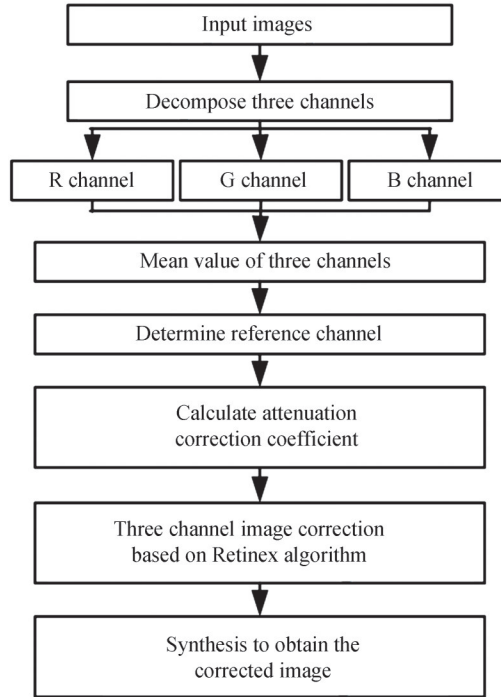


图5 颜色校正算法流程

Fig.5 Flow chart of color correction algorithm

$$\begin{cases} \bar{R}_{\text{Ret}} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I_R(i, j) \\ \bar{G}_{\text{Ret}} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I_G(i, j) \\ \bar{B}_{\text{Ret}} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I_B(i, j) \end{cases} \quad (12)$$

式中, M, N 代表图像大小, I_R, I_G, I_B 分别表示 RGB 三通道图像。根据 RGB 像素均值确定大小居中的颜色通道作为参考通道, 设置参考通道的衰减校正系数值为 1, 另外两个颜色通道的衰减校正系数为其对应通道像素均值与参考通道像素均值的比值。

根据 Retinex 理论, 图像 $LF(i, j)$ 可分解成反射图像 $R_{LF}(i, j)$ 和入射图像 $L_{LF}(i, j)$

$$LF(i, j) = R_{LF}(i, j) \cdot L_{LF}(i, j) \quad (13)$$

Retinex 理论认为入射光线是影响图像整体亮度的主要因素, 而图像的本质特征由物体的反射性质决定。它便是根据原始图像 $LF(i, j)$ 估算出入射光线的照射分量 $L_{LF}(i, j)$, 然后去除入射光的影响, 从而恢复水下目标物体原有的颜色信息。通过将图像转换到对数域, 将其入射分量与反射分量进行分离, 利用原始图像减去入射分量从而得到图像的反射分量, 原理公式为

$$\log R_{LF}(i, j) = \log LF(i, j) - \log L_{LF}(i, j) \quad (14)$$

为了获取图像的入射分量信息, 采用高斯滤波器对原始图像低通滤波来实现对图像照射分量的估计

$$D_{LF}(i, j) = LF(i, j) \otimes F(i, j) \quad (15)$$

式中, $F(i, j)$ 代表高斯函数, \otimes 表示卷积运算。

在对数域用原始图像减去高斯滤波后的低频分量即入射分量的估计值便可以得到图像的高频分量, 考虑到水介质对光线的衰减特性, 此处引入衰减校正系数, 对所减去的低频分量进行衰减校正, 最后对获得的高频分量进行反对数变换得到图像的反射分量为

$$R_{LF}(i, j) = \exp(\log LF(i, j) - \delta \cdot \log D_{LF}(i, j)) = \exp(\log LF(i, j) - \delta \cdot \log(LF(i, j) \otimes F(i, j))) \quad (16)$$

式中, δ 表示当前通道图像的衰减校正系数。最后利用自适应直方图均衡化算法对反射分量图像进行增强

处理,进一步改善图像的细节信息和局部对比度。将获取的水下图像三通道分解后,结合求得的三通道图像的衰减校正系数,利用Retinex算法分别求出RGB三个颜色通道图像的反射分量,并将求得的三通道反射分量进行合成获取颜色校正后的水下图像。

2.2 基于暗原色先验的水下图像清晰度增强

根据Jaffe-McGlamery水下图像退化理论,将式(11)简化后可知,水下图像的成像是线性过程,可描述为真实图像信息与水体透射率的乘积,再加上水下背景散射光分量。具体为

$$I_{LF}(u) = J_{LF}(u)t_{LF}(u) + A(1 - t_{LF}(u)) \quad (17)$$

式中, u 为图像中的像素点; $I_{LF}(u)$ 是待增强图像; $J_{LF}(u)$ 是增强处理后的清晰图像; A 是水下均匀光成分; $t_{LF}(u)$ 是水体透射率。

水下图像的增强处理过程其实就是对图像 $J_{LF}(u)$ 的求解过程,可分为四步。

Step1:获取水下暗原色图像。由暗原色先验理论可知,任意彩色图像中总会存在暗淡的局部图像,且RGB三通道中必存在一个低灰度值的颜色通道。水下图像的暗原色可定义为

$$J_{LF}^{dark}(u) = \min_{v \in \Omega(u)} \left(\min_{c \in \{R, G, B\}} J_{LF}^c(v) \right) \quad (18)$$

式中, $J_{LF}^{dark}(u)$ 表示图像在像素点 u 处的暗通道值(值趋近于0); $\Omega(u)$ 为滤波窗口(以像素 u 为中心的局部区域); J_{LF}^c 表示输入图像的 C 颜色通道。

Step2:估算水下均匀光成分。TAN R T和FATTAL R认为,一幅图像中最亮的点可很好的表示环境的背景光强,其核心思想是选取原始图像中光照强度最强点的像素值作为估算的环境均匀光成分^[17-18]。然而,由于水环境的复杂性和水体的散射吸收特性,使得该方法无法准确的估算出水下均匀光成分。为了获得准确性和鲁棒性较好的水下均匀光成分估计值,本文提出利用衰减校正系数和暗原色先验相结合的方法对均匀光成分进行估算。根据上文中求得的水下图像衰减校正系数,利用其比例关系对输入的彩色图像进行灰度转换。对转换得到的灰度图像中的每个像素点进行局部区域的方差计算,生成灰度方差图像。在输入图像中选择在对应暗原色图中亮度大小前0.1%的像素点作为预计水下均匀光估计点。从估计点中选取在相应的灰度方差图像中方差值最低的像素点,将其灰度值作为水下均匀光成分估算值。

Step3:计算和优化水下透射率。获得水下均匀光成分估算值后,对式(16)归一化处理,水下散射成像模型变换为

$$\frac{I_{LF}(u)}{A} = \frac{J_{LF}(u)}{A} t_{LF}(u) + 1 - t_{LF}(u) \quad (19)$$

根据暗原色先验假设,水下透射率在任意局部区域为一常数值,其预估值定义为 $\tilde{t}_{LF}(u)$ 。对水下成像模型进行最小值滤波可得

$$\min_{v \in \Omega(u)} \left(\min_{c \in \{R, G, B\}} \left(\frac{I_{LF}^c(v)}{A} \right) \right) = \tilde{t}_{LF}(u) \min_{v \in \Omega(u)} \left(\min_{c \in \{R, G, B\}} \left(\frac{J_{LF}^c(v)}{A} \right) \right) + 1 - \tilde{t}_{LF}(u) \quad (20)$$

由前述暗原色先验理论知,水下清晰图像的暗原色值 $J_{LF}^{dark}(u)$ 趋近于0。即

$$\min_{v \in \Omega(u)} \left(\min_{c \in \{R, G, B\}} \left(\frac{J_{LF}^c(v)}{A} \right) \right) \rightarrow 0 \quad (21)$$

结合式(20)可得水体透射率的预估值

$$\tilde{t}_{LF}(u) = 1 - \min_{v \in \Omega(u)} \left(\min_{c \in \{R, G, B\}} \left(\frac{I_{LF}^c(v)}{A} \right) \right) \quad (22)$$

Step4:增强水下图像。在求得水下均匀光成分和水体透射率后,可根据水下散射成像模型对水下图像进行增强

$$J_{LF}(u) = \frac{I_{LF}(u) - A}{t_{LF}(u)} + A \quad (23)$$

2.3 通用数据集对比实验

为了验证以上算法的效果,本文采用通用水下图像增强数据集,选取直方图均衡化算法(HE)^[19]、暗原

色先验算法(Dark Channel Prior, DCP)^[20]以及基于图像融合(Image Fusion, IF)^[21]的增强算法作为对比算法,通过几幅不同水下图像的增强处理效果对算法性能进行评估。本文算法与几种算法的增强效果对比如图6所示。

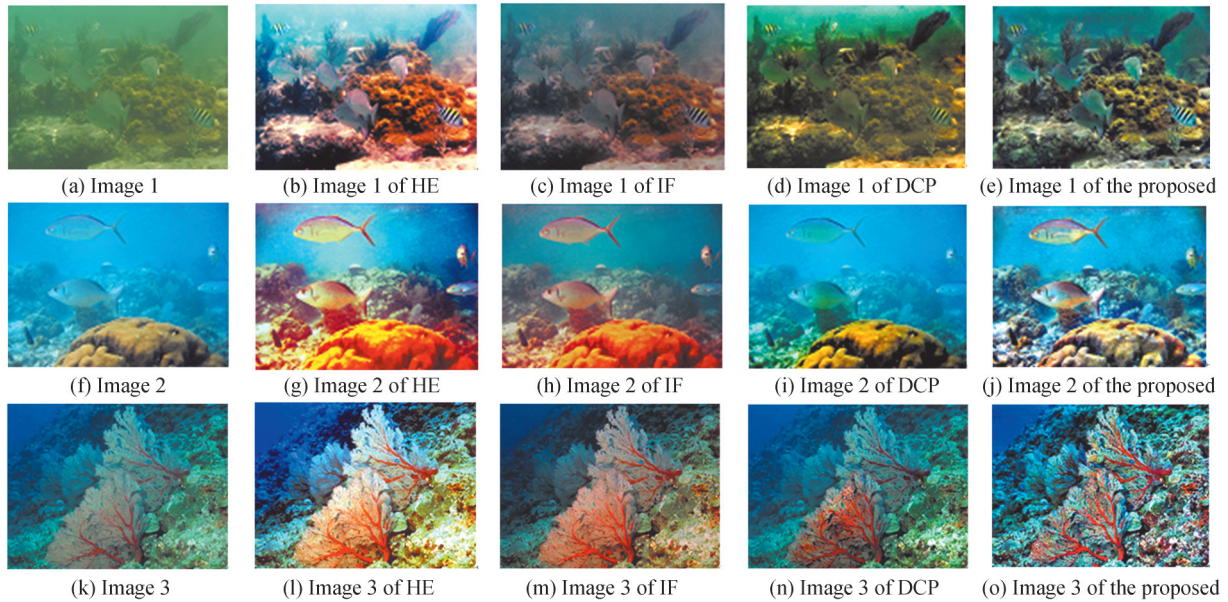


图6 图像增强效果

Fig.6 Enhanced effects of images

由以上实验结果可看出,由于水下光线的传播受到水介质的吸收散射影响,导致水下成像图像呈现蓝绿色调,偏色现象严重。采用HE算法处理后,图像背景干扰信息的对比度也得到增强,使得图像细节信息丢失,当背景光存在的较大突变时会造成显著的过增强问题。IF算法对图像亮度和纹理的恢复具有一定效果,但处理后的图像依然存在颜色失真的问题。DCP算法处理后的水下图像依旧存在偏色现象,且图像自身的模糊问题也未能被有效的去除。相比而言,本文提出的算法可有效校正水下图像的偏色现象,提高图像的对比度,且纹理细节信息也更加突出。

为进一步说明实验效果,选取信息熵、Sift特征点与Canny特征点的数量作为不同算法的评价标准。信息熵表示图像灰度分布的聚集特性,反映了图像平均信息量的多少,信息熵越大则图像清晰度越高。特征点数量是一种能够间接的反映图像质量的评价指标,高质量清晰图像比低质量模糊拥有更多的有效特征点,所以特征点数量越多图像质量越高。对图6中的实验结果进行分析,不同算法性能的评价结果如表1所示。

表1 不同算法的对比结果

Table 1 the comparison result of different algorithms

The image	The evaluation criterion	The original image	HE	IF	DCP	The proposed
1	Information entropy	6.83	7.28	7.24	7.14	7.47
	Sift	154	1 848	985	1 361	2 079
	Canny	18 666	67 714	53 259	60 471	67 963
2	Information entropy	7.70	7.39	7.45	7.27	7.88
	Sift	582	1 547	961	916	2 004
	Canny	27 392	67 934	56 943	32 463	68 557
3	Information entropy	7.15	7.23	7.32	7.19	7.70
	Sift	5 954	13 488	8 625	10 072	17 203
	Canny	130 335	211 820	167 198	150 484	211 804

根据表1可知,采用IF算法和DCP算法处理后的图像的Sift特征点和Canny特征点数量有所增加,但相比于HE算法和本文算法特征点数量较少,说明其对于水下图像具有一定的增强作用但处理效果不够理想。HE算法与本文算法相比,虽然增强处理后的图像的特征点数量接近但是信息熵值较低,这也验证了HE算法对于处理背景光存在较大突变的水下图像的过增强现象。综上,本文算法对于水下图像的颜色失真和细节纹理模糊问题都有明显的改善。

3 3D点云与2D RGB图像的对准叠加

将以上增强处理后的左、右水下图像转换为空气中的图像,进而计算目标三维点云^[11]。再提取匹配点的RGB三个通道的值,利用左、右匹配点的RGB信息计算获得标准RGB值(即空间目标点对应的彩色信息值)。根据获取的RGB彩色信息对生成的三维点云作色彩对准叠加处理得到RGB-D数据。具体处理过程如图7所示。

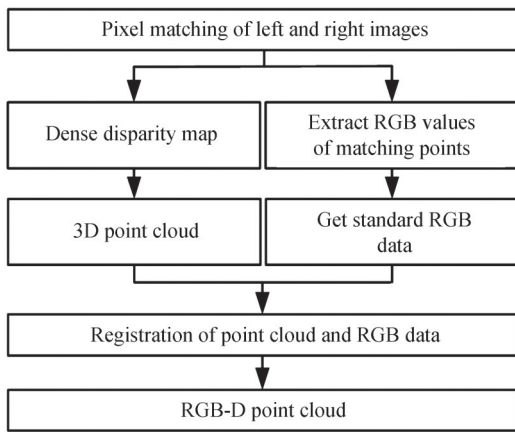


图7 3D点云与RGB色彩对准叠加流程图

Fig.7 Flow chart of alignment and superposition of 3D point cloud and RGB color

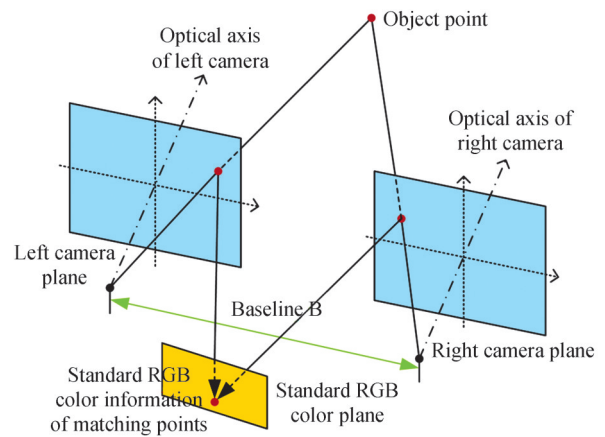


图8 3D点云与2D RGB数据对准叠加模型

Fig.8 Alignment superposition model of 3D point cloud and 2D RGB data

基于双目立体视觉的水下RGB-D相机利用左、右彩色相机获取RGB三通道信息从而形成彩色图像。对于空间中任意点,利用立体匹配算法获取其在左、右相机成像平面上对应的像素点,接着对左右成像图像上的匹配像素点的RGB三通道进行彩色信息提取,通过计算获得与目标点对应的标准RGB彩色信息。如图8所示为三维点云的RGB色彩对准叠加模型,可看出根据匹配点的视差信息,结合三角测量原理可以计算获得空间目标点的三维坐标信息 (X, Y, Z) ,再利用左右匹配像素点的彩色信息可计算得到与空间目标点对应的RGB彩色信息。最终融合三维坐标信息与彩色信息生成的RGB-D点云数据 (X, Y, Z, R, G, B) 。再对左、右成像平面上的匹配像素点进行RGB三通道彩色分离提取后,利用式(24)计算对应空间目标点的标准RGB信息。

$$\begin{cases} R = \frac{R_{\text{lef}} + R_{\text{rig}}}{2} \\ G = \frac{G_{\text{lef}} + G_{\text{rig}}}{2} \\ B = \frac{B_{\text{lef}} + B_{\text{rig}}}{2} \end{cases} \quad (24)$$

式中, $R_{\text{lef}}, G_{\text{lef}}, B_{\text{lef}}$ 与 $R_{\text{rig}}, G_{\text{rig}}, B_{\text{rig}}$ 分别为空间目标三维点在左、右相机图像上的匹配像素点分离后的RGB三通道值, R, G, B 为对应空间目标点的标准RGB彩色信息值。根据点云RGB色彩对准叠加模型中三维点云和彩色图像像素点之间的映射关系,便可将三维点云数据和二维彩色图像信息进行叠加融合,从而生成RGB-D图像数据。

4 样机研制及实验分析

4.1 工程样机研制

研制的水下RGB-D相机内部电气结构示意图如图9所示。采用的相机为MER2-1220-32U3C,分辨率 $4\,024 \times 3\,036$,像元尺寸 $1.85\ \mu\text{m}$ 。镜头为LM5JC10M,焦距12 mm。双目立体视觉系统的基线距离为350 mm。系统整体由左、右两个相机控制单元与中央计算单元构成,主要包含TX2、相机和电源等。相机采用外部供电,通过TX2发送脉冲的方式,实现两相机同时连续拍摄,并利用内部USB 3.0传输线将拍摄的图片传输至中央计算单元的TX2完成点云计算,样机与上位机通过水密接插件通信。

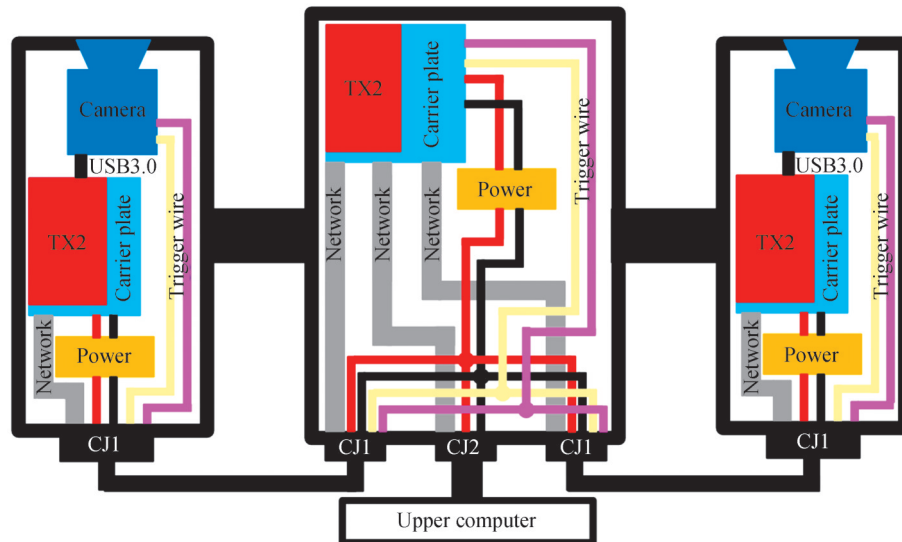


图9 系统电气结构示意图

Fig.9 System electrical structure diagram

图10为研制的工程样机的实物图,由中央计算单元与左、右相机控制单元3个密封主体构成,外部采用赛钢材质的夹具固定。密封壳体采用耐压、耐腐蚀的铝合金材料,光学窗口采用蓝宝石玻璃,最大设计工作水深为1 000 m。图11、图12分别为中央计算单元和相机控制单元的内部结构图。中央计算单元和相机控制单元的TX2、电源模块和网口分别固定在内部支架的两侧,其中中央计算单元拥有三个网口。相机控制单元拥有一个网口且与USB接口固定在支架同一侧,网口与TX2相连,相机控制单元的TX2通过USB与相机连接,用于对相机进行供电和数据传输。中央计算单元具备三个水密接插件接口,其中两个接口用于与左、右相机控制单元的连接通信,一个接口用于连接上位机进行数据传输。

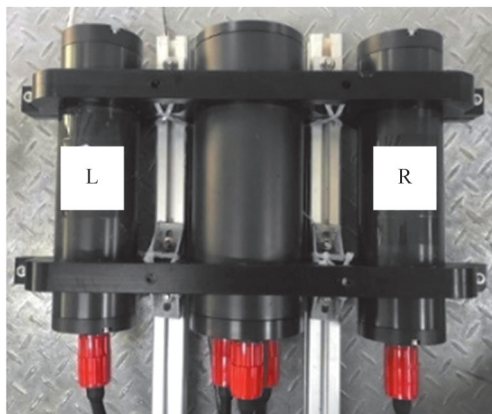


图10 工程样机实物图

Fig.10 Physical engineering prototype

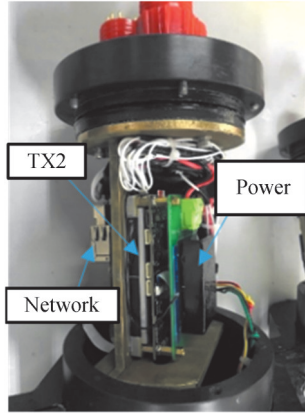


图 11 中央筒内部结构
Fig.11 Internal structure of the central cylinder

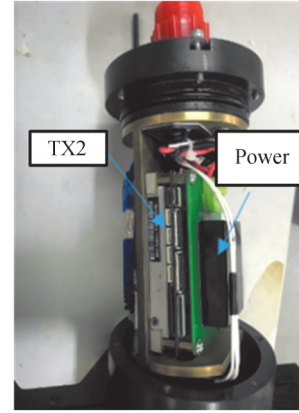


图 12 两侧筒内部结构
Fig.12 Internal structure of cylinder on both sides

4.2 系统标定

先利用迭代标定法获取空气中的相机标定参数,再根据水下标定算法^[22]对相机进行水下标定。水下标定结果见表2, f_x 和 f_y 为相机焦距参数; u_0 和 v_0 为相机主点坐标; k 和 P 为相机畸变系数; R 和 T 为左右相机的相对位姿转换矩阵; d 为相机光心到分界面的距离; n 为分界面法线; μ 为水的折射率。

表 2 左、右相机的标定参数
Table2 Left and right camera calibration parameters

Parameters	Left camera	Right camera
f_x, f_y	(6 801.773, 6 801.983)	(6 766.651, 6 767.138)
u_0, v_0	(2 032.804, 1 456.358)	(2 146.771, 1 495.447)
$(k_1, k_2, k_3, p_1, p_2)$	(-0.028, 0.132, 0.001, 0.001, -0.465)	(-0.027, 0.137, 0.001, -0.001, -0.485)
d	60.682 3	59.714 0
n	(0.002 1, -0.026 2, 0.999 3)	(0.007 4, -0.024 5, 0.999 3)
μ	1.333	
R	(-0.009 4, 0.001 1, 0.014 8)	
T	(-349.974 0, -4.178 4, -10.381 8)	

4.3 基于图像增强的水下三维重建实验

实验在实验室水池(3 m×3 m×3.5 m)中进行,如图 13所示。样机置于水池上方,并浸入水中,垂直向下采集位于水池底部的目标物。实验采用的水体为一般市用自来水,由于水介质的散射特性,水下图像不可避免存在色偏及图像质量降低等问题。图 14(a)、(c)分别为水下 RGB-D 相机采集到的左、右水下图像,

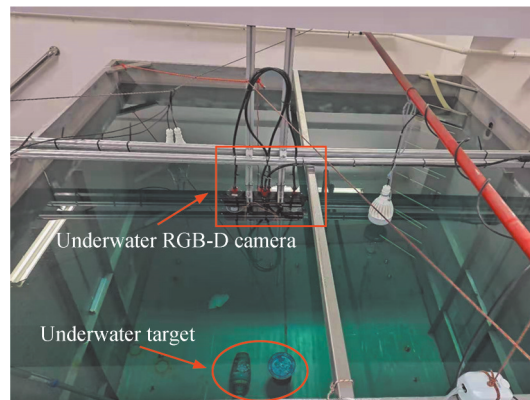


图 13 实验场景
Fig.13 Experimental scene

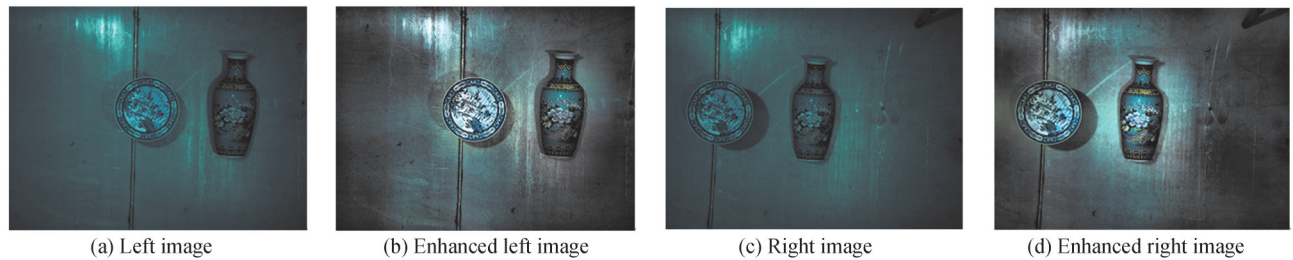


图14 原始水下图像与增强处理后的图像

Fig.14 Original underwater images and enhanced images

可看出原始图像整体偏暗对比度较低且存在偏色现象。图14(b)、(d)是经过图像增强处理后的效果图。从增强处理前后的图像对比中可看出,提出的水下图像增强算法不仅能校正水下图像的颜色失真,而且图像的清晰度也得到了提高。

图15分别为利用图像增强前后的左、右水下图像进行三维重建获得的三维点云图。从图15(a)可看出,未增强前的图像的三维重建效果不理想,表面存在较多缺失。与之相比,图15(b)中,原本存在缺失的区域可准确的重建出来。实验结果表明,本文提出的水下图像增强算法可提高水下三维重建的效果。

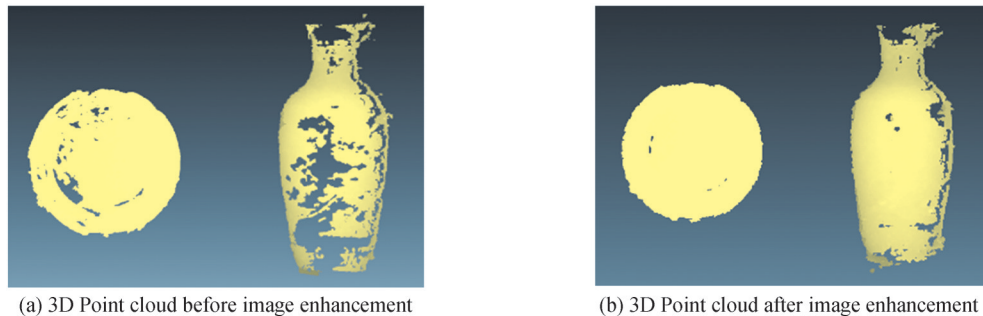


图15 增强前后的效果对比

Fig.15 Effect comparison of enhancement algorithms

4.4 三维点云RGB色彩对准叠加实验

在以上实验的基础上,将增强处理后得到的3D点云与二维RGB颜色数据进行对准叠加,获得RGB-D图像数据,图16(a)是计算得到的水下物体的初始3D三维点云,图16(b)是该三维点云与提取的标准RGB彩色数据叠加配准后的结果,图16(c)、图16(d)是局部放大图。

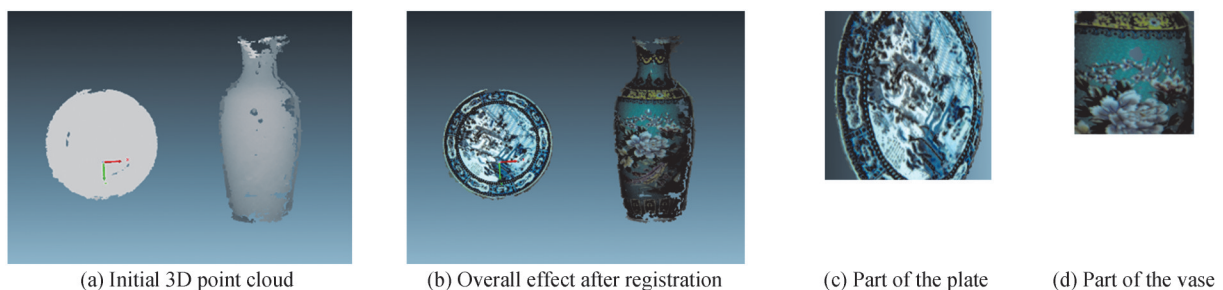


图16 目标物RGB色彩对准叠加效果

Fig.16 Registration effect of 3D point cloud and RGB color

从图16可看出3D点云图像与2D彩色RGB图像对准叠加获得RGB-D图像,既有二维信息,也有3D信息,这对于水下目标识别尤为重要。模型可有效的将采集到的二维彩色RGB图像与三维点云进行对准叠加,获得更具应用价值的RGB-D点云数据。

4.5 水下三维重建精度分析实验

在水池底部放置圆形陶瓷盘作为目标物(图17(a)),实验时,将目标陶瓷盘摆放在水下不同位置(图17

(b)),与样机距离大约3 m。采集多组含陶瓷盘的水下图像,利用提出的水下图像增强算法对图像进行增强处理,然后计算获取陶瓷盘的三维点云。陶瓷盘直径的名义值为259.3 mm,名义值与实验测量值之间的误差和标准差如表3所示。由表3可看出,陶瓷盘直径尺寸测量的标准差为2.6 mm,具有较好的测量精度。

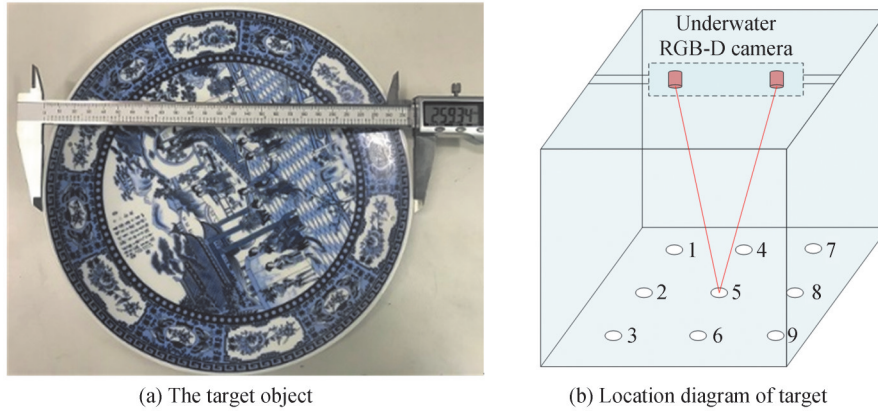


图17 精度测量实验

Fig.17 Precision measurement experiment

表3 测量数据

Table 3 The measurement data

Number	Experimental value/mm	Error/mm	RMS/mm
1	262.8	3.5	
2	261.1	1.8	
3	262.7	3.4	
4	261.9	2.6	
5	260.5	1.2	
6	261.1	1.8	2.6
7	261.4	2.1	
8	262.2	2.9	
9	262.4	3.1	
10	261.6	2.3	

5 结论

本文研制了一款融合3D点云和二维RGB图像为一体的RGB-D水下相机工程样机。针对水下图像对比度低、颜色失真等问题,提出了一种基于颜色校正和暗原色先验的水下图像增强算法,基于双目立体视觉原理,通过左右相机对应点匹配计算,获得了水下目标的三维点云,建立了3D点云与2D彩色RGB图像的对准叠加模型,实现了对目标三维深度信息和彩色RGB信息的叠加融合。进行了一系列实验及理论验证,包括:基于图像增强的水下三维重建实验表明,提出的水下图像增强算法不仅能校正水下图像的颜色失真,而且可提高水下三维重建的效果,为之后获取更高质量的RGB-D数据奠定基础;三维点云RGB色彩对准叠加实验表明,构建的点云RGB色彩对准叠加模型可将颜色校正后的颜色数据与3D点云数据配准生成高质量的水下RGB-D数据;最终的水下三维重建精度分析实验表明,基于以上理论研制的水下RGB-D样机具有较好的实际水下三维重建效果。

参考文献

- [1] DING Ying, FAN Jingtao, SONG Tianyu. Optimal forward design method for the binocular stereo vision inspection system [J]. Chinese Journal of Scientific Instrument, 2016, 37(3): 650-657.
丁莹, 范静涛, 宋天喻. 双目立体视觉检测系统正向最优化设计方法研究[J]. 仪器仪表学报, 2016, 37(3): 650-657.
- [2] WU Qingyang, HUANG Haotao, CHEN Shunzhi, et al. Research on 3D imaging technology of light field based on structural light marker[J]. Infrared and Laser Engineering, 2020, 49(3): 255-260.

- 吴庆阳, 黄浩涛, 陈顺治, 等. 基于结构光标记的光场三维成像技术研究[J]. 红外与激光工程, 2020, 49(3): 255-260.
- [3] CHIANG J Y, CHEN Y C. Underwater image enhancement by wavelength compensation and dehazing [J]. IEEE Transactions on Image Processing, 2012, 21(4): 1756-1769.
- [4] GHANI A. Underwater image quality enhancement through integrated color model with Rayleigh distribution [J]. Applied Soft Computing, 2014, 27: 219-230.
- [5] CARLEVARIS B N, MOHAN A, EUSTICE R M. Initial results in underwater single image dehazing [C]. Oceans, IEEE, 2010.
- [6] HITAM M S, YUSSOF W, AWALLUDIN E A, et al. Mixture contrast limited adaptive histogram equalization for underwater image enhancement [C]. International Conference on Computer Applications Technology, IEEE, 2013.
- [7] WANG Yongxin, DIAO Ming, HAN Chuang. Underwater image enhancement algorithm based on iterative histogram equalization with conventional light source [J]. Acta Photonica Sinica, 2018, 47(11): 1101002.
王永鑫, 刁鸣, 韩闯. 基于迭代直方图均衡化的常规光源水下成像增强算法 [J]. 光子学报, 2018, 47(11): 1101002.
- [8] KWON Y H, CASEBOLT J B. Effects of light refraction on the accuracy of camera calibration and reconstruction in underwater motion analysis [J]. Sports Biomechanics, 2006, 5(2): 315-340.
- [9] AGRAWAL A, RAMALINGAM S, TAGUCHI Y, et al. A theory of multi-layer flat refractive geometry [C]. IEEE Conference on Computer Vision & Pattern Recognition, 2012.
- [10] ZHANG Wenming, DENG Qianxue, ZHANG Qiang, et al. Non-Parallel system underwater image transformation model [J]. Acta Photonica Sinica, 2015, 44(2): 158-162.
张文明, 邓茜雪, 张强, 等. 基于非平行系统的水下图像转化模型 [J]. 光子学报, 2015, 44(2): 158-162.
- [11] WU X, TANG X C. Accurate binocular stereo underwater measurement method [J]. International Journal of Advanced Robotic Systems, 2019, 16(5): 1-13.
- [12] WEI Jingyang. High precision underwater 3D reconstruction method based on binocular stereo vision [D]. Harbin: Harbin Institute of Technology, 2017.
魏景阳. 基于双目立体视觉的水下高精度三维重建方法 [D]. 哈尔滨: 哈尔滨工业大学, 2017.
- [13] LANG E H, MCCANN J J. Lightness and Retinex theory [J]. JOSA, 1971, 61(1): 1-11.
- [14] LANG E H. The Retinex theory of color vision [J]. Scientific American, 1977, 237(6): 108-128.
- [15] JAFFE J S. Computer modeling and the design of optimal underwater imaging systems [J]. IEEE Journal of Oceanic Engineering, 1990, 15(2): 101-111.
- [16] MCGLAMERY B L. A computer model for underwater camera systems [C]. SPIE, 1980.
- [17] TAN R T. Visibility in bad weather from a single image [C]. 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2008: 1-8.
- [18] FATTAL R. Single image dehazing [J]. ACM Transactions On Graphics (TOG), 2008, 27(3): 1-9.
- [19] HITAM M S, AWALLUDIN E A, YUSSOF W N J H W, et al. Mixture contrast limited adaptive histogram equalization for underwater image enhancement [C]. 2013 International Conference on Computer Applications Technology (ICCAT), IEEE, 2013: 1-5.
- [20] XIE Kun. Underwater image enhancement algorithm based on dark channel prior and water turbidity recognition [D]. Xiamen: Xiamen University, 2018.
谢坤. 基于暗通道先验和水体浑浊度识别的水下图像增强算法 [D]. 厦门: 厦门大学, 2018.
- [21] ZHANG C, ZHANG X, TU D. Underwater image enhancement by fusion [C]. International Workshop of Advanced Manufacturing and Automation. Springer, Singapore, 2017: 81-92.
- [22] ZHANG C, ZHANG X, TU D, et al. On-site calibration of underwater stereo vision based on light field [J]. Optics and Lasers in Engineering, 2019, 121: 252-260.

Underwater RGB-D Camera Based on Binocular Stereo Vision

ZHUANG Sufeng, JI Yong, TU Dawei, ZHANG Xu

(School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China)

Abstract: Just as humans use two eyes to obtain 2D and 3D visual information of the scene to move and work freely in the objective world, robots also need RGB 2D color images and 3D point cloud information of the scene and target when carrying out target recognition and fine work. At present, Microsoft Kinect, Intel R200, R461, and other RGB-D cameras on the market can meet the needs of robot applications in

the air. Underwater robots also need RGB-D vision sensors to complete underwater target recognition and fine operation tasks, but so far there are no reports of commercial products of underwater RGB-D cameras. At present, the 3D measurement technology based on binocular stereo vision technology in the air has been relatively mature. However, the particularity of the underwater environment brings difficulties to the development of underwater RGB-D cameras. Firstly, due to the absorption and scattering of water medium and suspended particles in water, the captured underwater image has more serious problems than the image in the air, such as color distortion, low contrast, fuzzy details, which affects the matching accuracy of corresponding points in the stereo matching process and the authenticity of RGB image color. Therefore, for underwater RGB-D camera, underwater image enhancement processing is particularly important. The image enhancement technology studied by many scholars provides high-quality image guarantee for underwater stereo vision. In addition, when the industrial camera is applied to the underwater scene, the camera must be encapsulated in a waterproof shell. The propagation of light needs to pass through three media of "water-glass-air" in turn, and finally image in the camera image plane. However, due to their different refractive indices, light is refracted at the interface of different media, so that the object-image relationship of underwater camera imaging does not meet the imaging principle in air. Therefore, the traditional camera model is invalid. Researchers around the world have proposed that building an underwater refraction model is the fundamental solution to the problem of underwater 3D reconstruction. Although the research on underwater RGB-D technology is rare, the above research has paved the way for underwater RGB-D camera. In order to meet the needs of underwater target recognition and fine operation of an underwater robot, an RGB-D underwater camera engineering prototype based on binocular stereo vision integrating 3D point cloud and 2D RGB image are developed, and the related technologies are studied. It mainly includes: Firstly, considering the underwater multi-layer refraction effect, the underwater camera imaging and binocular stereo imaging models are established to convert the underwater image into the image in the air to eliminate the influence of refraction on the reconstruction accuracy. Then the corresponding points are matched to obtain the 3D point cloud; Secondly, according to the principle of underwater image imaging and Retinex theory, a color correction algorithm based on the attenuation correction coefficient of the underwater image is proposed to correct the color of the underwater image, and then the underwater uniform light component estimation method is improved based on the dark primary color prior theory to enhance the definition of the underwater image. Several common image enhancement algorithms are selected as a comparison, and the information entropy, the number of SIFT feature points, and canny feature points are used as evaluation indexes to evaluate the effect of different algorithms; Finally, the RGB values of the three channels of the matching points are extracted, and the standard RGB values are calculated by using the RGB information of the left and right matching points. The alignment superposition model is established, and the 3D point cloud and 2D color data are superimposed and fused to obtain underwater RGB-D image data. The experimental results show that the proposed underwater image enhancement algorithm can not only correct the color distortion of underwater image, but also improve the effect of underwater 3D reconstruction, which lays a foundation for obtaining higher quality RGB-D data; the established point cloud and RGB color registration model can register the color corrected color data with 3D point cloud data to generate high-quality underwater RGB-D data; the underwater RGB-D engineering prototype has good 3D measurement accuracy, and the system reconstruction error at 3M is 2.6 mm, which has a good actual underwater 3D reconstruction effect.

Key words: Binocular stereo vision; Imaging model of the underwater camera; Underwater image enhancement; Underwater 3D reconstruction; Alignment overlay model of point clouds and colors; Underwater RGB-D camera

OCIS Codes: 330.1400; 150.0155; 100.2980; 100.3020; 110.3010