Mechatronics 41 (2017) 58–66

Contents lists available at ScienceDirect

Mechatronics

journal homepage: www.elsevier.com/locate/mechatronics

# Design and implementation of an omnidirectional vision system for robot perception

Qing Shi[a,c,*], Chang Li[a], Chunbao Wang[b], Haibo Luo[c], Qiang Huang[a], Toshio Fukuda[a]

[a] *Intelligent Robotics Institute, School of Mechatronical Engineering, Beijing Institute of Technology, Beijing 100081, China*
[b] *Shenzhen Institute of Geriatrics, Shenzhen Second People's Hospital, Shenzhen 518035, China*
[c] *Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University), Fuzhou 350121, China*

## ARTICLE INFO

*Article history:*
Received 30 December 2015
Revised 21 October 2016
Accepted 20 November 2016

*Keywords:*
Robotics
Mechatronic system
Omnidirectional vision
Stereo vision

## ABSTRACT

To meet the demand of surrounding detection of a humanoid robot, we developed an omnidirectional vision system for robot perception (OVROP) with 5 Degrees of Freedom (DOFs). OVROP has a modular design and mainly consists of three parts: hardware, control architecture and visual processing part (omnidirectional vision and stereovision). As OVROP is equipped with universal hardware and software interfaces it can be applied to various types of robots. Our performance evaluation proves that OVROP can accurately detect and track an object with 360° field of view (FOV). Besides, undistorted omnidirectional perception of surroundings can be achieved through calibrations of both monocular and stereo cameras. Furthermore, our preliminary experimental results show that OVROP can perceive a desired object within 160 ms in most cases. As a result, OVROP can provide detailed information on surrounding environment for full-scope and real-time robot perception.

## 1. Introduction

Visual feedback plays an important role to prompt the sensing capability of the robot. Many vision systems, e.g., visual guided walking (KHR-2) [1], stereovision tracking [2] and remote meetings [3], have been applied to study humanoid robots. However, these directional perception systems only provide a limited field of view (FOV), which is usually no more than 120°. Humanoid robots cannot obtain omnidirectional information of surroundings through these systems, but such panoramic perceptual information plays an important role in detecting unknown situation behind the robot. Additionally, omnidirectional vision can provide a comprehensive view and allows more accurate navigation of mobile robots [4]. High-resolution omnidirectional imaging and modeling methods [5], omnidirectional image-based control law [6], preliminary 3D reconstruction adopting rotation estimation [7] have been extensively studied for the navigation of mobile robots. With the rapid development of advanced technology, many compact and high-performance LIDARs have been developed for robots. For example, a LIDAR called VLP-16 (Velodyne LiDAR, Japan) [8] can detect obstacles with 360° FOV, and it is compact and can obtain 3D distance. But comparing with cameras, LIDARs are usually built at high cost and they only provide grayscale images of limited resolution.

To realize panoramic perception for robots, various commercial cameras have emerged and been widely used, such as Hemispheric Camera Q25 (MOBOTIX, Germany) [9] and Ladybug (Point Grey, Canada) [10]. Compact design of Q25 allows complete room monitoring with only one camera. But it only perceives in a cone scope, the vertex of which is the fixed position of the camera (e.g., a Q25 is installed on the ceiling of a house). Thus, it is not easy for a Q25 to directly obtain the 360° observation information for a robot. Installed with 6 compact cameras, Ladybug series products are able to almost completely monitor their surroundings. Taking the Ladybug 3 for an example, as 5 cameras are positioned at the optical center of a plane and the remaining camera points at the normal direction plane, 6 cameras can easily obtain the omnidirectional images [11]. However, the low frame rate of Ladybug 3 to grab image ($\leq$15 FPS) makes it difficult for real-time perception. Furthermore, Ladybug cameras cannot provide stereo vision for detailed perception.

Many previous studies attempted to use catadioptric panoramic stereovision to address such problems [12–18]. Catadioptric cameras can acquire a single panoramic image containing the 360° FOV information. However, images captured by a catadioptric lens are prone to distortion, and such problem occurs especially in 3D reconstruction [18]. Then a high resolution lens is employed to reduce information loss resulting from severe distortion, but the low

---

[*] Corresponding author.
*E-mail address:* shiqing@bit.edu.cn (Q. Shi).

http://dx.doi.org/10.1016/j.mechatronics.2016.11.005
0957-4158/

frame rate hinders real-time perception. Furthermore, FOV of catadioptric panoramic stereovision system is restrained by the reflecting mirror, hence the scene above the system cannot be seen (or be significantly distorted).

To overcome these limitations, in this study we propose to use multiple distributed cameras instead of a catadioptric camera for robot perception. We developed an omnidirectional vision system for robot perception (OVROP). OVROP consists of three monocular cameras and one stereo camera. In general, each monocular camera is used to detect and track surrounding objects (target or obstacle), whereas the stereo camera is to locate the expected target and acquire more detailed information with the help of 3D image processing. As a result, OVROP shares advantages of both omnidirectional vision and stereovision, and provides detailed information of surrounding environment. Equipped with its own driving and control system, OVROP is a completely independent device. Additionally, a modularly designed hardware and software enable OVROP to be compatible with various types of robots. Regarding the real-time perception requirement, we consider 5FPS (processing time: 200 ms) as the minimum frame rate for our system based on application of many robots [19,20]. Besides, human reaction time is around 0.25 s, meaning perception greater than 4 FPS is sufficient to detect reactions to a robots action [21].

Our previous conference paper [22] described an initial prototype of OVROP and preliminary performance evaluation. In this paper, we followed scientific approaches and conducted experimental tests on omnidirectional perception and tracking, and the detailed calibration procedures. Additionally, we have improved the perception performance with subdivision method and optimized 3D reconstruction algorithm. Main advantages of the proposed system are as below: (a) undistorted panoramic perception of surrounding environment; (b) real-time 3D reconstruction with stereo vision; (c) flexible application to various robots. More specifically, the present system is able to obtain distinctive image features and capture video sequences by integrating all distributed cameras. It enlarges perception scope since each camera can move separately. Stereo vision facilitates detailed extraction of 3D features and depth information for a specific target. Additionally, we plan to design general hardware and software interfaces to make this visual device to be compatible with more robots.

The rest of the paper is organized as follows. Section 2 presents a description of hardware design and perception procedure of OVROP. Section 3 describes the calibration procedure of both monocular and stereo cameras. In Section 4, we experimentally and analytically study the perception capability of OVROP. The last section presents conclusions of this paper.

## 2. System description

As demonstrated in Fig. 1, OVROP consists of a control board, five motors and four cameras. The control board is used for initial image processing and driving motors. One motor is used to drive the whole device, and the other four drive the pitch rotation of four cameras respectively. An arm control board is used for image pre-processing and data transmission between control unit and cameras. The main computer (such as industrial computer) deals with image processing and sends valid information to the arm control board.

### 2.1. Hardware design

The mechanical structure of OVROP mainly consists of two parts: device and connection interface (Fig. 2). The device ( Diameter × Height: 200 mm × 50 mm) is composed of one plate (radius: 90 mm) and five holders for installing motors and cameras. Four cameras are installed on the plate at 90° to each other, and that enables OVROP to provide a panoramic view for robot. Driven by a motor, the whole device is rotated around the yaw axis ($\theta_b$). Each camera is separately driven by a motor around the pitch axis ($\theta_s, \theta_m^1, \theta_m^2, \theta_m^3$), and the moving range of each motor is within $-20^\circ \sim 40^\circ$. The connection part is used to fix the device with other platforms, such as a humanoid robot or a mobile robot. To the first prototype, we have designed a hat-shaped (Radius of the maximum : 230 mm; Height: 145 mm) appearance to connect a humanoid robot head. Furthermore, the connection part is designed to be flexible to connect a variety of other kinds of robots using four bolts.

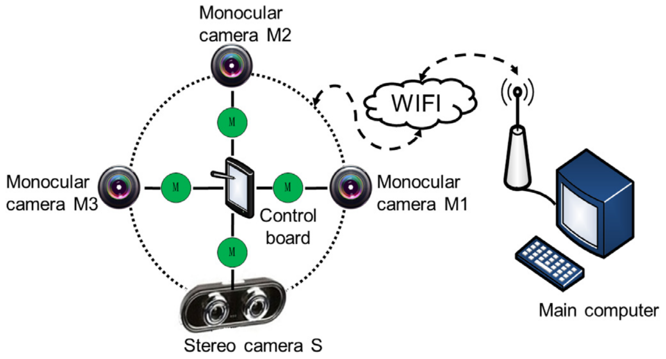

**Fig. 1.** The OVROP controlled by a main computer via WIFI is composed of an arm control board, five stepper motors and four cameras (three monocular cameras: M1, M2, M3, and one stereo camera: S). Four stepper motors are used to actuate four cameras M1, M2, M3 and S, whereas the fifth motor is hidden by the plate and it is used to rotate the whole body of OVROP.

Regarding the motor to actuate four cameras, stepper motor is chosen due to its low cost and high reliability. Given the pulse frequency $f$, the step angle $\varphi$ and the subdivision multiple $x$, the output torque $T_o$ of stepper motor can be derived from Eqs. (1) and (2). Since $T_o$ is inversely proportional to $f$, a smaller pulse frequency allows for a bigger output torque of stepper motor. A compact stepper motor with high output torque ($T_o = 0.06N \cdot m$): PG15S-D20 (NMB, USA) was chosen for OVROP. Furthermore, the maximum required torque to rotate the camera was calculated to be $T_m = 0.043N \cdot m$ using Eq. (3). Obviously $T_o > T_m$, the selected stepper motor meets system requirement.

$$n = f \cdot \frac{1}{360/\varphi \cdot x} \cdot 60 \tag{1}$$

$$p = k \cdot T_o \cdot n \tag{2}$$

$$T_m = k_l \cdot l_b \cdot \overline{D} + k_b \cdot (\nu \cdot n)^{\frac{2}{3}} \cdot \overline{D}^3 \times 10^{-7} \tag{3}$$

where $n$ is the velocity of bearing, $k$ the constant coefficient, $p$ the output power, $k_l$ the load coefficient, $l_b$ the bearding load, $\overline{D}$ the mean diameter of bearing, $k_b$ the bearing coefficient, $\nu$ the kinematic viscosity.

Regarding the motor to rotate the whole device of OVROP, high speed and output torque are the key parameters. The maximum output torque of the DC motor should be able to drive the OVROP to rotate at an ideal angular velocity (e.g., $2\pi$ rad $\cdot s^{-1}$). This means it should meet the following equation:

$$P_b \geq T_b \omega \tag{4}$$

where $P_b$ is the output power of the motor to rotate the whole device, $T_b$ the required torque, $\omega$ the angular speed. In this study, we chose DC motor: EC-i40 (Maxon, Sweden) for its high rotation speed and high torque to drive the whole device.

The arm control board used in this study mainly refers to a mini PC platform called PcDuino. As shown in Table 1, a CPU of