Infrared Physics & Technology 62 (2014) 115–123

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

A stereo range computation method using thermal infrared and visible cameras

Wen Chen, Li Cao *, Zijian Wu, Shengguo Huang

College of Civil Aviation, Nanjing University of Aeronautics & Astronautics, Nanjing, Jiangsu, PR China

HIGHLIGHTS

- A multimodal stereo ranging system composed of a thermal camera and a visible camera is built.
- A simple and reliable range computation method is presented when two different cameras are placed in parallel configuration.
- The procedure to identify the parameters required for the range computation method is described.
- The effect of the relative position of a certain target and reference points is also analyzed when only two reference points are used to identify the parameters.

ARTICLE INFO

Article history:
Received 16 April 2013
Available online 19 November 2013

Keywords:
Binocular stereo
Range computation
Thermal camera
Visible camera
Parallel camera configuration

ABSTRACT

Binocular stereo vision can provide geometric position information of the target, which enables one to track the moving object precisely. Although visible images are full of details of geometry and texture, it is difficult to detect moving objects in poor visibility. Appropriate fusion of infrared and visible images can combine the complementary information and obtain a better description of the scene, which will help in target detection and target localization. Considering the physical differences between thermal infrared cameras and visible cameras, a stereo ranging method in parallel camera configuration is proposed in this paper. The rotation center of a pan-tilt device is used as the origin of the coordinate system in the method. It could not only locate the object in poor visibility, but also determine the parameters essential in range computation in practical work, which is very difficult in traditional calibration methods. Furthermore, the effect of the relative position of a certain target and reference points is analyzed. Experiments also proved the validity of the proposed method.

1. Introduction

Monocular vision is simple because it could avoid image fusion, which means it can meet the requirement of real-time in visual surveillance. However, it can only provide two-dimensional information of the scene. On the other side, binocular vision can provide three-dimensional (3D) information of the object by shooting the object at the same time by two cameras placed in different positions. Thus, binocular vision makes it possible to precisely track the moving object. However, it is a huge challenge to detect and track the moving object by two visible cameras in poor visibility, such as night vision, fog, rain, and other inclement weather.

With the rapid development of science and technology, the application of multi-sensor technology is increasingly widespread. Multiple vision sensors are now widely used in visual surveillance, improving the robustness and accuracy of visual surveillance system in all-weather conditions. A target ranging system based on visual and active infrared imaging in night vision was proposed in [1]. An active infrared detection system is usually consisted of an active infrared transmitter and a passive infrared sensor (receiver). The transmitter emits a beam of light into the detection zone. The light, which is reflected by the background, returns to the receiver that constantly monitors the detection zone. Therefore, the effective range of an active infrared system is dependent on the power of the transmitter. Furthermore, the transmitter and its power supply are not convenient in field. By contrast, a passive infrared detection system can work without transmitters. Only one thermal camera is needed in the passive system. The thermal camera creates images based on differences in surface temperature by detecting infrared radiation (heat) that emanates from objects and their surrounding environment. It can be used in imaging when there is insufficient visible light to see such as in night vision, as well as special

* Corresponding author. Tel.: +86 13601585169.
E-mail address: caoli@nuaa.edu.cn (L. Cao).

1350-4495/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.infrared.2013.11.003

camouflage detection. However, the thermal image is poor in spatial resolution and texture which means a texture or an edge in a visible image is often missing in the thermal image. Using visible-infrared camera pairs is now attracting more attention because the combination of the information from these two sensors performs well in complementary situations. Although related work has been developed in visual–thermal fusion, most of them are focused on enhanced vision [2–4]. A multimodal stereo vision system consisted of a thermal camera and a visible camera is able to make full use of grayscale and temperature to get the depth information of the object in poor visibility. Therefore, it contributes to precise tracking in all weather conditions [5,6].

The intrinsic parameters of the lens (focal length and principle point), the extrinsic parameters of the lens (relative position of the optical centers), and the disparity information are indispensable in stereo ranging. Therefore, the precision of ranging results greatly depends on the precision of the parameters calibrated. In a typical stereo vision system composed of two visible cameras, the calibration is usually carried out with the help of a certain calibration board such as a checkerboard pattern. With a matter of placing a checkerboard pattern in front of the camera, these parameters can be carefully calibrated. However, most of the existing calibration methods are complicated and time-consuming [7–12]. It usually costs several hours to calibrate the parameters for just one camera. Furthermore, once the focal length changes or one of the cameras moves, it is necessary to re-calibrate the intrinsic/extrinsic parameters. So they are not appropriate in field.

It is obvious that calibrating these parameters is in the first place to compute the range in the multimodal stereo vision system composed of a thermal camera and a visible camera of which intrinsic parameters are different. Meanwhile, the calibration method must be simple and reliable to be suitable for practical application. Furthermore, a thermal image is very different from a visible image. If the calibration method based on the calibration board is adopted, some extra care must be taken to ensure the calibration board looks similar in each modality. In [13], the multimodal stereo calibration was carried out with a heated calibration board based on the typical calibration method. As mentioned above, although the calibration results can meet the requirement of precision, the procedure is too complicated to be applied in field. In [14], the thermal camera was calibrated with the help of a pair of calibrated visible cameras. However, it required ideal environmental conditions.

In this paper, the range is computed to provide position information of the target for the tracking system, so operation speed is more important than precision. To solve these problems, a stereo ranging method in parallel camera configuration is proposed in this paper. The rotation center of a pan-tilt device is used as the as the origin of the coordinate system in the method. Based on this method, the line distance between the target and the rotation center of the pan-tilt device can be computed as well as the angle. The procedure is also shown to identify the parameters used in this method. The method is simple and easy to calculate. Only geometry information of the pan-tilt devices and several reference points (at least two points) are needed to determine the parameters in the proposed method. The effect of the relative position of a certain target and reference points is also analyzed when only two reference points are used to identify the parameters. The region in which better ranging result can be gotten is provided in the experiment. Finally, experiments proved that the results of the proposed method could meet the requirement of object tracking and locating in all weather conditions.

2. Computing range

In stereo vision, the relationship between a 3D point and its image projection is determined by the camera model. The simplest model of a camera is the pinhole camera model. In this paper, pinhole model is used as the camera model for both visible camera and thermal camera. In addition, the position of the image plane and the position of the optical center are switched in the coordinate system described in this paper for convenience (see Fig. 1). As shown in Fig. 1, I is the image plane, O is the optical center of the lens, $O_o(u_o,v_o)$ is the intersection of the optical axis and the image plane(the principle point), $p(u,v)$ is the projection of the spatial point $P(x_c,y_c,z_c)$ to image coordinates, and f is the focal length, which is precisely the distance from the optical center to the image plane for a idealized pinhole camera. Usually, the optical center O is set as the origin of the camera coordinate system (CCS), the optical axis is set as Z-axis, and the optical axis is perpendicular to the image plane. Based on this assumption, the relationship between a 3D point P and its image projection p is given by:

$$z_c\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & u_o \\ 0 & f & v_o \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}. \tag{1}$$

2.1. Stereo ranging principle in parallel camera configuration

According to the relationship of the optical axes of the right and left lenses, there are two methods of binocular stereoscopic camera with controllable vergence: toed-in configuration and parallel configuration. In toed-in configuration, these optical axes are made to cross at some point during the shooting; while in parallel configuration, the optical axes are parallel to each other. The latter is relatively easier in computing range so the calibration is easier. Therefore, this study computed range using a visible camera and a thermal camera which were arranged in parallel configuration.

An example of parallel camera configuration is shown in Fig. 2(a). The camera coordinate system of the left camera C_1 is $O_1x_1y_1z_1$, and the camera coordinate system of the right camera C_2 is $O_2x_2y_2z_2$. Let P be the target point, then p_1 and p_2 are the projections of P to the image plane in the left and right camera view respectively. The intersection of O_1p_1 and O_2p_2 is P, and it is the one which is only confirmed. In the parallel configuration, each axis in $O_1x_1y_1z_1$ is parallel to the correspondence axis in $O_2x_2y_2z_2$. The distance from O_1 to O_2 in X-axis is b, the interval of O_1 and O_2 in Y-axis is e, and the interval of O_1 and O_2 in Z-axis is d (when O_1 is above O_2, $e > 0$; when O_2 is behind O_1, $d > 0$.). In the configuration described above, set the left camera coordinate system

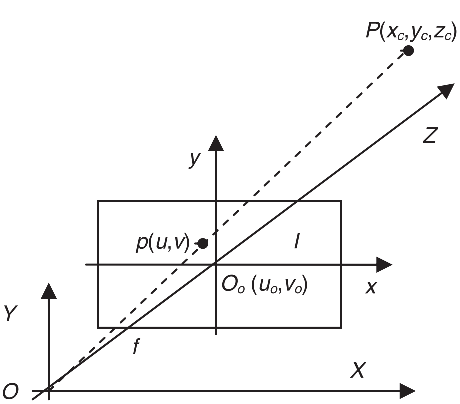

Fig. 1. Projection model.