https://doi.org/10.52326/ic-ecco.2021/CE.01

Localization of Objects in 3D Space by Using a Stereoscopic Video System

Valerii FRATAVCHAN [1], Tonia FRATAVCHAN [1], Roman DRUCHUK [1], Victor ABABII [2]
[1] Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine
[2] Technical University of Moldova, Chisinau, Republic of Moldova
Corresponding Author: Valerii Fratavchan, v.fratavchan@chnu.edu.ua

***Abstract*— This article describes two implementation models of an artificial vision for specific tasks of localization of objects in three-dimensional space using a stereoscopic system of two cameras.**

The periscopic model provides means for objects localization relative to the position of the stereoscopic system. This can be used for both detecting the position of bodies in space and for performing the technical system orientation with an artificial vision. The orthogonal model is capable of executing a high-precision dynamic localization of objects in space.

Keywords*— *artificial vision; stereoscopic video system; absolute and relative coordinates; localization and identification of objects; periscope video system; orthogonal video system.

I. Introduction

The implementation of computer vision is a part of many cyber-physical systems. Such systems perform the functions of monitoring the external environment and enable mobile devices to orient themselves to the terrain. The concept of artificial vision provides for the use of video units to obtain information about the exterior medium. It is often enough to have one video device to model artificial vision tasks. A fully functioning artificial vision is an integral part of a holonic artificial intelligence system and it is a useful aid to perform objects localization and identification tasks. The system must use at least two video devices to provide these functions by analogy with biological organisms [1-4]. This article suggests two strategies for artificial vision implementation, according to its tasks and additional requirements.

II. Periscopic Model of Artificial Vision

The periscopic model is a computer analogue of the biological organism's vision. The system consists of two closely-spaced video cameras. It is preferable to locate them on the same level, so that their centroidal observation axes are parallel (Fig. 1).

The distance between the cameras is defined as the periscopic base. The stereoscopic periscopic video system can be adjusted either by pointing both cameras at a sufficiently remote object or by using the setup plan (Fig. 1). Such a stereoscopic system is effective at distances at which there is a relative shift of objects in the images of two cameras. The variation of the distance to the object with respect to the offset value can be experimentally determined and tabulated in advance. A significant problem is the necessity of identifying objects in simultaneous images from two cameras in order to localize objects. This is possible when the object falls into the field of view of both cameras. In addition, the ability to localize several objects at the same time is an advantage of the periscopic method.

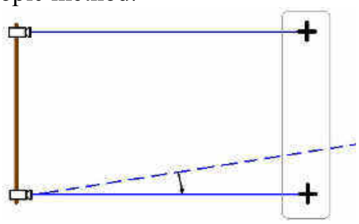

Figure 1. Periscopic video system model.

Thus, the work of the periscopic stereoscopic system consists of the following steps:

- machinal development of periscope system (Fig. 2) and its adjustment (Fig. 3);

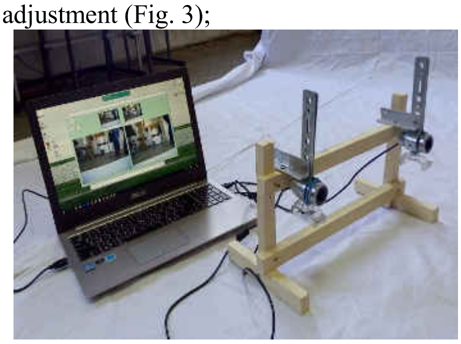

Figure 2. The development of periscopic video system.

https://doi.org/10.52326/ic-ecco.2021/CE.01

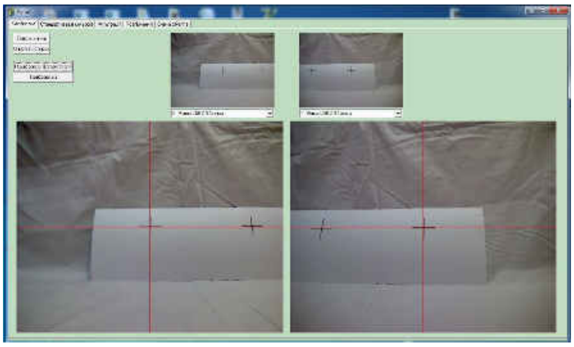

Figure 3. The adjustment of periscopic video system.

- tabling the variation of the distance to the object with respect to the offset value in the images from the two cameras (it is necessary for the cameras to have identical engineering data to get correct calculations);
- dynamical obtain of synchronous still images from two cameras (Fig. 4) and object identification (Fig. 5);

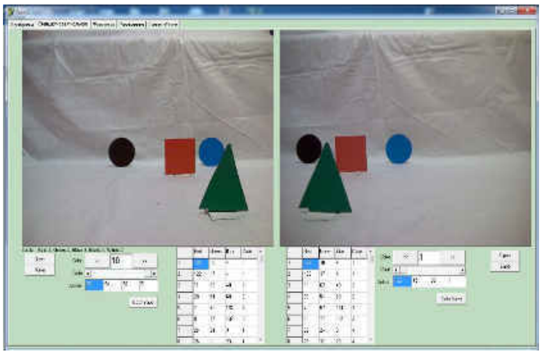

Figure 4. The obtain of synchronous still images from two cameras.

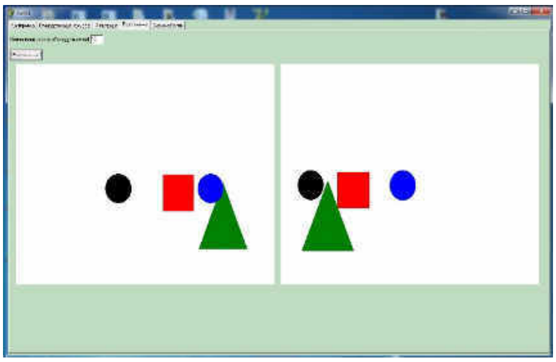

Figure 5. The obtain of synchronous still images object identification.

- objects localization and their placement on the "topographic scheme". As this takes place, one of the cameras is considered as the main one and determines the object's frontal coordinate, while the object's offset in another image determines the perspective coordinate according to the table (Fig. 6).

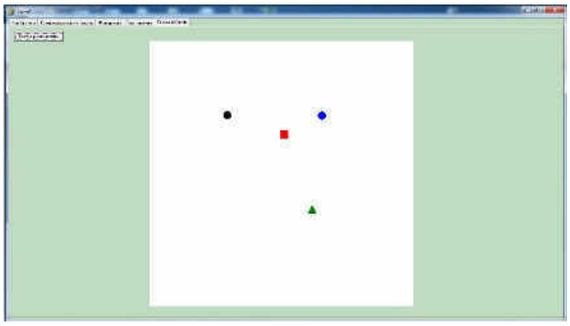

Figure 6. Localization scheme of identified objects.

III. Orthogonal Stereoscopic Model of an Artificial Vision

The periscope system is effective for localizing objects placed close to the artificial vision system. The orthogonal model can be used to observe and localize objects to high accuracy. This takes into account the fact that the video cameras can be placed as far apart from each other as from the "central organ" of the artificial intelligence system. To simplify the mathematical model, it makes more sense to place cameras in the same horizontal plane to intersect the central observation e right angle. The axes' cross-point is considered the center of the relative coordinate system (the third axis is placed perpendicular to the first two). We consider two model variations.

A. Orthogonal Model with Parallel Projecting

Let's assume that there is a fixed Cartesian reference system in a 3D space. Located on the x and y coordinate axes, cameras are directed toward the coordinate origin. The 3D space projection planes are at the right angles with the plane of the cameras' location and cut the x and y axes. Cameras are located at a far distance from the coordinate origin, thus we assume that projection rays are vertical to the projection plane (parallel projection) (Fig. 7). In this case, the geometric centers of the cameras' view windows will align with the center of the coordinate system in 3D space, while the field of view has the form of a rectangular parallelogram and which is limited to the size of the camera window.

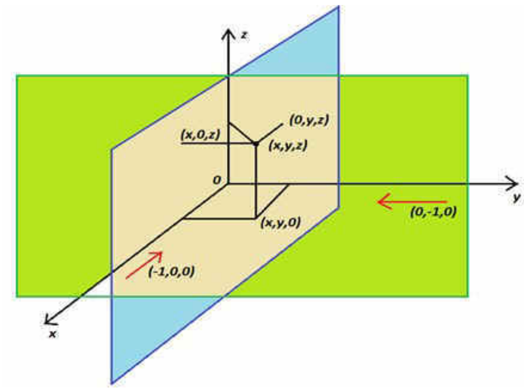

Figure 7. The model of 3D coordinate system and projective

https://doi.org/10.52326/ic-ecco.2021/CE.01

planes.

Triaxial coordinates of "moving" point (x, y, z) in space are calculated using the formulas:

$$x = X1 - Cx1$$

$$y = X2 - Cx2$$

$$z = -Y2 + Cy2,$$

where: $X1$ – the point coordinate x in the first camera's view window;

$X2$ – the point coordinate x in the second camera's view window;

$Y2$ – the point coordinate y in the second camera's view window;

$Cx1$ – the center point coordinate x in the second camera's view window;

$Cx2$ – the center point coordinate x in the second camera's view window

$Cy2$ – the center coordinates y of the second camera's view window.

B. Orthogonal model with perspective projection

The parallel projection model ignores the resizing of the object and angular values in relation to the distance between the object and the camera. Such dependence is taken into account in the mathematical model of localization, based on the conception of perspective projection. The location of the coordinate systems on the projective planes is analogous to the model of parallel projection. The difference between these models is that cameras are located on coordinate axes at relatively small focal distances while perspective projection (we assume that both focal distances are the same to simplify the model). Projective perspective coordinate transformations are shown in Fig. 8.

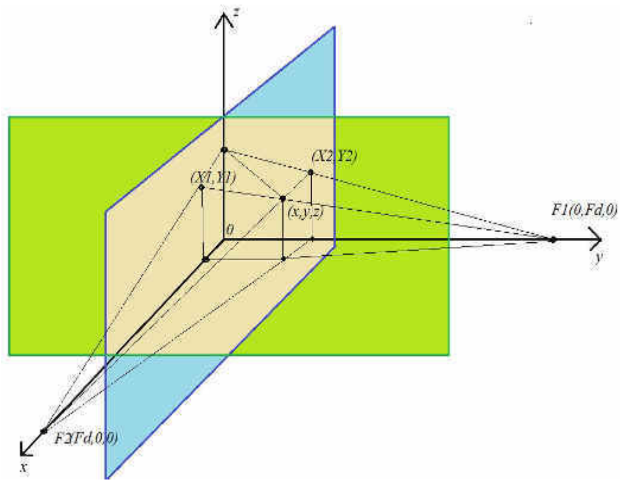

Figure 8. The model of stereometrical perspective projective transformations.

In Figure 8, there are the coordinates and the position of an observable object in the still images from the first and the second cameras after matching up 3D space coordinate systems and coordinate systems of still images.

The projective plane contains all information necessary for coordinate's localization (Fig. 9).

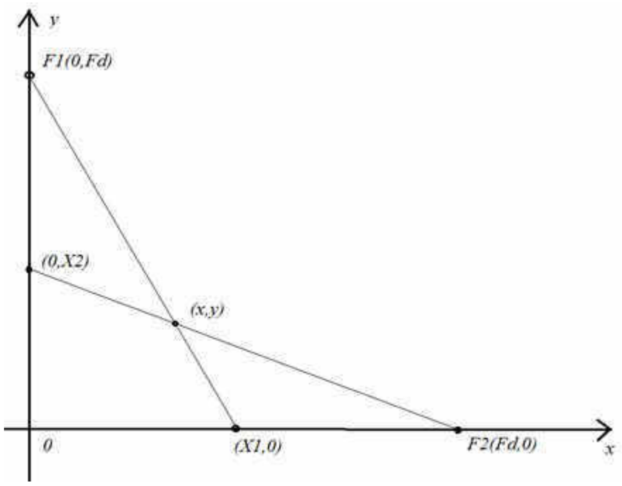

Figure 9. Projection of localization model on a plane xOy.

To determine the coordinates (x, y) it is necessary to perform the following actions:

1) Determine two intercept equations of a line (all wanted coordinates are known):

$$\begin{cases} \dfrac{x}{Fd} + \dfrac{y}{X2} = 1 \\ \dfrac{x}{X1} + \dfrac{y}{Fd} = 1 \end{cases} \quad (1)$$

2) Pass on to the general equations of a line:

$$\begin{cases} X2 * x + Fd * y = X2 * Fd \\ Fd * x + X1 * Y = X1 * Fd \end{cases} \quad (2)$$

3) Solve simultaneous equations using Cramer's rule:

$$x = \frac{X2 * Fd * X1 - Fd^2 * X1}{X1 * X2 - Fd^2}$$

$$y = \frac{X2 * Fd * X1 - Fd^2 * X2}{X1 * X2 - Fd^2} \quad (3)$$

The coordinate z is calculated straight from the 3D localization model, according to the still image from any video camera, for instance, from the ratio:

https://doi.org/10.52326/ic-ecco.2021/CE.01

$$\frac{z}{Y1} = \frac{Fd - x}{Fd}, \tag{4}$$

so

$$z = Y1 * \frac{Fd - x}{Fd}, \tag{5}$$

Note: the cases when there are zero values among coordinates $(X1, Y1)$ and $(X2, Y2)$ should be analyzed individually.

CONCLUSION

This study shows two models of objects localization in 3D space using a stereoscopic system of two video cameras. A perspective model can be used to simulate the artificial vision of the "biological" type. The model does not guarantee high localization accuracy but it allows to estimate the mutual arrangement of several "familiar" objects, and it helps to perform "terrain orientation" while placing an artificial vision system on a mobile autonomous platform, thus, determining its position within a certain area of 3D space. In the presence of remote video cameras, an orthogonal model can perform real-time high precision localization of moving objects in certain field of view. The accuracy of the object's localization in the orthogonal model dominantly depends on the accuracy of the geolocation of the involved cameras and their mutual orthogonal orientation. The GPS localization results can be reasonable for localization of relatively remote cameras. The cameras orientation can be easily and accurately done using theodolite or dip compass.

Consequently, both video location models are not complicated for technical implementation, they are reasonably priced and accessible for completing practical tasks using artificial vision systems.

REFERENCES

[1] D.Rodgers, J.Adams, Matematichescie osnovy mashinnoi grafici [Mathematical bases of computer graphics], M., 1980, 240 p.
[2] D.Rodgers, Algoritmichescie osnovy mashinnoi grafici [Algorithmically bases of computer graphics], M., 1989, 560 p.
[3] A.Angel, Interactivnaia mashinnaia grafica [Interactive computer graphics], M: «Williams», 2001, 590 p.
[4] R. Honsales, R.Woods., Tsifrovaia obrabotca izobragenii [Digital images preparation] — M.: "Technosphere", 2006. — 1072 p.