Optics and Lasers in Engineering 54 (2014) 269–274

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

# Modeling of binocular stereo vision for remote coordinate measurement and fast calibration

Jian Wang [a,*], Xiangjun Wang [b], Feng Liu [b], Ying Gong [c], Huanhuan Wang [a], Zhe Qin [c]

[a] Tianjin Key Lab for Advanced Signal Processing, Civil Aviation University of China, Tianjin 300300, China
[b] MOEMS Education Ministry Key Laboratory, Tianjin University, Tianjin 300072, China
[c] Sino-European Institute of Aviation Engineering, Civil Aviation University of China, Tianjin 300300, China

ARTICLE INFO

*Article history:*
Received 30 April 2013
Received in revised form
27 July 2013
Accepted 30 July 2013
Available online 22 October 2013

*Keywords:*
Events sensing
Binocular stereo vision
Parameters calibration
Airport surface surveillance

ABSTRACT

This research is intended for a super-size 3 dimensional scaling events sensing and positioning system based on binocular stereo vision. A model for computing the coordinates of targets in the physical world coordinate system is proposed and simulated. Constitution of a reference point for remote coordinate measurement is made for this purpose on condition that coordinates of both digital photogrammetric workstations are given, with the knowledge of the distance between the two workstations and the angles between optical axes of cameras and the base line connecting the workstations. A calibration method is presented. Applications of the results could be found in airport surface surveillance and runway incursion monitoring. Position data obtained in real time for aircraft and vehicles on the airport surface could be presented to an Advanced Surface Movement Guidance and Control System (A-SMGCS).


## 1. Introduction

Increasing traffic on airport surface causes air traffic control to become more complex than ever before. Traffic volumes in major airports of China will increase at a higher annual rate and runway incursion is considered as one of the most critical safety issues of airports. To alleviate air traffic controllers from heavy workload, some airports are equipped with surface movement radars (SMR) and additional means, such as multi-lateration [1,2] systems based on signals from Mode-S transponders on board aircraft, and automatic dependent surveillance broadcasting (ADS-B). Multi-lateration and ADS-B are working in a cooperative mode and subjected to interference caused by reflections, but non-cooperative targets cannot be detected by these systems, thus increasing the risk of a runway incursion. Primary radar technology suffers from such limitations as shadowing effects or multiple reflections. High prices and maintenance cost also limit the usage in most airports. For these reasons, a new non-cooperative positioning method is required, either as an additional sensor to cover blind areas of radars in major airports or as an alternative to surface movement radars in small airports. The principle of binocular stereo vision has found many applications [3–5], but mainly indoors and the range is normally within several meters. Few applications are reported for positioning large moving targets on airport surfaces.

The proposed scheme implements remote coordinate measurement for targets based on binocular stereo vision technology. Quite different from technologies now used in airport surveillance, sensors used for this scheme are video cameras with high resolution charge-coupled device, well arranged to detect targets according to their pixel positions on each camera. With system parameters known, position of each target would be computed with the model proposed. Positions of targets could be then sent in ASTERIX format (a standard for the exchange of radar data) to any A-SMGCS system [6].

## 2. Modeling for target positioning in a 3-dimentional scene

Modeling of finding the coordinates of targets of small size but long range has been made by our group in 2006 [7], where cameras are in a horizontal plane and the *Y* position on camera plane is neglected. Usual configurations for distance measuring and positioning using cameras require that cameras be arranged approximately frontal parallel and as close to horizontally aligned as possible [8]. Positions of targets in units of pixels, or points, on an image plane are the essential information for coordinate measurement. Binocular stereo vision also requires finding correspondences between points seen by one camera and the same points seen by the other.

* Corresponding author. Tel.: +86 22 24092827, +86 13602158556; fax: +86 22 24092827.
*E-mail address:* caucwang@263.net (J. Wang).

0143-8166/$ - see front matter © 2013 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.optlaseng.2013.07.021

In airport surface surveillance applications, where targets to be measured are large and the distance between a target and a camera is more than 1 km, cameras are required to be raised at a height and two cameras would not be always in the same level. Therefore, a new model for establishing a relation between coordinates of a target in physical units of meters and its position on an image plane in camera's units of pixels is critical in the attempt to measure distances of targets from the reference point in the three-dimensional world. With system parameters of cameras and correspondences known, locations of targets can be computed using the model established in this article. The search for corresponding points is not included in this article.

### 2.1. Formula of finding coordinates of targets

A pinhole camera model [9] is most widely used in the field of computer vision, transforming videos from cameras into a new representation for achieving particular goals in different kinds of applications. With a pinhole camera model, a projective transform of the point $P$ with coordinates ($X_w$, $Y_w$, $Z_w$) in a physical world to the point $p$ with coordinates ($X$,$Y$) on an image plane is given by the following basic matrix in terms of homogeneous coordinates [10,11].

$$z_c\begin{pmatrix}X\\Y\\1\end{pmatrix}=M\begin{pmatrix}X_w\\Y_w\\Z_w\\1\end{pmatrix} \quad (1)$$

where $M$ is a 3-by-4 matrix, consisting of the intrinsic parameters of a camera and the geometric parameters that constitute a rotation matrix and a translation vector. A rotation matrix is described in the form of a matrix multiplication. Our application requires two dimensional rotations around axes-$y$ and -$x$, with angles denoted $\alpha$ and $\beta$ respectively. A translation vector is the shift from the origin of the camera's coordinate system to the origin of a physical world coordinate system, and is expressed by components $S_x$, $S_y$, and $S_z$ in three axes. With basic forms of rotation matrix and translation vector [9], the matrix $M$ can be written in a form as follows:

$$M=\begin{pmatrix} f\cos\alpha & 0 & -f\sin\alpha & S_x\cos\alpha-S_z\sin\alpha\\ f\sin\alpha\sin\beta & f\cos\beta & f\cos\alpha\sin\beta & S_x\sin\alpha\sin\beta+S_y\cos\beta+S_z\cos\alpha\sin\beta\\ \sin\alpha\cos\beta & -\sin\beta & \cos\alpha\cos\beta & S_x\sin\alpha\cos\beta-S_y\sin\beta+S_z\cos\alpha\cos\beta \end{pmatrix} \quad (2)$$

where $f$ is the focal length of a camera, the same as the distance from the pinhole aperture to the image plane, or CCD in most cases, in a camera.

To begin with, an ideal case is considered where the optical axis is strictly at a right angle with respect to the image plane. By further investigation of the geometry, the first and second element of the transformation vector can be set to zeros, and the third element is the slant distance from the origin of the physical coordinate system to the center of the image plane of a camera, denoted by $R$,

$$M=\begin{pmatrix} f\cos\alpha & 0 & -f\sin\alpha & 0\\ f\sin\alpha\sin\beta & f\cos\beta & f\cos\alpha\sin\beta & 0\\ \sin\alpha\cos\beta & -\sin\beta & \cos\alpha\cos\beta & R \end{pmatrix} \quad (3)$$

where $R=\sqrt{S_x^2+S_y^2+S_z^2}$.

Substituting matrix $M$ into formula (1), it can be written in a linear form, with three equations connecting physical world coordinates of a target and the observed location on an image plane.

$$z_cX=f\cdot X_w\cos\alpha-f\cdot Z_w\sin\alpha \quad (4)$$

$$z_cY=f\cdot X_w\sin\alpha\sin\beta+f\cdot Y_w\cos\beta+f\cdot Z_w\cos\alpha\sin\beta \quad (5)$$

$$z_c=X_w\cos\alpha\cos\beta-Y_w\sin\beta+Z_w\cos\alpha\cos\beta+S' \quad (6)$$

Considering an application of target movement positioning on an airport surface, *i.e.*, a two-dimensional planar surface, the coordinate $Y_w$ of a target is insignificant. Collecting these three equations and eliminating $Y_w$ and $z_c$, an arbitrary scale factor, it is possible to express the relationship between coordinates ($X_w$, $Z_w$) and system parameters of one digital photogrammetric workstation, in which a camera is included.

$$(X\sin\alpha-Y\cos\alpha\sin\beta-f\cos\alpha\cos\beta)X_w+(X\cos\alpha-Y\sin\alpha\sin\beta-f\sin\alpha\cos\beta)Z_w+XS'\cos\beta=0 \quad (7)$$

Scheme diagram of binocular stereo vision is shown in Fig. 1. For a typical configuration for applications using a binocular stereo vision model, at least two digital photogrammetric workstations are necessary. When two workstations are installed, two equations with different system parameters and intrinsic parameters will be combined, forming an epipolar geometry. When two equations are combined, solving for $X_w$ and $Z_w$ yields,

$$X_w=\frac{X_1X_2(S_2\cos\alpha_1-S_1\cos\alpha_2)+X_2(Y_1'+f_1')S_2\sin\alpha_1-X_1(Y_2'+f_2')S_1\sin\alpha_2}{(X_1X_2+Y_1'Y_2'+f_1'f_2'+Y_1'f_2'+Y_2'f_1')\sin(\alpha_1-\alpha_2)+(X_1Y_2'+X_1f_2'-X_2Y_1'-X_2f_1')\cos(\alpha_1-\alpha_2)} \quad (8)$$

$$Z_w=\frac{-X_1X_2(S_2\sin\alpha_1-S_1\sin\alpha_2)+X_2(Y_1'+f_1')S_2\cos\alpha_1-X_1(Y_2'+f_2')S_1\cos\alpha_2}{(X_1X_2+Y_1'Y_2'+f_1'f_2'+Y_1'f_2'+Y_2'f_1')\sin(\alpha_1-\alpha_2)+(X_1Y_2'+X_1f_2'-X_2Y_1'-X_2f_1')\cos(\alpha_1-\alpha_2)} \quad (9)$$

All parameters on the right side of the expressions are denoted with subscripts, valued either 1 or 2, indicating corresponding parameters in photogrammetric workstation 1 or 2. $X_i$ and $Y_i$ are the lateral and the longitudinal positions of a target on an image plane. $Y_i'$ is a horizontal component of $Y_i$, with a value of $Y_i\sin\beta_i$ and $f_i'$ is a horizontal component of $f_i$, with a value of $f_i\cos\beta_i$. $\alpha_i$ is a rotation angle around the axis-$y$ and is measured counterclockwise, viewed from the $y$-axis. And $S_i$, in a form of $R_i\cos\beta_i$, is the distance in a horizontal plane between the reference point $P$ and the foot of the tower where the camera is mounted, and is expressed as the square root of sum of the square of components $S_x$ and $S_z$.

With all parameters given, measured or calibrated, coordinates ($X_w$, $Z_w$) of a target, or a target, in an airport surface plane can be calculated, thus realizing remote coordinate measurements of targets on an airport surface.