Scientific Journals
of the Maritime University of Szczecin

Zeszyty Naukowe
Akademii Morskiej w Szczecinie

2017, 52 (124), 103–111
ISSN 1733-8670 (Printed)
ISSN 2392-0378 (Online)
DOI: 10.17402/251

Received: *23.10.2017*
Accepted: *18.11.2017*
Published: *15.12.2017*

A method of determining inland vessel position using a single stationary, non-metric camera

Artur Kujawski[1]✉, Grzegorz Stępień[2]

Maritime University of Szczecin
[1] Faculty of Economics and Transport Engineering
11 H. Pobożnego St., 70-507 Szczecin, Poland, e-mail: a.kujawski@am.szczecin.pl
[2] Faculty of Navigation
1-2 Wały Chrobrego St., 70-500 Szczecin, Poland, e-mail: g.stepien@am.szczecin.pl
✉ corresponding author

Key words: inland shipping, projective transformation, image analysis, non-metric camera, UAV

Abstract

This article presents the results of research on determining inland vessel position using a single stationary, non-metric camera, which was positioned on the Długi Bridge in Szczecin. Studies included a comparison of graphical methods (i.e. geometric transformation and the bunch of rays method) and analytical methods (i.e. 2D to 2D transformation and 3D to 2D transformation). The research material was a collection of video images of an inland vessel navigating the West Odra River between the Długi and Kolejowy bridges in Szczecin. The results of the research were compared to reference points determined using tachymetric surveys and material acquired by an Unmanned Aerial Vehicle (UAV) and the onboard GNSS satellite receiver with Real Time Kinematic (RTK) correction.

Introduction

Determination of the position of a moving object (vehicle, ship or aircraft) is required in navigation to ensure safe movement from a starting point to a destination. In this publication, the authors aim to research the accuracy of the determination of the location of floating objects in the area between the Długi and Kolejowy bridges in Szczecin, using a single non-metric camera (CCTV). Simultaneously, the authors claim that positioning with a non-metric camera is possible, with an accuracy comparable to that of GNSS satellites used for navigation (without RTK correction) using a suitable distribution of points for the transformation.

The realization of such a task, and verification of the hypothesis, requires finding solutions to some research problems, identified in the following questions:

- What is the essence of inland navigation and what accuracy should it provide?
- How can the position of a vessel be determined using a non-metric CCTV camera and what are the metric processing methods required for its transformation?
- How should image processing methods be used for inland navigation to ensure adequate accuracy?

Generally, in navigation, a vehicle's position is determined by the observation of passing landmarks, counting motion parameters from a known position and by the determination of lines of position. A line of position is the set of all possible positions, based on a measured value of a physical quantity (line parameter) which is the same at any point on the line. It follows that the position line may be a straight line or a curve on the surface of the globe (ellipsoid or geoid) or a spatial figure when referring to positions in three dimensions (below or above sea level).

According to the method of determining the position of a vessel, navigation can be divided into: satellite navigation, using GPS receivers; terrestrial navigation, using land-based visibility, where geographic coordinates are based on land-based observations; radar navigation, observation of coastlines and objects; pilotage navigation, used for port approaches where route guidance is based on the identification of passing buoys, ponds and leading lights; dead reckoning navigation; celestial observation; radio navigation, positioning based on radio direction finding, and inertial navigation, mainly used in submarines (Piątek, 2011). In the cases of inland and limited waters, the most commonly used navigation methods are terrestrial, radar, pilotage and dead reckoning.

Determining the position of a vessel using photogrammetry

Photogrammetric methods, based on the metric processing of images, may also be used to determine a moving object's position. A photogrammetric, or non-metric, camera can be treated as a free measurement station. Generally, in similarity transformations, three angles of rotation are used to determine the external orientation e.g. Tait-Bryan or Euler methods (Baranowski, 2013; Stępień, Zalas & Ziębka, 2017). Images, represented by a set of pixels, are recorded in their own internal sensor coordinate system. In this case, the relationship between the coordinates read from the image in the pixel system and the coordinates in the external reference system is described by the relationship (Kurczyński, 2014):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{1}$$

where:

X, Y, Z – coordinates in the primary coordinate system (without rotation),

X_0, Y_0, Z_0 – translation vector (movement of the primary coordinate system),

λ – scale change factor,

ω – rotation around the X axis, the roll angle,

φ – rotation around the Y axis, the pitch angle,

κ – rotation around the Z axis, the yaw angle,

x, y, z – coordinates in the secondary (oblique) coordinate system.

The X and Y axes form a terrain plane and the Z axis is associated with the height. In practice, transformation (1) is very often shown in the form:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \lambda \cdot \mathbf{A} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2}$$

where:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{bmatrix} \begin{bmatrix} \cos\kappa & \sin\kappa & 0 \\ -\sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

a_{11}, a_{12}, a_{13}, …, a_{33} – factors of the **A** rotation matrix.

In order to determine the external orientation of the sensor and, then, the position of the shaded object, transformation (2) is written in the form:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \left(\lambda \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \right) \cdot \mathbf{A} \tag{4}$$

Transformation (4) is described as "Total Free Station" (Free Station or TFS). Free Station was used for GNSS measurements.

Relationship (2) with rotation matrix **A** written in parametric form, (3), is very convenient to use because it does not require knowledge of the angular elements of the camera's sensor. This situation applies to non-metric CCTV cameras, where the external orientation of the camera in the field coordinate system is not specified. If it is necessary to determine the orientation of the camera, formula (4) can be used. All reference points are required to accomplish the transformation in both reference systems. In the case of the space to plane transformation, formula (2) is reduced to parametric form (4) (Zalas et al., 2016):

$$x = \frac{AX + BY + CZ + D}{EX + FY + GZ + 1}, \quad y = \frac{HX + IY + JZ + K}{EX + FY + GZ + 1} \tag{5}$$

where: A, B, C, D, E, F, G, H, I, J, K – transformation factors.

Relation (4) is referred to as Direct Linear Transformation (DLT) and describes the relationship between the XYZ space and its image (center projection) expressed in the pixel plane of the image. The transformation does not retain similarity (plane coordinate $z = 0$). This transformation requires a knowledge of six points in both reference systems. It is used in the development of non-metric video images.

The benefits of using this method are that knowledge of the elements of the external and internal orientation of the sensor is not required and the mathematical relationships are linear. This can produce images of unknown orientation, which can only be interpreted by means of photomaps in which points with known coordinates in the terrain system are specified (pixels in the coordinate system have known coordinates). The limitation of this transformation is that it can only be applied to flat areas ($Z = 0$).

In practice, relationship (4) is further simplified to form (Jue, 2008):

$$X = \frac{Ax + By + C}{Dx + Ey + 1}, \; Y = \frac{Fx + Gy + H}{Dx + Ey + 1} \tag{6}$$

where: A, B, C, D, E, F, G, H – transformation factors.

Projective transformation (6) expresses an analytical relationship between two planes and requires four pairs of homologous points in both coordinate systems. Transformations (5) and (6) have no angular constraints in the analytical form. In the case of image transformation using relation (6), *XZ* and *YZ* transformations on *XY* (Kaczyński & Butowtt, 2010) are not used because of the significant change in the scale and area of the image in the image associated with the inclination angle (pitch) sensor. However, in this article it was decided to use vertical to horizontal transformations in an analytical sense.

In order to use the non-metric camera for measurement purposes, a camera calibration is often performed which fixes the camera and main image point locations, as well as distortion parameters. Radial and tangential distortion coefficients are determined by different computational models. Theoretical distortion can be described by a polynomial (Kurczyński, 2014):

$$\begin{aligned} \Delta x_r &= (x - x_0) \cdot (K_1 r^2 + K_2 r^4 + K_3 r^6 + \cdots) \\ \Delta x_t &= P_1 \left(r^2 + 2(x - x_0)^2\right) + 2P_2 (x - x_0) \\ \Delta y_r &= (y - y_0) \cdot (K_1 r^2 + K_2 r^4 + K_3 r^6 + \cdots) \\ \Delta y_t &= P_2 \left(r^2 + 2(x - x_0)^2\right) + 2P_1 (x - x_0)(y - y_0) \end{aligned} \tag{7}$$

where:
Δx_r, Δy_r – radial distortion,
Δx_t, Δy_t – tangential distortion,
K_1, K_2, K_3, … – radial distortion factors,
P_1, P_2, P_3, … – tangential distortion factors,
r – radius of the determined point – distance of the point from the main point of the image.

Characteristics of the research area

Analysis of dangerous incidents between the years 1990 and 2015 in the Odra River between Ognica and the maritime boundary clearly indicates that the largest number of incidents are collisions between vessels and bridges. When the number of dangerous events involving bridges is broken down into individual bridges in the lower section of the Odra River, it follows from the analysis that the largest number are collisions with the Długi and Kolejowy bridges in Szczecin (Figure 1).

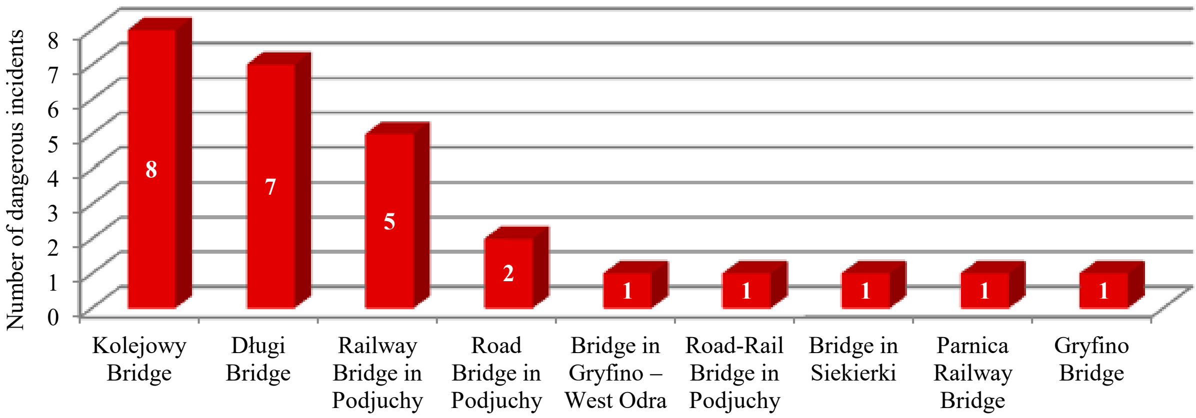

Figure 1. Number of dangerous incidents in the Lower Odra River. Years 1990–2015

Table 1. Details of bridges on the West Odra River in Szczecin

	Bridge	Kilometre of the West Odra River [km]	Width of the nautical channel [m]	Minimum clearance at highest navigable water [m]
1	Długi	35.95	17.5	3.78
2	Kolejowy	35.59	10	3.79

The research area covers the West Odra River between the Długi Bridge and the Kolejowy Bridge near the railway station in Szczecin. Details of both bridges are given in Table 1.

One CCTV camera (CCTV camera Funkwerk FAC 7864 with 1/3" progressive scan CMOS sensor, 1.3 MP. Max. res. 1280×960 px) was placed on each bridge. The visibility of the two cameras is different. The cameras are pointing in the same direction but they do not show the same area. The image of the camera shows two different areas of the West Odra river. The camera mounted on the Kolejowy Bridge shows a narrow section of the crossing between the pillars of the Długi Bridge. The camera installed on the tower of the Długi Bridge

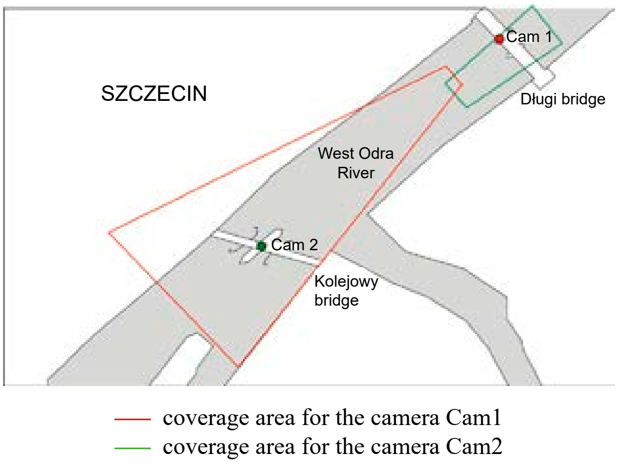

Figure 2. Visibility of CCTV cameras on the Długi Bridge and Kolejowy Bridge in Szczecin

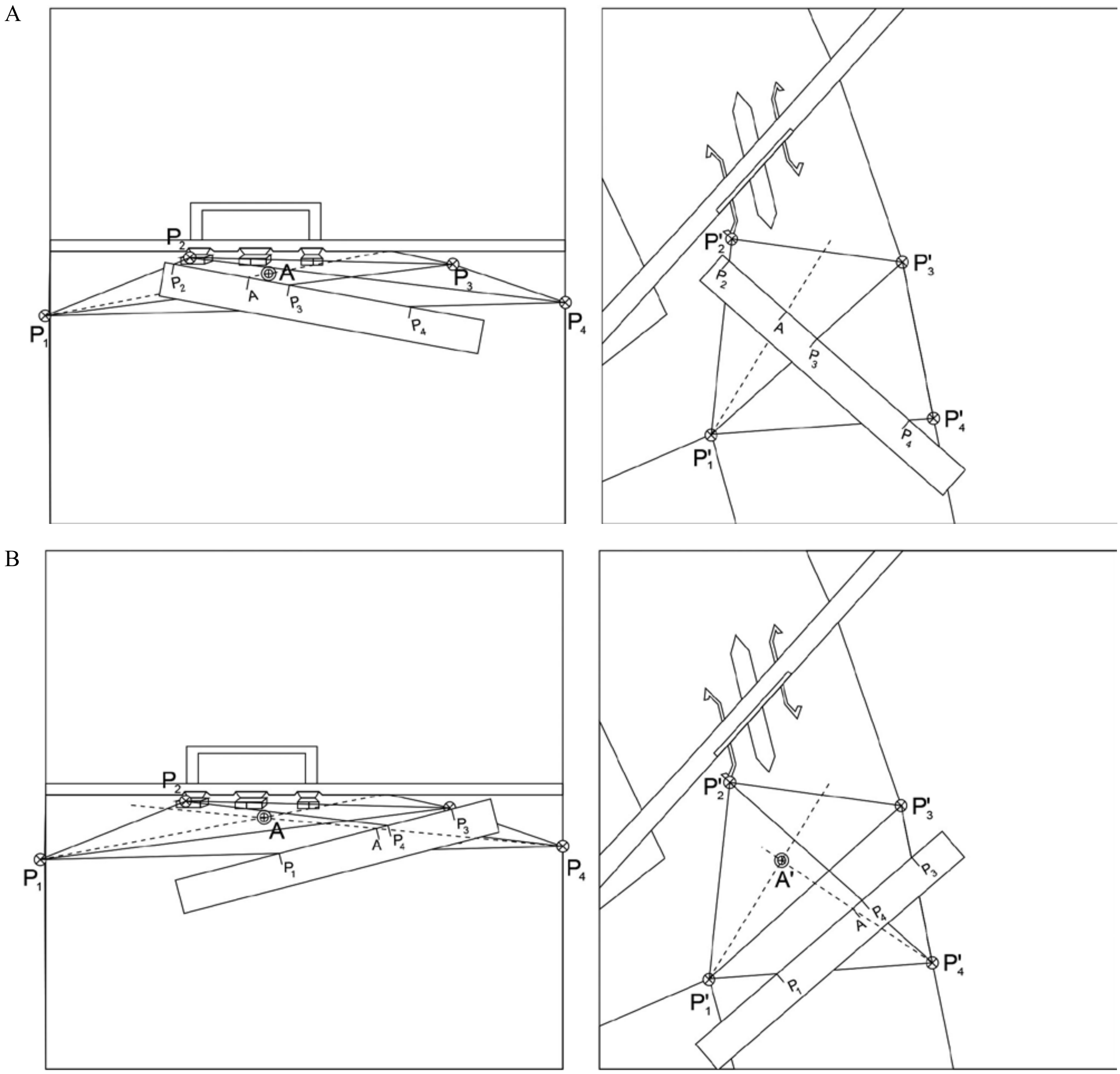

Figure 3. Two steps of the homologous rays method to determine the location of point A'

covers a much wider field of view and is directed at both crossings under the Kolejowy Bridge. The difference in the area covered between the bridges is shown in Figure 2.

Proposed methods

Analytical:

- Determining the camera orientation and local calibration (transition from tachymetry to points of rotation);
- 2D to 2D projection;
- 3D to 2D projection (DLT).

Graphical:

- homologous rays method.

The homologous rays method (paper strip method) uses the bifurcation principle of four points. The distances between four arbitrary points, lying on one line, remain unchanged in the transformation of the central projection. This method requires the following:

- the projection plane on which the processed images lies is a surface;
- the positions of a minimum of four points, which are not part of a common line and lie on the same plane in the image, are known;
- the quadrangle formed by joining these points should not have too sharp angles;
- the quadrilateral surface should be as large as possible;
- the quadrangle should be close in the camera image.

The mapping of point A (Figure 3A) for the selected view from the camera takes place in the following steps:

1. Designation of four known landmarks (P_1, P_2, P_3, P_4) on the camera image.
2. Plotting the diagonals of the resulting quadrangle.
3. Determining a straight line from point P_1 passing through point A.
4. Marking four lines from P_1 to P2, P3, P4 and A on the paper strip.
5. Mapping lines from the paper strip on the 2D image so that three of them overlap and the fourth line defines the line containing point A'.
6. Repeat steps 3 to 5 for the next selected point (e.g. P_4).

Experimental

The research experiment includes the following works:

1. The photogrammetric geodetic control network of the GNSS receiver was determined using the RTK correction system.
2. Three flights of an unmanned aerial vehicle (DJI Phantom 3 Professional) were performed.
3. Tachymetric measurements of the points identified in the CCTV camera image were made.
4. Transformations of the points measured by tachymetry into the PL-2000 rectangular state coordinate system were performed.
5. Video samples, from the non-metric camera, containing images of inland vessels passing the West Odra River were accumulated.
6. Trajectories of the passing vessels were established by moving object image analysis (Kujawski, 2015).
7. Non-metric camera calibration was performed.
8. A digital orthophotomap was developed based on photographs and measured points (Figure 4).
9. Transformation of coordinates from the camera coordinate system (pixel coordinates) to the PL2000 coordinate system was performed.
10. The positioning accuracy of an inland vessel was determined on the basis of a known trajectory (points measured with a GNSS receiver mounted on an inland vessel) and selected transformation methods.

Figure 4. Digital orthophotomap developed on the basis of UAV flights with photopoints measured using a GNSS receiver with RTK correction

Measurement of the geodetic control network (photopoints), using a GNSS receiver with RTK correction system, enabled the development of a PL-2000 digital orthophotomap. Flights were made from 30, 50 and 70 m altitude; in this way more than 400 photographs were obtained. An orthophotomap was developed in AgiSoft PhotoScan software with a 3 cm (midrange) pixel resolution in the XY plane. The accuracy of the study, verified at the control points with sub-pixel values (0.7 m), was obtained. The orthophotomap was developed according to the following steps:

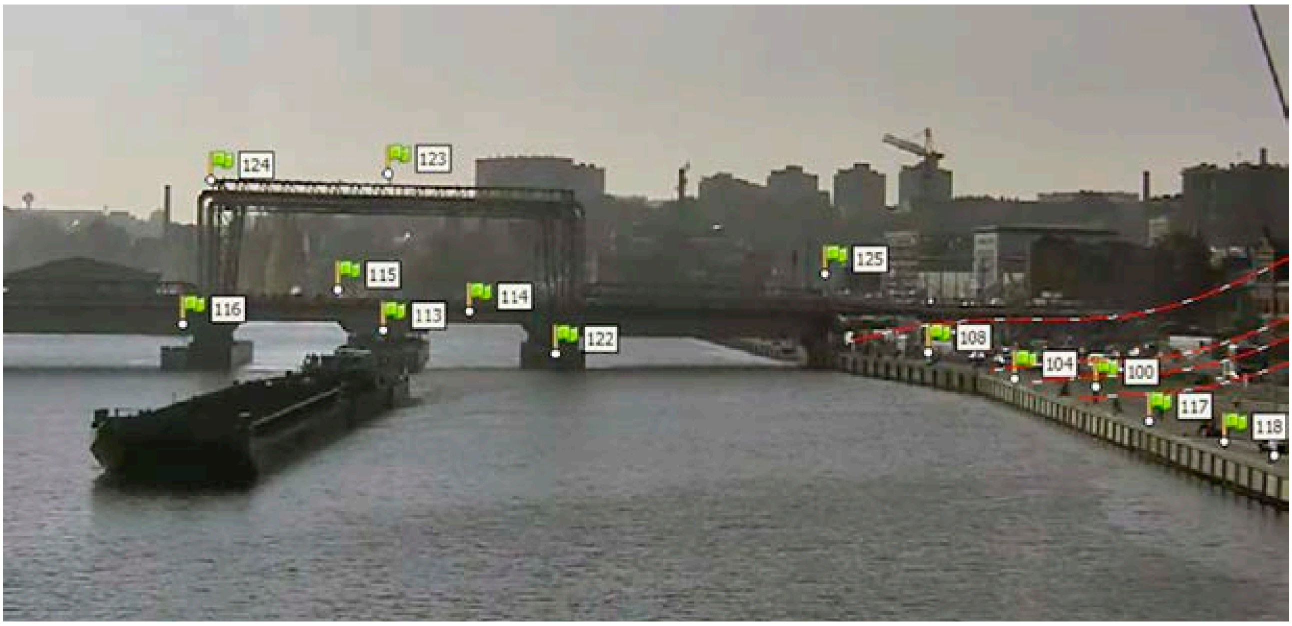

Figure 5. Calibration of the non-metric camera based on the points measured by an electronic tachymeter

1. Preparation of the project.
2. Aerial photography with self-calibration of the non-metric camera.
3. Generation of a dense cloud of points.
4. Development of a numerical area coating model.
5. Orthorectification of photos.
6. Creation of a digital orthophotomap.

The orthophotomap was used as a reference for point identification using the non-metric camera on the Długi Bridge. Thus, the identified points are assigned pixel coordinates in the local system of the camera, in addition to coordinates in the PL-2000 coordinate system. The selected points were also measured in the local system by means of an electronic tachymeter, with an accuracy of 0.01 m. These points were transformed using the TFS (4) transformation into the PL2000 system. Known points were used to calibrate the camera on the Długi Bridge and to calculate the transformation factors of the points from the camera system to the PL-2000 system in which the vessel position was determined (Figure 5).

Figure 6. Camera calibration based on incomplete data

Afterwards, the undistorted images were exported into the AgiSoft PhotoScan software, (Figure 6). It was found that the calculation of distortion parameters (6) based on incomplete data results in deformation of the image instead of improvement. Greater deformation was confirmed by a tenfold decrease in the accuracy of the transformation relative to points from the original image. The raw images were used for further work.

The next step was to transform the image using the homologous rays burst method and projective and DLT transformations. The homologous rays burst method (also known as the paper strip method) was difficult to perform and did not produce satisfactory results because of an unfavorable distribution of transformation points. It was impossible to establish points lying outside the main quadrangle. The DLT transformation also proved to be unworkable; no solution to equation (5) was found. For vertically positioned cameras, it is assumed that the approximate height values are equal to the horizontal distance to the orthophotomap. Only the projective transformation was usable due to its analytical form (6). For this transformation, points approximately forming the plane were used, mainly consisting of geometric means arranged on the margins of the poles. A graphical approach to the projective transformation did not produce satisfactory results and caused a large deformation of the camera image.

The next step was to compare the received coordinates, resulting in a new trajectory of the inland vessel, with the satellite-based trajectory of the satellite receiver. A detailed discussion and presentation of these results are presented in the next chapter.

Discussion

The coordinates of known points that formed the trajectory of the inland vessel were used to evaluate the transformations carried out and at the same time constituted control points. According to the law of moving Gaussian mean errors (Hausbrandt, 1971) the error of established control points was determined as the average error of the function:

$$m_{pk} = \sqrt{m_p^2 + m_i^2} \qquad (8)$$

where:

- m_{pk} – total position error of the control point,
- m_p – position error of the control point resulting from the satellite receiver,
- m_i – identification error of the satellite receiver as a result of image analysis.

A m_{pk} error of 2 m was assumed based on specifics of the GNSS receiver. Due to the ambiguous position of the GNSS receiver in the CCTV image, the m_i error was taken as 2–3 pixels. The average error in the changing distance of the vessel was assumed to be 1 m. For such values, the m_{pk} total position error of the control point was found to be ± 2.2 m.

Control points were divided into four areas. The main area was in the section of the plane determined by the points used for the transformation. The areas, along with the designated trajectories, are shown in Figure 7. The transformation of points in individual areas is presented in Table 2.

Table 2. Accuracy of coordinate transformation in individual areas

No.	Number of resolved points	Average control point error (standard deviation) [m]
I	22	1.6
II	9	11.6
III	13	11.8
IV	19	6.8

Analysis of the data from Table 2 and Figure 7 clearly indicates that the accuracy of area I is within the limits of the control points, while in other areas it is significantly exceeded. In area I, high accuracy is due to the fact that the control points are located within the area of the transformation (determined plane). In areas II–IV, we are dealing with extrapolation and a significant decrease in accuracy. For Area IV, the accuracy achieved is better than in areas II and III, which can be explained by a higher resolution of the terrain of the image (one pixel represents a smaller area), which makes the image content easier to interpret and identify. Area II could be expected to be more accurate; however, the inland vessel made a sudden turn in this area which caused a sudden twist and difficulty in correctly indicating the location of the satellite receiver. The method used in the area of the transformed plane gave results at the accuracy level of the satellite receiver mounted on the vessel. This confirms the hypothesis that positioning with a non-metric camera is possible with the accuracy provided by GNSS satellite navigation

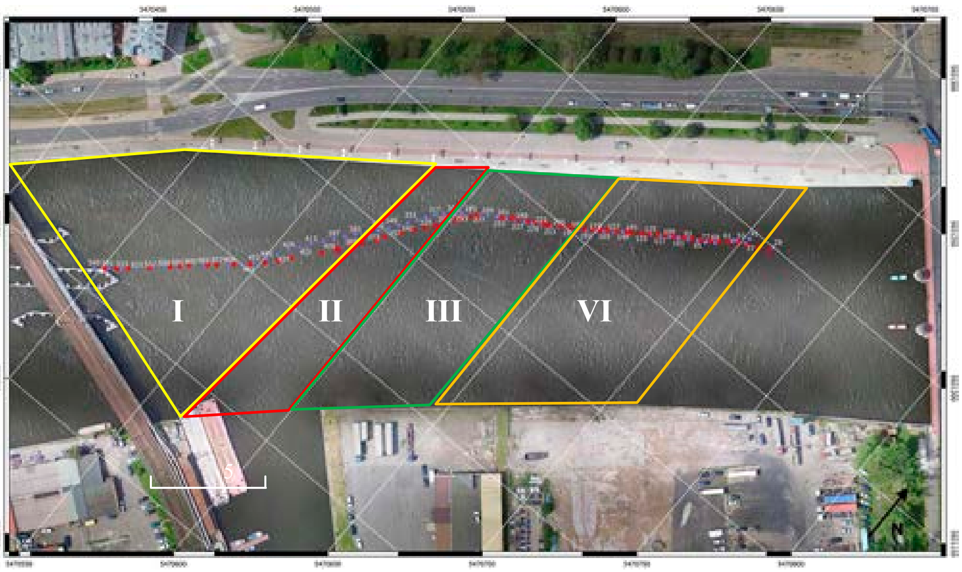

Figure 7. Digital orthophotomap with established trajectories: from the GNSS receiver (blue) and from the CCTV camera (red)

receivers (without RTK correction) at the correct positioning of the transformation points. At the same time, an examination of the different methods and variants of the transformation has helped to achieve the goal of this work.

Conclusions

This article proposes a solution to the problem of tracking moving objects using a single non-metric CCTV camera. The affirmed hypothesis was that identifying the position of a vessel on the West Odra River in Szczecin using a single non-metric camera (mounted on the Długi Bridge) is possible within the accuracy of traditional positioning systems. The following specific conclusions can be drawn:

- calibration of the camera based on incomplete (fragmentary) data has reduced accuracy and causes deformation of the images, instead of improving accuracy and improving the cartographic image. Performing a full-frame transformation from the entire image range is recommended, as well as further camera calibration using multiple vessel trajectories within the camera's tracking area;
- the analytical projection transformation of the *xz* plane onto the *xy* plane is possible and produces good results in the case of the proposed algorithm. In the well-known photogrammetric literature (Kaczyński & Butowtt, 2010) it is suggested not to perform such transformations. It was found that the transformation of a vertical plane to a horizontal one, in the case of an image, yields no results, as opposed to analytical transformations;
- projective transformation gives satisfactory results within the accuracy of GNSS receiver localization (without real-time RTK correction), including error identification during tracking of the receiver on the images, this applies to the area within the plane of the photopic array, despite their extremely unfavorable distribution (they formed a narrow section of the plane);
- extrapolation of planes II to IV produces a five-fold decrease in accuracy, despite the fact that the camera is a short distance away, this is related to an increase in the distance from the photopoints and the plane used for the transformation;
- in the case of a vessel moving away from the camera, no drop in accuracy was observed, despite the fact that the spatial resolution decreased (the pixel size increased);
- the DLT transformation was impossible to calculate, for a vertically positioned camera, an attempt was made to determine the altitude as a distance (the z axis became horizontal), but this did not allow for the determination of transform coefficients (no solution was found);
- the graphical transformation method based on homologous points (the so-called paper strip method) failed because of the extremely unfavorable homologous system (most of the imaging area was flowing water where markers cannot be identified);
- in order to increase the area of high accuracy, it is suggested that photo points be distributed in the whole area of interest (as large as possible), e.g. by deploying a photogrammetric survey buoy, or pointing the CCTV camera into the wharf on the left side of the Odra River;
- integration of photogrammetric methods (UAV flight resulting in an orthophotomap) with geodetic methods (tachymetric and GNSS measurements) and their corresponding processing through the proposed algorithm allows for the positioning of the object using the non-metric camera within the accuracy of the traditional measurements;
- the proposed algorithm enables the detection of the sudden degradation of the GNSS system mounted on a moving object;
- the authors expect that the accuracy achieved will also be possible in the case of the use of orthophotomaps with a 0.1 m spatial resolution, available on the national geoportal (www geoportal.gov.pl);
- the algorithm proposed in this paper has been proven and the proposed positioning method provides an alternative, or complimentary, positioning system with an accuracy comparable to that of traditional positioning systems.

Acknowledgments

This research outcome has been achieved under the grant No. 8/S/IIT/2017 financed by a subsidy from the Ministry of Science and Higher Education for statutory activities.

References

1. Baranowski, L. (2013) Equations of Motion of a Spin-Stabilized Projectile for Flight Stability Testing. *Journal of Theoretical and Applied Mechanics* 51 (1), pp. 235–246.
2. Hausbrandt, S. (1971) *Rachunek wyrównawczy i obliczenia geodezyjne*. Warszawa.
3. Jue, L. (2008) Research on close-range photogrammetry with big rotation angle. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII (Part B6b), pp. 11–14.

4. Kaczyński, R. & Butowtt, J. (2010) *Fotogrametria*. Warszawa.
5. Kujawski, A. (2015) Inland waterway vessels tracking using Closed Circuit Television Video. *Scientific Journals of the Maritime University of Szczecin, Zeszyty Naukowe Akademii Morskiej w Szczecinie* 44 (116), pp. 135–140.
6. Kurczyński, Z. (2014) *Fotogrametria*. Warszawa: Wydawnictwo Naukowe PWN.
7. Piątek, Z. (2011) *Nawigacja morska w pytaniach i odpowiedziach*. Warszawa: Almapress.
8. Stępień, G., Zalas, E. & Ziębka, T. (2017) New approach to isometric transformations in oblique local coordinate systems of reference. *Geodesy and Cartography* doi: 10.1515/geocart-2017-0017.
9. Zalas, E., Sanecki, J., Klewski, A. & Stępień, G. (2016) Determining the spatial orientation of remote sensing sensors on the basis of incomplete coordinate systems. *Scientific Journals of the Maritime University of Szczecin, Zeszyty Naukowe Akademii Morskiej w Szczecinie* 45 (117), pp. 29–33.