VILNIUS TECH
Vilnius Gediminas Technical University

Geodesy and Cartography
ISSN 2029-6991 / eISSN 2029-7009
2023 Volume 49 Issue 1: 25–36
https://doi.org/10.3846/gac.2023.16529

UDC 528.711.114

IMPLEMENTATION OF THE SMARTPHONE CAMERA IN THE MEASURES OF NARROW STREET FACADES

Mohammed ALDELGAWY*

Civil Engineering Department, Faculty of Engineering, Fayoum University, Fayoum, Egypt

Received 15 February 2022; accepted 17 February 2023

Abstract. This paper aims to perform metric measurements of narrow street façades using single image captured by smartphone's camera. Since tight area accompanied by narrow street limits object to camera distance, object lines perpendicular to façade do not appear in image and consequently their vanishing point (VP) is hard to detect. Accordingly, semi-automated MATLAB® application was designed depending only on two orthogonal VPs. Novelty of work comes from using smartphone as a cost and time efficient tool for measurements, depending only on two VPs, and applying image line refinement approach exploiting detected VPs. Three single images were captured by three different smartphones. Then, undistorted single images were formed after calibrating cameras. Image lines for horizontal and vertical object lines were extracted semi-automatically. Two VPs were detected applying two models: Model-I solves for vanishing points' Cartesian coordinates, whereas Model-II solves for angle coordinate peaks of histogram. Image line refinement approach was applied before applying cross-ratio using one horizontal and one vertical reference lines to calculate object lengths of 46 check lines (horizontal and vertical). Proposed models provided reliable and comparable results. Applying line refinement approach improved solution with best overall accuracy of 0.010 m and 0.011 m for Model-I and Model-II, respectively.

Keywords: single image, metric measurements, façade observation, vanishing point, cross-ratio, smartphone's camera, semi-automated line extraction.

Introduction

The task of metric measurements for the purpose of façade observation is one of the highly demanded tasks that is frequently required by surveyors. For narrow streets lying in crowded areas, using traditional land surveying instruments such as total station or theodolite is not the best choice. Besides the land surveying is a time-consuming option, the narrow width of street confines the angle of view of instrument's telescope and thereby prevents the instrument from observing façade points with high altitudes. In addition, the streets that are crowded of people impedes the setting-up of land surveying instrument for long time. In contrary, using the smartphone's camera for façade observation delivers an adequate tool to get over the difficulties accompanied by the ordinary land surveying techniques. The smartphone's camera is free to rotate such that the high-elevated points are possible to be captured and the time taken in site is only the seconds required for image to be captured. Moreover, the use of smartphones, which are currently available at high resolution and affordable prices, will result in high accuracy and low cost of the observation. Additionally, utilizing smartphone' camera facilitates the image capturing process as there is no need for special expensive metric camera nor professional user to take the image. Consequently, the overall time required for façade observation is shortened.

In this paper, we present an adapted semi-automated metric measurements technique for observing façades of buildings located in narrow streets using smartphone's camera. The proposed techniques should be faster and more convenience than the traditional land surveying one. The main goal of the current research is to perform metric measurements from single image captured by smartphone's main rear camera. The proposed solution is based on detecting only two (not three) orthogonal vanishing points (VPs) on an undistorted single image. A MATLAB® application was developed for gathering metric information from the single image. First, the camera is calibrated and the captured single image is undistorted before applying a semi-automated approach for raw image lines extraction. Then, the two orthogonal vanishing points are robustly detected. Furthermore, a refinement approach based on the robustly detected two orthogonal vanishing

*Corresponding author. E-mail: *mas00@fayoum.edu.eg*

points is applied to the extracted raw image lines. Each raw image line is adapted by modifying its endpoint that is nearest to vanishing point to be lying along the line between the farthest endpoint and vanishing point keeping the image line's length the same. Meanwhile, the user is simply required to measure the width and height of an accessible rectangular window on the façade using a measuring tape. The measured width is exploited as the reference length for horizontal object lines, and measured height is exploited as the reference length vertical object lines. The cross-ratio theory is applied for calculating the object length of any horizontal or vertical line belonging to the façade utilizing the horizontal and vertical reference object lines, respectively. Three smartphones' cameras of different brand-names (iPhone_7_Plus, Samsung Galaxy Note8, and HUAWEI_nova_7i) were utilized for data analysis and comparison.

Dissimilar to most of published researches that rely on three orthogonal vanishing points detection such as Li et al. (2010), Vouzounaras et al. (2014), Arslan (2018), the current research relies only on two orthogonal vanishing points detection. These two vanishing points are the vanishing points for horizontal and vertical object lines belonging to the façade in question. Usually, lines perpendicular to building's façade usually do not appear in images captured in narrow streets and, by consequence, the vanishing point related to these perpendicular lines is hard to be assigned. This is caused by one of two reasons. The first reason is that the depth of building (plane perpendicular to façade) is occluded by the other buildings that are next to the building in question. The other reason is that the short object to camera distance (resulting from the small width of narrow street) does not supply a sufficient space for the camera to be moved away enough from the building in purpose to capture such perpendicular lines. In this accordance, the presented research depends only on the two orthogonal vanishing points for the horizontal and vertical object lines appearing in the façade plane.

1. Related work

Many papers covered the topic of metric measurements from single image including the applications for cross-ratio and vanishing points. One of the early research projects was introduced by Debevec et al. (1996), where a methodology to render and model architectural views using photos was developed combining geometry and image-based techniques. The cross-ratio variances assuming normal distribution case were formulated in Liu and Chuang (2002). A method to identify cross-ratios with minimum error variances was introduced relying on geometric error analysis. Simulated results showed that the presented approach led to estimate the minimum error variances better than other methods. Lee and Nevatia (2003) integrated various sources including aerial and ground scene photos for performing hierarchical representation of 3D buildings located in urban areas. The results showed that the proposed technique was an efficient and low-cost one. An approach based on cross-ratio invariability for perspective projection was presented by Zhang et al. (2003) for calibrating camera radial distortion. Results obtained by both simulated real data emphasized the simplicity along with efficiency of the proposed approach. In Aguilera et al. (2005), the authors introduced an approach for vanishing points detection taking into account the presence of mini-segments resulted from automatic line extraction which leads to unfavorable intersection cases. The introduced approach consists of three main steps: clustering mini-segments, estimation of vanishing points, and computation applying a re-weighted least-squares solution (L.S.S.) support by an M-estimator. A technique for computer vision systems' localization error analysis exploiting cross-ratios was proposed by Chuang et al. (2007). An applied method for selecting images' point features was developed by authors. An approach for extracting geometric information from uncalibrated single image was introduced in Wang et al. (2010). Cross-ratio theory was applied to calculate the line segments' length on planar surface knowing reference lines' length. Both indoor and outdoor images were utilized for verification. In Li et al. (2010), the authors presented a method to estimate the camera focal length along with three orthogonal vanishing points simultaneously from single images depending on image polar coordinate system. Compared to previous methods, the proposed method required less running time and computer memory. A study to directly extract vehicle speeds from video footages utilizing cross-ratio theory was introduced by Wong et al. (2014). In addition, a control study using a calibrated Doppler radar was performed for validation by the authors. A fully automated approach for 3D models creation was introduced by Vouzounaras et al. (2014). The proposed approach combined the extracted image line segments for the identification of vanishing points, image orientation, indoor planes, and outdoor planes. An invariant on the projective plane and its use for reducing the complexity of projective transformations was proposed by Erdnüß (2018). The camera centers and 3D structure of scene were detected through the proposed approach. Arslan (2018) presented a study for extracting metric data from single image. Two methods were examined by the author. The first method relied on cross-ratio, while the second method relied on robust camera calibration and statistical estimators. The results revealed that the second method provided higher accuracy than the first one.

2. Coordinate systems

Both object and image coordinate systems are defined in this section. In Figure 1, the object space axes are the (X,Y,Z) axes perpendicular to façade, parallel to façade width, parallel to façade height, respectively. Additionally, the image coordinate system axes are denoted as (u,v). The upper left corner of image represents the image coordinate system's origin. The vanishing point for image lines corresponding to horizontal object lines (lines parallel

to Y-Axis) is denoted as (v_y). Whereas, the vanishing point for image lines corresponding to vertical object lines (lines parallel to Z-Axis) is denoted as (v_z). The principal point is denoted as (pp).

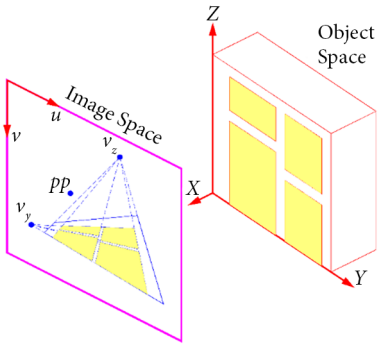


Figure 1. Object coordinate system vs. image coordinate system

3. Research methodology

The smartphone's camera is first calibrated in order to create the image undistortion array, then metric measurements are done on undistorted single image. Two MATLAB® applications were developed: one application for camera calibration and the other one for metric measurements. A summary for the research methodology is given in the flowchart illustrated in Figure 2.

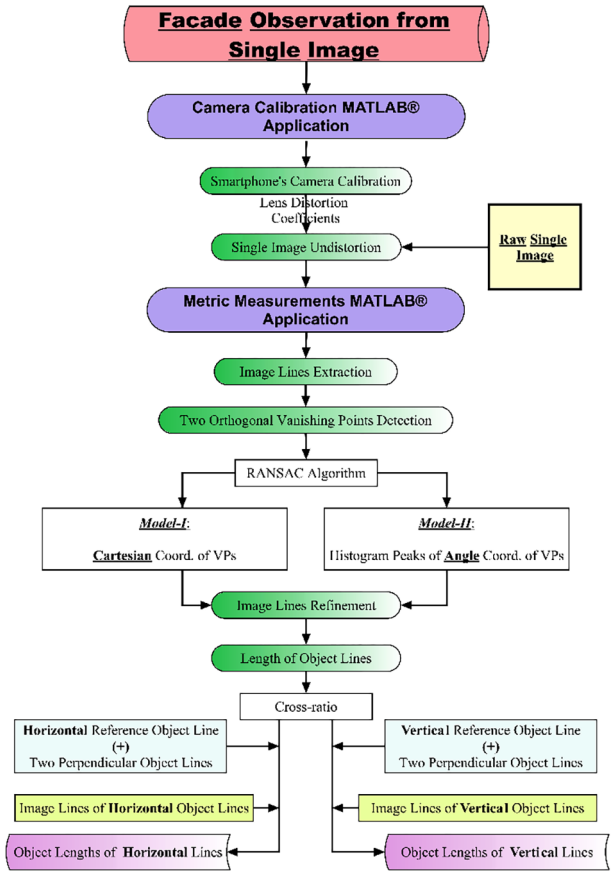


Figure 2. Flowchart for research methodology

3.1. Camera calibration MATLAB® application

A camera calibration MATLAB® application was created for lens distortion coefficients calculation and image undistortion array generation. A screenshot for the application is given in Figure 3 illustrating the different functions of application.

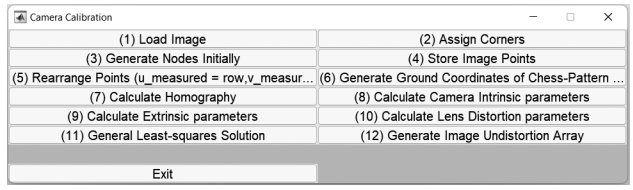


Figure 3. MATLAB® application created for camera calibration

3.1.1. Smartphone's camera calibration

The smartphone's camera is calibrated applying Zhang's technique for camera calibration using chessboard pattern (Zhang, 2000; Burger, 2019; Aldelgawy & Abu-Qasmieh, 2021a, 2021b). A number of images changing the angles of view is captured for a well-calibrated chessboard pattern. Then, the image nodes of chessboard pattern are used to establish the calibration model. The pixel image point coordinates (u, v) in terms of metric ground point coordinates (X, Y, Z) are given as follows (Zhang, 2000; Burger, 2019):

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \lambda A \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \tag{1}$$

with (λ) is scale factor, $\begin{bmatrix} R & T \end{bmatrix}$ is extrinsic parameters matrix, $R = \begin{bmatrix} R_1 & R_2 & R_3 \end{bmatrix}$ is rotation matrix, and $T = -R\begin{bmatrix} X_O & Y_O & Z_O \end{bmatrix}^T$ translation vector. The vector (X_O, Y_O, Z_O) represents the metric ground coordinates vector of camera perspective center. In addition, the pixel camera intrinsic parameters matrix (A) is given as follows:

$$A = \begin{bmatrix} \alpha & \gamma & u_o \\ 0 & \beta & v_o \\ 0 & 0 & 1 \end{bmatrix},$$

where (u_o, v_o) are the pixel image coordinates of (pp), (α, β) are the image axes scale factors, and (γ) denotes the skewness coefficient.

For the chess pattern (2D case), putting $(Z = 0)$ in equation (1) leads to:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \tag{2}$$

with $\left(H = \lambda A \begin{bmatrix} R_1 & R_2 & T \end{bmatrix}\right)$ represents homography matrix.

Then, Equation (2) is solved for (H) prior to calculating both intrinsic and extrinsic parameters.

3.1.2. Single image undistortion

The lens distortion effect is considered this sub-section. Based on previous researches such as Zhang (2000), Burger (2019), Aldelgawy and Abu-Qasmieh (2021a), only the radial lens distortion is taken into account through estimating the three radial lens distortion coefficients (k_1,k_2,k_3). Accordingly, the image undistortion array is generated. Knowing the point ground coordinates of chessboard node (X,Y), the normalized (ideal) image coordinates (x,y) are given as follows:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} R_1 & R_2 & T \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \quad (3)$$

Then, the undistorted (ideal) pixel image coordinates (u,v) are:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (4)$$

The relationship between the observed (distorted) pixel image coordinates $(\tilde{u},\tilde{v})$ and undistorted ones (u,v) is given as follows (Burger & Burge, 2016; Burger, 2019; Wang et al., 2006):

$$\tilde{u} = u + \Delta_{ur},$$
$$\tilde{v} = v + \Delta_{vr}, \quad (5)$$

where $(\Delta_{ur},\Delta_{vr})$ are the radial lens distortion components in (u,v) directions, respectively that are calculated as follows:

$$\Delta_{ur} = u(k_1.(x^2+y^2) + k_2.(x^2+y^2)^2 + k_3.(x^2+y^2)^3),$$
$$\Delta_{vr} = v(k_1.(x^2+y^2) + k_2.(x^2+y^2)^2 + k_3.(x^2+y^2)^3). \quad (6)$$

Then, given the original image array (I), the undistorted image array (I') is:

$$I'(u,v) = I(\tilde{u},\tilde{v}). \quad (7)$$

3.2. Metric measurements MATLAB® application

An application was developed by MATLAB® for metric measurements (Figure 4). The developed application relies basically on extracting image lines semi-automatically, then applying a refinement approach on the extracted lines based on the firmly detected two orthogonal vanishing points. The procedure used for the developed application is summarized in the following lines.

3.2.1. Image lines extraction

Image lines are extracted from the undistorted single image. First, image noise is reduced applying a median filter. Then, the raw image line segments are extracted. The developed application is provided with the "Group of Lines" extraction option which allows the user to draw a window on the desired zone in image and manually input the maximum number of raw line segments that are required to be extracted inside the window. Furthermore, the application is provided with the "Line-by-Line" line extraction option which enables the user to pick up each single raw line segment separately by drawing a line close to the demanded line segment. Then, the raw line segment in question is automatically extracted within a selected buffer width (Figure 5a). This option was the one utilized in the proposed research since it allows the line segments to be extracted more precisely. The "Canny" edge detector is applied for edge detection (Canny, 1986; Rong et al., 2014) and, thereby, Hough transform approach (Hough, 1962; Di et al., 2003; Hassanein et al., 2015) is applied for raw line segment extraction (Figure 5b). Finally, the image lines are extracted by detecting their endpoints. Each pair of intersected raw line segments are extended (or trimmed) to their point of intersection simply by a mouse-click on them. Thus, the endpoints of image lines are detected and the image lines are extracted (Figure 5c). This technique for image line extraction helps to reduce the runtime as it allows to exactly extract the desired line segment and ignore the non-desired ones. In addition, the used technique improves the accuracy of line extraction since the user is able to select specific lines in the image.

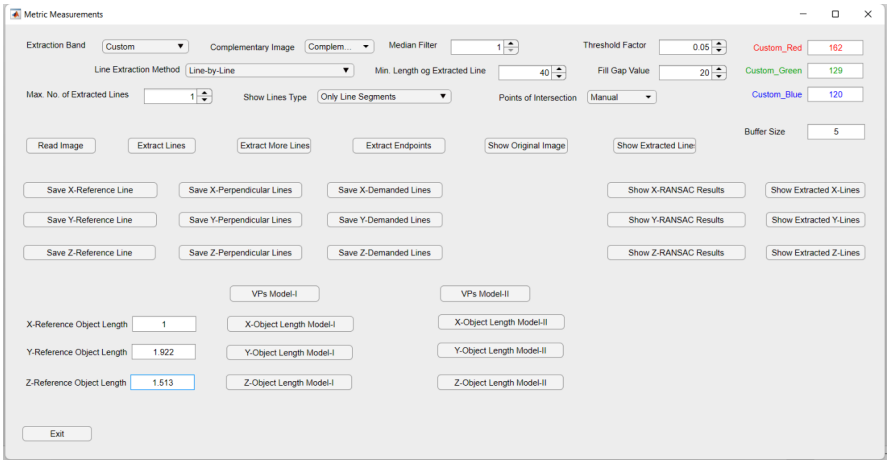


Figure 4. Application developed by MATLAB® for metric measurements

a) b) c)

Figure 5. Samples for extracted image lines:
a – line selection buffer; b – raw line segments; c – image lines extended to their endpoints

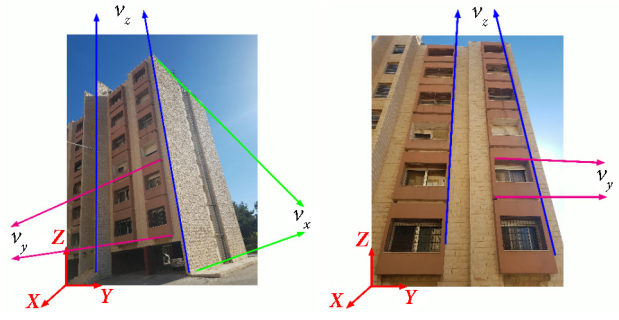
3.2.2. Two orthogonal vanishing points detection

The step of locating vanishing points is the most critical step in the single image metric measurements process (Arslan, 2018). As prementioned, the three orthogonal vanishing points are hard to be detected since the lines perpendicular to building's façade usually do not appear in images that are captured in narrow streets (Figure 6). Accordingly, the solution presented in this research relies on detecting only two orthogonal vanishing points. The two orthogonal vanishing points (v_y) and (v_z) for horizontal and vertical object lines, respectively are detected. Given the image line parameters (r,θ), then the Cartesian coordinates (u_v,v_v) for any of the vanishing points are:

$$(u_v - u_o)\mathrm{Cos}\,\theta + (v_v - v_o)\mathrm{Sin}\,\theta = r, \tag{8}$$

where (r) is the perpendicular distance from principal point (pp) to image line, and (θ) denotes the angle included by the line perpendicular to image line and the u-axis.

First, the outliers are excluded by applying a RANdom SAmple Consensus (RANSAC) approach (Fischler & Bolles, 1981; Oh & Jung, 2012; Wildenauer & Hanbury, 2012) on image lines corresponding to horizontal and vertical object lines. These outliers could be the object lines on the façade that are not perfectly horizonal or vertical in nature, or they could be the image lines that are not correctly extracted due to image noise. Afterwards, two models (Model-I and Model-II) are used for vanishing points detection. Coordinates of (VPs) detected by RANSAC approach are considered as initial values for detecting (VPs). In Model-I we depend on the Cartesian coordinates of vanishing points, while in Model-II we depend on the peaks of histogram of angle coordinate and, by consequence, the polar coordinates. A detailed explanation for both models is given below.



a) b)

Figure 6. Vanishing points detection: a – three orthogonal vanishing points; b – two orthogonal vanishing points

Model-I:

This model is mainly based on calculating the Cartesian coordinates (u_{vy},v_{vy}) and (u_{vz},v_{vz}) of the two orthogonal vanishing points (v_y) and (v_z). Equation is the equation that relates the image line parameters and vanishing point's coordinates. Given (n_y) image lines corresponding to horizontal object lines and (n_z) image lines corresponding to vertical object lines, we solve for the four unknowns $(u_{vy},v_{vy},u_{vz},v_{vz})$ using the following minimization equations:

$$\sum_{i=1}^{n_y}\left\|(u_{vy} - u_o)\mathrm{Cos}\,\theta_i + (v_{vy} - v_o)\mathrm{Sin}\,\theta_i - r_i\right\|^2, \tag{9}$$

$$\sum_{j=1}^{n_z}\left\|(u_{vz} - u_o)\mathrm{Cos}\,\theta_j + (v_{vz} - v_o)\mathrm{Sin}\,\theta_j - r_j\right\|^2. \tag{10}$$

Additionally, we have one constraint equation between Cartesian coordinates of the two (VPs) according to Li et al. (2010), Caprile and Torre (1990) as follows:

$$(u_{vy} - u_o)(u_{vz} - u_o) + (v_{vy} - v_o)(v_{vz} - v_o) + \alpha^2 = 0. \tag{11}$$

A constrained least-squares solution (L.S.S.) (Chapter 9 of (Mikhail & Ackermann, 1982)) is applied for Equations (9) to (10) with the constraint Equation (11).

Model-II:

This model is mainly based on calculating the peaks of histogram of angle coordinates of the two orthogonal vanishing points (v_y) and (v_z). We first solve for the angle coordinate (ϕ_v) of vanishing point, then solve for the radial coordinate (ρ_v). As a consequence, the coordinates of vanishing points are represented in the polar form (ρ_v,ϕ_v) in place of Cartesian form (u_v,v_v) with coordinates (ρ_v,ϕ_v) are:

$$\rho_v = \sqrt{u_v^2 + v_v^2}, \tag{12}$$

$$\phi_v = \mathrm{Tan}^{-1}(v_v / u_v). \tag{13}$$

In order to solve for the angle coordinates (ϕ_{vy},ϕ_{vz}) of the two orthogonal vanishing points, distribution of the

intersection of bundle of convergent lines represented by each vanishing point obeys elliptical Gaussian distribution. In this elliptical Gaussian distribution, the length of minor axis is very short with respect to the major one. Moreover, the angle included between the ellipse's major axis and the line from the pp to VP (radial coordinate of VP) is very small (Li et al., 2010). Accordingly, the angle coordinate of each of the two orthogonal VPs is obtained by looking for the peak of histogram of angle coordinates of the intersection points of image lines corresponding to horizontal and vertical object lines.

Then, we solve for the radial coordinates (ρ_{vy},ρ_{vz}) of the two orthogonal vanishing points. The polar form of Equation (8) is:

$$\rho_v.(\mathrm{Cos}(\phi_v-\theta))=r\,. \tag{14}$$

Given (n_y) and (n_z) image lines, we have the following minimization equations to solve for the two unknowns (ρ_{vy},ρ_{vz}):

$$\sum_{i=1}^{n_y}\left\|\rho_{vy}.(\mathrm{Cos}(\phi_{vy}-\theta_i))-r_i\right\|^2; \tag{15}$$

$$\sum_{j=1}^{n_z}\left\|\rho_{vz}.(\mathrm{Cos}(\phi_{vz}-\theta_j))-r_j\right\|^2. \tag{16}$$

Furthermore, we have one constraint equation between the polar coordinates of the two vanishing points adopted from Equation (11) as follows:

$$\rho_{vy}.\rho_{vz}.(\mathrm{Cos}(\phi_{vy}-\phi_{vz})+\alpha^2=0\,. \tag{17}$$

A constrained least-squares solution (Chapter 9 of Mikhail and Ackermann (1982)) is applied for Equations (15) to (16) with the constraint Equation (17).

Then, Cartesian coordinates of the two orthogonal VPs are given by:

$$u_{vy}=\rho_{vy}.\mathrm{Cos}\,\phi_{vy}\,; \tag{18}$$

$$v_{vy}=\rho_{vy}.\mathrm{Sin}\,\phi_{vy}\,; \tag{19}$$

$$u_{vz}=\rho_{vz}.\mathrm{Cos}\,\phi_{vz}\,; \tag{20}$$

$$v_{vz}=\rho_{vz}.\mathrm{Sin}\,\phi_{vz}\,. \tag{21}$$

3.2.3. Image line refinement

A new approach for refining the extracted raw image lines is applied in this sub-section. The previously detected two orthogonal vanishing points are utilized for the refinement process. In other words, the extracting raw image lines are to be refined in accordance with the detected vanishing point coordinates. For each of the raw image lines of horizontal or vertical object lines, the raw image line is manipulated by modifying the position of its endpoint nearest to the vanishing point. A line is extended from the farthest endpoint of raw image line to the related vanishing point. Then, the nearest endpoint is moved to lie on the extended line such that the length of the refined image line is equal to that of the extracted raw one. Such manipulation should enhance the accuracy resulted by the metric measurement process. In Figure 7, the line (*1-2*) is the extracted raw image line, the line (*1-2′*) is the refined image line, and the point (*VP*) is the vanishing point. Applying the image line refinement approach resulted in modifying the endpoint of the extracted raw line that is nearest to the vanishing point (endpoint (*2*)) to the refined endpoint (*2′*). Consequently, the extracted raw image line (*1-2*) was modified to the refined image line (*1-2′*).

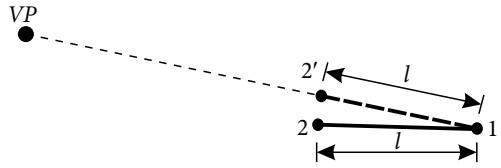


Figure 7. Image line refinement

3.2.4. Length of object lines

The length of any horizontal or vertical object line is calculated applying cross-ratio theory on the refined image lines. The following is the methodology for calculation length of any vertical object line. For any horizontal object line, the object length can be calculated likewise. In Figure 8, given the vanishing point for vertical object lines (v_z), object length of vertical reference line (*AB*), image line of vertical reference line (*ab*), and two image lines (*ae, bf*) of object lines perpendicular to reference line in the object space at its two endpoints (*A, B*). Then, the length of demanded object line (*CD*) is calculated as follows. First, we extend the two perpendicular image lines (*ae, bf*) till they intersect the image line in question (*cd*) at points (*a′, b′*), respectively. The object length (*A′B′*) is equal to object length (*AB*) since both object lines (*AB*) and (*CD*) are parallel (vertical) and the two object lines (*AE, BF*) are perpendicular to line (*AB*). Then, the cross-ratio condition is applied for points (C,A',B',V_Z) (Liu & Chuang, 2002; Arslan, 2018):

$$\frac{cb'.a'v_z}{cv_z.a'b'}=\frac{CB.AV_Z}{CV_Z.AB},$$

with (V_Z) is the vanishing point for vertical lines in the object space. Then,

$$CB=\frac{cb'.a'v_z}{cv_z.a'b'}.AB. \tag{22}$$

Similarly, applying cross-ratio condition for points (D,A',B',V_Z), we get:

$$DB=\frac{db'.a'v_z}{dv_z.a'b'}.AB. \tag{23}$$

Then, $CD=\left|DB-CB\right|$, or:

$$CD=\frac{a'v}{a'b'}.\left|\frac{db'}{dv_z}-\frac{cb'}{cv_z}\right|.AB\,. \tag{24}$$

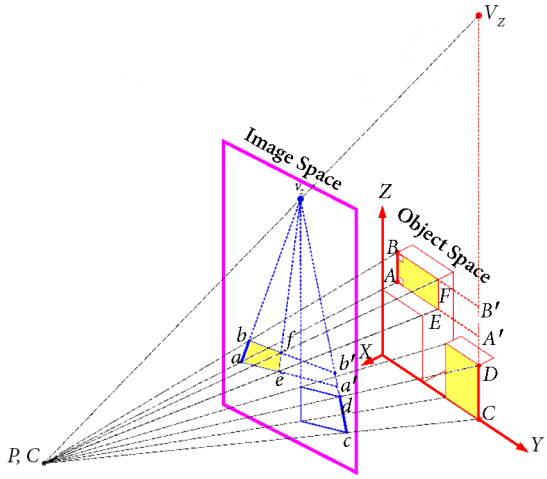


Figure 8. Length calculation of object line using cross-ratio

a) b)

Figure 9. Image lines: a – image lines for horizontal object lines; b – image lines for vertical object lines

4. Experimental work

The dataset was created using the main rear cameras of three smartphones of different brand-names that are denoted as follows. Camera-I: iPhone_7_Plus, Camera-II: Samsung Galaxy Note8, and Camera-III: HUAWEI_nova_7i. The image resolution of each of the three cameras was 12 mega pixels. Twenty images for chessboard pattern of 30×30 mm size were taken by each smartphone's camera for the purpose of camera calibration. Thereafter, a single image for the façade under study was captured by each camera and then undistorted. The cameras distance to facade was about 8 m in order to simulate the real case of narrow side street. The building's façade under study was about 9 m in width by 24 m in height. A recently calibrated SOKKIA SET05N total station was used for results validation. The distance measuring accuracy for the total station instrument was ±(2 mm + 2 ppm), whereas the angle measuring accuracy was 5". One horizontal and one vertical reference object lines were used for metric measurements along with image lines of two object lines perpendicular to each of reference object lines at their endpoints. In addition, 23 horizontal check lines with object lengths of 1.922 to 2.818 m besides 23 vertical check lines with object lengths of 1.513 to 3.194 m were used for results verification. Image lines for both of horizontal and vertical object lines appearing in the image captured by Camera-II, as a sample for other cameras, are illustrated in Figure 9a and 9b, respectively. In the shown figures, the image reference lines are shown in blue, image perpendicular lines are shown in yellow, image check lines are shown in red, and image endpoints for lines are shown in blue. Four models were applied for metric measurements as follows:

Model-I-A: based on the Cartesian coordinates of vanishing points using the extracted raw image lines (without refinement).

Model-I-B: based on the Cartesian coordinates of vanishing points applying image line refinement approach on the extracted raw image lines.

Model-II-A: based on the peaks of histogram of angle coordinates of vanishing points using the extracted raw image lines (without refinement).

Model-II-B: based on the peaks of histogram of angle coordinates of vanishing points applying image line refinement approach on the extracted raw image lines.

First, the three smartphones' cameras were calibrated utilized the MATLAB® application for camera calibration described earlier. The format size along with intrinsic and lens distortion parameters for each camera are demonstrated in Table 1. Afterwards, each single images of façade captured by each camera were undistorted (Figure 10) utilizing the obtained calibration radial and decentring lens distortion coefficients shown in Table 1.

Table 1. Camera parameters

Camera	Camera-I	Camera-II	Camera-III
Format size (pixels)	3024×4032	3024×4032	3000×4000
α (pixels)	3302.563	3126.647	2929.776
β (pixels)	3324.837	3135.546	2952.027
γ (pixels)	–3.251	–3.080	–2.525
u_o (pixels)	1517.219	1453.758	1503.039
v_o (pixels)	2012.123	2022.513	1973.367
k_1	0.1424	0.1097	0.1026
k_2	–0.6872	–0.3564	–0.2099
k_3	1.0831	0.2975	0.1268

Fifty-three image lines corresponding to horizontal object lines along with fifty-eight image lines corresponding to vertical object lines were extracted in each image utilizing the metric measurements MATLAB® application. The results obtained by applying the RANSAC algorithm are illustrated in Table 2.

a) b) c)

Figure 10. Undistorted single images of the three cameras: a – Camera-I; b – Camera-II; c – Camera-III

Table 2. Results obtained applying RANSAC algorithm

Number of lines	Camera-I		Camera-II		Camera-III	
	Horizontal lines	Vertical lines	Horizontal lines	Vertical lines	Horizontal lines	Vertical lines
Total image lines	53	58	53	58	53	58
Accepted image lines	35	34	40	41	37	35
Omitted image lines	18	24	13	17	16	23

The histograms for angle coordinates of the two orthogonal VPs (ϕ_{vy},ϕ_{vz}) for the three cameras used for Model-II-A and Model-II-B are displayed in Figure 11 to Figure 13 where the left column of figures represents histograms of (ϕ_{vy}) and the right column of figures represents histograms of (ϕ_{vz}). Moreover, the values of image Cartesian and polar coordinates for the two orthogonal vanishing points for all models are illustrated in Table 3 to Table 5. The shown Cartesian coordinates are calculated with respect to image origin (the upper left corner), while the polar coordinates are calculated with respect to the principal point. Even though, the values of Cartesian and radial coordinates for Model-I-A and Model-II-A along with Model-II-B and Model-II-B are notably different to each other, the angle coordinates are close. This conforms to the elliptical Gaussian distribution that is previously mentioned (Li et al., 2010). However, the calculated object lengths were not significantly affected by those different values of Cartesian and radial coordinates of VPs. It is worth noting that the values of polar coordinates for VPs applying Model-I-A and Model-I-B are those values obtained by solving for Cartesian coordinates, then the radial and angle coordinates were calculated according to Equations (12) and (13), respectively. On the other hand, the values of angle coordinates for VPs applying Model-II-A and Model-II-B are the peak values for the angle coordinates' histograms illustrated in Figure 11 to Figure 13. Then, after solving for radial coordinates, the values of Cartesian coordinates were obtained according to Equation (18) to (21).

The values of mean error (μ) and standard deviation for errors (σ) obtained applying Model-I-A, Model-I-B,

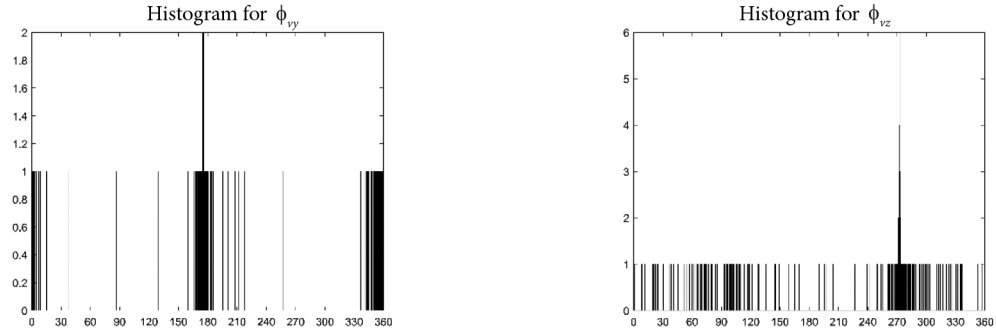


Figure 11. Histograms for angle coordinates of two orthogonal VPs for Camera-I

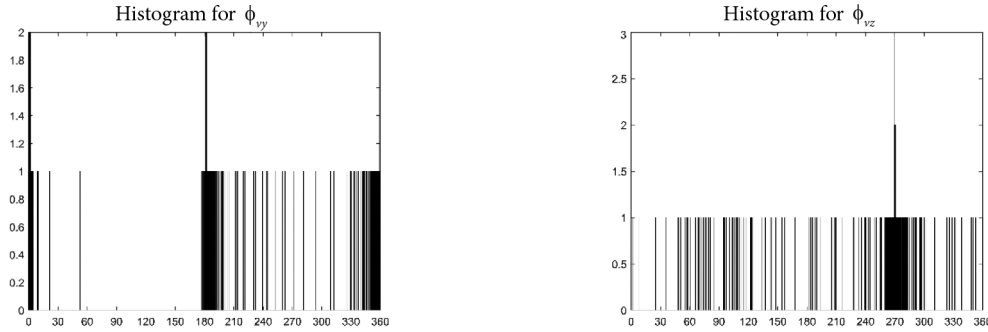


Figure 12. Histograms for angle coordinates of two orthogonal VPs for Camera-II

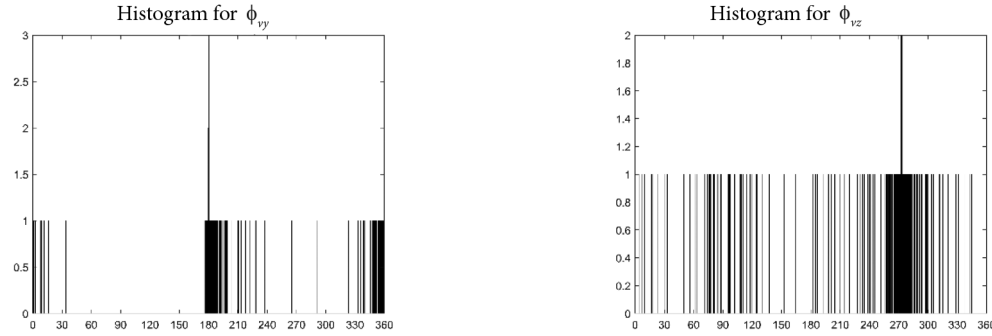


Figure 13. Histograms for angle coordinates of two orthogonal VPs for Camera-III

Table 3. Image cartesian and polar coordinates of vanishing points for Camera-I

Model	Cartesian coordinates				Polar coordinates			
	u_{vy} (pixels)	v_{vy} (pixels)	u_{vz} (pixels)	v_{vz} (pixels)	$\varnothing_{vy}$ (degrees)	ρ_{vy} (pixels)	$\varnothing_{vz}$ (degrees)	ρ_{vZ} (pixels)
Model-I-A, Model-I-B	–18 575.6	2605.6	1718.5	–3257.0	175.4627	20 145.5	272.8390	4274.1
Model-II-A, Model-II-B	–16 496.5	2686.2	1720.1	–3296.3	174.6866	18 081.0	272.8352	4313.4

Table 4. Image cartesian and polar coordinates of vanishing points for Camera-II

Model	Cartesian coordinates				Polar coordinates			
	u_{vy} (pixels)	v_{vy} (pixels)	u_{vz} (pixels)	v_{vz} (pixels)	$\varnothing_{vy}$ (degrees)	ρ_{vy} (pixels)	$\varnothing_{vz}$ (degrees)	ρ_{vZ} (pixels)
Model-I-A, Model-I-B	84 868.9	3836.1	1487.8	–2463.6	1.2796	83 391.1	269.8612	4437.4
Model-II-A, Model-II-B	224 036.8	3899.6	1496.3	–2525.0	0.4958	222 546.6	269.9716	4498.9

Table 5. Image cartesian and polar coordinates of vanishing points for Camera-III

Model	Cartesian coordinates				Polar coordinates			
	u_{vy} (pixels)	v_{vy} (pixels)	u_{vz} (pixels)	v_{vz} (pixels)	$\varnothing_{vy}$ (degrees)	ρ_{vy} (pixels)	$\varnothing_{vz}$ (degrees)	ρ_{vZ} (pixels)
Model-I-A, Model-I-B	–27 825.5	2000.6	1785.0	–3065.7	179.9467	29 328.6	273.2027	5046.9
Model-II-A, Model-II-B	–34 260.4	2035.7	1727.8	–3291.3	179.9001	35 757.4	272.5110	5269.7

Model-II-A, and Model-II-B for Camera-I to Camera-III are displayed in Table 6 to Table 8. The tables demonstrate the values of (μ,σ) for horizontal object lines, vertical object lines, and overall (horizontal and vertical) object lines. The number of check lines (n) and range of object line lengths are illustrated in the tables as well. Furthermore, the values of standard deviations for the overall object lines for all applied models are illustrated in Figure 14. As can be noticed from the tables and figure, the maximum absolute value of mean error for the three cameras was 0.006 m ((Camera-I, Model-II-B, horizontal object lines) & (Camera-III, Model-II-B, horizontal object lines)). The is a good indicator for non-presence of bias. In addition, the values of standard deviation for errors for all of the three cameras varied from 0.008m (Camera-I, Model-I-B, vertical object lines) to 0.017 m ((Camera-III, Model-II-A, horizontal object lines) & (Camera-III, Model-II-A, overall object lines)). These are considered as reasonable values especially when comparing with the previous researches such as Wang et al. (2010), Arslan (2018). Generally, the results obtained by Model-I were slightly better than those obtained by Model-II. However, both models led to results that are relatively close to each other. Moreover, results obtained for all cameras applying Model-I-B, Model-II-B are better than those obtained applying Model-I-A, Model-II-A, respectively. In other words, applying the image

Table 6. Results for Camera-I

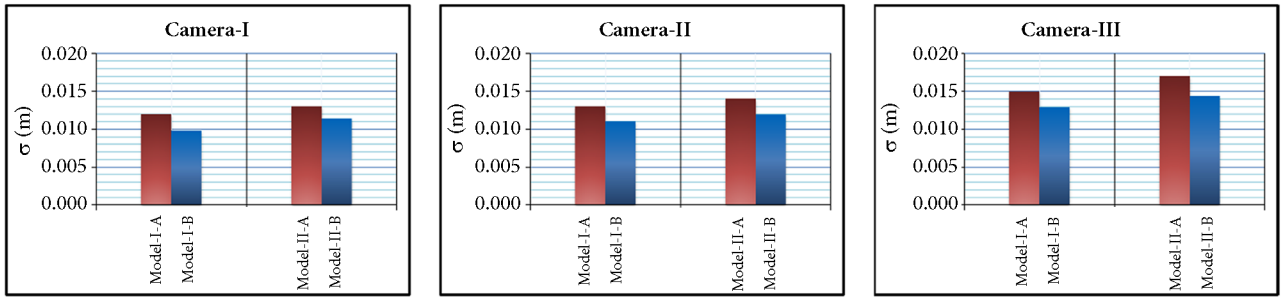
Model	Horizontal object lines				Vertical object lines				Overall object lines			
	n	length range (m)	μ (m)	σ (m)	n	length range (m)	μ (m)	σ (m)	n	length range (m)	μ (m)	σ (m)
Model-I-A	23	1.922 to 2.818	–0.004	0.013	23	1.513 to 3.194	0.003	0.010	23	1.513 to 3.194	–0.003	0.012
Model-I-B			–0.003	0.011			0.004	0.008			0.001	0.010
Model-II-A			–0.004	0.014			–0.002	0.012			0.003	0.013
Model-II-B			–0.006	0.012			0.003	0.009			–0.002	0.011

Table 7. Results for Camera-II

Model	Horizontal object lines				Vertical object lines				Overall object lines			
	n	length range (m)	μ (m)	σ (m)	n	length range (m)	μ (m)	σ (m)	n	length range (m)	μ (m)	σ (m)
Model-I-A	23	1.922 to 2.818	0.002	0.014	23	1.513 to 3.194	0.004	0.011	23	1.513 to 3.194	0.002	0.013
Model-I-B			0.003	0.011			–0.005	0.009			–0.001	0.011
Model-II-A			–0.004	0.015			0.002	0.015			0.002	0.014
Model-II-B			0.001	0.011			0.004	0.013			0.002	0.012

Table 8. Results for Camera-III

Model	Horizontal object lines				Vertical object lines				Overall object lines			
	n	length range (m)	μ (m)	σ (m)	n	length range (m)	μ (m)	σ (m)	n	length range (m)	μ (m)	σ (m)
Model-I-A	23	1.922 to 2.818	0.003	0.016	23	1.513 to 3.194	–0.002	0.014	23	1.513 to 3.194	–0.003	0.015
Model-I-B			0.005	0.014			0.003	0.012			0.004	0.013
Model-II-A			–0.003	0.017			–0.001	0.016			0.004	0.017
Model-II-B			0.006	0.015			0.004	0.014			0.005	0.014



a) b) c)

Figure 14. Overall lines' standard deviations for all models: a – Camera-I; b – Camera-II; c – Camera-III

line refinement approach improved the obtained results. The standard deviations for errors for the total 46 lines applying Model-I-B and Model-II-B respectively were 0.010 m and 0.011 m for Camera-I, 0.011 m and 0.012 m for Camera-II, and 0.013 m and 0.014 m for Camera-III. In general, the best results were obtained using Camera-I, Camera-II, and Camera-III, respectively. This complies with the cameras calibration results demonstrated in Table 1 where the values of (α,β), which is equivalent camera focal length, are sorted descending from Camera-I to Camera-III. For the same object to camera distance, the image scale (and, by consequence, the calculated object coordinates' accuracy) is inversely proportional to the camera focal length. All in all, the obtained accuracy is realistic and reasonable especially when compared with traditional land surveying work where in most cases the desired accuracy is around 0.01 m.

Conclusions

This paper investigated the possibility of utilizing the low-cost smartphone's camera for performing metric measurements of building's façades existing in narrow streets. The proposed technique is an adapted semi-automated technique based on detecting only two orthogonal vanishing points on single undistorted image captured by smartphone's main rear camera. Three smartphones of different brand-names were utilized applying two models for vanishing points detection. Model-I solved for the Cartesian coordinates of vanishing points, whereas Model-II solved mainly for the peaks of histogram of angle coordinate. The angle coordinates of vanishing points calculated by both models were close to each other in contrast to Cartesian and radial coordinates. A novel image line refinement approach based on the detected two orthogonal vanishing points was applied. A total number of 46 object lines (23 horizontal and 23 vertical lines) with lengths of 1.513: 3.194 m were used for results validation. The results obtained by both models utilizing the three smartphones' cameras were close to each other with slightly better results for Model-I. The best results were obtained by Camera-I which is the camera with the least equivalent focal length. Additionally, the solution was improved applying the image line refinement approach where the overall standard deviations obtained by Camera-I were 0.010 m and 0.011 m by Model-I and Model-II, respectively. In general, the obtained accuracy was reasonable especially when compared with the accuracy that is usually obtained by the traditional land surveying work.

Future work will develop a methodology for dealing with more complicated façades such as façades with circular and irregular features. In addition, more attention will be paid for extracting image features full automatically. Moreover, stereoscoping with smartphone's camera will be studied.

Acknowledgements

The authors would like to thank Fayoum University for supporting this research project.

References

Aguilera, D. G., Lahoz, J. G., & Codes, J. F. (2005). A new method for vanishing points detection in 3D reconstruction from a single view. In *Proceedings of ISPRS Comission* (pp. 197–210). https://www.isprs.org/proceedings/xxxvi/5-w17/pdf/6.pdf

Aldelgawy, M., & Abu-Qasmieh, I. (2021a). Calibration of smartphone's rear dual camera system. *Geodesy & Cartography*, *47*(4), 162–169. https://doi.org/10.3846/gac.2021.13434

Aldelgawy, M., & Abu-Qasmieh, I. (2021b). Semi-automatic reconstruction of object lines using a smartphone's dual camera. *The Photogrammetric Record*, *36*(176), 381–401. https://doi.org/10.1111/phor.12388

Arslan, O. (2018). Accuracy assessment of single viewing techniques for metric measurements on single images. *IET Computer Vision*, *12*(5), 693–701. https://doi.org/10.1049/iet-cvi.2017.0549

Burger, W. (2019). *Zhang's camera calibration algorithm: In-depth tutorial and implementation* (Technical Report HGB16-05). Department of Digital Media, University of Applied Sciences Upper Austria, School of Informatics, Communications and Media.

Burger, W., & Burge, M. J. (2016). *Digital image processing: An algorithmic introduction using Java* (2nd ed.). Springer. https://doi.org/10.1007/978-1-4471-6684-9

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-8*(6), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

Caprile, B., & Torre, V. (1990). Using vanishing points for camera calibration. *The International Journal of Computer Vision*, *4*(2), 127–140. https://doi.org/10.1007/BF00127813

Chuang, J.-H., Kao, J.-H., Lin, H.-H., & Chiu, Y.-T. (2007). Practical error analysis of cross-ratio-based planar localization. In *Pacific-Rim Symposium on Image and Video Technology* (pp. 727–736). Springer. https://doi.org/10.1007/978-3-540-77129-6_62

Debevec, P., Taylor, C., & Malik, J. (1996). Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques* (pp. 11–20). https://doi.org/10.1145/237170.237191

Di, H., Wang, L., & Xu, G. (2003). A three-step technique of robust line detection with modified Hough transform. In *Proceedings of the SPIE* (Vol. 5286, pp. 835–838). https://doi.org/10.1117/12.538699

Erdnüß, B. (2018). Measuring in images with projective geometry. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. XLII-1, pp. 141–148). https://doi.org/10.5194/isprs-archives-XLII-1-141-2018

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381–395. https://doi.org/10.1145/358669.358692

Hassanein, A., Mohammad, S., Sameer, M., & Ragab, M. (2015). A survey on hough transform, theory, techniques and applications. *International Journal of Computer Science Issues*, *12*(1), 139–156.

Hough, P. (1962). *Method and means for recognizing complex* (US Patent 3,069,654, Ser. No. 17,7156 Claims). https://patents.google.com/patent/US3069654A/en

Lee, S. C., & Nevatia, R. (2003). Interactive 3D building modeling using a hierarchical representation. In *First IEEE*

International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis (pp. 58–65). IEEE.

Li, B., Peng, K., Ying, X., & Zha, H. (2010). Simultaneous vanishing point detection and camera calibration from single images. In *International Symposium on Visual Computing* (pp. 151–160). Springer. https://doi.org/10.1007/978-3-642-17274-8_15

Liu, J.-S., & Chuang, J.-H. (2002). A geometry-based error estimation for cross-ratios. *Pattern Recognition*, *35*(1), 155–167. https://doi.org/10.1016/S0031-3203(00)00174-6

Mikhail, E. M., & Ackermann, F. E. (1982). *Observations and least squares* (2nd ed.). University Press of America.

Oh, S. H., & Jung, S. K. (2012). Ransac-based or thogonal vanishing point estimation in the equirectangular images. *Journal of Korea Multimedia Society*, *15*(12), 1430–1441. https://doi.org/10.9717/kmms.2012.15.12.1430

Rong, W., Li, Z., Zhang, W., & Sun, L. (2014). An improved Canny edge detection algorithm. In *2014 IEEE International Conference on Mechatronics and Automation* (pp. 577–582), Tianjin, China. https://doi.org/10.1109/ICMA.2014.6885761

Vouzounaras, G., Daras, P., & Strintzis, M. (2014). Automatic generation of 3D outdoor and indoor building scenes from a single image. *Multimedia Tools and Applications*, *70*(1), 361–378. https://doi.org/10.1007/s11042-011-0823-0

Wang, J., Shi, F., Zhang, J., & Liu, Y. (2006). A new calibration model and method of camera lens distortion. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 5713–5718), Beijing, China. https://doi.org/10.1109/IROS.2006.282376

Wang, M., Liu, X., Zhen, Y., & Lu, Y. (2010). Plane geometric information extraction from single image based on cross ratio. *18th International Conference on Geoinformatics* (pp. 1–5), Beijing, China. https://doi.org/10.1109/GEOINFORMATICS.2010.5567625

Wildenauer, H., & Hanbury, A. (2012). Robust camera self-calibration from monocular images of Manhattan worlds. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2831–2838). IEEE. https://doi.org/10.1109/CVPR.2012.6248008

Wong, T., Tao, C., Cheng, Y., Wong, K., & Tam, C. (2014). Application of cross-ratio in traffic accident reconstruction. *Forensic Science International*, *235*, 19–23. https://doi.org/10.1016/j.forsciint.2013.11.012

Zhang, G., He, J., & Yang, X. (2003). Calibrating camera radial distortion with cross-ratio invariability. *Optics & Laser Technology*, *35*(6), 457–461. https://doi.org/10.1016/S0030-3992(03)00053-7

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *22*(11), 1330–1334. https://doi.org/10.1109/34.888718