ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-201701.00024

Epipolar geometry comparison of SAR and optical camera postprint

Authors: Li, Dong, Zhang, Yunhua

Date: 2017-01-04T00:00:00+00:00

Abstract

In computer vision, optical cameras are often utilized as the eyes of the computer. Replacing the camera with synthetic aperture radar (SAR) facilitates a microwave vision of the world. This paper presents a comparative analysis of SAR imaging and optical camera imaging from the perspective of epipolar geometry. The imaging models and epipolar geometries of the two sensor types are analyzed in detail. Their differences are illustrated, and their underlying unification is specifically demonstrated. It is anticipated that these findings may benefit researchers in the fields of computer vision and SAR image processing in constructing a computer SAR vision system, which aims to complement and enhance human vision through electromagnetic perception and interpretation of images.

Full Text

Epipolar Geometry Comparison of SAR and Optical Camera

Dong Li, Yunhua Zhang
Key Lab of Microwave Remote Sensing (MiRS), Center for Space Science and Applied Research
Chinese Academy of Sciences, Beijing, China
Email: {lidong, zhangyunhua}@mirslab.cn

Abstract– In computer vision, optical cameras are often used as the "eyes" of the computer. If we replace the camera with synthetic aperture radar (SAR), we enter a microwave vision of the world. This paper provides a comparison of SAR imaging and camera imaging from the viewpoint of epipolar geometry. The imaging models and epipolar geometry of the two sensors are analyzed in detail. Their differences are illustrated, and their unification is particularly demonstrated. We hope these insights may benefit researchers in the field of computer vision or SAR image processing to construct a computer SAR vision, which is dedicated to compensating and improving our human vision by electromagnetically perceiving and understanding images.

Keywords– Computer vision, computer SAR vision, epipolar geometry, imaging model, optical camera, synthetic aperture radar (SAR), 3D reconstruction.

1 Introduction

Human vision is an intelligent system composed of eyes and brain. The eyes capture image information, which is then submitted to the brain for analysis, learning, recognition, classification, reconstruction, and determination. This system enables us to dynamically interact with the outside world, and it is so natural that we often neglect its complexity. Computer vision aims to duplicate the ability of human vision by electronically perceiving and understanding images [1]. Optical cameras are often used as "eyes" to capture images, which are then processed by computer for recognition, analysis, and reconstruction of objects. This field has received intensive attention and achieved great development in recent decades, benefiting our daily lives. Nevertheless, it still has a long way to go to achieve the intelligence of human vision.

If we replace the optical camera with synthetic aperture radar (SAR), the vision of the computer becomes completely different. SAR acquires images of objects by actively transmitting electromagnetic waves with certain frequencies and polarizations. This active operating mode makes it independent of solar illumination and thus enables all-day imaging. SAR operates in the microwave region of the electromagnetic spectrum (usually between P-band and Ka-band), which can avoid the effects of clouds, fog, rain, and smoke, thus enabling almost all-weather continuous monitoring. The wave-object interaction excites a scattered wave that carries characteristic information about the object, such as reflectivity, shape, and orientation. By processing the scattering to synthesize a 2D high spatial resolution image, we can achieve perception of the object. Imaging SAR systems are usually mounted on moving platforms such as airplanes or satellites and operate in a side-looking geometry. Airborne and spaceborne SARs provide us with microwave visions of the world from aerospace. Using SAR as "eyes," the vision of the computer may be compensated and improved. With the launch and operation of many spaceborne and airborne SAR systems recently, the available high-resolution SAR datasets have increased dramatically, making the joint processing of multi-view SAR images for accurate understanding and apperception of objects possible.

The foundation of computer SAR vision requires us to construct the SAR imaging model and epipolar geometry first. The related models for cameras are not applicable because SAR employs slant range imaging geometry. In view of this, a concise imaging model and a rigorous epipolar geometry model of SAR were developed recently [2]. This paper is dedicated to providing further comparison of SAR imaging and camera imaging from the geometrical viewpoint. Section 2 first concisely presents the imaging model and epipolar geometry of camera. As a comparison, the corresponding SAR models are then introduced in Section 3. The differences are revealed and the consistency is indicated. Section 4 finally concludes the paper.

2.1 Imaging Model of Camera

Camera acts as a mapping from 3D space to 2D image, which involves four different coordinate systems: the world coordinate system O_w-$X_{wY_{wZ}}w$, the camera coordinate system O_c-$X_{cY_{cZ}}c$, the physical image coordinate O_p-xy, and the digital image coordinate system O_i-uv. The transformation from O_w-$X_{wY_{wZ}}w$ to O_c-$X_{cY_{cZ}}c$ usually involves a 3D rotation (R) and a translation (t) because of the geometrical misalignment between the two systems.

The pinhole camera model then determines the transformation from O_c-$X_{cY_{cZ}}c$ to O_p-xy:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c/Z_c \\ Y_c/Z_c \\ 1 \end{bmatrix}$$

where f is the focal length. The final digital image is a sampling of the physical image, which can be formulated as:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & s_\theta & u_0 \\ 0 & s_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where s_x and s_y are the scales in horizontal (u) and vertical (v) image directions, respectively, which relate to the resolutions of the image, and θ is the angle between u- and v-axes, which accounts for the fact that the pixel grid may not be exactly orthogonal (it is usually very close to $\pi/2$) [3].

Combining the transformations, the relation between O_w-$X_{wY_{wZ}}w$ and O_p-xy is obtained:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

where K is the intrinsic matrix accounting for camera sampling and optical characteristics:

$$K = \begin{bmatrix} fs_x & fs_\theta & u_0 \\ 0 & fs_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Equation (4) is usually simply expressed as:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

where λ is an arbitrary constant. Equation (6) is just the projective camera model, in which P is the camera matrix.

2.2 Epipolar Geometry of Camera

Consider the stereo system composed of two cameras shown in Fig. 1 [Figure 1: see original paper], where C_1 and C_2 are the optical centers of the cameras, and $I_1 \leftrightarrow I_2$ is the corresponding projective image pair. Given a pixel m_1 in I_1, it corresponds to a series of points $M_1, M_2, ...$ in 3D space, and these points lie on the line through C_1 and m_1. When these points are viewed by the second camera from a distinct position, they will be mapped to the line L_2 in I_2, and this line is the epipolar line of m_1. Given a pixel m_2 on L_2, there then always exists a constraint between m_1 and m_2. This relation is termed the epipolar geometry of the stereoscope, also known as the image geometrical warp function because it maps a pixel position in I_1 into a different pixel position in I_2 and forms the so-called image geometrical warp.

The epipolar geometry of camera has been extensively studied in computer vision, where the fundamental matrix and homography are the widely-used descriptions of epipolar geometry when the pinhole model is considered. Let the displacement from the first camera to the second be (R, t). Let M be the 3D point corresponding to pixels m_1 and m_2. Without loss of generality, we assume that M is expressed in the coordinate system of the first camera. We then have:

$$\lambda_1 \begin{bmatrix} m_1 \\ 1 \end{bmatrix} = [I \quad 0] \begin{bmatrix} M \\ 1 \end{bmatrix}, \quad \lambda_2 \begin{bmatrix} m_2 \\ 1 \end{bmatrix} = [R \quad t] \begin{bmatrix} M \\ 1 \end{bmatrix}$$

Eliminating M and the constants λ_1 and λ_2, we obtain the following fundamental equation:

$$m_2^T F m_1 = 0$$

where F is the fundamental matrix defined as:

$$F = K_2^{-T} [t]_\times R K_1^{-1}$$

and $[t]_\times$ is an antisymmetric matrix defined by t. The essential matrix E is defined as:

$$E = [t]_\times R$$

which accounts for camera position and orientation in the world coordinate system.

Based on the correspondence set $\{m_1^i \leftrightarrow m_2^i\}$ extracted by certain feature extractors, the fundamental matrix F can be robustly estimated, then the camera matrix P can be simply retrieved from F. Using methods such as triangulation, we can finally locate the 3D point M_i corresponding to m_1^i [3]. This is the task of 3D reconstruction from multiple images.

If point M is in a 2D planar scene $\pi = (n^T, d)^T$, then m_1 in I_1 is uniquely mapped to m_2 in I_2. Such mapping can be expressed by the following transformation:

$$m_2 = Hm_1$$

where H is the plane-induced homography, which can be easily obtained from (8) as:

$$H = K_2 \left(R - \frac{tn^T}{d} \right) K_1^{-1}$$

By fitting $\{m_1^i \leftrightarrow m_2^i\}$ to (11), an estimation of homography H is also achieved, based on which we can then geometrically align the image pair $I_1 \leftrightarrow I_2$. This is the task of image registration, which is the foundation of many applications such as object tracking and recognition, camera calibration and image reconstruction, as well as digital elevation model inversion and deformation mapping of the Earth' s surface.

3 SAR Imaging Model and Epipolar Geometry

The formulation above indicates that modeling epipolar geometry involves the imaging model of the sensor. SAR acquires images from slant range, thus the 3D points which are imaged to the same pixel m_1 in I_1 locate in a Doppler circle formed by the intersection of the range sphere and Doppler cone, as shown in Fig. 2 [Figure 2: see original paper]. These points are then imaged to a series of pixels in I_2 through the slant projection of the second SAR sensor and compose the epipolar line of m_1, which is no longer a simple straight line like L_1 in Fig. 1. Hence, describing SAR epipolar geometry with the fundamental matrix is inappropriate because it is only suitable for central projection. SAR images often cover large ground scenes with varied topography, thus approximating the 3D ground scene as a planar surface is inaccurate, making the plane-induced homography also inappropriate. In order to construct a rigorous description of SAR epipolar geometry, we seek another approach to achieve it directly from the SAR imaging model.

The imaging model maps a 3D point M to its projective pixel, so we can relate the two imaged pixels m_1 and m_2 of M under the two SAR projections by combining the two SAR imaging models. The imaging model should be accurate and concise to achieve rigorous and analytical epipolar modeling. Existing SAR imaging models can be generally attributed to two categories: the physical model and the empirical model. The physical model takes into account several aspects that influence the acquisition procedure based on the range-Doppler equations (RDEs):

$$R = \sqrt{(X_P - X_S)^2 + (Y_P - Y_S)^2 + (Z_P - Z_S)^2}$$

$$f_d = -\frac{2}{\lambda} \frac{(X_P - X_S)V_X + (Y_P - Y_S)V_Y + (Z_P - Z_S)V_Z}{R}$$

where R is the distance from object (X_P, Y_P, Z_P) to antenna phase center (APC) (X_S, Y_S, Z_S), (V_X, V_Y, V_Z) is the velocity of the platform, λ is the wavelength of the transmitted wave, and f_d is the Doppler frequency. Theoretically, we can obtain an accurate SAR epipolar model from the RDEs of the two SARs, but the model may not be concise because of the complex nonlinearity of the RDEs, which impacts further application.

The empirical model is used when system parameters, imaging geometry, and physical models are unavailable, and polynomial and rational polynomial functions are often employed:

$$u = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{k=0}^{p} a_{ijk}X^iY^jZ^k}{\sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{k=0}^{p} b_{ijk}X^iY^jZ^k}, \quad v = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{k=0}^{p} c_{ijk}X^iY^jZ^k}{\sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{k=0}^{p} d_{ijk}X^iY^jZ^k}$$

which are in fact extensions of homography and collinearity equations obtained from central projection:

$$u = \frac{a_1X + a_2Y + a_3Z + a_4}{a_9X + a_{10}Y + a_{11}Z + 1}, \quad v = \frac{a_5X + a_6Y + a_7Z + a_8}{a_9X + a_{10}Y + a_{11}Z + 1}$$

However, different from the central projection of camera, the equivalent projection center for slant range SAR imaging is not fixed—SAR is a variable focus system or multi-central projection system—thus the empirical model is also inaccurate.

3.1 Concise Imaging Model of SAR

Here we consider the general SAR imaging geometry shown in Fig. 3 [Figure 3: see original paper]. Rigorous modeling of SAR epipolar geometry requires the imaging model to be concise and accurate. To achieve this, we also construct four coordinate systems: the global coordinate system O-XYZ, the platform coordinate system O'-$X'Y'Z'$, the imaging coordinate system o-xyz, and the image coordinate system o_i-uv. O-XYZ is similar to the world coordinate system O_w-$X_{wY_{wZ}}w$. We assume the radar moves along a straight track of height H parallel to the ground plane XOY, which indicates that the influence from Earth' s curvature and track curvature is neglected, thus we mainly focus on airborne SAR systems. Nevertheless, it may also hold for spaceborne SAR systems if we can compensate those non-ideal influences using high-precision platform-borne GPS and INS beforehand. O'-$X'Y'Z'$ is used to characterize the attitude of the platform (its counterpart is the camera coordinate system O_c-$X_{cY_{cZ}}c$), where O' is located at (T_X, T_Y, H) in O-XYZ representing the initial APC, X' denotes the flight direction of the platform, Z' is parallel to Z, and Y' is orthogonal to X' and Z'.

If the flight direction X' is deviated from X by angle β, the transformation between O'-$X'Y'Z'$ and O-XYZ can thus be expressed as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ H \end{bmatrix}$$

We further consider an imaging geometry where the antenna has a squint angle of α which anti-clockwise rotates Y' to the incidence plane. This is a special characteristic of SAR imaging. For squint SAR, we should compensate the scattering to the zero Doppler centroid first. For convenience, the imaging coordinate system o-xyz is further defined, where o is located at $(T_X, T_Y, 0)$ in O-XYZ, z is parallel to Z, y is parallel to the ground projection of the antenna boresight, and x is orthogonal to y and z. The compensation results in an anti-clockwise rotation of α from X' to x, thus the relation between o-xyz and O'-$X'Y'Z'$ is:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}$$

By combining (16) and (17), we can obtain the transformation from O-XYZ to o-xyz:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T$$

where R is the rotation matrix and T is the translation vector, and ϕ $(= \alpha + \beta)$ denotes the anti-clockwise rotation from X to x.

After these definitions, SAR imaging can be modeled as a geometrical projection from the ground plane to the slant plane. Let C be a 3D point within the radar beam with coordinates (X, Y, Z) and (x, y, z) in O-XYZ and o-xyz, respectively. After slant projection, C is mapped to C' (x_p, y_p, z_p). From the projection geometry in Fig. 3, we can easily obtain:

$$x_p = x, \quad y_p = y\cos\theta, \quad z_p = z + y\sin\theta$$

where θ is the local radar incidence angle related to the position and height of C. Thus the relation between the slant projective plane and ground plane can be written as:

$$\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & 0 \\ 0 & \sin\theta & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The counterpart of this transformation in the camera model is (2).

The final SAR image is the sampling of the projective plane. The image is defined in the image coordinate system o_i-uv, and the transformation between pixel (u, v) and projection (x_p, y_p, z_p) can then be expressed as:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & t_x \\ 0 & s_y & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}$$

where (t_x, t_y, H) is the location of o_i in o-xyz, and s_x and s_y are the scales related to the azimuth and range pixel sizes Δ_a and Δ_r, respectively:

$$s_x = \frac{1}{\Delta_a} = \frac{2V}{\lambda L_e} \cdot \frac{1}{PRF} \cdot s_a, \quad s_y = \frac{1}{\Delta_r} = \frac{2B}{c} \cdot s_r$$

where L_e is the effective antenna aperture, B is the bandwidth of the transmitted signal, s_a is the azimuth oversampling rate related to the PRF of the system, s_r is the range oversampling rate, and c is the velocity of light. The transformation in (21) is similar to (3) of the camera model. However, the pixel grid in SAR images can generally be kept exactly orthogonal, thus the retinal distortion is neglected here.

By combining (18), (20), and (21), we have:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = SMR \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Equation (23) finally relates a 3D point to its projective pixel in a SAR image. Under the assumption that the radar moves along a track parallel to the ground plane, the RDEs are in fact consistent with (23) because no approximation was used when deriving the relation. Nevertheless, we decompose the complex RDEs into the multiplication of three simple matrices of physical significance based on transformations among four different coordinate systems, which helps us achieve a concise and accurate result. The model involves the system parameters s_x and s_y, the imaging geometry parameters ϕ, T_X, T_Y, t_x, and t_y, as well as the object parameter θ. It is interesting to observe that (23) is similar to the linear camera model of (6): here R as well as T_X, T_Y, $s_x t_x$, and $s_y t_y$ correspond to (1) denoting the transformation from world coordinate system to camera coordinate system by rotation and translation, M corresponds to (2) which indicates the transformation from camera system to physical image coordinate system by pinhole model, and S corresponds to (3) denoting the transformation from image physical system to pixel coordinate system by digital sampling. Therefore, the obtained model enables us to geometrically unify SAR imaging and camera imaging. Besides this, from (23) one can see that the point position (X, Y) is explicitly related to the pixel position (u, v), but the point elevation Z is implicit in the local incidence θ, thus the model may also enable a flexible strategy to model the epipolar geometry and reconstruct the object.

3.2 Rigorous Epipolar Geometry of SAR

The side-looking geometry of SAR makes the epipolar geometry description in terms of fundamental matrix and homography inappropriate, so we turn to construct the SAR epipolar geometry directly from the imaging model. This kind of epipolar geometry description is less used for cameras because the fundamental matrix and homography are both adequate, but it facilitates modeling the rigorous epipolar geometry for SAR from the developed concise imaging model.

We consider a general stereoscopic configuration here. Let I_1 and I_2 be an image pair acquired by different SAR systems from different imaging geometries. For a 3D point (X, Y, Z) in O-XYZ, if its two projective pixel positions in I_1 and I_2 are (u_1, v_1) and (u_2, v_2), respectively, according to (23) we obtain:

$$\lambda_1 \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = S_1 M_1 R_1 \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

where the subscript 1 indicates the parameters of I_1. Based on (23) and (24), for pixel (u_2, v_2) of I_2 we can have:

$$\lambda_2 \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = S_2 M_2 R_2 \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

where the subscript 2 indexes the parameters of I_2. By eliminating the object position (X, Y, Z), we can then relate the two pixels. Equation (25) can be further rearranged as (26), shown at the top of the next page, where $A = S_2 M_2 R_2 R_1^{-1} M_1^{-1} S_1^{-1}$, $\Delta\phi$ $(= \phi_2 - \phi_1)$ is the rotation between the two imaging systems, B_x and B_y are the projections of B_X $(= T_{X1} - T_{X2}$, denoting the initial separation between the two imaging systems in X-direction) and B_Y $(= T_{Y1} - T_{Y2}$, denoting the initial separation between the two imaging systems in Y-direction) in x- and y-directions of the second SAR imaging system, and they denote the along-track and cross-track baselines, respectively. Equation (26) is an affine transformation which models the epipolar geometry of a general SAR stereo. Interestingly, it is similar to the plane-induced homography in (12): here $S_1 M_1$ and $S_2 M_2$ correspond to the intrinsic matrices K_1 and K_2 of the two cameras, while A as well as t_u and t_v correspond to the rotation and translation $R - tn^T/d$ between the two cameras. However, different from the fundamental matrix and homography which are independent of the object, the SAR epipolar geometry in (26) is object-dependent because the local radar incidences θ_1 and θ_2 vary with each 3D point. Nevertheless, (20) and (23) show that, for slant range imaging, the position and elevation of a 3D point are wrapped into the imaged pixel and related to the local incidence. Therefore, given the imaging parameters and pixel correspondences, we can obtain a retrieval of incidence from (26), then the 3D geometry of the object may be achieved based on the projective pixel positions. Therefore, image reconstruction in computer SAR vision seems more straightforward than that in computer optical vision, as detailed in [2].

4 Conclusions

The slant range imaging of SAR makes the equivalent projection center unfixed. SAR is therefore a variable focus system or multi-central projection system. This paper is dedicated to providing a comparison of SAR imaging and camera imaging from the geometrical point of view. A unified expression of camera imaging model and SAR imaging model is obtained. Nevertheless, the side-looking geometry makes the SAR imaging model vary with each object position, besides introducing an extra squint-related rotation. We thus cannot use a fixed model to express the mapping from 3D space to 2D SAR image. The epipolar geometry models the relation between the two pixels of a 3D point projected by a stereoscope. The central projection of camera enables a pixel position in the first image to correspond to a straight line in the second image. Such relation is described in terms of the fundamental matrix. It can also be expressed by the homography if the considered scene is planar. However, these two descriptions of epipolar geometry are both inappropriate for SAR because slant range imaging makes the epipolar line non-straight, and SAR images often cover large ground scenes with varied topography that cannot be approximated as planar surfaces. Hence, we turn to construct the rigorous SAR epipolar geometry directly from the imaging model. Nevertheless, its unification with the plane-induced camera homography is also clear. The obtained SAR epipolar geometry also varies with each object position, i.e., we cannot use a fixed homography to model the geometric warping of two SAR images. This makes image registration more difficult. However, such object-dependent epipolar geometry is welcome for the retrieval of 3D geometry of objects. Hence, 3D reconstruction of objects in computer SAR vision thus seems more straightforward than that in computer optical vision.

The consistency in imaging model as well as epipolar geometry indicates the geometrical unification of SAR imaging and camera imaging in a sense. Nevertheless, this does not mean that we can unify the two sensors physically. Our focus in this paper is on the geometrical part of imaging. In fact, the side-looking geometry also impacts the physical appearance of SAR images. The detailed comparison of SAR imaging and camera imaging from the electromagnetic scattering and signal point of view was presented in [4]. We hope such geometrical consistency will benefit researchers in the field of computer optical vision or SAR image processing to construct a computer SAR vision. As an interdisciplinary field, computer SAR vision is dedicated to improving and compensating human vision by electromagnetically perceiving and understanding images, which will make our view of the world more colorful.

Acknowledgements

This work was supported by the Youth Innovation Promotion Association, Chinese Academy of Sciences under the grant no. Y42133A39S.

References

[1] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 3rd ed. Toronto, Canada: Thomson, 2008.

[2] D. Li and Y. Zhang, "A rigorous SAR epipolar geometry modeling and application to 3D target reconstruction," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 6, no. 5, pp. 2316-2323, Oct. 2013.

[3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[4] Y. Zhang and J. Jiang, "Why optical images are easier to understand than radar images? -From the electromagnetic scattering and single point of view," in *Proc. PIERS*, Guangzhou, China, 2014, pp. 1411-1415.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.