IJAMML 1:1 (2014) 93-122 October 2014
ISSN: 2394-2258
Available at http://scientificadvances.co.in

A NOVEL 3D RECONSTRUCTION APPROACH BASED ON 1D RANGE DATA SEQUENCE AND STEREO VISION PERCEPTIONS

Jie Zhao[a*], Yanchun Zhang[b] and Yongmei Qi[a]

[a]College of Electronic and Information Engineering, Hebei University, Baoding, 071000, P. R. China

[b]School of Engineering and Science, Victoria University, Melbourne, VIC 8001, Australia

Abstract

The three-dimensional reconstruction technology is an important tool to analyze and study external shape and space status of special objects. The two-dimensional image series of an object, obtained by optical image sensor, can be used to perform three-dimensional reconstruction. But there exist cases that some objects are covered by some media or located inside some media. Hence, we can only analyze the distance sequences from one-dimensional sensors to obtain three-dimensional information of the target. However, because of little information and weak correlation with one-dimensional data, it is very difficult to restore the complete three-dimensional information. A novel method is proposed to perform three-dimensional reconstruction by mining one-dimensional range data sequences. This method can be widely used in most of the one-dimensional sensor systems. Finally, simulation experiments on two typical radar systems show the feasibility and efficiency of our proposed method.

[*]Corresponding author.
E-mail address: jzhao_hbu@163.com (Jie Zhao).

2010 Mathematics Subject Classification: 68.
Submitted by Shuaiqi Liu.
Received September 15, 2014

Keywords: 3D reconstruction, range data sequence, euclid update, bundle adjustment.

---

1. Introduction

With the tremendous development of information science, 3D perception has become a state-of-art technology for current intricate equipments to touch the real world. Visual image, the most important way of capturing things, is fully used to extract 3D information from the scene. Nowadays, with the advance of computer vision technique, a series of 3D reconstruction theories, such as key-points extraction [1, 2], image matching [2-4], projective geometry [5], self-calibration [6, 7], and 3D model [8] become mature and a lot of vision systems have already been applied to various fields. However, because of the increasingly complex circumstances and diverse purposes of applications, single medium is harder to fulfill the necessity of higher precision and quality of the information. Thus, there is no doubt that a promising trend is to associate various sensor modalities to obtain real 3D information.

To apperceive the concrete object, several media may be utilized, such as light ray, sound wave, laser, electromagnetic wave, and ultrasonic, etc. These resources have been exploited in various systems, such as vision system, sonar, laser, radar, and ultrasonoscope etc. Actually, all of them can only capture the lower-dimensional profiles of 3D objects and can be classified into two categories. The first category is essentially a 3D-to-2D projective measure, which is exclusive to vision systems. It mainly attributes to the work of the camera lens which makes the 3D scene projected on the 2D optical image. The others can be wholly ranked as the 3D-to-1D measure. The sonar, laser, radar, and ultrasound systems are 1D range sensing modalities, which apperceive the distance of the target via the time-delay of the wave transmitted. These measures are of exclusive potentials which the vision system cannot be applied to in some areas, such as underwater, far distance, and inner-object detection.

Accordingly, as the theory and methodology of 3D reconstruction has matured in computer vision, how to build a homologous theoretical framework of 3D reconstruction of 1D sensing system [9] remains a problem to be solved.

Associating data extracted from different low-dimensional profiles is a feasible method to recover the 3D information of an object based on the geometric theory of 3D reconstruction [10]. In the technique of 3D reconstruction based on image sequence which is widely researched in the last two decades, the pictures of an object or a scene are captured frame-by-frame from different locations. Then, as illustrated in Figure 1, a series of processes are implemented, such as key-points extraction, image matching, initial projective reconstruction, structure from motion, euclid update, and bundle adjustment. It has a charming advantage, which is the camera does not require intricate calibration procedures owing to the self-calibration process involved in this system. As a result, the reconstruction becomes so flexible that the motion and intrinsic parameters of the camera are unnecessary to be known and fixed. Therefore, a question to ask is if it is possible that this advanced theory and system can be applied to 1D sensing modalities? The answer is positive.

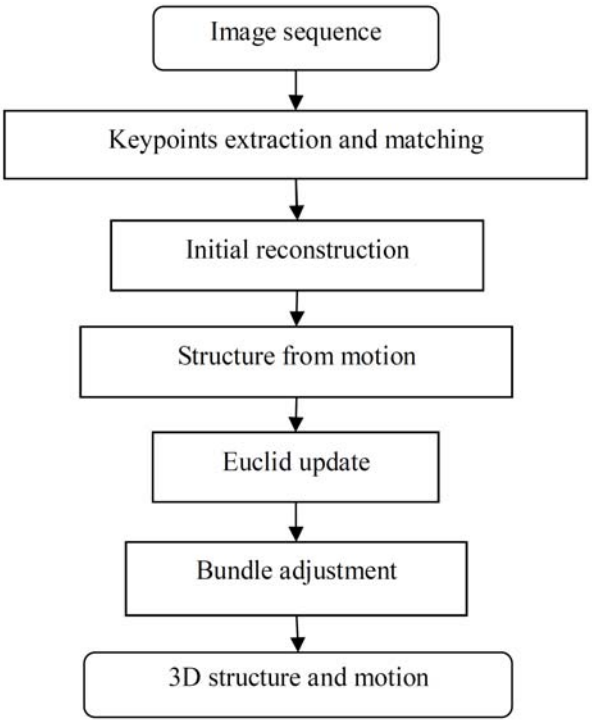


Figure 1. The processes of 3D reconstruction.

In the radar, sonar, laser, and ultrasound systems, a single array of the sensor can transmit and receive the wave reflected from the object. As Figure 2 has shown, compared with the 2D imaging captured by the camera, the 1D sensor can obtain the 1D range image, which represents the accumulated amplitude of the target on the range direction. Then during the motion of the target or sensor, a sequence of 1D range images can be captured. Different sensor modalities have different features. In the vision, the feature points can be extracted based on the colour and luminance information of the image; in the 1D image, it could be the scatterers that have instinct reflected amplitude. For example, in a radar system, the scattering center model is taken as a steady feature of the radar target and it is widely used in the automatic recognition of the target [11, 12]. There are some mature techniques to extract the

scattering centers from the radar 1D range image with high accurate range data [13, 14]. In addition, the key-points matching [15, 16] process is also a tough problem in 1D sensor. During the target moving, the feature points migrate with different distance locations in the 1D range image sequence as Figure 2 has shown. Like the object tracking from video stream in vision domain, the 1D feature points among the sequence of the 1D range images can be associated pulse-by-pulse using the tracking method too [17]. Via these preliminaries, the 1D range data of the collected feature points make it possible to recover the 3D structure and motion of the unknown target utilizing geometric approaches.

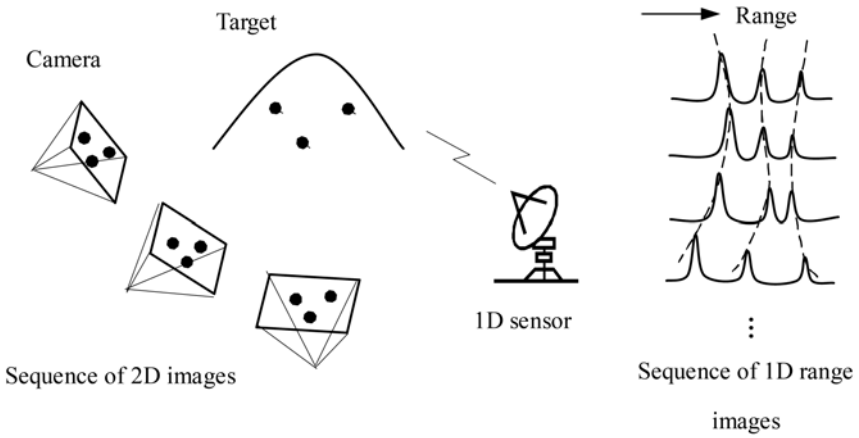


Figure 2. Comparison of 2D optical and 1D sensing system.

Nowadays, the 1D-to-3D topic has not been a virgin land since Stuff proposed a reconstruction algorithm based on the assumption of rigid-body of the radar targets [18]. Then Ferrerre, inspired by the works of Tomasi and Kanade in computer vision, proposed an improved algorithm based on "Factorization" [19, 20]. Though these existing algorithms provide a possibility of reconstructing the 3D shape and motion of unknown target using the 1D sensor, the geometric theory of this 1D-to-3D case is still not explicit enough. In this paper, referring to the 3D reconstruction theory in computer vision, a generalized skeleton of 3D

reconstruction based on 1D range image sequence is proposed. Firstly, based on the geometric analysis of the 1D sensor, we propose a model of 1D-to-3D projective geometry. Then, the initial affine reconstruction is obtained and the procedure of "structure from motion" (SFM) is built to finish the affine reconstruction of the target. After that, the euclid update is realized by utilizing the geometric invariance of the rigid target. Finally, the method of bundle adjustment is proposed with the non-linear near-field range model involved. Compared with Ferrerre's method, these systematic methods we proposed are applicable to the reconstruction by using incomplete range data. In addition, the accuracy of the reconstruction can be improved with bundle adjustment implemented in our algorithm where, meanwhile, the near-field non-linear error can be suppressed effectively.

2. The Reconstruction of 1D-to-3D

2.1. Geometric model of the 1D sensing modality

Naturally, almost all the perception measures can merely capture the low-dimensional profiles of 3D targets. We denote the 3D homogeneous coordinates of the target as $\widetilde{\boldsymbol{X}}_i = [x_i, y_i, z, 1]^T$, then the captured low-dimensional profile $\widetilde{\boldsymbol{m}}_i$ is basically the degenerated form of the real 3D structure. The degeneracy can be mathematically represented as a transform as shown in [5]

$$\widetilde{\boldsymbol{m}}_i = \boldsymbol{P}\widetilde{\boldsymbol{X}}_i. \tag{1}$$

In the case of computer vision, $\widetilde{\boldsymbol{m}}_i = [u, v, 1]^T$ is the 2D homogeneous coordinates of the imaged point, and transform $\boldsymbol{P} = \boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]$ is a 3×4 matrix (containing the intrinsic $\boldsymbol{K}$ and extrinsic parameters $\boldsymbol{R}$ and $\boldsymbol{t}$ of the camera), which attains the degeneracy from 3D to 2D.

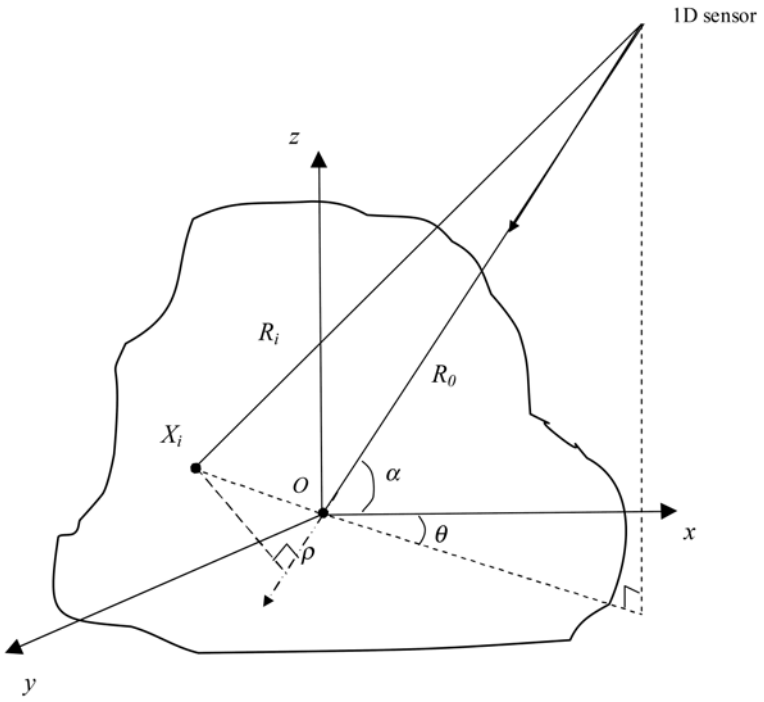


Figure 3. Geometric model of the 1D sensing.

In 1D sensing modality, the degeneracy occurs from 3D to 1D. As illustrated in Figure 3, a rigid target has a relative motion under the detection of the 1D sensor. Then, we place the target in the coordinate system $Oxyz$ with the target's centroid on the origin O. Here, we assume the base-distance between the sensor and target centroid is R_0 and the two angles of the view are θ and α, then the 1D range of the 3D point X_i has the form of

$$\begin{aligned} R_i &= \sqrt{(R_0 + \rho_i)^2 + \|\boldsymbol{X}_i\|^2 + \rho_i^2} \\ &= \sqrt{R_0^2 + 2R_0\rho_i + \|\boldsymbol{X}_i\|^2}, \end{aligned} \tag{2}$$

where $\|\boldsymbol{X}_i\| = \sqrt{x_i^2 + y_i^2 + z_i^2}$ is the distance between $\boldsymbol{X}_i$ and origin O. ρ_i is the 1D projected range of X_i on the view line of the sensor, which has the form of

$$\rho_i = \boldsymbol{c} \cdot \boldsymbol{X}_i, \tag{3}$$

where $\boldsymbol{c}$ is a vector of the view aspect which has the form as

$$\boldsymbol{c} = [C_1, C_2, C_3]^T = [\sin\theta\cos\alpha, \cos\theta\cos\alpha, \sin\alpha]^T. \tag{4}$$

From (2), we can see that the 1D perception of range is nonlinear. However, when the base-distance R_0 is long enough to satisfy the far-field hypothesis, the 1D range of R_i will approach the following linear relation:

$$R_i = \rho_i + R_0. \tag{5}$$

Then, corresponding to the degeneracy of (1), we have

$$R_i = \boldsymbol{p}\widetilde{\boldsymbol{X}}_i \text{ with } \boldsymbol{p} = \left[\boldsymbol{c}^T | R_0\right]. \tag{6}$$

Here, the transform matrix $\boldsymbol{P}$ becomes a 1×4 vector which makes the 3D target degenerate to 1D range information.

The low-dimensional profiles are required to be associated together to extract the 3D information of the target. For example, in computer vision, several images are captured at different positions. Similarly, in the 1D perception cases, various aspects can be obtained by the consecutive pulses detected during the motion of the target or the sensor. Assuming the target is detected L times, we use the subscript $j(j = 1 \sim L)$ to represent the items at the j-th pulse. Assuming the 1D range data of N key-points in L view aspects are extracted and associated as $R_{ji}(j = 1 \sim L, i = 1 \sim N)$, then we can construct the 1D range data R_{ji} as an $L \times N$ range matrix.

$$\widetilde{\varphi} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1N} \\ R_{21} & R_{22} & \cdots & R_{2N} \\ \vdots & \vdots & & \vdots \\ R_{L1} & R_{L2} & \cdots & R_{LN} \end{bmatrix}. \tag{7}$$

Meanwhile, assuming the structure and rotation matrices are

$$\boldsymbol{S} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_N] \text{ and } \boldsymbol{C} = [\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_L]^T. \tag{8}$$

Then from (6), we can obtain the 1D-to-3D geometry as

$$\widetilde{\varphi} = \boldsymbol{P}\widetilde{\boldsymbol{S}}, \tag{9}$$

where $\widetilde{\boldsymbol{S}} = \left[\widetilde{\boldsymbol{X}}_1, \widetilde{\boldsymbol{X}}_2, \cdots, \widetilde{\boldsymbol{X}}_N\right] = \begin{bmatrix} \boldsymbol{S} \\ \mathbf{1}_{1\times N} \end{bmatrix}$ and $\boldsymbol{P} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_L]^T = [\boldsymbol{C} \,|\, \boldsymbol{R}_0]^T$ can be deduced ($\mathbf{1}_{1\times N}$ denotes a $1 \times N$ vector which is full of elements of 1; and $\boldsymbol{R}_0 = [R_{01}, R_{02}, \cdots, R_{0L}]^T$).

2.2. Initial affine reconstruction

According to (9), the procedure of the 1D-to-3D reconstruction is used to solve the structure and motion parameters $\boldsymbol{S}$ and $\boldsymbol{P}$ when the 1D range data $\widetilde{\varphi}$ is given. In computer vision, the reconstruction starts from a group of initial projective reconstructions obtained from two views by fixing the two camera matrices as $\boldsymbol{P}_1 = [\boldsymbol{I} \,|\, 0]$ and $\boldsymbol{P}_2 = [[e']_\times \boldsymbol{F} \,|\, e']$. In the 1D-to-3D cases, a low-level modality of reconstruction also needs to be initiated.

First of all, the base-distance $\boldsymbol{R}_0$ can be estimated by calculating the motion of the target's centroid, whose coordinates can be generally selected as $\left[\frac{1}{N}\sum_{i=1}^{N} x_i, \frac{1}{N}\sum_{i=1}^{N} y_i, \frac{1}{N}\sum_{i=1}^{N} z_i\right]^T$. Without loss of generality, the centroid can be deemed as the original O of the target coordinates system. Thus, once given a completed matrix $\widetilde{\varphi}$, the base-distance of the target $\boldsymbol{R}_0$ has the following estimation:

$$R_{0j} = \frac{1}{N}\sum_{i=1}^{N} R_{ij} \ (j = 1 \sim L). \tag{10}$$

Then, a pure rotation model of the target's motion can be obtained by removing the translation motion from the matrix $\tilde{\varphi}$ as follows:

$$\varphi = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1N} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2N} \\ \vdots & \vdots & & \vdots \\ \rho_{L1} & \rho_{L2} & \cdots & \rho_{LN} \end{bmatrix} = \tilde{\varphi} - \begin{bmatrix} R_{01} & R_{01} & \cdots & R_{01} \\ R_{02} & R_{02} & \cdots & R_{02} \\ \vdots & \vdots & & \vdots \\ R_{0L} & R_{0L} & \cdots & R_{0L} \end{bmatrix}. \tag{11}$$

Then, according to (3), we have

$$\varphi = \boldsymbol{CS}. \tag{12}$$

Equation (12) provides a constraint condition to estimation $\boldsymbol{C}$ and $\boldsymbol{S}$, which has the $L \times 3$ and $3 \times L$ matrix form, respectively. It is a simple method of implementing the singular value decomposition (SVD) to decompose the matrix φ and reduce the SVD of φ to rank 3. The reduced SVD has the following form where $\sum$ is the eigenmatrix containing the three main eigenvalue of φ.

$$\varphi = \boldsymbol{V} \sum \boldsymbol{U}^T. \tag{13}$$

Then, we can obtain an estimation

$$\boldsymbol{C}_A = V \sum, \quad \boldsymbol{S}_A = U. \tag{14}$$

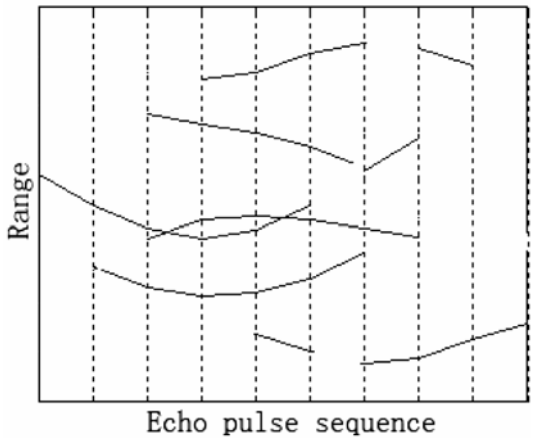
Essentially, $\boldsymbol{S}_A$ and $\boldsymbol{C}_A$ belong to one kind of affine reconstruction of $\boldsymbol{S}$ and $\boldsymbol{C}$. They have affine ambiguity that is given by an arbitrary 3×3 matrix $\boldsymbol{M}$. Also, we have another group affine reconstruction as follows that accords with the condition of (12).

$$\boldsymbol{C}_A = \boldsymbol{CM}, \quad \boldsymbol{S}_A = \boldsymbol{M}^{-1}\boldsymbol{S}. \tag{15}$$

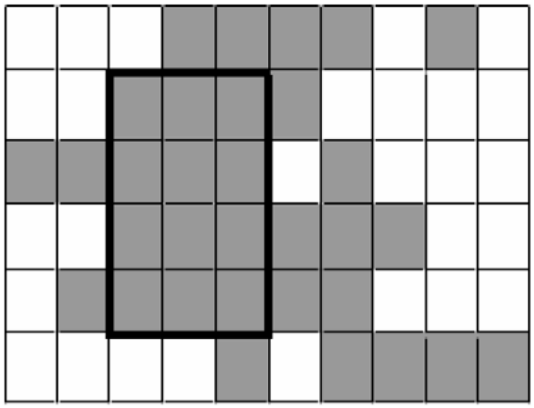
Here, the estimated parameters of $\boldsymbol{R}_0$ and $\boldsymbol{C}_A$ can be combined as the affine motion matrix of $\boldsymbol{P}_A = [\boldsymbol{C}_A \mid \boldsymbol{R}_0]^T$. $\boldsymbol{P}_A$ and $\boldsymbol{S}_A$ can be selected as a group of affine reconstructions of the target's structure and motion.

2.3. Structure from motion

In practice, the key-points are not always visible in every detection. Typically, in computer vision, a point is usually visible in one view and then missed in the latter views due to the phenomenon of obscuration. Concerning this problem, the process of "structure from motion" (SFM) can be taken after the initial reconstruction is estimated. SFM is a typical method in computer vision to gradually estimate the structure and motion parameters when images are added frame-by-frame.



(a) Scatterers' 1D range traces.



(b) Incomplete 1D range matrix.

Figure 4. Model of the complicated 1D range data.

In 1D perception, due to the specific characteristics of the scatterers, their 1D range histories extracted from range images do not hold their completeness during the motion of the target. Figure 4(a) gives an illustration of the incomplete 1D range trace having different emergency and limited durations within a large angular motion, and some of them have missing sections which are not extracted. 1D projected range matrix $\tilde{\varphi}$ has an incomplete formation as Figure 4(b) illustrated, in which the gray elements represent the known data in the matrix and the white elements represent the unknown data. The initial affine reconstruction proposed in Subsection 2.2 can only be applied to the completed part of the original data as the black box showed in Figure 4(b). Then, the rest elements the data can be utilized to recover more scatterers and motion paths using the SFM method, which has a novel application in the 1D-to-3D case.

Let us assume that an affine reconstruction has been initiated as $\boldsymbol{S}_A$ and $\boldsymbol{P}_A$ with N scatterers and L pulses. Then according to (6), we can construct the following equation:

$$R_{ij} = \boldsymbol{p}_{Aj}\tilde{\boldsymbol{X}}_{Ai} = x_{Ai}C_{A1j} + y_{Ai}C_{A2j} + z_{Ai}C_{A3j} + R_{0_j}. \tag{16}$$

As Figure 4(b) has shown, the motion and scatterers of the added data outside the initial box can be extrapolated by the following two main steps:

Firstly, according to (16), an unknown scatterer's affine structure $\boldsymbol{X}_{Ai}$ can be estimated from more than three groups of known motion $\boldsymbol{p}_{Aj}$ s. We assume bundles of range data belonging to one unknown scatterer are collected as vector $\boldsymbol{R}_{\{J\}i}$, where class $\{\boldsymbol{J}\}$ denotes the pulses with known motion. Then $\boldsymbol{X}_{Ai}$ has the following least squares estimation:

$$\boldsymbol{X}_{Ai}^T = (\boldsymbol{R}_{i\{J\}} - \boldsymbol{R}_{0\{J\}})\boldsymbol{C}_{A\{J\}}{}^T(\boldsymbol{C}_{A\{J\}}\boldsymbol{C}_{A\{J\}}{}^T)^{-1}, \tag{17}$$

where $\boldsymbol{R}_{0\{J\}}$ and $\boldsymbol{C}_{A\{J\}}$ are the sub-vector and sub-matrix of $\boldsymbol{R}_0$ and $\boldsymbol{C}_A$, respectively, containing merely the pulses belonging to class $\{\boldsymbol{J}\}$.

Secondly, similarly, an unknown pulse of motion $\boldsymbol{p}_{Aj}$ can be estimated from more than four groups of known scatterers' structure of $\boldsymbol{X}_{Ai}$ s. Also, we assume bundles of range data belonging to one unknown pulse are collected as vector $\boldsymbol{R}_{j\{I\}}$, where class $\{I\}$ denotes the scatterers with known structure. Then $\boldsymbol{p}_{Aj}$ has the following least squares estimation:

$$\boldsymbol{p}_{Aj} = (\widetilde{\boldsymbol{S}}_{A\{I\}}{}^{T}\widetilde{\boldsymbol{S}}_{A\{I\}})^{-1}\widetilde{\boldsymbol{S}}_{A\{I\}}{}^{T}\boldsymbol{R}_{j\{I\}}, \tag{18}$$

where $\widetilde{\boldsymbol{S}}_{A\{I\}}$ is the sub-matrix of $\widetilde{\boldsymbol{S}}_{A}$ containing merely the scatterers belonging to class $\{I\}$.

With the two steps implemented circularly above, more abundant data can be used to reconstruct more scatterers' structures and motion paths.

2.4. Geometry invariance and euclid update

Actually, the low-level reconstructions such as projecting or affine cannot provide the real profile of the target's structure and motion. Usually, euclid reconstruction is essential to recognizing the target. In computer vision, camera calibration is an essential task to obtain the intrinsic and extrinsic parameters of the camera, which is important to attain the euclid update of 3D reconstruction. In 1D sensing systems, the specific perceptive mechanics make the transformation from 1D to 3D simpler. However, as Subsection 2.2 presented, only the affine reconstruction has been obtained based on the geometric condition of (12). Thus, finding new conditions of the moving target is indispensable to realizing the euclid update.

In computer vision, the improvement in camera self-calibration techniques makes it possible to update the reconstruction to euclid without knowing any of the parameters of the camera and its location. In the geometric model presented in Subsection 2.1, this geometry

invariance is actually derived from the features of the view parameter. According to (4), view vector c is a unit vector that has an invariable property as follows:

$$\boldsymbol{c}_j^T \cdot \boldsymbol{c}_j = 1. \tag{19}$$

Thus, for the motion matrix $\boldsymbol{C}$ involving L view aspects, it deduces L equations as follows:

$$\text{diag}(\boldsymbol{C}\boldsymbol{C}^T) = \mathbf{1}_{L\times 1}, \tag{20}$$

where $\text{diag}(\boldsymbol{X})$ represents the vector formed by the diagonal elements of $\boldsymbol{X}$.

Then according to (15), we have the following deduction based on the geometry invariance of (20):

$$\boldsymbol{C}\boldsymbol{C}^T = \boldsymbol{C}_A\boldsymbol{M}^T\boldsymbol{M}\boldsymbol{C}_A^T = \boldsymbol{C}_A\boldsymbol{W}\boldsymbol{C}_A^T, \tag{21}$$

where $\boldsymbol{W} = \boldsymbol{M}^T\boldsymbol{M}$ is a 3×3 symmetric matrix. Next, we construct a vector including the 6 upper triangular elements of $\boldsymbol{W}$ as $w = [W_{11}, W_{12}, W_{13}, W_{22}, W_{23}, W_{33}]^T$ (W_{ij} is the i-th row, j-th column element of matrix W), and form an $L\times 6$ matrix as

$$\boldsymbol{B} = \begin{bmatrix} C_{A1}\cdot C_{A1} \\ 2C_{A1}\cdot C_{A2} \\ 2C_{A1}\cdot C_{A3} \\ C_{A2}\cdot C_{A2} \\ 2C_{A2}\cdot C_{A3} \\ C_{A3}\cdot C_{A3} \end{bmatrix}^T, \tag{22}$$

where $C_{Ak}(k = 1, 2, 3)$ represents the k-th column vector of C_A and "$\cdot$" means the Hadamard product. Thus, we can apply (22) to (20) and then form the following linear equation:

$$\boldsymbol{B}\boldsymbol{w} = \mathbf{1}_{L\times 1}. \tag{23}$$

When $L \geq 6$, the least-square estimation of $\boldsymbol{w}$ has the form of

$$\boldsymbol{w} = \left(\boldsymbol{B}^T\boldsymbol{B}\right)^{-1}\boldsymbol{B}^T\mathbf{1}_{L\times 1}. \tag{24}$$

Then matrix $\boldsymbol{W}$ can be resolved from $\boldsymbol{w}$ and the affine translation matrix $\boldsymbol{M}$ can be resolved as the Cholesky decomposition of $\boldsymbol{W}$.

Finally, the euclid reconstruction can be updated from the affine reconstruction via the following transformations according to (15):

$$\boldsymbol{C}_E = \boldsymbol{C}_A\boldsymbol{M}^{-1}, \quad \boldsymbol{S}_E = \boldsymbol{M}\boldsymbol{C}_A. \tag{25}$$

The estimated parameters of $\boldsymbol{R}_0$ and $\boldsymbol{C}_E$ can be combined as the euclid motion matrix of $\boldsymbol{P}_E = [\boldsymbol{C}_E \mid \boldsymbol{R}_0]^T$. $\boldsymbol{P}_E$ and $\boldsymbol{S}_E$ contain the expected information of the 3D shape and motion of the target with only a rotational ambiguity, i.e., for an arbitrary orthogonal translation $\boldsymbol{R}$, it also has $\boldsymbol{C}_E = \boldsymbol{C}\boldsymbol{R}^T$ and $\boldsymbol{S}_E = \boldsymbol{R}\boldsymbol{S}$.

2.5. Bundle adjustment

As Subsection 2.3 proposed, SFM algorithm is actually an extrapolation method which may cause many cumulated errors in the step-by-step estimations. Accordingly, it is essential to utilizing the optimization techniques in order to make the reconstruction more accurate. In computer vision, the bundle adjustment (BA) technique is well implemented to optimize the whole parameters of structure and motion effectively [21, 22]. In particular, the lens distortion model of the camera is involved in the BA and successfully achieves the nonlinear optimization by using the Levenberg-Marquardt (LM) algorithm. Similarly, in the system of 1D perception, optimization is also required to shrink the cumulated errors produced after the SFM. Meanwhile, according to the geometry model of (2), the nonlinear model also exists in the 1D sensing systems especially for the near-field background. Here, we propose a feasible BA method for the 1D-to-3D reconstruction aiming at improving the quality of the reconstructed results.

Based on the 1D sensing geometry, the geometry model of (2) is always satisfied with a given euclid reconstruction $\boldsymbol{S}_E$ and $\boldsymbol{P}_E$ involving any rotation and translation ambiguity. However, the ambiguities will cause the reconstruction parameters to have infinite solutions. As a result, once the optimization is implemented directly by using $\boldsymbol{S}_E$ and $\boldsymbol{P}_E$, the optimized solutions cannot be converged. Focusing on this problem, new optimized parameters with fixed solutions should be selected. According to (23), we know that, the euclid reconstruction can be completely described by $\boldsymbol{S}_A$, $\boldsymbol{P}_A$, and $\boldsymbol{w}$, which are unique for any rotation transform in $\boldsymbol{M}$. Here, we try to use $\boldsymbol{S}_A$, $\boldsymbol{P}_A$, and $\boldsymbol{w}$ to describe Equation (2), where variables of ρ_{ij} can be presented as $\rho_{ij} = \boldsymbol{c}_{Aj}\boldsymbol{X}_{Ai}$ and $\|\boldsymbol{X}_i\|$ (the distance between $\boldsymbol{X}_{Ei}$ and the original O) as

$$\|\boldsymbol{X}_{Ei}\|^2 = \left\|\boldsymbol{X}_{Ai}^T\boldsymbol{M}^{-1}\right\|^2 = \boldsymbol{X}_{Ai}^T\boldsymbol{M}^{-1}\boldsymbol{M}^{-T}\boldsymbol{X}_{Ai} = \boldsymbol{X}_{Ai}^T\boldsymbol{W}_1\boldsymbol{X}_{Ai}, \tag{26}$$

where $\boldsymbol{W}_1 = \boldsymbol{M}^{-1}\boldsymbol{M}^{-T} = \boldsymbol{W}^{-1}$ is also a 3×3 symmetric matrix. Similarly, we form the 6 upper triangular elements of $\boldsymbol{W}_1$ as vector $\boldsymbol{w}_1$, and construct the following vector from the affine coordinates $\boldsymbol{X}_{Ai} = [x_{Ai},\ y_{Ai},\ z_{Ai}]^T$.

$$\boldsymbol{d} = \left\lfloor x_{Ai}^2,\ 2x_{Ai}y_{Ai},\ 2x_{Ai}z_{Ai},\ y_{Ai}^2,\ 2y_{Ai}z_{Ai},\ z_{Ai}^2 \right\rfloor. \tag{27}$$

Then, from (26), we have

$$\|\boldsymbol{X}_{Ei}\|^2 = \boldsymbol{d}\boldsymbol{w}_1. \tag{28}$$

After constraining the new added parameter $\boldsymbol{w}_1$, the mutual inverse condition of $\boldsymbol{W}\boldsymbol{W}_1 = \boldsymbol{I}$ can be applied to construct the following six equations, denoted as $l_i = 0\,(i = 1, \cdots, 6)$, where

$$\begin{cases} l_1 = w_1 w_{I1} + w_2 w_{I2} + w_3 w_{I3} - 1, \\ l_2 = w_2 w_{I2} + w_4 w_{I4} + w_5 w_{I5} - 1, \\ l_3 = w_3 w_{I3} + w_5 w_{I5} + w_6 w_{I6} - 1, \\ l_4 = w_1 w_{I2} + w_2 w_{I4} + w_3 w_{I5}, \\ l_5 = w_1 w_{I3} + w_2 w_{I5} + w_3 w_{I6}, \\ l_6 = w_2 w_{I3} + w_4 w_{I5} + w_5 w_{I6}. \end{cases} \tag{29}$$

Herein, w_i and w_{1i} are the i-th element of $\boldsymbol{w}$ and $\boldsymbol{w}_1$, respectively.

In conclusion, the parameters need to be optimized including $\boldsymbol{S}_A$, $\boldsymbol{C}_A$, $\boldsymbol{R}_0$, $\boldsymbol{w}$, and $\boldsymbol{w}_1$, which are under the constraints of Equations (2), (23), and (29). We use the penalty algorithm to collect all the constraints and build the following minimized function:

$$\min_{\{X_{Ai}, c_{Aj}, R_{0j}, \boldsymbol{w}, \boldsymbol{w}_I\}} \{ \sum_{i=1,\ldots,N; j=1,\ldots,L} \left\| \sqrt{R_{0j} - 2R_{0j} c_{Aj}^T \boldsymbol{X}_{Ai} + \boldsymbol{d}\boldsymbol{w}_I} - R_{ij} \right\|^2 + \sigma_1 \sum_{k=1}^{L} \|\boldsymbol{B}_k \boldsymbol{w} - 1\|^2 + \sigma_2 \sum_{i=1}^{6} \|l_i\|^2 \}. \tag{30}$$

Herein, $\boldsymbol{B}_k$ represents the k-th row vector of $\boldsymbol{B}$; σ_1 and σ_2 are the penalty coefficients that determine the sanctions of the two conditions. Using the LM algorithm [22], we implement the BA based on the minimized function (27). The initial value of the parameters can be the elementary results after the SFM proposed in Subsection 2.3. Finally, the optimized results should be euclid update as shown in Subsection 2.4.

3. Experimental Results

Among the 1D sensing systems, radar is a typical application which has lots of merits, such as high imaging resolution, long distance detection, and high signal-noise-ratio (SNR). In a radar system, the

feature points of the radar target can be well presented by the scattering centers, some point-like scatterers, which provide the main contribution to the electromagnetic scattering of the radar target. Currently, there are already some mature techniques to extract the range data of the scattering centers in the radar's 1D range images. Thus, it becomes an effective application of our proposed 1D-to-3D reconstruction, which can recover the 3D structure and motion information of the radar target.

Based on the application above, the 1D-to-3D reconstruction algorithm proposed in this paper is evaluated under two scenarios. One is radar detection for the moving target using ground-based radar systems; another is the detection of ground 3D information using moving air-borne radar systems. In both cases, the relative motions between the radar and target involve non-cooperative 3D motions. In our experiments, the complicated 1D range data are simulated and the algorithms we proposed are inspected.

3.1. Simulation 1

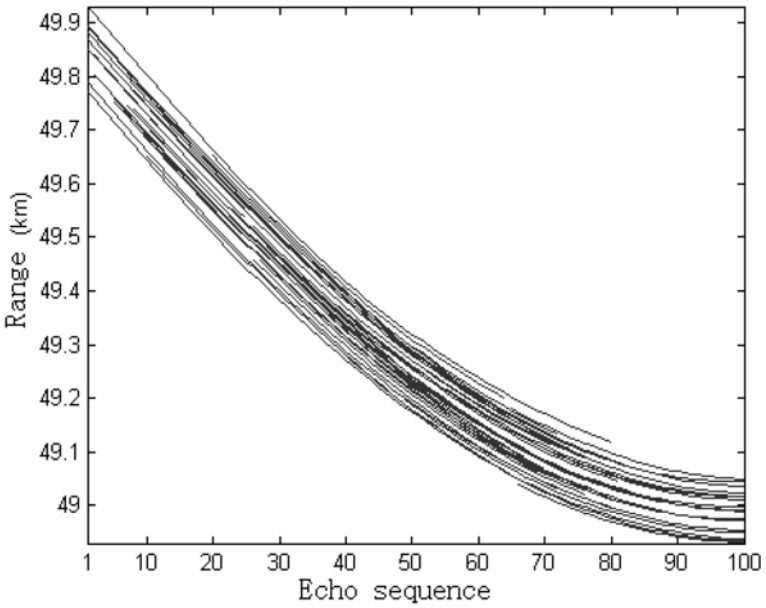
In Simulation 1, the 3D reconstruction for the non-cooperative radar target based on a ground-based radar system is tested. Firstly, as Figure 6(a) illustrated, the structure of the F-22 aircraft is represented as 48 3D points which stand for the scattering centers of the target. Assuming the length of the detecting pulses is $L = 100$, we build the motion model of the target as follows:

(1) Set the pure rotation path of the view angles (α, θ) as the solid curve illustrated in Figure 6(e) where the motion scope of the rotations is below 15^{o}.

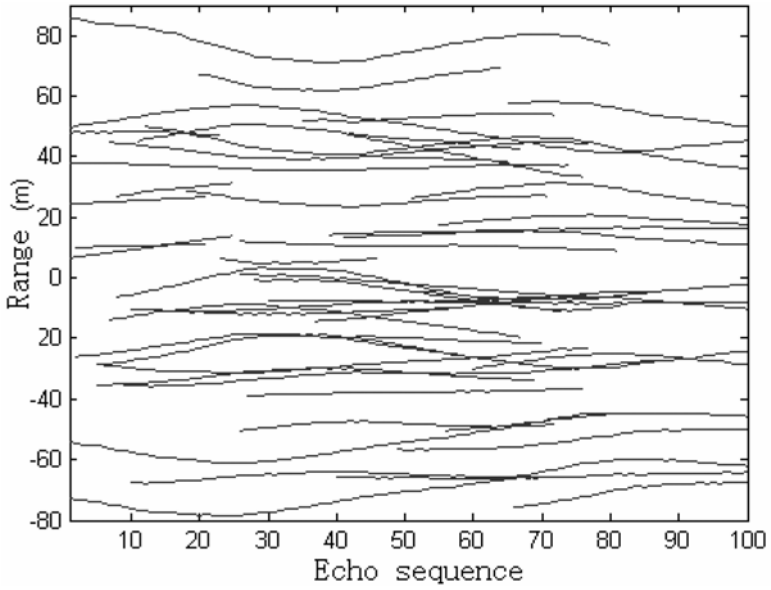
(2) Set the translation motion between the radar and target as the following model:

$$R_{0j} = R_{00} - \Delta R \sin(j\pi / 2L), \quad j = 1, 2, \cdots, L, \tag{31}$$

where the initial range is $R_{00} = 50\text{km}$ and changing scope is $\Delta R = 1\text{km}$.



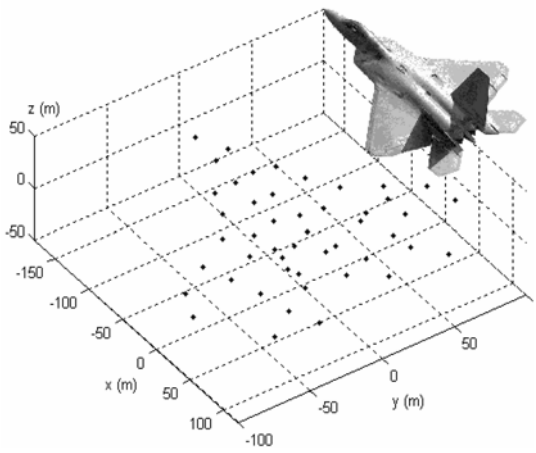
(a) 1D range traces of the scattering centers.



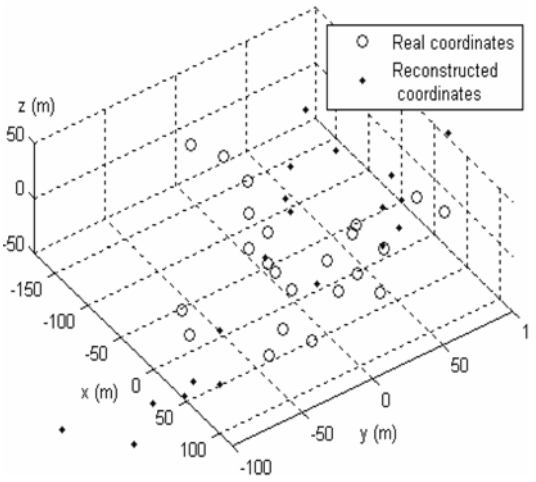
(b) 1D range traces with the translation motion removed.

Figure 5. Simulation of the complicated 1D range data.

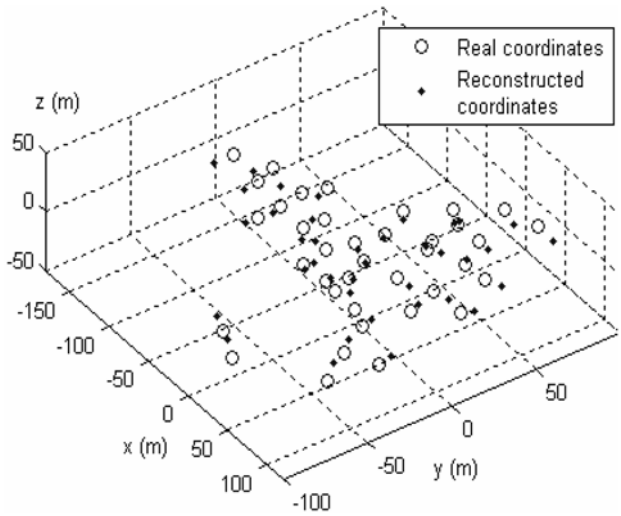
Based on the motion model above, the range data of the scattering centers, which should be extracted from the radar range image in practice, are simulated in our experiments. Figure 5(a) illustrates the simulated range traces of the 48 scattering centers in which some range errors are added. Also, the emergence and the duration of each scattering center's trace are randomly placed. It is clearer to see the simulated range traces in Figure 5(b), where the translation of the target's motion is removed.



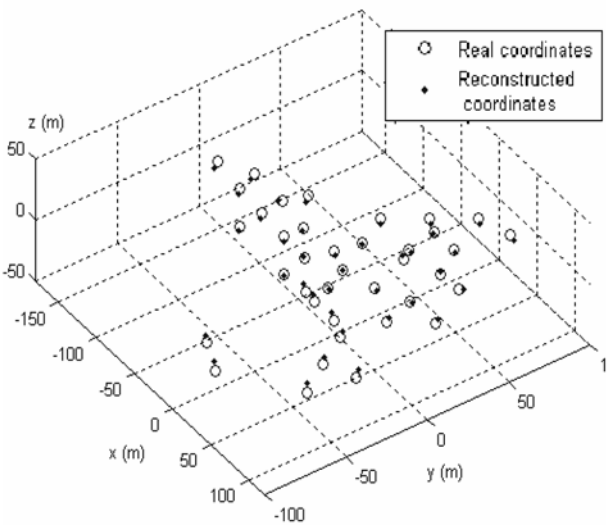
(a) F-22 aircraft model.



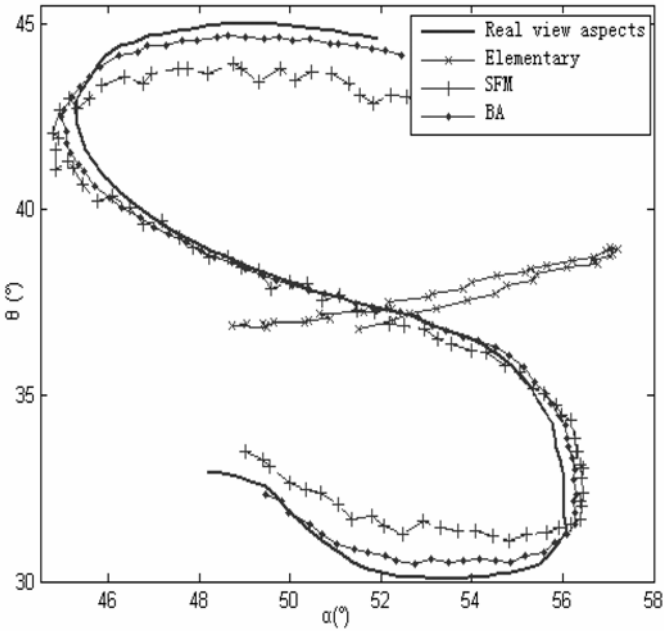
(b) Reconstructed target of the elementary reconstruction.



(c) Reconstructed target of the SFM reconstruction.



(d) Reconstructed target of the BA reconstruction.



(e) Real view aspects of the target and the reconstructed of elementary, SFM and BA reconstruction.

Figure 6. Reconstructed results of the target's structure and motion.

Using the simulated range data, the algorithms of the 1D-to-3D reconstruction proposed in our paper are implemented. Three groups of results of reconstruction are presented in our experiments including:

(1) Elementary reconstruction: We choose the completed part of the original range data and then obtained the initial affine reconstruction as Subsection 2.2 proposed; then updated it to euclid according to Subsection 2.4.

(2) SFM reconstruction: We used the results of initial affine reconstruction to finish the SFM reconstruction as Subsection 2.3 proposed; then updated it to euclid according to Subsection 2.4.

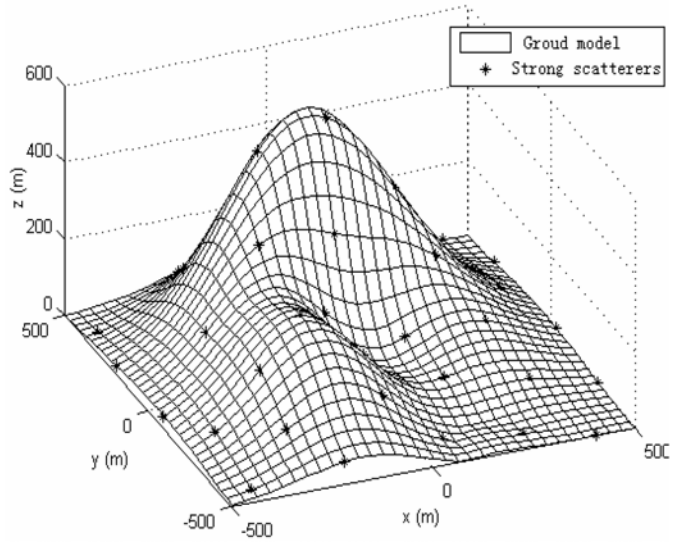
(3) BA reconstruction: We implemented the BA for the reconstructed parameters using the algorithm proposed in Subsection 2.5.

The experimental results are given in Figure 6. As Figure 6(b) and Figure 6(e) have shown, in the elementary reconstruction, less 3D points and motion paths are reconstructed and the reconstructed structure and motion have hug errors compared with the real target model. These poor results mainly attribute to the limited abundance of the range data applied. Thus, it severely impacts the accuracy of the euclid update.

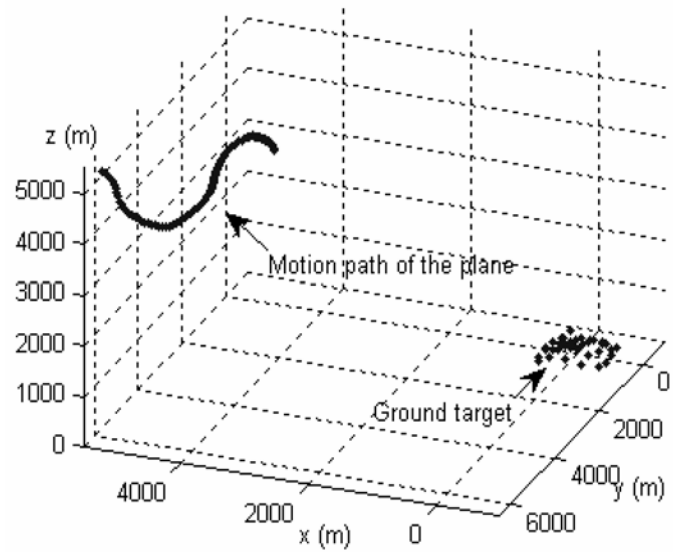
As Figure 6(c) and Figure 6(e) have shown, after the SFM reconstruction, more 3D points and motion paths are reconstructed and, with the abundance of the range data improved, the results of the reconstruction can basically embody the structure and motion characters of the target.

In our experiments, the BA algorithm is successfully implemented with the reconstructed results shown in Figure 6(d) and Figure 6(e). The results verify that, with the SFM and, especially, the BA algorithm applied, the effectiveness and accuracy of the reconstruction can be greatly promoted.

3.2. Simulation 2



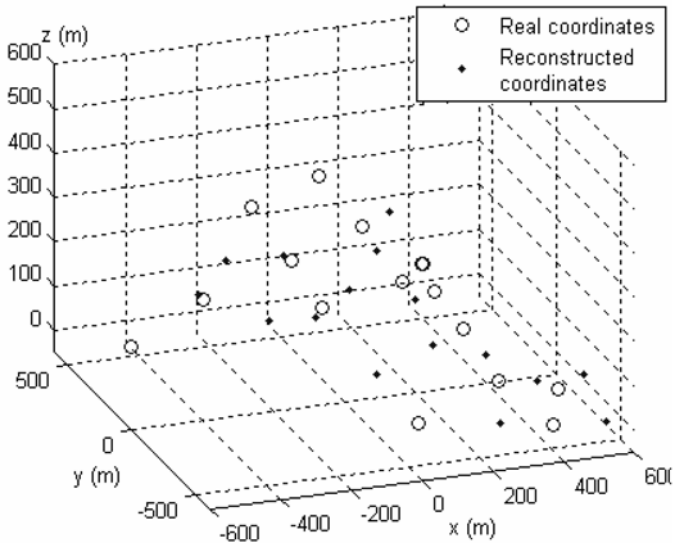
(a) Terrain model.



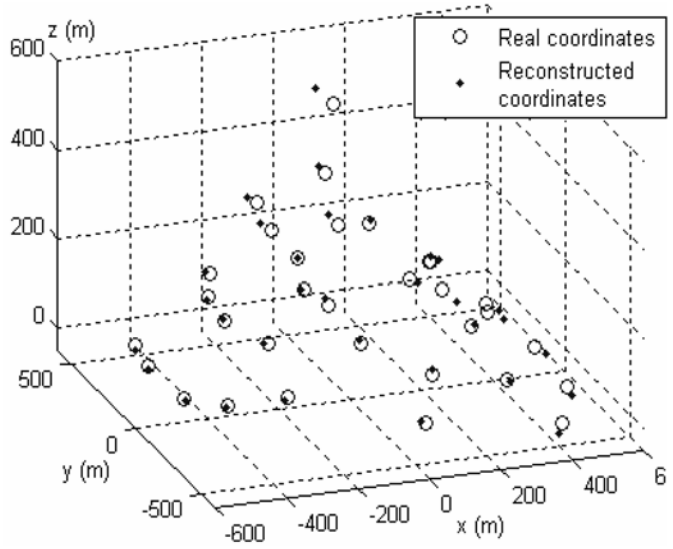
(b) Motion track model of plane.

Figure 7. Targets on the ground and motion track of plane.

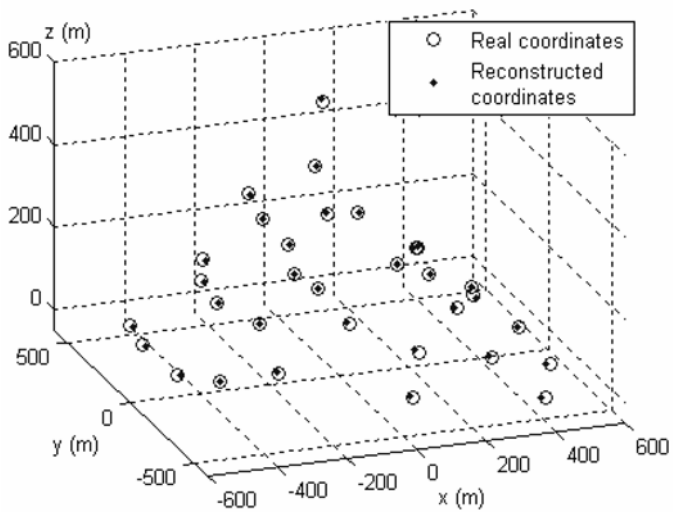
In Simulation 2, another extended application is simulated where the 3D ground information is reconstructed by the airborne radar system. In this case, the terrain is simulated as the curved surface shown in Figure 7(a) where 36 strong scatters exist. The radar system is carried by a maneuvering plane whose motion is simulated based on the model of Simulation 1, too. Based on this motion model, the maneuvering 3D motion trace of the plane is illustrated in Figure 7(b). As is calculated, the maximum near-field range error (i.e., $\max_{i=1\ldots N;\, j=1\ldots L} \left|R_{ij} - (R_{0j} + \rho_{ij})\right|$) is about 27m. In this case, the far-field hypothesis cannot be fulfilled due to a larger size of the ground target and the relatively low height of the plane.



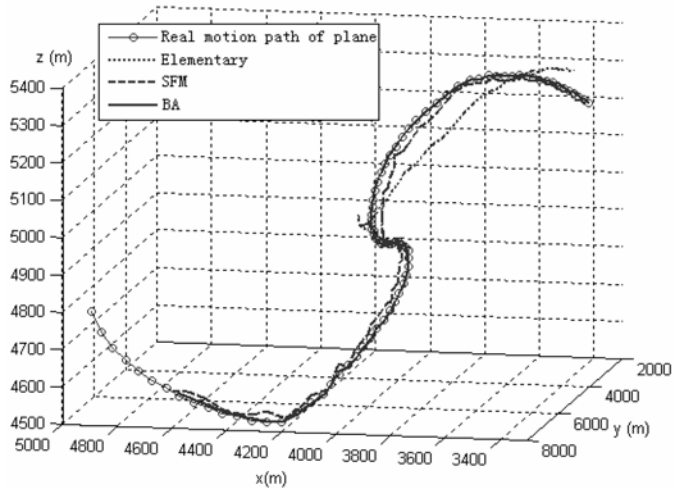
(a) Reconstructed target of the elementary reconstruction.



(b) Reconstructed target of the SFM reconstruction.



(c) Reconstructed target of the BA reconstruction.



(d) Real motion path of the plane and the reconstructed of elementary, SFM and BA reconstruction.

Figure 8. Reconstructed results of target structure and motion path.

Using the same approaches of Simulation 1, the complicated range data are simulated and the 3 groups of experiments of reconstruction (including elementary, SFM, and BA reconstruction) are implemented.

Figure 8 gives the reconstructed results of target structure and motion path of the plane for the 3 groups of experiments, from which we can see the similar improvement of our proposed algorithms. Especially with the near-field model involved in the BA algorithm, the precision of the reconstruction is enhanced greatly. That contributes to the successful applied extension for 3D ground information reconstruction using the airborne radar system.

4. Conclusion

In this paper, we proposed a theoretic extension of 3D reconstruction based on 1D sensing modality. As in the geometric model we analyzed, the 1D-to-3D reconstruction is a special case compared with the 2D-to-3D reconstruction in computer vision. Besides, we exploit key-points extraction and matching, initial reconstruction, structure from motion, euclid update, and bundle adjustment. In the experiments, two typical applications in radar system are simulated. From the experimental results we see that, with the 1D-to-3D reconstruction method applied, the 3D perception of the radar system can be realized without the constraint of motion of the target and the radar-carrier. In addition, this technique we proposed can be applied to most of the 1D sensing systems. The application of it for other 1D sensing systems, such as sonar, laser, ultrasonic etc, will be investigated in our future works.

Acknowledgement

This work is supported by the research and development projects of Science and technology of Hebei Province [Project No.06242188D-2], Natural Science Foundation of Hebei Province [Project No. F2007000221], Main Item of Natural Science Research Plans of The Education Department of Hebei Province [Project No.ZH2006012], and Action Projects of National Science and Technology Personnel Service Enterprise [Project No.2009GJA20043].

References

[1] C. Harris and M. Stephens, A combined corner and edge detector, In Proceedings of the fourth Alvey Vision Conference (1988), 147-151.

[2] C. Choi and H. I. Christensen, Robust 3D visual tracking using particle filtering on the special euclidean group: A combined approach of keypoint and edge features, 31(4) (2012), 498-519.

[3] G. Richard, J. Brown and C. Geoffrey et al., A pointwise smooth surface stereo reconstruction algorithm without correspondences, Image and Vision Computing 30(9) (2012), 619-629.

[4] Z. H. Zhou, J. W. Zhao and F. L. Cao, Surface reconstruction based on extreme learning machine, Neural Computing and Applications 23(2) (2013), 283-292.

[5] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2000.

[6] M. Pollefeys and L. Van Gool, Self-Calibration from the Absolute Conic on the Plane at Infinity, Proceedings of Computer Analysis of Images and Patterns, Lecture Notes in Computer Science, Springer-Verlag, 1997, 1296: 175-182.

[7] W. Triggs, Auto-calibration and the absolute quadric, IEEE Conference on Computer Vision and Pattern Recognition (1997), 609-614.

[8] A. T. Fleury, F. C. Trigo and F. P. R. Martins, A new approach based on computer vision and non-linear kalman filtering to monitor the nebulization quality of oil flames, 40(12) (2013), 4760-4769.

[9] H. H. Qi and M. S. Altinakar, Simulation-based decision support system for flood damage assessment under uncertainty using remote sensing and census block information, Natural Hazards 59(2) (2011), 1125-1143.

[10] M. Favalli, A. Fornaciai and I. Isola et al., Multiview 3D reconstruction in geosciences, Computers & Geosciences 44 (2012), 168-176.

[11] Buddendick Hermann and Thomas F. Eibert, Incoherent scattering-center representations and parameterizations for automobiles, IEEE Antennas and Propagation Magazine 54(1) (2012), 140-148.

[12] J. X. Zhou, H. Z. Zhao and Z. G. Shi et al., Global scattering center model extraction of radar targets based on wideband measurements, IEEE Transactions on Antennas and Propagation 56(7) (2008), 2051-2060.

[13] S. Y. Shin, H. Lim and N. H. Myung, Estimating three-dimensional scattering centers of a target using high resolution techniques, Microwave Conference APMC, (2007), 1-4.

[14] A. Fasoula and P. Genderen, Modelling of extended objects using sparse multi-aspect high range resolution radar data set, IET Radar Sonar Navig. 5(7) (2011), 756-768.

[15] S. Shimizu and H. Fujiyoshi, Keypoint recognition with two-stage randomized trees, IEICE Transactions on Information and Systems E95-D (7) (2012), 1766-1774.

[16] T. Simon and D. Tom, Binary histogrammed intensity patches for efficient and robust matching, International Journal of Computer Vision 94(2) (2011), 241-265.

[17] L. Savy and C. Gaie, Model-based classification of aircraft range profiles using data association algorithms, IEEE Radar Conference (2007), 822-827.

[18] M. Stuff, P. Sanchez and M. Biancala, Extraction of three-dimensional motion and geometric invariants, Multidimensional Systems and Signal Processing 14 (2003), 161-181.

[19] M. Ferrerre, G. Arnold and M. Stuff, Shape and motion reconstruction from 3D-to-1D orthographically projected data via object image relations, IEEE Transaction on Pattern Analysis and Machine Intelligence 31(10) (2009), 1906-1912.

[20] G. H. Wang and Q. M. Jonathan, Quasi-perspective projection model: Theory and application to structure and motion factorization from uncalibrated image sequences, International Journal of Computer Vision 87(3) (2010), 213-234.

[21] B. Triggs, P. McLauchlan and R. I. Hartley, Bundle Adjustment-A Modern Synthesis Vision Algorithm: Theory and Practice, Springer-Verlag, Page 298-375, 2000.

[22] J. More, The Levenberg-Marquardt Algorithm, Implementation and Theory, Numerical Analysis, Lecture Notes in Mathematics, 1977.

■