A Robust Recognition Algorithm for Encoded Targets in Close-range Photogrammetry

REN-BO XIA, JI-BIN ZHAO, WEI-JUN LIU, JIAN-HUANG WU*, SHENG-PENG FU, JUN JIANG AND JIA-ZHI LI
Shenyang Institute of Automation
Chinese Academy of Sciences
Shenyang, 110016 P.R. China
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences
Shenzhen, 518055 P.R. China

In this paper, a robust recognition algorithm for encoded targets in close-range photogrammetry is proposed. Firstly, Canny detector is used to detect edges from an input image. Secondly, the least squares method is employed to fit ellipses to the set of data points yielded by the edge detector. Three restriction criteria based on length, shape, and embedding are applied to restrict the set of candidate ellipses. The initial parameters of a candidate ellipse are modified according to the information of the encoded band pattern. Finally, the identification number of the encoded target is obtained through a certification and interpretation of the arrangement of the bit segments surrounding this encoded target. The proposed algorithm has been applied to a close-range photogrammetric system, and its robustness is validated by real measuring experiments and industrial applications.

Keywords: ellipse extraction, close-range photogrammetry, encoded target, target recognition, restriction criteria

1. INTRODUCTION

Close-range photogrammetry is one of the most important methods to model large-scale three-dimensional objects and scenes [1]. It is common to use easily identifiable encoded targets for multi-view matching of unordered image sets because there are seldom enough feature points on the surface of an object. There are many kinds of encoded targets. Fig. 1 shows the structure of an encoded target which is invariant to rotation, scale, and distortion. Therefore it is the most widely accepted in industrial measurement. The central dot is surrounded by an encoded band with bit segments at equally spaced angular intervals. Each of the bit segments can be either white or left empty (black). Each of targets is encoded with a unique ID. Once encoded targets are recognized correctly, the multi-view matching can be solved quickly and reliably. Therefore, the recognition of encoded target,
as the initial process of reconstructing 3D models from 2D image sequences, plays an important role and directly affects the accuracy of the resulting reconstruction.

Until now, many algorithms [2-5] have been introduced for identifying such kind encoded target shown in Fig. 1. The basics of those algorithms can be described as: edges are first extracted from images and eligible edges that pass through all restriction tests are chosen as the candidate edges of central ellipses to be fitted. The estimation of encoded band patterns can be then carried out directly according to the parameters of the fitting ellipses. Finally, those targets are decoded according to the bit segments of the encoded band pattern. Unfortunately, the above algorithms are sensitive to noise, resolution of central ellipses and shooting angles of views. Particularly, correct results cannot be guaranteed if the following two situations occur. In the first situation, due to discretization error, the parameters of central ellipse might be incorrect when the number of pixels on an eligible edge is not enough to determine accurately the shape of the ellipse. Thus the encoded band pattern estimated by central ellipse is unbelievable and is not suitable to use for decoding the target. In the second situation, the value of one bit that has a value of zero may be determined mistakenly as one due to noise when initial search point is just located in the middle of a bit segment, and vice versa. It is therefore really dangerous to rely on arbitrary selection of initial point for searching bit segments (BP) in decoding process.

In this paper, a robust recognition algorithm is proposed to automatically identify the encoded targets from input images with complex environments. The rest of the paper is organized as follows. In section 2, the recognition algorithm for encoded targets is described in detail. The experimental results are shown in section 3. Finally, a conclusion is given in section 4.

2. RECOGNITION OF ENCODED TARGETS

The proposed algorithm consists of three steps. First, eligible central ellipses are extracted from input images. Second, the corresponding encoded band patterns are obtained from the extracted ellipses, and then are used to modify the parameters of the extracted ellipses. Finally, the encoded targets are decoded through a certification and interpretation of the arrangement of the bit segments surrounding this encoded target.

2.1 Ellipse Extraction for Encoded Target

As we know, the central white dot of the encoded target is imaged to be an ellipse in a
perspective projection. First, edges are extracted by Canny detector which is widely used in computer vision to locate intensity changes and to find object boundaries in an image [6]. Edges produced by Canny detector often contain discontinuities. Different linking techniques have been presented in order to close open edges. In this work, an iterative linking procedure developed in our early work [7] is adopted directly. Next, ellipse fitting is performed using the method of Fitzgibbon et al. [8]. The method is computationally efficient. Then, candidate ellipses are obtained by discarding all ineligible ellipses that cannot pass through restriction tests.

The designed restriction criteria including length criteria, shape criteria, and embedding criteria are described as follows,

(1) **Length criteria:** The edge length of an eligible ellipse should meet the following conditions

\[ L_{\text{low}} \leq L \leq L_{\text{up}} \]  

(1)

where \( L_{\text{low}} \) and \( L_{\text{up}} \) are respectively the minimum length threshold and the maximum length threshold. They are set empirically in advance. If the edge length of an ellipse is too short or too long, the ellipse will be rejected. Further test is not needed any more.

(2) **Shape criteria:** A long and narrow ellipse is considered as an outlier. So the long half-axis \( a \) and short half-axis \( b \) of an ellipse must conform to the following constraint

\[ \frac{a}{b} \leq R_{\text{axis}} \]  

(2)

in which \( R_{\text{axis}} \) is threshold ratio of \( a \) to \( b \). It is generally set as a value greater than two. In addition, ellipticity, as one of the most important individual shape predictor, is defined as the mean distance of points on edges to the fitting ellipse in this paper. The ellipticity of an eligible ellipse must satisfy the following conditions

\[ \varepsilon \leq \varepsilon_{\text{ellipticity}} \]  

(3)

where \( \varepsilon_{\text{ellipticity}} \) is threshold ellipticity. Ideally, \( \varepsilon = 0 \) if pixels on edge are located accurately at the fitting ellipse. But it is impossible due to discretization error.

(3) **Embedding criteria:** An ellipse is called as an embedding ellipse if it contains other ellipses. Embedding ellipse should be discarded. There exist three relative position relations between any two ellipses \( e_i \) and \( e_j \):

- \( e_i \) is an embedding ellipse if \( e_i \) contains \( e_j \). In this case, \( e_j \) is considered as eligible ellipse, as shown in Fig. 2 (a).
- \( e_i \) is an embedding ellipse if \( e_j \) contains \( e_i \), and then \( e_i \) is considered as eligible ellipse, as shown in Fig. 2 (b).
- \( e_i \) and \( e_j \) is both not an embedding ellipse if \( e_j \) does not contain \( e_i \) and \( e_i \) does not contain \( e_j \), and then \( e_i \) and \( e_j \) is both considered as eligible ellipse, as shown in Fig. 2 (c).
Therefore, an embedding matrix can be defined as follows,

$$ M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix}_{n \times n} $$

(4)

where $m_{ij}$ represents the location relationship between ellipse $e_i$ and ellipse $e_j$. $m_{ij}$ is given as follows

$$ m_{ij} = \begin{cases} 1 & \text{if } e_j \in e_i \\ 0 & \text{if } e_i \notin e_j \text{ and } e_j \notin e_i, \text{ or } i = j \\ -1 & \text{if } e_i \in e_j \end{cases} $$

(5)

All of the embedding ellipses can be found by computing the matrix $M$ according to the parameters of ellipses. The $i$th ellipse $e_i$ is rejected if $m_{ij} = 1$, or the $j$th ellipse $e_j$ is rejected if $m_{ij} = -1$. Except embedding ellipses, the rest of ellipses are all candidate ellipses. The matrix $M$ is obtained from Eq. (5) by only computation the upper triangular part of $M$ because $M$ is an anti-symmetric.

### 2.2 Modification of Parameters of Candidate Ellipse

The candidate central ellipse usually can not reflect the real shape. So the position of the encoded target can not be located accurately if the size of ellipse is too small, i.e. there are not enough pixels on the edge of ellipse. The positioning error will be amplified further and even false recognition is likely to occur if the encoded band pattern is estimated directly by using the fitting ellipse and the designed size of the encoded target. Therefore, the parameter modification for candidate ellipse is required to identify accurately the encoded targets.

A candidate central ellipse is shown in Fig. 3. The modification procedure for candidate ellipse is described as follows,

**Step 1:** Compute the internal boundary ellipse $C$ and the outer boundary ellipse $A$ according to the parameters of ellipse $D$ and the size proportion in design. Then compute the middle ellipse $B$ between $A$ and $C$.

**Step 2:** Compute the average intensity $T_f$ for the pixels in the internal region of the central ellipse $D$. $T_f$ is then used as the foreground intensity threshold.
Step 3: Compute the average intensity $T_b$ for the pixels in the region enclosed by the central ellipse $D$ and the internal boundary ellipse $C$. $T_b$ is then used as the background intensity threshold.

Step 4: Define the segmentation threshold $T$ as the mean value of $T_f$ and $T_b$, that is $T = (T_f + T_b)/2$.

Step 5: Set the intensity of each pixel in the internal region of the central ellipse $D$ as $T_b$.

Step 6: Perform thresholding on the internal region of the ellipse $A$ with threshold $T$, and obtain a pixel set $S$ whose elements are with value 1 in the corresponding binary map.

Step 7: Rasterize the ellipse $B$ using the technique developed in [9] and obtain the corresponding edge $E_B$ from input image. Construct the union of the set $S$ and $E_B$ as a new set $E_{SB}$.

Step 8: Fit an ellipse to the set $E_{SB}$ and denote this fitting ellipse by $B'$. Replace the initial ellipse $B$ with $B'$.

Step 9: Iteratively perform the steps 7 and 8 and update the set $B'$. The skeleton of encoded band is well represented by $B'$ after 3-4 iterations.

Step 10: Re-estimate the parameters of the central ellipse $D$ following the size proportion of encoded target in design and get the modified ellipse $D'$.

2.3 Decoding of the Encoded Targets

Each encoded target has a completely different pattern and therefore has a unique ID. Decoding refers to the process of ascertaining the ID of each encoded target. To do this, a decoding scheme is developed as follows,

Step 1: Estimate the new internal boundary ellipse $C'$ and the new outer boundary ellipse $A'$ according to the modified central ellipse $D'$ and the size proportion of encoded target in design, meanwhile, calculate the middle ellipse $B'$ between $A'$ and $C'$. Rasterize the ellipse $B'$ and get the corresponding edge $E_{gb}$.

Step 2: Linearly filter the intensity of each pixel within the set $E_{gb}$ in the edge-normal direction [5].

As shown in Fig. 4, assume that $P_{gb}$ is a pixel on the edge $E_{gb}$ and $OP_{gb}$ is a line through the center $O$ and pixel $P_{gb}$. Let $P_{C'}$ and $P_{A'}$ be the intersection points of the line $OP_{gb}$ and the ellipse $C'$ and the ellipse $A'$ respectively. Search for all pixels on the line segment $P_{C'}P_{A'}$ and sort them with respect to their intensities. Set the middle intensity value as the intensity value for $P_{gb}$. The above scheme is very necessary for reducing the negative effect of image noise.
Step 3: Compute the coordinates of points on the unit circle corresponding to the ellipse $B'$ by the following inverse transform equation

$$x' = \begin{bmatrix} 1/a \\ 1/b \end{bmatrix} \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} (x - x_o)$$

(6)

where $x_o$ is the coordinate of center of ellipse $B'$, $x$ is the coordinate of point $P_B'$ on ellipse $B'$ and $x'$ is the coordinate of point $P_B'$ in unit circle coordinate system. $\alpha$ is the orientation angle of ellipse $B'$.

Step 4: Binarize the pixels on the contour of the unit circle and get a point set $S_b$.

Step 5: Search for the optimal initial point $P_{opt}$ inside set $S_b$ to solve for bit segments in decoding process.

As shown in Fig. 5, assume that $X$ are the index numbers of pixels of the set $S_b$ and $Y$ are the corresponding intensities of pixels of the set $S_b$. Most existing algorithms choose arbitrarily a point as the initial search point for finding the bit segments. To do that is fairly dangerous. For example, the value of one bit that has a value of ‘0’ may be determined mistakenly as ‘1’ due to intensity fluctuation when initial search point is just located in the middle of a bit segment, and vice versa. So the selection of initial search point must be analyzed beforehand in detail. Intuitively, the optimal point should be a point whose right neighbor and left neighbor have very different intensities. Unfortunately, it is difficult to locate precisely this point because the location where intensity signal changes abruptly cannot be observed obviously due to noise.

Thus, a search strategy to find the optimal point is proposed as follows,
(a) Given a point on the unit circle, find its right $k$-nearest neighbor $\{P_{i+k}, \ldots, P_{i+2}, P_{i+1}\}$ and its left $k$-nearest neighbor $\{P_{i-1}, \ldots, P_{i+k-1}, P_{i+k}\}$. Compute the average intensity of the set $\{P_{i+k}, \ldots, P_{i+2}, P_{i+1}\}$ as $T_{pre}$ and the average intensity of the set $\{P_{i-1}, \ldots, P_{i+k-1}, P_{i+k}\}$ as $T_{post}$. The value range for $k$ is defined as

$$1 \leq k \leq \lfloor N/2 \rfloor$$  (7)

where $N$ is the total number of pixels on the unit circle, the notation $\lfloor \rfloor$ means to the nearest integer in the direction of negative infinity. Generally, the bigger $k$ is, the closer the found point is to the optimal point, as well as the longer the compute time is. Empirically, a compromise and reliable selection for parameter $k$ can be adaptively chosen as follows,

$$k = \begin{cases} 7 & \text{if } \lfloor N/2 \rfloor \geq 7 \\ \lfloor N/2 \rfloor & \text{if } \lfloor N/2 \rfloor < 7 \end{cases}$$  (8)

(b) Denote $T_{abs}$ by the absolute difference $T_{pre}$ and $T_{post}$ such that $T_{abs} = |T_{pre} - T_{post}|$.

(c) The point with maximum $T_{abs}$ is regarded as the optimal point $P_{opt}$.

Step 6: Starting at the angle of $\theta$ shown in Fig. 6, the unit circle is divided into 12 bit segments with angular step of 30°. The average intensity value corresponding to each bit segment is used to determine whether the bit segment is a ‘1’ or a ‘0’. To decode the pattern, the binary code is read clockwise. Each bit is considered as the first bit in turn. This means that there are 12 binary numbers to be considered for a 12-bit code. The number corresponding to the code pattern is the lowest among these 12 numbers. For instance, among these 12 binary numbers of the pattern shown in Fig. 7 (a), 000101101001, has the lowest value ($000101101001_2 = 361_{10}$), so the ID of this encoded pattern is labeled as number 361. Last but not least, an encoded target is considered as non-encoded target and must be discarded if its ID doesn’t correspond to any entry in the look-up table built in advance.

Step 7: Output the ID of the recognized encoded target.

3. EXPERIMENTAL RESULTS

In this section, three experiments were conducted to demonstrate the robustness of the
The code of the proposed algorithm was written in VC++ and implemented on a PC (AMD Athlon X2 Dual Core, 2.4 GHz and 2 GB RAM). All images were obtained by the Nikon D200 digital SLR camera with resolution of 3872 × 2592 pixels. In our experiments, we generally choose the parameters as follows: \( L_{\text{low}} = 10 \), \( L_{\text{up}} = 500 \), \( R_{\text{axis}} = 2.5 \) and \( \varepsilon_{\text{ellipticity}} = 0.2 \).

3.1 Recognition Under Perspective Distortion

The first experiment was designed to evaluate the recognition ability of the proposed algorithm under perspective distortion. Fig. 7 shows partly the recognition process on test image. Fig. 7 (a) shows a local image region containing an encoded target. The edge map obtained by Canny detector is illustrated in Fig. 7 (b). The internal boundary ellipse, middle ellipse, and outer boundary ellipse of the encoded band are depicted respectively by blue, red, and green marker in Fig. 7 (c). The ID of the encoded target is recognized as number ‘361’ shown in Fig. 7 (d). This result is consistent with the real value of the encoded target.

3.2 Recognition Under Noise Environment

In the second example, the robustness of the proposed algorithm is further tested under Gaussian noise environment. Fig. 8 (a) shows an initial input image. The recognized results of initial image are illustrated in Fig. 8 (b). To create more legible text annotation, the IDs of encoded targets are marked by red numbers with white background. The recognition of non-encoded targets is not the focus of this work, so the IDs of non-encoded targets are just marked by green numbers with little fonts. As we can see, thirty-five encoded targets from the initial image are completely identified by the proposed algorithm. Deliberately corrupting an image with noise allows us to test the resistance of the proposed algorithm to noise. Fig. 8 (c) shows the recognized results of encoded targets in the initial image stained by Gaussian noise with \( \sigma = 8 \), which are the same as given in Fig. 8 (b). The effect of noise to the proposed algorithm is further assessed by increasing \( \sigma \). As shown in Fig. 8 (d), though the most of encoded targets are recognized accurately by the proposed algorithm, the encoded target with IDs ‘423’ and ‘427’ cannot be identified when the initial image is heavily polluted with the Gaussian noise of \( \sigma = 20 \) and the value of parameter \( k \) is set as 4. This is because the optimal initial point \( P_{\text{opt}} \) in heavy noise image is incorrectly
located so that ‘423’ is determined mistakenly as ‘427’ due to intensity fluctuation when $k$ is small. As a result, there exist two ‘427’, one true and one false, in the same image. In our procedure, two encoded targets with same ID in the same image will be considered as invalid and not reserved for later process because they are completely equivalent and cannot be judged to be true or false due to no prior information is known. The encoded target with ID ‘423’ and ‘427’ are re-identified by increasing the value of parameter $k$ to 7, as shown in Fig. 8 (e). The recognized results using ground truth are given in Fig. 8 (f). This example illustrates that the proposed algorithm is robust to some extent with respect to...
noise if the value of \( k \) is reasonably chosen. Certainly, a large value of \( k \) requires long computation time, see Table 1, where \( t \) is the average computation cost to find the optimal initial point for each encoded target.

### 3.3 An Application for Railway Liquid Tanker

An important advantage of the close-range photogrammetry is that this technique is appropriately used for measuring large-scale objects such as building, large casting object, car, railway liquid tanker and so on. This experiment demonstrates a real application for measurement of railway liquid tanker to assess the capability of the proposed algorithm for dealing with a relatively large object in natural environment. Encoded targets are placed directly on the structure of the tanker. This means that each target has a specific place on the tanker. These targets are magnetic so that they can be placed and removed easily. All encoded targets have to be photographed from different directions and also different heights. These targets are not only necessary as measuring points but are also used to make multi-view matching easier by using macros. Homography matrix can be estimated via corresponding encoded targets from image pairs. The correspondences of non-encoded targets from multiple views can be directly recovered from the homography matrix. Once enough encoded and non-encoded targets are recognized and matched the software bundles these points from multiple views, and all positions are calculated. As a result, a point cloud is created. Using these points, the 3D surface of railway liquid tanker can be obtained via a surface-fitting algorithm.

Figs. 9 (a) and (b) show the recognized results of two images among 156 images acquired in this experiment. The recognized results are then used as an input of a close-range photogrammetric system developed by our team to reconstruct the 3D coordinates of en-

![Fig. 9. An application for railway liquid tanker.](image)
coded targets and non-encoded targets, as shown in Fig. 9 (c). Fig. 9 (d) shows the recovered surface of railway liquid tanker by using the well-known NURBS fitting method. The volume error of railway liquid tanker is estimated and compared with the international standard error. The final result shows that our mean value of volume error (± 0.25%) is very close to that (± 0.2%) provided by international organization. This result shows that the proposed recognition method can provide an effective alternative solution for practical industry measurement.

4. CONCLUSION

In the field of close-range photogrammetry, alongside the accuracy of measurement results, robustness of measurement procedures is indispensable for industrial application. In this paper, we have presented a robust recognition algorithm for encoded target. The robustness of the proposed algorithm is tested and evidenced under perspective distortion and noise environments. The proposed algorithm has been embedded into a close-range photogrammetric system and has the potentiality of being used in industry measurements. Furthermore, we expect to make the proposed algorithm more accurate and efficient and apply our method to other measurement of objects.

REFERENCES

Ren-Bo Xia (夏仁波) received his B.S. and M.S. degrees in Aerospace Engineering and Mechanics from Harbin Institute of Technology and Engineering Mechanics from Harbin Institute of Technology, Harbin, China, in 2000 and 2002 respectively, and his Ph.D. degree in Mechanical Electronic Engineering from Graduate University of Chinese Academy of Science, Beijing, China, 2006. He has been an Associate Research Fellow of Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China, since January 2009. His research interests are in computer vision and pattern recognition.

Ji-Bin Zhao (趙吉賓) received his Bachelors in Mechanical Engineering at Hefei University of Technology in 1996, his Masters from Shandong University in 2000 and his Ph.D. from Graduate School of the Chinese Academy of Science in 2004, respectively. His research interests include computer-aided design, rapid prototyping and reverse engineering. He is an Associate researcher in Shenyang Institute of Automation, Chinese Academy of Science.

Wei-Jun Liu (劉偉軍) received his Bachelors and Master in Mechanical Engineering at Shenyang University of Technology in 1992 and 1995, respectively, his Ph.D. from Dalian University of Technology in 1998. His research interests include rapid manufacturing and reverse engineering. He is a researcher in Shenyang Institute of Automation, Chinese Academy of Science.

Jian-Huang Wu (吳劍煌) received his Ph.D. degree from Shenyang Institute of Automation, Chinese Academy of Sciences in 2007. He is currently an Assistant Professor at the Center for Human-Computer Interaction, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His research interests include medical visualization, geometric modeling and computer graphics.

Sheng-Peng Fu (付勝鵬) received his B.E. degree from Shandong University of Logistics Engineering in China. He is pursuing his Ph.D. degree in Shenyang Institute of Automation (SIA), Chinese Academy of Sciences. His research interests include CAD/CAM, multi-axis NC machining and compute vision.

Jun Jiang (姜軍) received his B.E. degree from Harbin Institute of Technology (HIT), China, in 2007. He is pursuing his Ph.D. degree in Shenyang Institute of Automation (SIA), Chinese Academy of Sciences. His research interests include robotics, theory of control, aviation and navigation.

Jia-Zhi Li (李家智) received his B.E. degree from Yanshan University of Automation in China. He is pursuing his M.S. degree in Shenyang Institute of Automation (SIA), Chinese Academy of Sciences. His research interests include CAD/CAM, multi-axis NC machining and compute vision.