Proceedings of the 5th World Congress on Intelligent Control
and Automation, June 15-19, 2004, Hangzhou, P.R. China

第五届全球智能控制与自动化大会
2004 年 6 月 15 - 19 日，中国杭州

Ellipse Detection Based on Improved-GEVD Technique*

Zhong-Gen Yang, Gui-Xiang Jiang, Lei Ren
Department of Electronic Engineering
Shanghai Maritime University
Shanghai, 200135,China
zgyang@cen.shmtu.edu.cn

Abstract - The standard Generalized Eigen Value Decomposition (GEVD) is a popular ellipse detection technique whose statistical analysis is given to prove its disadvantages of very big estimation bias and MSE. It is also proved that the effective measurement to improve the performance of ellipse detection is whitening the data noise and regularizing data observation. This theoretic analysis has strongly supported the Hartley's regularization method. Then, an improved-GEVD algorithm was developed. The theoretical analysis and computer simulation experiments have demonstrated that the proposed technique has the advantages that it is intrinsically able to whiten the data noise and to regularize the data observation so as to output a non-biased estimation of ellipse parameter with very small MSE. Furthermore, the computation complex is largely simplified.

Index Terms - Robot vision, Ellipse detection, GEVD, Regularization.

I. Introduction

Ellipse extraction plays an important role in image recognition and robot vision, because ellipse is the most simple and most popular curve geometrical prime. A typical application example is the man-made target recognition and tracking in the underwater environment completed by an Autonomous Underwater Vehicle [1]. The key step in ellipse extraction is ellipse detection, i.e. ellipse fitting. In the long time, it interested the academic field [2, 3, 4, 5, 6, 7].

The traditional method is the standard Generalized Eigen Value Decomposition (GEVD) technique in which the optimal estimation of ellipse parameter vector is the eigen vector corresponding to the smallest generalized eigen value of the matrix pencil composed of data ACF matrix and parameter constraint matrix. But, it has been shown that the practical performance of the standard GEVD (SGEVD) is poor, especially when it is used to fit a digital ellipse. Hartley[8] has intuitively considered that the main reason is the too big condition number of the data ACF matrix. Based on this viewpoint, he has suggested a regularization measurement. Although the performance is greatly improved, the theoretical proof isn't given out. The statistical analysis of GEVD ellipse fitting algorithm has also not been seen yet.

By means of the statistical performance analysis, we has proved in theory that the poor performance of SGEVD is resulted from the color noise contained in the data vector and the big condition number of the data ACF matrix, and has pointed out that the two basic measurements to improve the performance are to whiten the data noise and to regularize the data ACF matrix. As the regularization transformation suggested by Hartley has strong regularization ability and can whiten data vector in some degree, so the GEVD algorithm fusing Hartley's regularization technique (abbreviated to HGEVD) has a good performance. Thus, our analysis has strongly supported HGEVD in theory and in practice. In the same time, we also has pointed out that after Hartley's regularization transformation the data noise is still color so that HGEVD can just give a biased estimation of the ellipse parameter vector. Then we develop an improved GEVD algorithm that intrinsically has the ability to whiten the data noise and to regularize the data ACF matrix on the condition of not carrying any whitening transformation or regularization transformation. In addition, the dimension number of the optimization procedure is reduced from 6 to 2. So, IGEVD can give estimation slightly super over or the same as HGEVD, on the condition that the computation complex and realization cost have been greatly reduced. The experiments have demonstrated our analysis.

II. GEVD-Based Ellipse Fitting

It is well known that a conic can be described by the following algebraic equation

$$\mathbf{D}^T\boldsymbol{\theta} = 0 \tag{1}$$

where $\mathbf{D} = [x^2 \quad xy \quad y^2 \quad x \quad y \quad 1]^T$ is the data vector, $\boldsymbol{\theta} = [a_0 \quad a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5]^T$ is the parameter vector, superscript T, denotes transposition. Ellipse is a conic constrained by

$$\boldsymbol{\theta}^T\mathbf{C}'\boldsymbol{\theta} = 1 \tag{2}$$

where $\mathbf{C}' = \begin{bmatrix} \mathbf{C} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{bmatrix}$。

* This work is supported by project 01G02 of Science and Technology Development Fund Program for Universities in Shanghai.

0-7803-8273-0/04/$20.00 ©2004 IEEE

Ellipse fitting is a procedure that fits ellipse equation to the given 2D point set $\{\mathbf{p}_i = [x_i \quad y_i]^T \mid i = 1,2,\cdots,N\}$ by means of a least square optimization so as to minimize the cost function

$$J = \boldsymbol{\theta}^T \mathbf{R}_{\mathbf{D}}^{(s)} \boldsymbol{\theta} - \lambda(\boldsymbol{\theta}^T \mathbf{C}' \boldsymbol{\theta} - 1) \tag{3}$$

where
$$\mathbf{R}_{\mathbf{D}}^{(s)} = (\sum_{i=1}^{N} \mathbf{D}_i \mathbf{D}_i^T)/N \tag{4}$$

is the ACF matrix of data observation, superscript (s) denotes sample average operation.

It is easy to prove that the minimization of (3) leads us to the well-known standard GEVD procedure, in which the optimal estimation $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$ is the generalized eigen-vector corresponding to the minimal generalized eigen-value $\lambda_{\min} = \hat{\boldsymbol{\theta}}^T \mathbf{R}_{\mathbf{D}}^{(s)} \hat{\boldsymbol{\theta}}$ of the matrix pencil $(\mathbf{R}_{\mathbf{D}}^{(s)}, \mathbf{C}')$, where

$$\mathbf{R}_{\mathbf{D}}^{(s)} \hat{\boldsymbol{\theta}} = \lambda_{\min} \mathbf{C}' \hat{\boldsymbol{\theta}} \tag{5}$$

III. Statistical Analysis of GEVD algorithm

A. Statistical model

Suppose that the noisy observation of a 2D point $\mathbf{p}_{ti}$ is

$$\mathbf{p}_i = \mathbf{p}_{ti} + \mathbf{n}_{pi} \tag{6}$$

where $\mathbf{n}_{pi}$ is the additive observation noise, which is an i.i.d. Gaussian white noise with zero mean and is independent on the truth-value $\mathbf{p}_{ti}$, so that

$$\begin{cases} E(\mathbf{n}_{pi}\mathbf{n}_{pj}^T) = \sigma_n^2 \delta_{ij} \mathbf{I} \\ E(\mathbf{n}_{pj}) = \mathbf{0} \\ E(\mathbf{p}_{ti}\mathbf{n}_{pj}^T) = \mathbf{0} \end{cases} \tag{7}$$

where $E(\ )$ denotes mathematical expectation , $\delta_{ij} = \begin{cases} 1 & \forall i = j \\ 0 & else \end{cases}$, $\mathbf{I}$ is the unit matrix. So, the corresponding noisy data vector can be described by

$$\mathbf{D}_i = \mathbf{D}_{ti} + \mathbf{n}_{\mathbf{D}i} \tag{8}$$

where $\mathbf{D}_{ti}$ is the truth value of the data vector and the additive noise vector is

$$\mathbf{n}_{\mathbf{D}i} = \begin{bmatrix} (x_{ti} + n_{xi})n_{xi} \\ x_{ti}n_{yi} + y_{ti}n_{xi} \\ (y_{ti} + n_{yi})n_{yi} \\ n_{xi} \\ n_{yi} \\ 0 \end{bmatrix} \tag{9}$$

with the mean

$$\mu_{\mathbf{n}_\mathbf{D}} = E(\mathbf{n}_{\mathbf{D}i}) = \sigma_n^2 [1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0]^T .$$

By means of the above equations, we have

$$\mathbf{R}_{\mathbf{D}}^{(s)} = \mathbf{R}_{\mathbf{D}t}^{(s)} + \Delta\mathbf{R}_{\mathbf{D}}^{(s)} \tag{10}$$

where $\mathbf{R}_{\mathbf{D}t}^{(s)}$ is the truth value of $\mathbf{R}_{\mathbf{D}}^{(s)}$ and

$$\Delta\mathbf{R}_{\mathbf{D}}^{(s)} = \mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(s)} + \mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(s)},$$
$$\begin{cases} \mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(s)} = [\sum_{i=1}^{N} (\mathbf{D}_{ti}\mathbf{n}_{\mathbf{D}i}^T + \mathbf{n}_{\mathbf{D}i}\mathbf{D}_{ti}^T)]/N \\ \mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(s)} = (\sum_{i=1}^{N} \mathbf{n}_{\mathbf{D}i}\mathbf{n}_{\mathbf{D}i}^T)/N \end{cases} \tag{11}$$

Their mathematical expectations are

$$\mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}} = E(\mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(s)}) = \begin{bmatrix} \mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(11)} & \mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(12)} \\ (\mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(12s)})^T & \mathbf{0} \end{bmatrix}$$

$$\mathbf{R}_{\mathbf{D},\mathbf{n}_\mathbf{D}}^{(12)} = \sigma_n^2 \begin{bmatrix} \mu_x & \mu_y & 0 \\ \mu_y & \mu_x & 0 \\ \mu_x & 3\mu_y & 0 \end{bmatrix}, \begin{cases} \mu_x = E(x) \\ \mu_y = E(y) \end{cases}$$

$$\mathbf{R}_{\mathbf{n}_\mathbf{D}} = E(\mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(s)}) = \begin{bmatrix} \mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(11)} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(11)} \end{bmatrix}$$

$$\mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(11)} = \sigma_n^4 \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{R}_{\mathbf{n}_\mathbf{D}}^{(22)} = \sigma_n^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus the data noise is a non-independent color noise with nonzero mean and with a non-diagonal ACF matrix.

B. Statistical performance analysis

The estimation error of the parameter vector $\boldsymbol{\theta}$ is $\delta\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_t$ where $\boldsymbol{\theta}_t$ is the truth-value of $\boldsymbol{\theta}$. It is can be proved that $\mathbf{R}_{\mathbf{D}t}\boldsymbol{\theta}_t = \mathbf{0}$ and $\mathbf{R}_{\mathbf{D}t}^{+}\boldsymbol{\theta}_t = \mathbf{R}_{\mathbf{D}t}^{+}\mathbf{C}'\boldsymbol{\theta}_t = \mathbf{0}$. So, by means of matrix disturbing theory, it is can be proved that

$$\delta\boldsymbol{\theta} = -\mathbf{W}\mathbf{R}_{\mathbf{D}t}^{+}\Delta\mathbf{R}_{\mathbf{D}}^{(s)}\boldsymbol{\theta}_t \tag{12}$$

where

$$\begin{aligned} \mathbf{W} &= [\mathbf{I} + \mathbf{R}_{\mathbf{D}t}(\Delta\mathbf{R}_{\mathbf{D}}^{(s)} - \lambda_{\min}\mathbf{C}')]^{-1} \\ &\approx [\mathbf{I} + \mathbf{R}_{\mathbf{D}t}(\Delta\mathbf{R}_{\mathbf{D}} - \lambda_{\min}\mathbf{C}')]^{-1} \end{aligned} \tag{13}$$

is the weight matrix. Therefore, the estimation bias is

$$E(\delta\boldsymbol{\theta}) = -\mathbf{W}\mathbf{R}_{\mathbf{D}t}^{+}\mathbf{R}_{\mathbf{n}_\mathbf{D}}\boldsymbol{\theta}_t \tag{14}$$

Its strength is

$$\begin{aligned} \|E(\delta\boldsymbol{\theta})\|/\|\boldsymbol{\theta}_t\| &\le w\|\mathbf{R}_{\mathbf{D}t}^{+}\| \cdot \|\Delta\mathbf{R}_{\mathbf{D}}\| \\ &= wK(\mathbf{R}_\mathbf{D})(1 + \rho\sqrt{SNR})/SNR \end{aligned} \tag{15}$$

where

$$\begin{aligned} w &= 1/(1 - \|\mathbf{R}_{\mathbf{D}t}^{+}\|\|\Delta\mathbf{R}_{\mathbf{D}} - \lambda_{\min}\mathbf{C}'\|) \\ &= 1/[1 - K(\mathbf{R}_\mathbf{D})\|\Delta\mathbf{R}_{\mathbf{D}} - \lambda_{\min}\mathbf{C}'\|/\|\mathbf{R}_{\mathbf{D}t}\|] \end{aligned} \tag{16}$$

Here the condition number of matrix $\mathbf{R}_\mathbf{D}$ is defined as

$$\begin{aligned} K(\mathbf{R}_\mathbf{D}) &= K(\mathbf{R}_{\mathbf{D}t}) \\ &= \|\mathbf{R}_{\mathbf{D}t}^{+}\| \cdot \|\mathbf{R}_{\mathbf{D}t}\| = (\lambda_{\max} - \lambda_{\min})/(\lambda_{s\min} - \lambda_{\min}) \end{aligned} \tag{17}$$

where $\lambda_{\max}$, $\lambda_{\min}$ and $\lambda_{s\min}$ are the maximal, minimal, and second minimal eigen values of matrix $\mathbf{R}_\mathbf{D}$, respectively. The SNR of the data vector is defined as

$$SNR = \|\mathbf{R}_{\mathbf{D}t}\|/\|\mathbf{R}_{\mathbf{n}_\mathbf{D}}\| \tag{18}$$

The relative correlation coefficient between data and noise is defined as

$$\rho = \|\mathbf{R}_{D,n_D}\| / \sqrt{\|\mathbf{R}_{n_D}\| \|\mathbf{R}_{Dr}\|} \quad (19)$$

After mathematical derivation, it can be proved that the indeterminacy of the parameter estimation is

$$\sigma_{\delta\theta} / \|\theta_t\| = \sqrt{tr(\Sigma_{\delta\theta})} / \|\theta_t\| \le 2w * K(\mathbf{R}_D)\sqrt{(1+\rho^2)/SNR} / \sqrt{N} \quad (20)$$

where $\Sigma_{\delta\theta}$ is the variance matrix of the parameter estimation error. It can also be seen that the mean of the ellipse fitting error is

$$\begin{aligned} E(\lambda_{\min}) / \left[N\|\theta_t\|^2\right] &= \theta_t^T \Delta\mathbf{R}\theta_t / \left[N\|\theta_t\|^2\right] \le \|\Delta\mathbf{R}\| / N \\ &\le \left(\|\mathbf{R}_{n_D}\| + \|\mathbf{R}_{D,n_D}\|\right) / N \\ &= \|\mathbf{R}_{n_D}\|\left(1 + \rho\sqrt{SNR}\right) / N \end{aligned} \quad (21)$$

and the indeterminacy of the ellipse fitting error is

$$\begin{aligned} \sigma_{\lambda_{\min}} / \|\theta_t\| &\le \|\mathbf{R}_{n_D}\| \sqrt{4(1+\rho^2)SNR + 1} / N \\ &\approx 2\|\mathbf{R}_{n_D}\| \sqrt{(1+\rho^2)SNR} / N \end{aligned} \quad (22)$$

If the data vector set is de-correlated so that $\mathbf{R}_{D,n_D} = \mathbf{0}$, in turn that $\rho = 0$, then the strength of the estimation bias, the indeterminacy of the parameter estimation and the indeterminacy of the ellipse fitting error can be greatly reduced. Furthermore, when $\Delta\mathbf{R}_D^{(s)} = \sigma_{n_D}{}^2\mathbf{I}$ replaced to (12),

$$\delta\theta = -\sigma_{n_D}{}^2 \mathbf{R}_{Dr}{}^+ \theta_t = \mathbf{0} \quad (23)$$

Therefore, if the additive noise of the data vector set is pre-whitened, the parameter vector estimation can be unbiased.

From the above analysis, the following conclusions can be made.

- The standard GEVD (SGEVD) technique of ellipse fitting cannot give us an unbiased estimation of the parameter vector. Both of bias strength and estimation indeterminacy are directly proportional to the condition number $K(\mathbf{R}_D)$ of data ACF matrix. The higher the data's SNR is and the larger the N is, the smaller the estimation indeterminacy is.
- The lower the noise power is and the larger the N is, the smaller the error strength and its indeterminacy of the ellipse fitting are.
- Pre-whitening processing of the data vector can not only efficiently eliminate the estimation bias of the parameter vector to give us an unbiased estimation, but also greatly reduce the fitting error strength, the parameter estimation indeterminacy and the fitting error indeterminacy.
- The regularization to greatly reduce the condition number $K(\mathbf{R}_D)$ of data ACF matrix or the denoising filtering to increase the data's SNR can efficiently reduce both of the fitting error and its indeterminacy.
- The increase of the point number N helps the decrease of the estimation indeterminacy as well as the fitting error indeterminacy to enhance the estimation precision and the fitting precision.
- The effect of the constraint on the optimization procedure is embodied in the weight matrix $\mathbf{W}$, in other words, in the weight factor w. Because $w > 1$ and both of $\Delta\mathbf{R}_D$ and $\lambda_{\min}$ are infinitesimal, therefore w is slight bigger than 1. So, the fitting precision will be little reduced.

III. HARTLEY'S REGULARIZATION GEVD ALGORITHM

A great deal of practice has proved that when the SGEVD algorithm is employed to fit a digital ellipse, its performance is very poor. Hartley, when analyzing its reason, has understood that the big difference in the scatter of values between the homogeneous coordinates of 2D digital points must lead the condition number of the data ACF matrix extremely big so that the estimation bias and the estimation indeterminacy are both insufferably big. He has suggested that the pre-whitening processing of 2D point set can be utilized to efficiently reduce the condition number to a rather small level. This is a rather simple and effect regularization procedure. Specifically, he used the linear regularization transformation of the 2D point set

$$\tilde{\mathbf{p}}_i = \mathbf{L}\left(\mathbf{p}_i - \left(\sum_{i=1}^{N} \mathbf{p}_i\right) / N\right) \quad (24)$$

so that the regularized 2D point $\tilde{\mathbf{p}}_i$ has zero mean and unit variance $\mathbf{R}_{\tilde{p}} = \Sigma_{\tilde{p}} = \mathbf{L}\Sigma_p\mathbf{L}^T = \mathbf{I}$. So, the regularization transformation matrix must be satisfied to

$$\Sigma_p = \mathbf{L}^{-1}\mathbf{L}^{-T} \quad (25)$$

From the mapping relationship from the point set to the data vector set $\mathbf{D}$, we have the regularization transformation for the data vector set

$$\tilde{\mathbf{D}} = \mathbf{L}_D\mathbf{D} \quad (26)$$

where the regularization transformation matrix for the data vector can be analytically determined from $\mathbf{L}$ and $\overline{\mathbf{p}}$.

It is can be shown that, after regularization transformation, the condition number $K(\tilde{\mathbf{R}}_{\tilde{D}})$ of the ACF matrix $\tilde{\mathbf{R}}_{\tilde{D}}$ of the data vector is in the order of $O(1)$ so that the estimation bias strength and indeterminacy can be greatly reduced. It is should be noticed that, when $\hat{\tilde{\theta}}$ is the generalized eigen vector corresponding to the minimal generalized eigen value $\tilde{\lambda}_{\min} = \hat{\tilde{\theta}}^T\tilde{\mathbf{R}}_{\tilde{D}}\hat{\tilde{\theta}}$ of the matrix pencil $(\tilde{\mathbf{R}}_{\tilde{D}}, \mathbf{C}')$, that is, when

$$\tilde{\mathbf{R}}_{\tilde{D}}\hat{\tilde{\theta}} = \tilde{\lambda}_{\min}\mathbf{C}'\hat{\tilde{\theta}} \quad (27)$$

the corresponding optimal estimation of the parameter vector is $\hat{\theta} = \mathbf{L}_D{}^T\hat{\tilde{\theta}}$.

Actually, the optimal estimation of the ellipse parameter vector does not be the generalized eigen of the matrix pencil $(\tilde{\mathbf{R}}_{\tilde{D}}, \mathbf{C}')$ at all, because

$$\begin{aligned} \mathbf{R}_D\hat{\theta} &= \left(\mathbf{L}_D{}^{-1}\tilde{\mathbf{R}}_{\tilde{D}}\mathbf{L}_D{}^{-T}\right)\left(\mathbf{L}_D{}^T\hat{\tilde{\theta}}\right) \\ &= \tilde{\lambda}_{\min}\mathbf{L}_D{}^{-1}\mathbf{C}'\hat{\tilde{\theta}} = \tilde{\lambda}_{\min}\mathbf{L}_D{}^{-1}\mathbf{L}_D{}^{-T}\mathbf{C}'\hat{\theta} \ne \lambda\mathbf{C}'\hat{\theta} \end{aligned},$$

which has explained the reason why the Hartley's regularized

GEVD (HGEVD) can have the effect ultimately different from the SGEVD.

It must be pointed out that, because Hartley's whitening is made for the 2D point set rather than for the noise component of the data vector, its parameter vector estimation is still biased. In addition, its computation complex is a little higher.

IV. Improved GEVD (IGEVD) Ellipse Fitting

In section 2, it is proved that the method to give an unbiased estimation is to whiten the noise component of data vector. But, the prerequisite that requires its mean and variance must be known a priori is very difficult to be satisfied in practice.

In order to overcome this difficulty, we firstly centralize the data vector, i.e. (1) so that

$$\overset{o}{\mathbf{D}}^T \boldsymbol{\theta} = \begin{bmatrix} \overset{o}{\mathbf{D}}_1^T & \overset{o}{\mathbf{D}}_2^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} = 0 \qquad (28)$$

where

$\overset{o}{\mathbf{D}}_1 = \left[(\overset{o}{x^2}) \;\; 2(\overset{o}{xy}) \;\; (\overset{o}{y^2}) \right]^T$, $\overset{o}{\mathbf{D}}_2 = \left[2\overset{o}{x} \;\; 2\overset{o}{y} \right]^T$, $\boldsymbol{\theta}_1 = [a_0 \;\; a_1 \;\; a_2]^T$

and $\boldsymbol{\theta}_2 = [a_4 \;\; a_5]^T$, in which the superscript " $_o$ " denotes centralization operation. It is also concluded that

$$a_5 = -\left(\overline{\mathbf{D}}_1^T \boldsymbol{\theta}_1 + \overline{\mathbf{D}}_2^T \boldsymbol{\theta}_2 \right) \qquad (29)$$

where the above bar sign represents averaging operation. Based on (28), the fitting error for the centralized data is

$$\begin{aligned} \varepsilon &= \sum_{i=1}^{N} \left(\overset{o}{\mathbf{D}}_i^T \boldsymbol{\theta} \right)^2 / N \qquad (30) \\ &= \boldsymbol{\theta}_1^T \mathbf{R}_{11} \boldsymbol{\theta}_1 + 2\boldsymbol{\theta}_1^T \mathbf{R}_{12} \boldsymbol{\theta}_2 + \boldsymbol{\theta}_2^T \mathbf{R}_{22} \boldsymbol{\theta}_2 \end{aligned}$$

where $\mathbf{R}_{ij} = (\sum_{k=1}^{N} \overset{o}{\mathbf{D}}_k \overset{o}{\mathbf{D}}_k^T)/N$. Thus, we have the minimized criterion function

$$J_1 = \boldsymbol{\theta}_1^T \mathbf{R}_{11} \boldsymbol{\theta}_1 + 2\boldsymbol{\theta}_1^T \mathbf{R}_{12} \boldsymbol{\theta}_2 + \boldsymbol{\theta}_2^T \mathbf{R}_{22} \boldsymbol{\theta}_2 - \lambda \left(\boldsymbol{\theta}_1^T \mathbf{C} \boldsymbol{\theta}_1 - 1 \right)$$

which is equivalent to the following cost function

$$J_2 = \boldsymbol{\theta}_1^T \mathbf{R}_{11} \boldsymbol{\theta}_1 + 2\boldsymbol{\theta}_1^T \mathbf{R}_{12} \boldsymbol{\theta}_2 + \boldsymbol{\theta}_2^T \mathbf{R}_{22} \boldsymbol{\theta}_2 - \lambda \left(\boldsymbol{\theta}_2^T \mathbf{C}_2 \boldsymbol{\theta}_2 - 1 \right) .$$

Matrix $\mathbf{C}_2$ must be chosen to make the constraint $\boldsymbol{\theta}_2^T \mathbf{C}_2 \boldsymbol{\theta}_2 = 1$ equivalent to the constraint $\boldsymbol{\theta}_1^T \mathbf{C} \boldsymbol{\theta}_1 = 1$, which means that we must chose

$$\mathbf{C}_2 = \mathbf{R}_{12}^T \mathbf{R}_{11}^{-1} \mathbf{C} \mathbf{R}_{11}^{-1} \mathbf{R}_{12} \qquad (31)$$

According to J_2, it is easy to know that the optimal estimation must be satisfied to

$$\begin{cases} \mathbf{R}'_{22} \hat{\boldsymbol{\theta}}_2 = \lambda_{\min} \mathbf{C}_2 \hat{\boldsymbol{\theta}}_2 \\ \hat{\boldsymbol{\theta}}_1 = -\mathbf{R}_{11}^{-1} \mathbf{R}_{12} \hat{\boldsymbol{\theta}}_2 \end{cases} \qquad (32)$$

where $\mathbf{R}'_{22} = \mathbf{R}_{22} - \mathbf{R}_{12}^T \mathbf{R}_{11}^{-1} \mathbf{R}_{12}$ is certainly not negative definite, $\lambda_{\min}$ is the minimal generalized eigen value of $\mathbf{R}'_{22}$ and constraint matrix $\mathbf{C}_2$, $\hat{\boldsymbol{\theta}}_2$ is the corresponding vector. This procedure is just the suggested IGEVD algorithm.

V. Statistical performance analysis of IGEVD

It is easy to prove that on the condition that the noise contained in 2D point set $\{\mathbf{p}_i\}$ is supposed to be suit for the assumption of (7), the noise component of the centralized data vector has a zero mean vector and a block-diagonal variance matrix, that is

$$E\left(\mathbf{n}_{\overset{o}{\mathbf{D}}} \right) = \mathbf{0} \text{ and } \Delta \mathbf{R}_{\overset{o}{\mathbf{D}}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{n}_{\overset{o}{D}_1}} & \mathbf{0} \\ \mathbf{0} & \sigma_n^2 \mathbf{I} \end{bmatrix} \qquad (33)$$

where $\sigma_n^2 \mathbf{I}$ is the variance matrix of the noise $\mathbf{n}_{pi}$ of the 2D point set $\{\mathbf{p}_i\}$. So, we can complete the pre-whitening operation on the "observation noise" of the centralized data vector by means of the pre-whitening transformation to the noise component contained in $\overset{o}{\mathbf{D}}_1$, that is, let $\begin{cases} \tilde{\mathbf{D}}_1 = \mathbf{W}_1 \overset{o}{\mathbf{D}}_1 \\ \tilde{\mathbf{D}}_1 = w_2 \overset{o}{\mathbf{D}}_2 \end{cases}$,

where w_2 is a scaling factor. This results in mapping relationship $\tilde{\mathbf{R}}_{ij} = \mathbf{W}_i \mathbf{R}_{ij} \mathbf{W}_j^T \;\; \forall ij \in \{11, 12, 22\}$, where $\mathbf{W}_2 = w_2 \mathbf{I}$, between the ACF matrixes of before- and after-transformation. Thus, $\tilde{\mathbf{R}}'_{22} = (w_2)^2 \mathbf{R}'_{22}$ is valid. It means that, after the noise pre-whitening process, IGEVD algorithm can still give an optimal solution of $\boldsymbol{\theta}_2$ which is the same as the original. Furthermore, after transformation, the optimal solution of $\tilde{\boldsymbol{\theta}}_1$ is $\hat{\tilde{\boldsymbol{\theta}}}_1 = -\tilde{\mathbf{R}}_{11}^{-1} \tilde{\mathbf{R}}_{12} \hat{\tilde{\boldsymbol{\theta}}}_2$, thus the corresponding optimal solution of $\boldsymbol{\theta}_1$ is certainly equal to

$$\hat{\boldsymbol{\theta}}_1 = \mathbf{W}_1^T \hat{\tilde{\boldsymbol{\theta}}}_1 = -\mathbf{W}_1^T \tilde{\mathbf{R}}_{11}^{-1} \tilde{\mathbf{R}}_{12} \hat{\tilde{\boldsymbol{\theta}}}_2 = -\mathbf{R}_{11}^{-1} \mathbf{R}_{12} \hat{\boldsymbol{\theta}}_2 .$$

Therefore, without noise pre-whitening process and by directly using the computation of (32), the IGEVD can output the same optimal solution as the one given by the IGEVD based on the noise pre-whitening process. As the latter can output an unbiased estimation, the former can also output an unbiased estimation of the parameter vector. It is also shown that the condition number of $\mathbf{R}'_{22}$ is in the order of $O(1)$, so IGEVD is intrinsically able to regularize data.

In general, the suggested IGEVD algorithm naturally has the ability to pre-whiten the data noise as well as to regularize the condition number of the data's ACF matrix on condition that no any pre-whitening or regularizing procedure is performed. Because it has the advantages given by both of noise pre-whitening and data regularization, it can output an unbiased estimation with very small error variance. In addition, the dimension number of parameter estimation is decreased from 6 to 2. So, IGEVD also has the advantages of fast computation and simplified realization.

VI. Computer Simulation Experiments

In order to analyze the practical performance of the developed algorithm by experiment, we analyze and compare the performances of SGEVD, HGEVD and IGEVD, by a great deal of computer simulations for the ellipse fitting, The experiment results presented here are for the typical ellipse example whose center is $(C_x, C_y) = (10, 5)$, whose long and short half axis length pair is $(a, b)=(25,12)$ and whose incline angle is $\alpha = 30^{o}$.

The statistical performances of SGEVD, HGEVD and IGEVD for different noise strength in whole ellipse case or a half of ellipse case are shown in Table 1. Typical results fitted from the points belong to a half of the ellipse degraded by the additive noise with variance σ^2=0.5 for SGEVD, HGEVD and IGEVD are displayed in Figure 1 in which all the reported numbers are the averaged relative error amplitudes.

The experiments in all cases have demonstrated that both of IGEVD and HGEVD have much better performances than the performances of SGEVD and that the performance of IGEVD is slightly better than or almost the same as HGEVD's. The running speed of IGEVD is much faster than HGEVD speed, which has the same speed-order as SGEVD. From these facts, the advantages of the proposed IGEVD algorithm are obviously revealed.

TABLE 1. AVERAGE RELATIVE ERRORS

		Whole ellipse cases			Half ellipse cases		
$\sigma^2=$		0.5	1.0	2.0	0.5	1.0	2.0
SGEVD	C_x	0.0605	0.1344	0.3505	0.5536	0.5271	2.0161
	C_y	0.0610	0.1368	0.3232	0.5665	0.6169	0.3032
	a	0.0307	0.0767	0.2055	0.4891	0.5043	0.0570
	b	0.0051	0.0045	0.0533	0.5301	0.5483	0.6166
	α	0.1132	0.2258	0.4103	0.5005	0.4747	0.5548
IGEVD	C_x	0.0001	0.0003	0.0003	0.0098	0.0070	0.0315
	C_y	0.0001	0.0001	0.0001	0.0006	0.0001	0.0064
	a	0.0010	0.0018	0.0037	0.0304	0.0580	0.0880
	b	0.0127	0.0253	0.0501	0.0302	0.0637	0.1231
	α	0.0011	0.0021	0.0043	0.0061	0.0184	0.0410
HGEVD	C_x	0.0002	0.0002	0.0002	0.0103	0.0110	0.1275
	C_y	0.0002	0.0001	0.0002	0.0009	0.0165	0.0047
	a	0.0040	0.0076	0.0150	0.0669	0.0861	0.0509
	b	0.0038	0.0060	0.0127	0.0876	0.1464	0.2186
	α	0.0014	0.0034	0.0071	0.0239	0.0442	0.0285

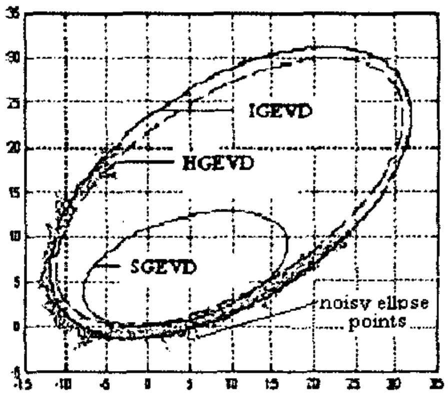

Fig 1. Typical fitting result

REFERENCES

[1]. Z.-G. Yang, J.-L. Qiao, C.-B. Zhao, X.-M. Luan, "A real-time underwater 3D vision system guiding AUV in underwater missions," *IEEE ACCV'95*, Vol. 3 of 3, pp. III-509-II-513, Dec. 1995.

[2]. Z.-G. Yang, J.-L. Qiao, C.-B. Zhao, X.-M. Luan, S.-H. Sun, "Robustly, exactly and real-timely extracting lines and conics from a gray-level image," *IEEE ICNNSP'95*, Vol. II, pp. II-1410-II-1413, Dec. 1995.

[3]. Y. C. Cheng, S. C. Lee, "A new method for quadratic curve detection using K-RANSAC with acceleration techniques," *Pattern Recognition* 1995, Vol. 28, No. 5, pp. 663-682, May 1995.

[4]. C. F. Olson, "Constrained Hough transform for curve detection," *Computer Vision & Image Understanding* 1999, Vol. 73, No. 3, pp. 329-345.

[5]. O. Chutatape, L. Guo, "A modified Hough transform for line detection and its performance," *Pattern Recognition*, 1999, Vol. 32, No. 2, pp. 181-192, Feb. 1999.

[6]. N. Bennett, R. Burridge, N. Saito, "A method to detect and characterize ellipses using the Hough transform," *IEEE Trans. Pattern Analysis and Machine Intelligence* 1999, Vol. 21, No. 7, pp. 652-657, July 1999.

[7]. A. Fitzgibban, M. Pilu, R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Trans. Pattern Analysis and Machine Intelligence* 1999, Vol. 21, No. 5, pp. 476-480, May 1999.

[8]. R. I. Hartley, "In defense of eight-point algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence* 1997, Vol. 19, No. 6, pp. 580-593, June 1997.