GRAPHICAL MODELS AND IMAGE PROCESSING
Vol. 58, No. 5, September, pp. 494–502, 1996
ARTICLE NO. 0041

NOTE

Assessing Error of Fit Functions for Ellipses

PAUL L. ROSIN*

Department of Computer Science and Information Systems, Brunel University, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom

Received August 17, 1995; revised April 3, 1996; accepted July 10, 1996

A variety of error of fit (EOF) functions have been proposed for use in the least-square fitting of ellipses. We describe four measures for assessing the suitability of such EOFs, quantifying their linearity, curvature bias, asymmetry, and overall goodness. These measures enable a better understanding to be gained of the individual merits of the EOF functions. © 1996 Academic Press, Inc.

1. INTRODUCTION

The fitting of ellipses to edge data is a common task in computer vision. In particular, this often arises in the context of industrial inspection since circular parts in the scene are projected into the image as ellipses. There are many algorithms for ellipse fitting, but in this paper we shall concentrate on minimization techniques rather than others such as the Hough transform voting method. Despite its sensitivity to non-Gaussian noise, least-squares (LS) fitting is probably the most widely used approach for estimating the ellipses' parameters, due to its computation efficiency and its high efficiency as an estimator. It operates by minimizing the sum of squares of some error term e_j measured at each data point $\mathbf{X}_j = (x_j, y_j)$. Thus for N points the parameters $\mathbf{P}$ of the best fit ellipse are obtained by

$$\min_{\mathbf{P}} \sum_{j=1}^{N} e_j^2.$$

Alternatively, techniques from robust statistics have recently become popular [7, 9] since they have a strong immunity to outliers (i.e., a high breakdown point) although they are often computationally expensive and their results can have lower accuracy. As an example, the least median of squares (LMedS) approach is commonly used. The best fit ellipse would then be obtained by

$$\min_{\mathbf{P}} \operatorname{med}_{j=1}^{N} e_j^2.$$

Thus, many different ellipse estimation techniques depend on a suitable error term, although there has been little comparison between them (but see [3, 4, 8]). The Euclidean distance from $\mathbf{X}_j$ to the ellipse boundary would be a good choice for e_j, but requires solving a quartic equation which may have up to four solutions, requiring the one with the minimum distance to be determined [10]. To avoid the complexity of evaluating the true Euclidean distance it is usual practice to approximate it by some measure—the error of fit (EOF) function—that is simpler to calculate. The simplest approximation is the algebraic distance [2] calculated from $\mathbf{X}_j$ to the ellipse

$$\mathrm{EOF}_1 = Q(x_j, y_j) = Ax_j^2 + Bx_jy_j + Cy_j^2 + Dx_j + Ey_j + F, \tag{1}$$

where $Q(x, y) = 0$ is the general equation of a conic, describing ellipses when $B^2 - 4AC < 0$. An advantage of (1) is that a closed form LS solution is available whereas the other EOFs described below must be minimized by iterative procedures. However, many more accurate approximations have been suggested, most prominently the algebraic distance inversely weighted by its gradient [1, 11, 13]

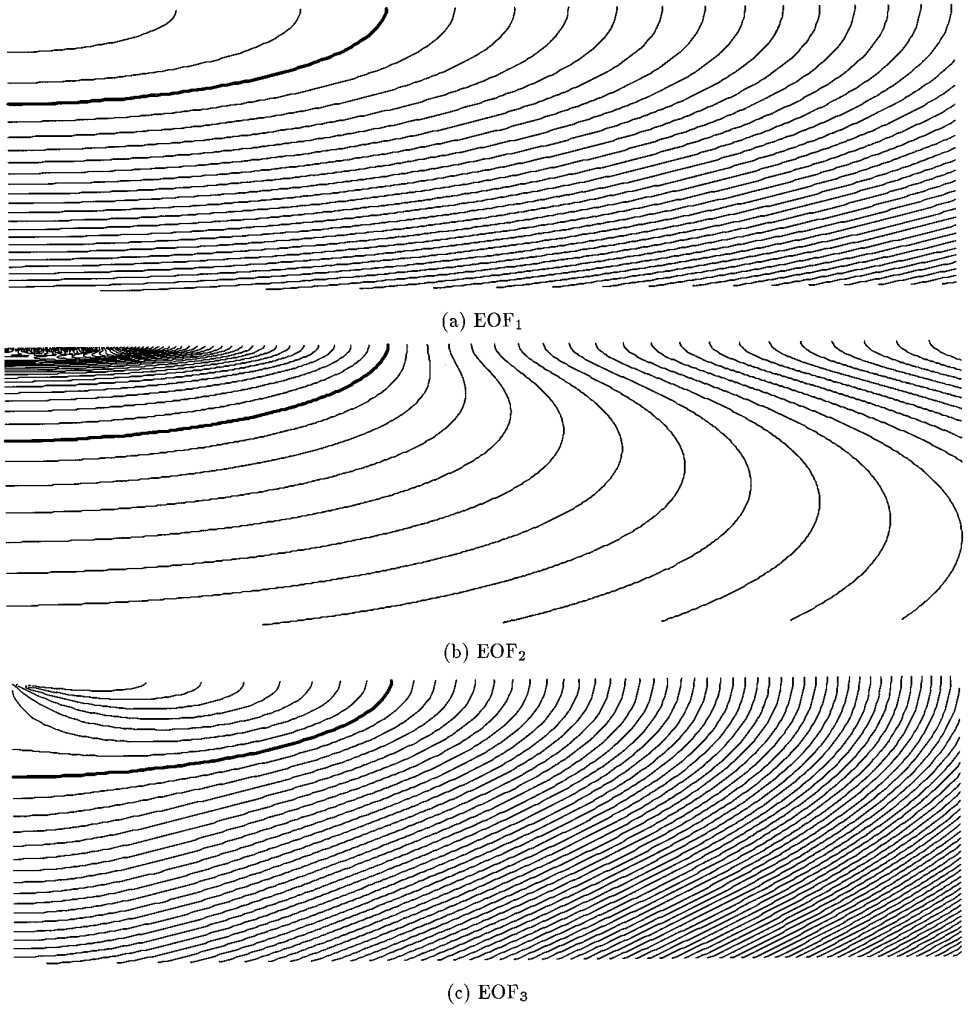
$$\mathrm{EOF}_2 = \frac{Q(x_j, y_j)}{|\nabla Q(x_j, y_j)|}. \tag{2}$$

Safaee-Rad *et al.* [10] provide an alternative weighting. A ray is drawn between $\mathbf{X}_j$ and the center of the ellipse $\mathbf{C}$, intersecting the ellipse at $\mathbf{I}_j$. The lengths of the bisected portions of the ray $m_j = \overline{\mathbf{CI}}_j$ and $n_j = \overline{\mathbf{I}_j\mathbf{X}}_j$ are determined. The approximate distance is then given by

$$\mathrm{EOF}_3 = m_j \frac{1 + n_j/(2a)}{1 + n_j/(2m_j)} Q(x_j, y_j), \tag{3}$$

* E-mail: Paul.Rosin@brunel.ac.uk.

1077-3169/96 $18.00



(a) EOF_1

(b) EOF_2

(c) EOF_3

FIG. 1. Iso-distance contours.

where a is the length of the semimajor axis of the ellipse. We discount the additional $1/a$ multiplying term given by Safaee-Rad *et al.* because we can ignore constant scalings of the error measure. Since n_j is normally much smaller than m_j or a they further approximate the distance by

$$\mathrm{EOF}_3 \approx \mathrm{EOF}_4 = m_j Q(x_j, y_j). \tag{4}$$

Nakagawa and Rosenfeld [6] used

$$\mathrm{EOF}_5 = n_j \tag{5}$$

alone as an error measure.

Stricker [12] proposed that an approximation to the distance from $\mathbf{X}_j$ to the ellipse E be determined by generating a new ellipse $\tilde{E}$ which passes through $\mathbf{X}_j$ and has the same center and orientation as E. Letting the length of the semimajor axis of ellipse $\tilde{E}$ be $\tilde{a}$, he takes the approximate distance as

$$\mathrm{EOF}_6 = \tilde{a} - a. \tag{6}$$

To calculate $\tilde{a}$ an initial estimate is made first to the approximate distance

$$d_{\mathrm{est}} = \frac{\tilde{a}_{\mathrm{est}} - a + \tilde{b}_{\mathrm{est}} - b}{2},$$

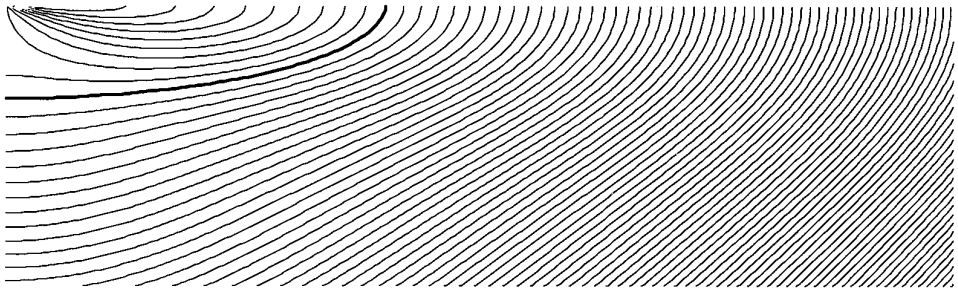
where

$$\tilde{a}_{\mathrm{est}} = \frac{|\mathbf{X}_j - \mathbf{F}_1| + |\mathbf{X}_j - \mathbf{F}_2|}{2}$$

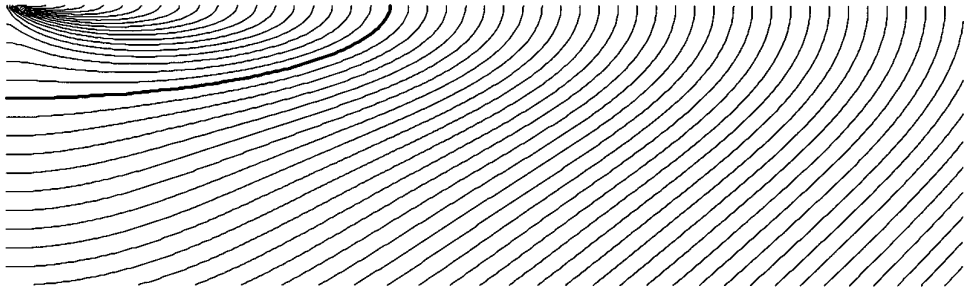
$$\tilde{b}_{\mathrm{est}} = \sqrt{\tilde{a}_{\mathrm{est}}^2 - a^2 + b^2}.$$

$\mathbf{F}_1$ and $\mathbf{F}_2$ are the focal points of E, and b is the length of the semiminor axis of E. The focal points of $\tilde{E}$ are[1]

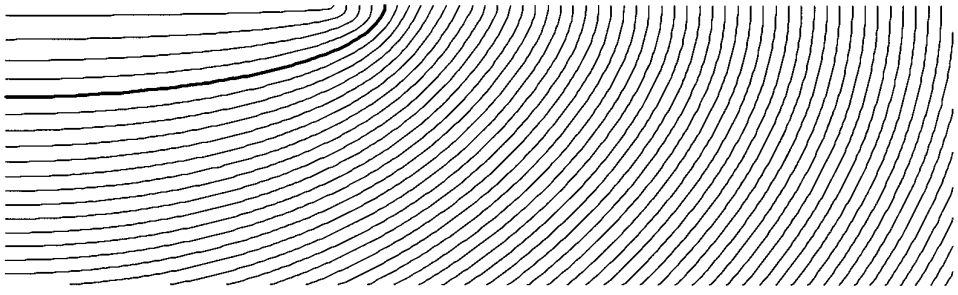
[1] Note the corrected expression for $\tilde{\mathbf{F}}_1$ and $\tilde{\mathbf{F}}_2$ provided by Stricker (personal communication, 1995).



(d) EOF_4



(e) EOF_5



(f) EOF_6

FIG. 1—*Continued*

$$\tilde{\mathbf{F}}_1, \tilde{\mathbf{F}}_2 = \frac{\mathbf{F}_1 + \mathbf{F}_2}{2} \pm \tilde{c}\,\frac{\mathbf{F}_1 - \mathbf{F}_2}{|\mathbf{F}_1 - \mathbf{F}_2|},$$

where

$$\tilde{c} = \sqrt{(a + d_{\text{est}})^2 - (b + d_{\text{est}})^2}$$

enabling the semimajor axis of $\tilde{E}$ to be calculated as

$$\tilde{a} = \frac{|\mathbf{X}_j - \tilde{\mathbf{F}}_1| + |\mathbf{X}_j - \tilde{\mathbf{F}}_2|}{2}.$$

2. ASSESSMENT MEASURES

Despite the variety of approximations to the distance from a point to the boundary of an ellipse little comparative analysis has been carried out.[2] Previously, to better understand the different nature of the approximations, we visualized the EOFs by plotting their iso-distance contours at regular intervals of the EOF [8]. Figure 1 shows the iso-distance contours for the different EOFs (the ellipse boundary is drawn bold) from which we can make several observations. First, EOF_1 demonstrates the so-called ''high curvature bias'' in that the spacing between the contours becomes wider at the pointed ends of the ellipse, i.e., near the regions of high curvature. This is overcompensated by EOF_2, which also exhibits a bowing out of the contours near the pointed ends of the ellipse. EOFs 3–5 also demonstrate substantial

[2] Gross and Boult [5] experimentally evaluated four EOFs for superquadrics using graphical plots comparing scaled summed EOF against the true RMS error, as well as showing one dimensional cross sections of the errors of fit.

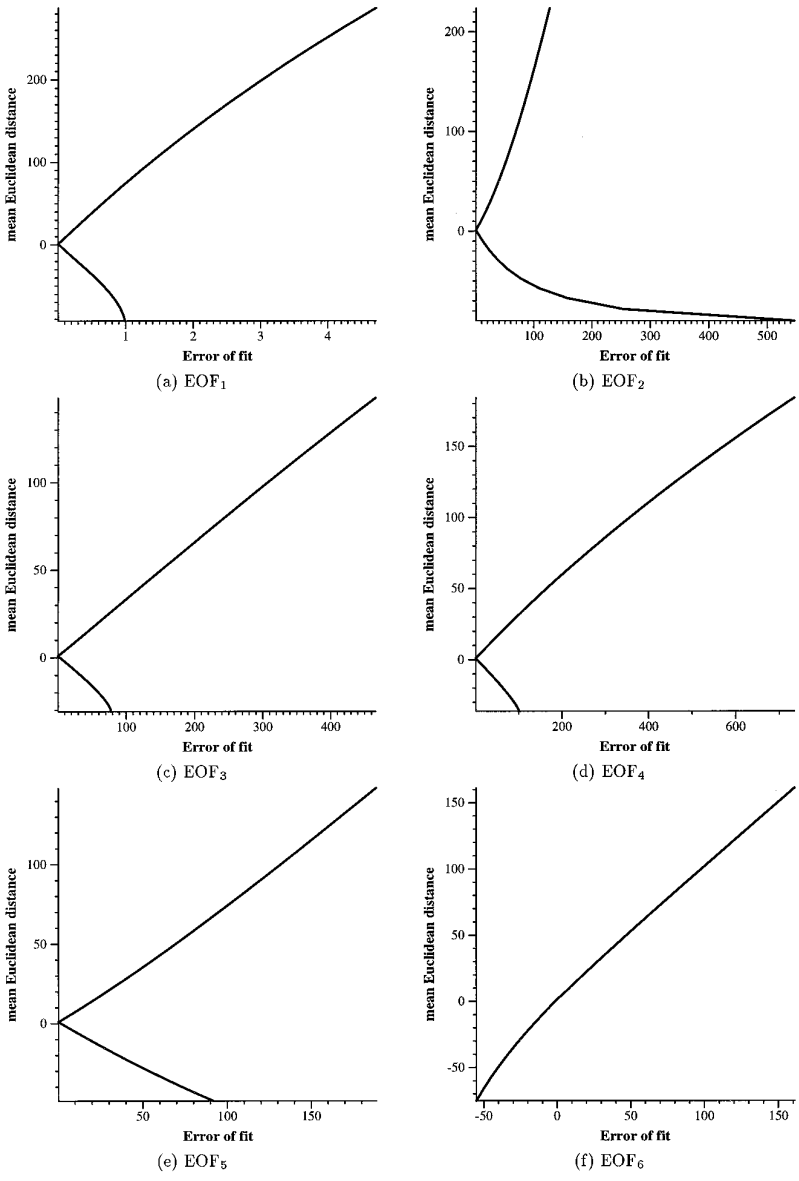


FIG. 2. Mean Euclidean distances of iso-distance contours of EOFs.

curvature bias—but only in the interior of the ellipse. Even EOF_6 has some curvature bias, again more obvious inside the ellipse. Second, it can be seen that the sign of the gradient of the measure perpendicular to the ellipse differs between the measures. With EOF_1 the contours get closer with increasing distance out from the ellipse, whereas for EOF_2 the contours drift apart. Moreover, we can see that for EOF_2 the magnitude of the gradient of the measure inside the ellipse is considerably larger than on the outside.[3]

These three factors have different effects on the ellipse fitting due to the varying weights they assign to the data. Curvature bias causes data near the ends of the ellipse to have more or less influence on the fit. In the case of EOF_1 their influence is reduced, often resulting in an overestimate of the eccentricity of the fitted ellipse. The second factor is the relationship between the EOF and the Euclidean distance as a function of the Euclidean distance. Ideally they should be linearly related; a constant scaling factor has no effect and can be ignored. A superlinear relationship causes the outlying data to have a stronger influence on the fit than a linear or sublinear relationship. Finally, the third factor is the asymmetry between the distance approximation inside and outside the ellipse. Using EOF_1 the data inside the ellipse has

[3] Artifacts from the plotting process have caused some contours to be missed near the center of the ellipse.

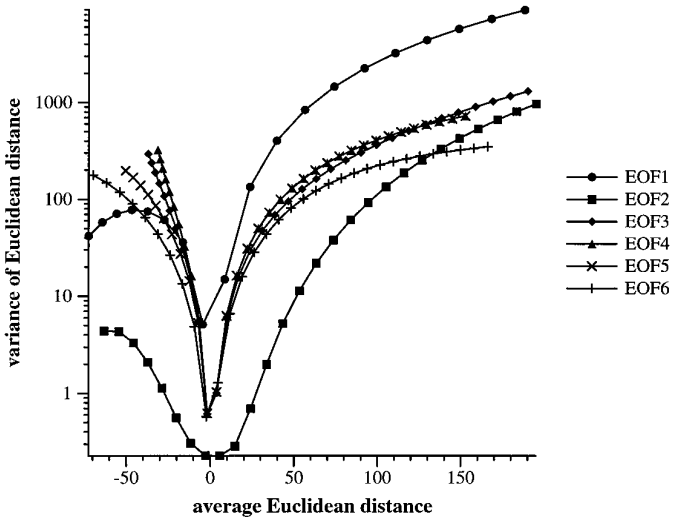


FIG. 3. Variance of Euclidean distances of iso-distance contours of EOFs plotted against their mean Euclidean distances.

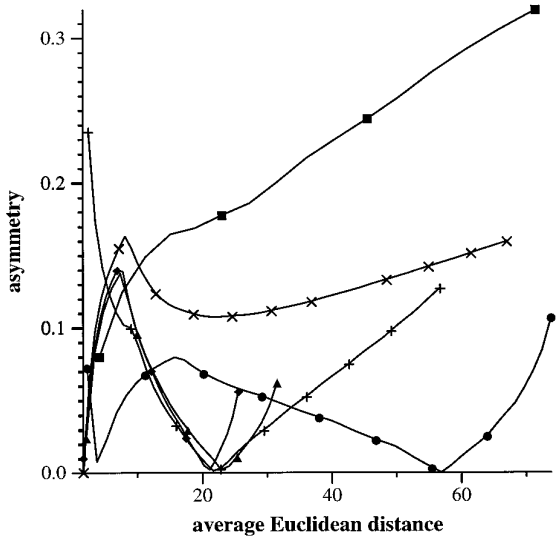


FIG. 4. Asymmetry of Euclidean distances of iso-distance contours of EOFs plotted against their mean Euclidean distances.

less influence than the exterior data, which could cause the size of the fitted ellipse to be overestimated. The opposite effect occurs with EOF_2 since the interior data has a disproportionate effect on the fitting.

Although visualizing the iso-distance contours is a useful tool for the qualitative analysis of EOFs, a quantitative assessment would be valuable for objectively comparing them. We develop assessment measures based on the three factors of deviation from the Euclidean distance described above. One consideration is that the degree of deviation may not be constant with increasing distance from the ellipse. Therefore it may be necessary to make the measures a function of the distances. Second, although we show the measures applied to all values $\mathbf{X} = (x, y)$ in the $\mathscr{R}^2$ plane, in practice we must discretely sample a subset of $\mathscr{R}^2$. This leads to the problem of which subset? One solution is to assume a noise model which, given an ellipse, specifies where the data is expected to occur. For instance, for noise with a Gaussian distribution $N(0, \sigma)$ we can weight the data in $\mathscr{R}^2$ by

$$w(d) = \frac{1}{\sigma\sqrt{2\pi}} e^{-d^2/2\sigma^2},$$

where d is the Euclidean distance to the ellipse boundary. Alternatively, for uniform noise $U(R)$ in the range $[-R, R]$ we can weight the data by

$$w(d) = \begin{cases} 1/(2R) & \text{if } -R < d < R \\ 0 & \text{otherwise.} \end{cases}$$

2.1. *Linearity*

To test for a linear relationship between the Euclidean distance values $d_{\mathbf{X}}$ and their approximations $e_{\mathbf{X}}$, it is natural to use the Pearson correlation coefficient

$$\rho = \frac{\Sigma_{\mathbf{X}}(e_{\mathbf{X}} - \bar{e}_{\mathbf{X}})(d_{\mathbf{X}} - \bar{d}_{\mathbf{X}})}{\sqrt{\Sigma_{\mathbf{X}}(e_{\mathbf{X}} - \bar{e}_{\mathbf{X}})^2 \Sigma_{\mathbf{X}}(d_{\mathbf{X}} - \bar{d}_{\mathbf{X}})^2}}. \tag{7}$$

The value of ρ lies between ± 1, although in our context negative values are unlikely, and we can think of $\rho \in [0, 1]$ with increasing values meaning a better linear correlation.

A problem with (7) is that equidistant values further from the ellipse will have greater effect on the measure than close values since as iso-distance contours become more distant from the ellipse they become longer. This can be compensated by weighting points in $\mathscr{R}^2$ by the length of the iso-distance contour through the point. Since determining the length is not straightforward we take an alternative approach which is to sample a set of iso-distance contours at regular intervals of the distance approximation. The iso-distance contour of the distance approximation at some value[4] E_i is at $d_{\mathbf{X}} \,|\, e_{\mathbf{X}} = E_i$, and the mean Euclidean distance along each contour is

$$\mu_i = \mathrm{E}[d_{\mathbf{X}} \,|\, e_{\mathbf{X}} = E_i]. \tag{8}$$

The correlation coefficient is then calculated between the iso-distance values and their Euclidean distance means

$$L = \frac{\Sigma_i w_i(E_i - \bar{E}_i)(\mu_i - \bar{\mu}_i)}{\sqrt{\Sigma_i w_i(E_i - \bar{E}_i)^2 \Sigma_i(\mu_i - \bar{\mu}_i)^2}}. \tag{9}$$

Thus each distance is now given equal weight in the assessment measure. In addition, we have included the term w_i

[4] To avoid confusion we assume distances inside and outside the ellipse are signed, or can be made so.

TABLE 1
Normalized Assessment Results with $N(0, 2)$ Noise Model

EOF	L	C	A	G	G'
1	1.000	1.000	1.000	1.000	1.000
2	0.987	0.011	3.977	0.808	0.841
3	1.000	0.127	2.042	0.998	0.967
4	1.000	0.146	2.808	0.994	0.965
5	1.000	0.127	2.396	1.012	0.976
6	1.000	0.106	4.393	0.996	1.001

Note. $a = 400$, $b = 100$.

for each contour corresponding to the weighting factor associated with the noise model, where $w_i = w(\bar{\mu}_i)$.

2.2. *Curvature Bias*

To measure the departure of the iso-distance contour from the desired constant Euclidean distance we use the variance of the underlying Euclidean distance

$$\sigma_i^2 = \mathrm{Var}[d_{\mathbf{X}} | e_{\mathbf{X}} = E_i]. \tag{10}$$

Since this is a local measure it must be combined over contours to give a global measure

$$C = \sum_i w_i \sigma_i^2. \tag{11}$$

The ideal error of fit should have no curvature bias which is obtained when $C = 0$.

2.3. *Asymmetry*

Assuming signed distances, the mean Euclidean distances along corresponding iso-distance contours inside and outside the ellipse are calculated

$$\mu_i^+ = \mathrm{E}[d_{\mathbf{X}} \,|\, e_{\mathbf{X}} = E_i]$$
$$\mu_i^- = \mathrm{E}[d_{\mathbf{X}} \,|\, e_{\mathbf{X}} = -E_i].$$

Asymmetry is calculated at each contour pair as the normalized difference in their mean Euclidean distances,

$$a_i = \frac{|\mu_i^+ - \mu_i^-|}{\mu_i^+ + \mu_i^-}.$$

Again, a weighted sum of the local measures over the contours is made to produce a global measure

$$A = \sum_i w_i a_i \tag{12}$$

which would equal zero for the ideal error of fit.

2.4. *Overall Goodness*

Rather than individually assess the three specific characteristics described above, sometimes it may be more convenient to produce a single overall assessment of the distance approximations, ignoring the details. One approach would be to produce a weighted sum of the three measures above. Instead, we use an alternative which avoids the need for deciding on suitable weights and use the squared difference between the approximation and the true Euclidean distance

$$G = \sum_{\mathbf{X}} w_{\mathbf{X}} (d_{\mathbf{X}} - S e_{\mathbf{X}})^2. \tag{13}$$

Since we can ignore uniform scalings of the distances, we allow the approximation to be scaled by some constant S, and choose S so as to minimize G. This is found when $\delta G / \delta S = 0$, yielding

$$S = \frac{\Sigma_{\mathbf{X}} w_{\mathbf{X}} e_{\mathbf{X}} d_{\mathbf{X}}}{\Sigma_{\mathbf{X}} w_{\mathbf{X}} e_{\mathbf{X}}^2}.$$

3. RESULTS

We show the results of generating and applying the assessment measures described in the previous section to the six error of fits given in Section 1. A quadrant of a single ellipse with semimajor axis $a = 400$ and semiminor axis $b = 100$ and centered at the origin was used. The Euclidean distance was found by plotting the ellipse into an image followed by performing a Euclidean distance transform. The distance was made signed by setting negative all nonzero distance values 4-way connected to the origin. The EOFs were generated and sampled at unit increments in one quadrant of size $1\,000 \times 300$. Iso-distance contours crossing the Y axis at 10 pixel intervals were detected, and the average and variance of the Euclidean distances (8) and (10) along the contours were calculated. The mean distances are

TABLE 2
Normalized Assessment Results with $N(0, 64)$ Noise Model

EOF	L	C	A	G	G'
1	1.000	1.000	1.000	1.000	1.000
2	0.877	0.041	8.404	2.771	0.099
3	1.005	0.054	1.475	0.309	0.310
4	1.003	0.067	1.270	0.282	0.288
5	1.004	0.056	2.997	0.584	0.262
6	1.004	0.042	1.547	0.218	0.179

Note. $a = 400$, $b = 100$.

TABLE 3
Normalized Assessment Results with $U(10)$ Noise Model

EOF	L	C	A	G	G'
1	1.000	1.000	1.000	1.000	1.000
2	1.000	0.008	2.870	0.509	0.162
3	1.000	0.117	1.833	0.964	0.744
4	1.000	0.152	1.685	0.943	0.727
5	1.000	0.123	2.111	1.025	0.728
6	1.000	0.108	1.852	0.957	0.874

Note. $a = 400$, $b = 100$.

shown in Fig. 2. The linearity and asymmetry aspects of the EOFs become clear on looking at these plots. Most EOFs exhibit either a sub- or superlinear relationship with Euclidean distance, while the degree of asymmetry is accentuated toward the center of the ellipse.

Figure 3 shows the variance of the Euclidean distances (10) for the iso-distance contours. So that variance plots for all the different EOFs can be plotted together with the same scales, they have been plotted against the mean Euclidean distance of each iso-distance contour rather than its EOF value. For much of the extent shown, EOF_1 has the worst, and EOF_2 the least, curvature bias. Since the curves cross, none of the EOFs has the lowest curvature bias over all $\mathscr{R}^2$.

To simplify calculation of the measures and ensure uniform sampling, for both the average and variance of the Euclidean distance values along the contours the EOF, values were resampled at unit intervals using linear interpolation. It is then straightforward to find the corresponding points μ_i^+ and μ_i^-. Figure 4 shows the asymmetry measures for the different EOFs. Again, they have been plotted against the mean Euclidean distance to facilitate comparison. It is difficult to discern clear trends, but we note that asymmetry fluctuates considerably with distance from the ellipse, except for EOF_2 which monotonically increases, reaching a substantial level.

Tables 1–4 give the assessment values for the Gaussian noise model with $\sigma = \{2, 64\}$ and uniform noise with $R = \{10, 160\}$. To facilitate comparison, all values have been scaled w.r.t. the corresponding EOF_1 value. Recall that we wish to maximize L and minimize the remaining assessment measures. To factor out the extreme asymmetry shown by EOF_2, we also show the overall goodness measure applied only to values outside the ellipse, denoted by G'. Considering Gaussian noise first, we see that close to the ellipse (i.e., for small σ) all measures have similar linearity. But even close there are large variations in curvature bias: EOF_1 performing relatively poorly and EOF_2 performing well. On the other hand, EOF_1 has the lowest degree of asymmetry, while EOF_2 and EOF_6 have the greatest degrees of asymmetry. However, the overall goodness assessment values G and G' of the various EOFs are similar.

With increased σ—effectively including a larger portion of $\mathscr{R}^2$—the differences become more noticeable, while some rankings change. EOFs 3–6 now do slightly better than EOF_1 while EOF_2 is significantly worse. EOF_2's result is degraded due to its marked asymmetry. EOF_1 still has the greatest curvature bias, but the bias of the remaining EOFs now appear comparable with each other. Again EOF_1 still has the least asymmetry. Now EOF_2 has a much worse asymmetry than the other EOFs, while EOF_6 is relatively low. The overall rating G assigns the best score to EOF_6 and good scores to EOF_3 and EOF_4. The poor rating given to EOF_2 is incurred by its asymmetry, as can be seen by the excellent score achieved with G' when only exterior values are considered. Applying the uniform noise model over different distances we can see similar effects to the Gaussian noise model.

The effect of changing the eccentricity e of the ellipse is investigated, where $e^2 = 1 - b^2/a^2$. The length of the minor axis was fixed at $b = 100$ and the length of the major axis varied as $a = \{100, 120, 150, 200, 300, 400\}$, giving $e = \{0, 0.55, 0.74, 0.86, 0.94, 0.96\}$. The various measures are plotted in Fig. 5 using the Gaussian noise model with $\sigma = 64$. It can be seen that the linearity of all the EOFs is roughly independent of eccentricity. The curvature bias of EOF_1 increases exponentially with eccentricity; the remaining EOF curvature biases also increase exponentially, but less dramatically. While the asymmetry of EOF_2, EOF_5, and EOF_6 increases with eccentricity, the asymmetry of EOF_1, EOF_3, and EOF_4 decreases. Finally, all the EOFs show increasing G and G' against eccentricity.

4. CONCLUSIONS

We have described four measures for assessing the suitability of EOF functions used for the fitting of ellipses. This allows us to compare different EOFs in an objective, quantitative manner. Applying the measures to six different EOFs we see that the choice of the most suitable EOF depends in part on the degree of noise present in

TABLE 4
Normalized Assessment Results with $U(160)$ Noise Model

EOF	L	C	A	G	G'
1	1.000	1.000	1.000	1.000	1.000
2	0.907	0.049	8.629	2.826	0.082
3	1.000	0.045	1.464	0.293	0.297
4	1.000	0.062	1.266	0.254	0.257
5	1.000	0.044	3.065	0.473	0.235
6	1.000	0.026	1.596	0.174	0.150

Note. $a = 400$, $b = 100$.

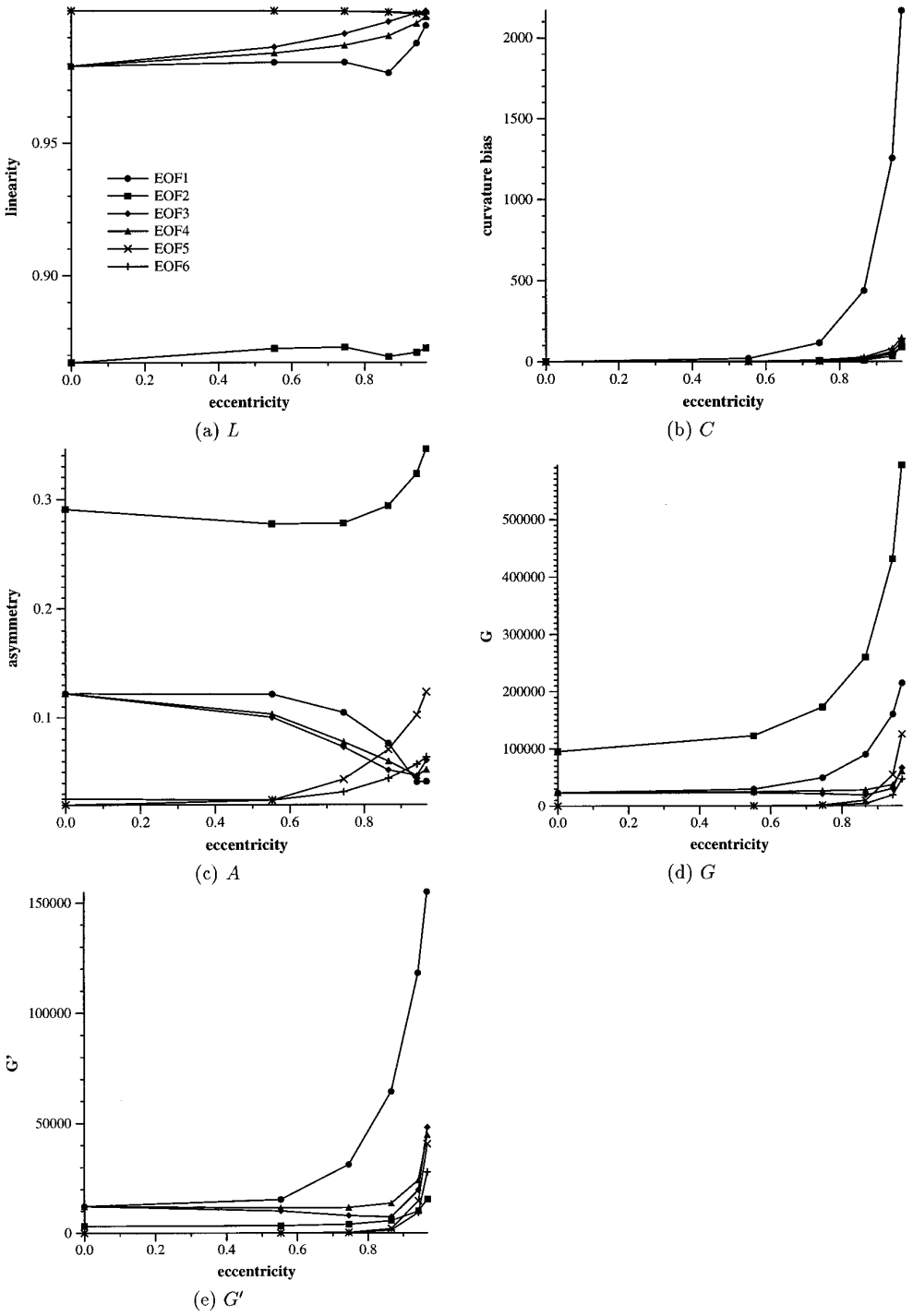


FIG. 5. Effect of eccentricity on measures.

the data with respect to the size of the ellipse, the nature of the data, and the fitting algorithm. For instance, curvature bias tends to have less effect if data describing the complete ellipses is available compared to data describing only small arcs of the ellipse. Likewise, the linearity of the EOF is more important if a nonrobust fitting technique such as LS is used rather than a robust algorithm such as LMedS.

In general, some EOFs have good properties close to the ellipse (e.g., EOF_2's overall assessment) but dramatically degrade at further distances. Thus we can conclude that for small amounts of noise EOF_2 is the overall best approxi-

mation to use, but that for larger amounts of noise EOF_6 is better. However, no single EOF was a winner in all categories: the individual assessment measures show that an EOF may satisfy one assessment criterion well but another one poorly (e.g., EOF_2 has low curvature bias but poor asymmetry).

REFERENCES

1. G. J. Agin, *Fitting Ellipses and General Second-Order Curves*, Technical Report CMU-RI-TR-81-5, The Robotics Institute, Carnegie-Mellon University, Pittsburgh, PA 1981.
2. A. Albano, Representation of digitized contours in terms of conic arcs and straight line segments. *Comput. Graphics Image Process.* **5,** 1974, 23–33.
3. A. W. Fitzgibbon and R. B. Fisher, A buyer's guide to conic fitting, in *British Machine Vision Conference* 1995, pp. 513–522.
4. W. Gander, G. H. Golub, and R. Strebel, *Fitting of Circles and Ellipses Least Squares Solution*, Technical Report TR-217, Institut für Wissenschaftliches Rechen, ETH, Zurich, Switzerland, 1994.
5. A. D. Gross and T. E. Boult, Error of fit for recovering parametric solids, in *Int. Conf. Computer Vision, Tampa, FL, 1988,* pp. 690–694.
6. Y. Nakagawa and A. Rosenfeld, A note on polygonal and elliptical approximation of mechanical parts, *Pattern Recognition,* **11,** 1979, 133–142.
7. P. L. Rosin, Ellipse fitting by accumulating five-point fits, *Pattern Recognition Lett.* **14,** 1993, 661–669.
8. P. L. Rosin, A note on the least squares fitting of ellipses, *Pattern Recognition Lett.* **14,** 1993, 799–808.
9. G. Roth and M. D. Levine, Extracting geometric primitives, *CVGIP: Image Understanding* **58,** 1993, 1–22.
10. R. Safaee-Rad, I. Tchoukanov, B. Benhabib, and K. C. Smith, Accurate parameter estimation of quadratic curves from grey level images, *CVGIP: Image Understanding* **54,** 1991, 259–274.
11. P. D. Sampson, Fitting conic sections to very scattered data: An iterative refinement to the Bookstein algorithm, *Comput. Vision Graphics Image Process.* **18,** 1982, 97–108.
12. M. Stricker, A new approach for robust ellipse fitting, in *Int. Conf. Automation, Robotics, and Computer Vision,* 1994, pp. 940–945.
13. G. Taubin, Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation, *IEEE Trans. PAMI* **13**(11), 1991, 1115–1138.