Dear Author

Please use this PDF proof to check the layout of your article. If you would like any changes to be made to the layout, you can leave instructions in the online proofing interface. Making your changes directly in the online proofing interface is the quickest, easiest way to correct and submit your proof. Please note that changes made to the article in the online proofing interface will be added to the article before publication, but are not reflected in this PDF proof.

If you would prefer to submit your corrections by annotating the PDF proof, please download and submit an annotatable PDF proof by clicking here and you'll be redirected to our PDF Proofing system.

Pattern Recognition xxx (xxxx) xxx

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Ellipse fitting by spatial averaging of random ensembles

Q1 Karl Thurnhofer-Hemsi[a,b,*], Ezequiel López-Rubio[a,b], Elidia Beatriz Blázquez-Parra[c], M. Carmen Ladrón-de-Guevara-Muñoz[c], Óscar David de-Cózar-Macías[c]

[a] Department of Computer Languages and Computer Science, Universidad de Málaga, Bulevar Louis Pasteur, 35, 29071, Málaga, Spain
[b] Biomedic Research Institute of Málaga (IBIMA), C/ Doctor Miguel Díaz Recio, 28, 29010, Málaga, Spain
[c] Department of Graphical Engineering, Design and Projects, Universidad de Málaga, C/ Doctor Ortiz Ramos, 29071, Málaga, Spain

ARTICLE INFO

Article history:
Received 24 February 2020
Accepted 28 April 2020
Available online xxx

Keywords:
Ellipse fitting
Geometric curve fitting
Ensemble methods
Spatial median
Robust estimation

ABSTRACT

Earlier ellipse fitting methods often consider the algebraic and geometric forms of the ellipse. The work presented here makes use of an ensemble to provide better results. The method proposes a new ellipse parametrization based on the coordinates of both foci, and the distance between them and each point of the ellipse where the Euclidean norm is applied. Besides, a certain number of subsets are uniformly drawn without replacement from the overall training set which allows estimating the center of the distribution robustly by employing the L1 median of each estimated focus. An additional postprocessing stage is proposed to filter out the effect of bad fits. In order to evaluate the performance of this method, four different error measures were considered. Results show that our proposal outperforms all its competitors, especially when higher levels of outliers are presented. Several synthetic and real data tests were developed and confirmed such finding.

1. Introduction

Curve fitting is the process of specifying a model based on a particular curve, such as a circle, ellipse or parabola that provides a good fit to a set of points [1–4]. Fitting has become an important research topic in the last years in many areas of science and technology. Also, this process is the first step for a large number of applications, for example, engineering applications that may include trend analysis, extrapolation or interpolation between data points.

Borges [5] presents a geometric fitting method based on least squares. This method is used to analyze rocks' properties and other materials. Mitchell and den Berg [6] use ellipse fitting to analyze electron diffraction in order to obtain different patterns from polycrystalline materials. And Liao et al. [7] use a method based on detection techniques and ellipse fitting for cell image segmentation.

In this work, a new method to fit an ellipse to a set of input points is presented. Our approach introduces ensemble methods into the curve fitting research field, which has not been previously proposed in the literature. This methodology differs from our previous works for other conic curves [4].

Our proposal is motivated by the observation that several alternative solutions can be obtained from the same base algorithm by slightly changing the set of input points. Therefore, an enhanced fit can be computed by robustly combining those different solutions, so that grossly erroneous fits have a small effect on the final solution. In order to attain such a goal, a parametrization of the ellipse is proposed which is amenable to the robust averaging of the combined solutions. The median is employed to carry out such averaging. The rationale behind this is that the median is resilient to outliers, so that some completely wrong solutions do not ruin the overall estimation. The research gap to be addressed here is how accurate ensemble averaging of ellipse fits can be, as compared to the individual fits by classic algorithms which do not consider ensembles. Given the previous successes of ensemble methods when applied to other research fields, and the lack of ensemble proposals for ellipse fitting, the significance of our investigation is guaranteed. The validity of our proposal is assessed by quantitative and qualitative studies of its performance, as compared to state of the art ellipse fitting algorithms.

The structure of the paper is: Section 2 presents state-of-art techniques related to ellipse fitting and ensembles. Section 3 describes the theory and the proposed algorithm and after, the results and discussion of the experiments are presented in Section 4. At the end, Section 5 reports the conclusions and possible further works.

* Corresponding author.
E-mail address: karlkhader@lcc.uma.es (K. Thurnhofer-Hemsi).
URL: http://www.lcc.uma.es/%7Eezeqlr/index-en.html (E. López-Rubio)

https://doi.org/10.1016/j.patcog.2020.107406
0031-3203/

2. Background

2.1. Fitting ellipse in literature

Fitting methods can be divided into algebraic and geometric fitting [8,9]. Thus, in an algebraic fitting, the curve is given by an implicit equation of a conic.

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{1}$$

Only if $B^2 - 4AC < 0$ then this equation defines an ellipse. If the value of a point (x, y) is replaced in the equation, the value obtained in the implicit equation is called an algebraic distance. Also, the solution is not always an ellipse in spite of the algorithms being efficient. On the other hand, geometric fitting tries to minimize the geometric distance, i.e. Ahn et al. [10] who minimize the sum of the orthogonal distances.

Many researchers have tried to obtain accurate and robust methods. They approximate the geometric distance by a function of the ellipse parameters. Yu et al. [11] have determined a geometric objective function considering that the sum of the distances from a point to the foci is constant. On the other hand, Kanatani and Rangarajan [12] have presented an algebraic method using the ellipse implicit equation with a stipulated constraint. This gives rise to an accurate algebraic solution by full error analysis. However, parabolas and hyperbolas are described by the general equation of the conic as a solution and the sensitivity to outliers. Liang et al. [13] used the alternating direction method of multipliers on an adaptation of direct least square fitting using the l_p-norm with $p < 2$.

Ten fundamental methods used in this paper are presented below. First, Taubin's method [14] is a nonlinear least squares problem that consists in a non-iterative curve fitting method. It is based on implicit representation to a dataset minimizing the approach mean square distance where statistical properties of noise are not considered. Second, Szpak's method [15] introduces an ellipse estimation procedure which endorses an optimization of the Sampson distance reached for a particular alternative of the Levenber-Marquardt algorithm. Third, Fitzgibbon's method [9,11] is an algebraic method that minimizes the algebraic distance. This method fits an ellipse including ellipticity constraint into the normalization factor. Fourth, Muñoz's method [1] is a robust multicriteria algorithm based on the mean absolute error. Fifth, Halir & Flusser's method [16] presents an approach based on a least squares minimization. This method is non-iterative and stable from a numerical point of view. It is simple and efficient even in the presence of a large amount of noisy data, hence recommended for an initial robust ellipse estimation. Sixth, Rosin's method [17] is an accurate and robust method based on least median of squares. The median of intrinsic parameters of the supposed ellipse is used to reckon the geometric parameters of the ellipse. Seventh, Prasad's method [18] proposes an efficient least squares ellipse fitting method without any constrained optimization that is stable under high levels of noise. Eighth, Köning's method [19] proposes a geometric fitting to an ellipse minimizing measured and fitted signal values' distance. This method supplies model parameters' best linear unbiased estimators and includes the statistical uncertainties. Ninth, Liang's method [20] proposes a linear combination of a subset of few data points to construct the ellipse parameters. Also, it employs the absolute residuals to reduce the contribution of extreme data points and binds the position error of data points to be robust against the worst case. Finally, Sobhani et al. [21] propose a method based on algebraic distance minimization which uses the robust Huber's function. Their algorithm detects inliers that can determine the ellipse parameters.

2.2. Ensemble methods

In literature, some methods to solve problems by ensembles of solutions have been described. However, most of them are applied to classification issues [22,23]. The accuracy of the ensemble is linked to the performance of each of the base algorithms used as well as the variety of their results. Choosing between more exact but less varied based learners and vice versa can actually be a rightful definition of distinct ensemble methods. In general, better outcomes are obtained since the process of blending and averaging makes the variance of the ensemble model decrease. The output of an ensemble depend on the individual achievement and the independence of the effect of each learner, i.e. high diversity and low error [24]. A common procedure to generate diversity is the use of different data subsets to train each base algorithm of the ensemble. In order to bring about different training data sets, some approaches such as random methods [24–27] and rotating (rotation forest) [28] are employed. Regarding the most well-known and useful ensemble methods are bagging, AdaBoost and LogitBoost [29]. Bagging (bootstrap aggregating) ensemble method was introduced in the mid-'90s by Breiman [30]. He uses a bootstrap sampling technique, guaranteeing improvements in terms of accuracy as it reduces the classification error variance. The second procedure is the AdaBoost method, the most extensively used boosting algorithm published by Freund and Schapire [31]. It entails the implementation of an adaptive re-sampling method with improved predictive behavior since it reduces both bias and variance [32]. The third kind is the LogitBoost, a boosting method presented by Friedman et al. [33], the idea of which is the bias and variance reduction [34]. LogitBoost process is considered as an expansion of the AdaBoost method [35,36]. Lopes [37] proposes a bootstrap process to estimate this variance for bagging, random forests and related algorithms in the context of classification.

Notwithstanding the above, Rosin [38] uses a method that stores five-point ellipse fits as an ensemble and then determines the fit using medians. The Rosin's method advantages comprehend the high amount of outliers permitted without requiring any random parameters and its reliability owing to the lack of problems with convergence as in others least-squares based methods.

Rosin uses the natural parametrization of the ellipse, based on the center coordinates, the length of the major and minor axes and the orientation. This means that with every five points subset an ellipse is fitted. When the outcome curve does not match an ellipse, the five-point subset is discarded. Taking all combinations of five points leads to an undesirable situation due to the great number of ellipses generated even if a small number of points are present. To accelerate this process, only a random sample of the original set of points that have not been employed yet is regarded. By doing this, each point of the initial data set is ensured to be included in one of the 5-tuple fits.

3. The SAREfit method

In this section our ellipse fitting method is detailed. Section 3.1 introduces our proposal, while Section 3.2 is devoted to the study of its robustness against outliers.

3.1. The algorithm

Let us consider an ellipse fitting algorithm φ that, given an input set of points T of size N, $T = \{\mathbf{x}_i \in \mathbb{R}^2 : i \in \{1, \ldots, N\}\}$ produces an output ellipse fit:

$$\begin{aligned} &\varphi : \mathbb{R}^{2N} \rightarrow \mathbb{R}^5 \\ &\varphi(T) = \boldsymbol{\psi} \end{aligned} \tag{2}$$

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

where $\boldsymbol{\psi} \in \mathbb{R}^5$ is a suitable parametric representation of the fitted ellipse which is obtained as the output of the algorithm. Now, if the overall training set $\mathcal{S}$ contains more than N samples, we may consider the 5-dimensional probability distribution of $\boldsymbol{\psi}$ when T is randomly drawn from $\mathcal{S}$ with replacement. Such multivariate distribution on $\mathbb{R}^5$ is expected to be unimodal, where the center of the distribution is close to the true ellipse that is to be estimated. In other words, the samples of $\boldsymbol{\psi}$ are expected to cluster around the point in $\mathbb{R}^5$ which represents the true ellipse under the considered ellipse parametrization. The less biased the algorithm, the closer the center of the distribution to the true ellipse. Depending on the proportion of outliers in $\mathcal{S}$, some of the realizations of $\boldsymbol{\psi}$ are noisy, i.e. some ellipse fits are bad estimates of the true ellipse. In order to overcome this difficulty, we propose to employ robust estimators of the center of the probability distribution of $\boldsymbol{\psi}$.

This robust estimation depends on the nature of the parametrization of the ellipse to be considered. As noted in [38], the algebraic parametrization of the ellipse is not adequate for accumulation of multiple fits. In this work, an alternative parametrization of the ellipse is proposed for this purpose:

$$\boldsymbol{\psi} = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma) = (\psi_1, \psi_2, \psi_3, \psi_4, \psi_5) \tag{3}$$

where (α_1, α_2) are the coordinates of the first focus of the ellipse, (β_1, β_2) are the coordinates of the second focus, and γ is the sum of the distances from each point of the ellipse to (α_1, α_2) and (β_1, β_2). This way, the ellipse is defined as the set of points $\mathbf{x} \in \mathbb{R}^2$ which satisfy the following equation:

$$\|(\alpha_1, \alpha_2) - \mathbf{x}\| + \|(\beta_1, \beta_2) - \mathbf{x}\| = \gamma \tag{4}$$

where $\|\cdot\|$ stands for the Euclidean norm. It must be highlighted that (3) represents an ellipse if and only if the following condition holds:

$$\|(\alpha_1, \alpha_2) - (\beta_1, \beta_2)\| < \gamma \tag{5}$$

i.e. the focal distance is smaller than the sum of distances parameter.

Let us consider K random realizations of the probability distribution of $\boldsymbol{\psi}$, which are associated to K subsets of size N uniformly drawn without replacement from $\mathcal{S}$, $Q = \left\{\boldsymbol{\psi}_j : j \in \{1, \ldots, K\}\right\}$. The fact that (α_1, α_2) and (β_1, β_2) are points on the plane suggests a robust procedure to estimate the center of the distribution of $\boldsymbol{\psi}$. The robust estimates of the coordinates of the two foci are computed as follows:

$$\left(\hat{\alpha}_1, \hat{\alpha}_2\right) = \arg \min_{(\alpha_1, \alpha_2) \in \mathbb{R}^2} \sum_{j=1}^{K} \left\| (\alpha_1, \alpha_2) - \left(\psi_{j,1}, \psi_{j,2}\right) \right\| \tag{6}$$

$$\left(\hat{\beta}_1, \hat{\beta}_2\right) = \arg \min_{(\beta_1, \beta_2) \in \mathbb{R}^2} \sum_{j=1}^{K} \left\| (\beta_1, \beta_2) - \left(\psi_{j,3}, \psi_{j,4}\right) \right\| \tag{7}$$

that is, $\left(\hat{\alpha}_1, \hat{\alpha}_2\right)$ is the spatial L1 median of the first estimated foci $(\psi_{j,1}, \psi_{j,2})$, while $\left(\hat{\beta}_1, \hat{\beta}_2\right)$ is the L1 median of the second estimated foci $(\psi_{j,3}, \psi_{j,4})$. The robustness of the estimation comes from the fact that the sum of Euclidean distances is considered in (6) and (7), rather than the sum of the squares of the Euclidean distances. The latter choice would have led us to the mean of the sets of estimated foci, and the mean is known to be non robust [39].

The estimate $\hat{\gamma}$ of the sum of distances parameter can also be robustly computed by the univariate median:

$$\hat{\gamma} = \arg \min_{\gamma \in \mathbb{R}} \sum_{j=1}^{K} \left| \gamma - \psi_{j,5} \right| \tag{8}$$

where $|\cdot|$ stands for the absolute value of a real number.

As done in (6) and (7), the selection of a robust estimator ensures that a small quantity of bad fits $\boldsymbol{\psi}_j$ does not affect the estimation too severely. A tentative estimation can be built by joining the results of (6), (7) and (8):

$$\hat{\boldsymbol{\psi}} = \left(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}\right) \tag{9}$$

An additional postprocessing stage is proposed in order to filter out the deleterious effect of the bad fits. First of all, the Euclidean distance in the natural parameter space is computed between the tentative estimation and the fits:

$$d_j = \left\| \hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_j \right\| \tag{10}$$

where $j \in \{1, \ldots, K\}$. Then the q-th quantile ω of the distances d_j is computed:

$$P(d < \omega) = q \tag{11}$$

where $q \in (0, 1)$ is a tunable parameter. After that, the final estimation is obtained as a robust estimate based on the best fits according to the distance d:

$$(\tilde{\alpha}_1, \tilde{\alpha}_2) = \arg \min_{(\alpha_1, \alpha_2) \in \mathbb{R}^2} \sum_{d_j < \rho} \left\| (\alpha_1, \alpha_2) - \left(\psi_{j,1}, \psi_{j,2}\right) \right\| \tag{12}$$

$$\left(\tilde{\beta}_1, \tilde{\beta}_2\right) = \arg \min_{(\beta_1, \beta_2) \in \mathbb{R}^2} \sum_{d_j < \rho} \left\| (\beta_1, \beta_2) - \left(\psi_{j,3}, \psi_{j,4}\right) \right\| \tag{13}$$

$$\tilde{\gamma} = \arg \min_{\gamma \in \mathbb{R}} \sum_{d_j < \rho} \left| \gamma - \psi_{j,5} \right| \tag{14}$$

$$\tilde{\boldsymbol{\psi}} = \left(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\beta}_1, \tilde{\beta}_2, \tilde{\gamma}\right) \tag{15}$$

The proposed algorithm reads as follows:

1. For values of j between 1 and K, repeat steps 2 and 3.
2. Draw uniformly at random a subset T of the training set $\mathcal{S}$, with replacement.
3. Run the base ellipse fitting algorithm φ with input T to obtain the output ellipse fit $\boldsymbol{\psi}_j$.
4. Apply Eqs. (6)–(8) to obtain the tentative estimation $\hat{\boldsymbol{\psi}} = \left(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}\right)$.
5. Compute the distances in the parameter space by (10) for all the fits.
6. Apply Eqs. (12)–(14) to obtain the final estimation $\tilde{\boldsymbol{\psi}} = \left(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\beta}_1, \tilde{\beta}_2, \tilde{\gamma}\right)$.
7. If $\tilde{\boldsymbol{\psi}}$ corresponds to an ellipse as described by (5), then return $\tilde{\boldsymbol{\psi}}$ and halt. Otherwise, run a backup ellipse fitting algorithm with input $\mathcal{S}$, then return the obtained output and halt.

The backup procedure of step 7 is included because sometimes the tentative estimation does not describe an ellipse. This might happen when the training set $\mathcal{S}$ does not correspond to an ellipse, or when the individual fits $\boldsymbol{\psi}_j$ have a poor quality. It is worth noting that circles, as a special case of ellipses, can be detected equally well as non circular ellipses. For a circle we have that the two foci (α_1, α_2) and (β_1, β_2) are the same. This means that the estimated foci of the base ellipse fits $(\psi_{j,1}, \psi_{j,2})$ and $(\psi_{j,3}, \psi_{j,4})$ concentrate around the center of the circle. Consequently the final estimations of both foci $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ and $\left(\tilde{\beta}_1, \tilde{\beta}_2\right)$ are close to the center of the circle, which is the correct solution. Fig. 1 depicts an example of circle fitting. It can be observed that both estimated foci coincide up to a distance around 10^{-9}.

3.2. Robustness analysis

In order to assess the robustness of the SAREfit algorithm against outliers, a formal analysis of the expected number of 'bad' ellipses generated in step 3 of the algorithm is carried out in this

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

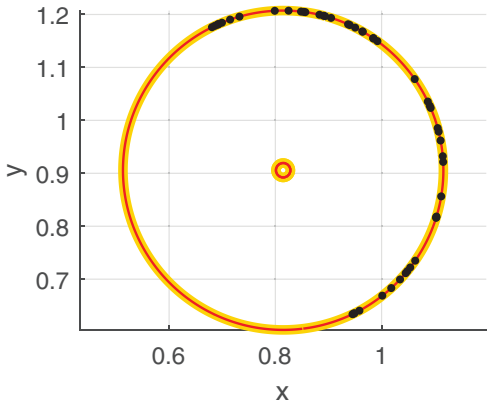


Fig. 1. Example of the SAREfit execution with a perfect circle (degenerate case).

subsection. This study is done with respect to the fraction of outliers in the training dataset and the robustness of the baseline algorithm. In other words, in this subsection, it is analyzed how many outlying input points can exist in the training set without preventing our approach to yield a reasonable output ellipse fit. Let us remember that the breakdown point of an estimator is the fraction of outlying input samples that must exist in order to make the estimator produce a broken estimation, i.e. an estimation with too large output parameters. On the other hand, the q-th quantile of a distribution is the value that leaves a fraction q of the samples below the quantile.

Let us assume that the baseline algorithm has a breakdown point of $\xi \in [0, 1]$. That is, the baseline algorithm fails in the fit of the ellipse whenever the proportion of outliers in the input data to the baseline algorithm is higher than ξ. Also, let $\zeta \in [0, 1]$ be the proportion of outlying (bad) samples contained in the training set S. Thus, the number of outlying (bad) samples B in a subset T with cardinal N of the training set S follows a binomial distribution:

$$B \sim Bi(n = N, p = \zeta) \tag{16}$$

where N is the number of samples in a subset T to be provided as input to the baseline algorithm.

The baseline algorithm will produce a bad output ellipse whenever the number of outlying samples in its input subset T is between ξN and N. Next, the probability of having between ξN and N outlying (bad) samples in a subset T of size N of the training set S can be estimated computing the sum of the probabilities of achieving between ξN and N successes with the above binomial probability distribution:

$$P(\xi N \leq B \leq N) = P(B \geq \xi N) \tag{17}$$

For $\xi N \leq N$, upper bounds for the lower tail of the distribution function can be derived. Equivalently, the upper tail bound is derived directly from the Chernoff bound [40]:

$$P(B \geq \xi N) \leq \exp\left(-N \cdot D(\xi \| \zeta)\right) \quad \text{if } \zeta < \xi < 1 \tag{18}$$

where $D(t\|s)$ denotes the relative entropy between the *Bernoulli*(t) and *Bernoulli*(s) distribution and is obtained as:

$$D(t\|s) = t \log \frac{t}{s} + (1-t) \log \frac{1-t}{1-s} \tag{19}$$

Combining Eqs. (18) and (19), the following bound is derived:

$$P(B \geq \xi N) \leq \left(\frac{\zeta}{\xi}\right)^{\xi N} \left(\frac{1-\zeta}{1-\xi}\right)^{(1-\xi)N} \quad \text{if } \zeta < \xi < 1 \tag{20}$$

Eq. (20) gives an upper bound of the probability $P(B \geq \xi N)$ that a bad ellipse is produced as the output of an execution of the baseline algorithm. Step 4 of our proposed algorithm (Section 3.1) employs medians to obtain the tentative estimation of the ellipse by computing medians of the parameters of the output ellipses coming from the baseline algorithm. Since the median has a breakdown point of 0.5, in order to avoid that step 4 of the algorithm generates a bad ellipse, it should be ensured that $P(B \geq \xi N) < 0.5$. From Eq. (20), this can be guaranteed if the following condition holds:

$$\left(\frac{\zeta}{\xi}\right)^{\xi N} \left(\frac{1-\zeta}{1-\xi}\right)^{(1-\xi)N} < 0.5 \quad \text{if } \zeta < \xi < 1 \tag{21}$$

Therefore, values of ζ and ξ which fulfill Eq. (21) should be employed in order to ensure the quality of the final ellipse fit.

In Fig. 2 the representation of Eq. (20) is depicted for a variety of values of ζ and ξ between 0 and 1. When ζ is increased,

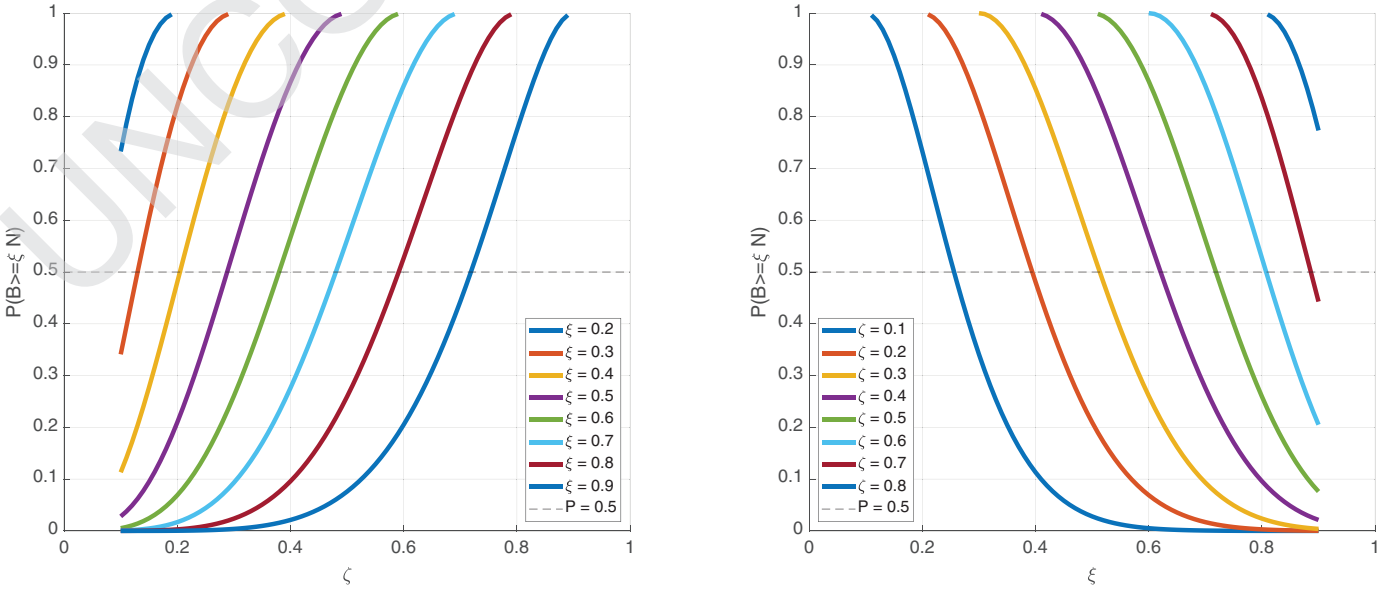


Fig. 2. Graphical analysis of the robustness of SAREfit method. Values of ζ and ξ are sampled in the interval (0,1) and Eq. (20) is evaluated. Horizontal dashed black line represents the breakdown point of the median.

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

i.e. there are more outliers in the dataset, the baseline algorithm tolerates a higher proportion of bad samples before reaching the breakdown point of 0.5. Likewise, the second graphic shows that if the tolerance of the baseline algorithm to cope with many outlying points is good enough, then the whole dataset S may contain a larger number of outliers while our proposal still works fine since there are more values of ζ below $P = .5$.

4. Experimental results

A set of experiments were designed in order to assess the robustness of our method. First, the evaluation metrics are described in Section 4.1 and the parameter optimization procedure is summarized in Section 4.2. The results on synthetic and real data samples are discussed in Sections 4.3–4.6. Finally, the problem of multiple ellipse fitting is discussed in Section 4.7.

4.1. Evaluation method

The proposed method, named as SAREfit[1], was compared with other state-of art algorithms for ellipse fitting, i.e. Muñoz [1], Taubin [14], Szpak [15], Fitzgibbon [9], Halir&Flusser [16], Rosin [17], Prasad [18], Köning [19], Liang [20] and Sobhani [21] methods. Default parameter values were used for all the methods, with the exception of Liang and Sobhani where their sparsity parameter λ was tuned for our experiments because their proposed value does not fit with our experimental setup. $\lambda = 0.2$ was chosen for both methods.

We have considered the use of performance profiles [41] with the intention of condensing the results of numerous tests. The aim of this representation is to build a probability cumulative distribution that measures the performance of each method. Thus, the best method will be the one with the nearest probability curve to the left-upper corner of the figure axes. A detailed description of this representation can be found in the literature.

Four different error measures were considered in order to assess the robustness of our method compared with the competitors:

1. Natural error: the natural parametrization (3) of the fitted ellipse is computed and then it is compared with the true values as

$$NaturalError = \sqrt{\sum_{i=1}^{5} (\boldsymbol{\psi}^i - \tilde{\boldsymbol{\psi}}^i)^2} \tag{22}$$

2. Algebraic error: the algebraic parametrization of the ellipse is considered, $\boldsymbol{\xi} = (A, B, C, D, E, F)$ associated to the general equation of a conic section. The true and predicted parameters are normalized and the error is computed as

$$AlgebraicError = \sqrt{\sum_{i=1}^{5} \left(\frac{\boldsymbol{\xi}^i}{||\boldsymbol{\xi}||} - \frac{\tilde{\boldsymbol{\xi}}^i}{||\tilde{\boldsymbol{\xi}}||} \right)^2} \tag{23}$$

3. Geometric errors: given the geometric parametrization $\boldsymbol{\zeta} = (x, y, a, b, \theta)$, where $\mathbf{c} = (x, y) \in \mathbb{R}^2$ is the center of the ellipse, a is the half length of the major axis, b is the half length of the minor axis, $a \geq b > 0$, and $\theta \in [0, \pi]$ is the angle of tilt, we computed the errors of the center, angle, major and minor semiaxes and area of the ellipse:

$$\epsilon_c = \|\mathbf{c} - \tilde{\mathbf{c}}\|, \ \epsilon_a = |a - \tilde{a}|, \ \epsilon_b = \left|b - \tilde{b}\right|, \ \epsilon_A = \left|A - \tilde{A}\right|, \ \epsilon_\theta = \left|\theta - \tilde{\theta}\right| \tag{24}$$

4. Root Mean Squared Orthogonal (RMSO) error: a set of M points belonging to the true ellipse are projected to the estimated ellipse in order to compute their orthogonal distances o_i [42]. Then, the root mean squared error of those measures is calculated as:

$$RMSOError = \sqrt{\frac{1}{M} \sum_{i=1}^{M} o_i^2} \tag{25}$$

The use of more than one performance measure of different nature is a good way to assess the quality of a new methodology. Depending on the application for which the fitting method will be used, the principal error measure should be different. Our algorithm is based on the natural parametrization, so the natural error is a bit more biased because most of the competitors use the algebraic or geometric parametrization. Nevertheless, when the aim of the fit is to obtain a good precision on the focal points, this error has to be the reference. On the other hand, algebraic and geometric errors are correlated, and it is for interest when we want to achieve a correct orientation and center of the ellipse, for example. Finally, the RMSO error is a general measure that can be used to compare different methods irrespective of their fitting criterion and the quality of the fitted ellipses [43]. In order to overcome problems with algebraic errors, we can make use of the RMSO error which is not associated with any parametrization, is invariant to transformations in Euclidean space and has less high curvature bias. It is also suitable when the true ellipse is not provided although might be affected by the presence of outliers.

All experiments were carried out on a 64-bit PC, with an Intel Core i7-4790, 3.6GHz CPU, 32 GB RAM and standard hardware. Matlab R2018b was used to implement the proposed method as well as to perform the comparisons with the competing methods, with no use of GPU resources.

4.2. Parameter fitting

SAREfit model is based on an ellipse fitting algorithm φ and a backup algorithm too, as described in the previous section. Due to its extreme simplicity, Halir&Flusser's method was chosen to be this algorithm φ, and Fitzgibbon to be the backup method. Both of them with their default parameters. Nevertheless, SAREfit includes its own parameters that need to be fitted:

- Subsampling factor (s_T): is the proportion of the input samples considered in the fitting procedure. size of the set S creates the input set T, as it is described in Section 3.
- Ensemble size (K): the number of distributions created from the whole samples, i.e., the number of ellipses fitted by the algorithm using different training sets.
- Quantile (q): the q-th quantile value used to filter the fits according to the natural error.

In Table 1 both the set of tested parameters and those found to be the best ones are summarized according to the natural, algebraic and RMSO errors. All the possible combinations of parameters were run in order to find a global minimum. 100 different initializations were executed incorporating between 5% and 20% of

Table 1
Parameters tested for the SAREfit algorithm and the optimal results according to the Natural, Algebraic and RMSO errors. The selected parameters are in bold.

Parameter	Tested values	Nat. Er.	Alg.Er.	RMSO Er.
s_T	{0.1,...,1} (step 0.05)	**0.15**	0.15	0.70
K	{15,...,150} (step 15)	**90**	90	15
q	{0,0.05,0.1,...,1} (step 10)	**0.10**	0.10	0

[1] The source code and demo of the proposed approach will be published in case of acceptance.

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

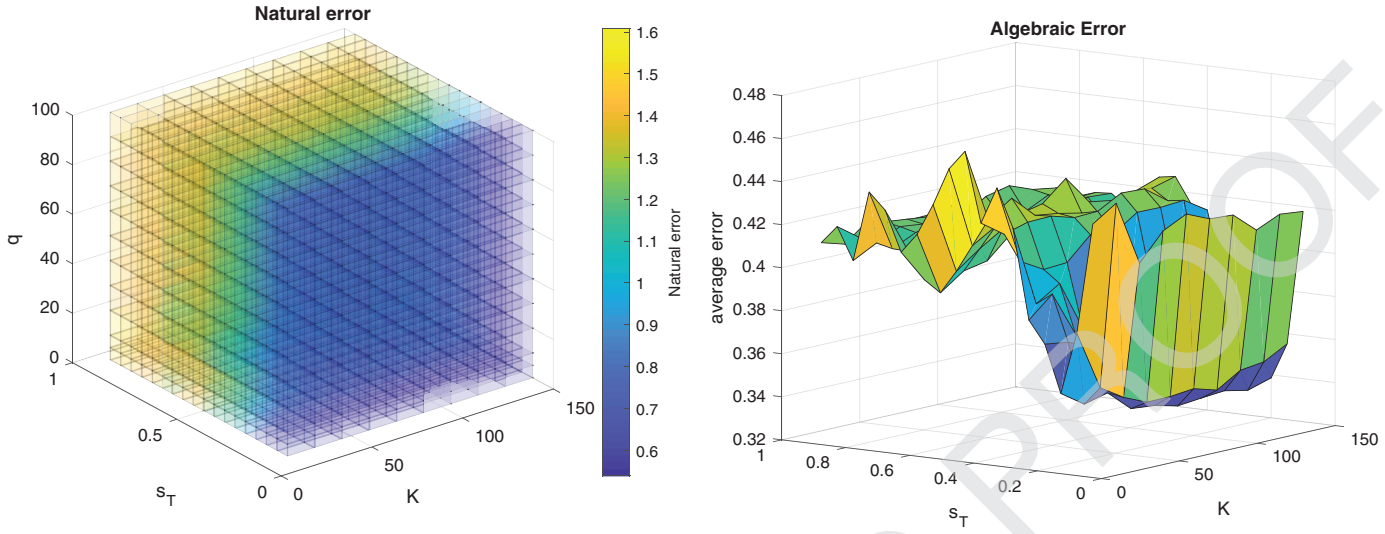


Fig. 3. Parameter optimization based on the average along 100 executions. Left: 4D representation of the Natural error obtained in the simulations (the darkest the better). Right: Algebraic errors varying the subsampling factor and the ensemble size, and fixing $q = 0.10$.

outliers randomly, and the average errors were calculated to have a precise estimation. $q = 0$ means that the best ellipse based on the error with respect to the tentative estimation $\hat{\boldsymbol{\psi}}$ is chosen. The best configuration was exactly the same for two types of errors. $s_T = 0.15$ indicates that 15% of the input points are enough to achieve good results. This can be interpreted as a way to remove outlier points from the training dataset. $K = 90$ fits seems to be enough to achieve stability across different initializations. Finally, the post-selection of the best 10% of ellipses with respect to the first estimation assures improvement of the outcome. These estimated parameters were used for both synthetic and real experiments. Fig. 3 depicts the detailed results of the optimization. The cube represents all possible combinations of the parameters, whose outcomes are plotted following the color bar. The best results are achieved for small values of Subsampling factors, which confirms our supposition about the good performance with the presence of outliers. The right side image represents one slice of the cube ($q = 0.10$), so the tendency of the other two parameters can be determined. We can observe a clear minimum at $s_T = 0.15$ and a stabilization of the error when the number of fits is increased.

4.3. Synthetic data

The first experiments deal with artificial datasets generated as follows:

1. Firstly, the center, the major and minor axes of a ellipse are chosen at random uniformly: $x,\ y \sim U(0,\ 1)$, $a \sim U(0.2,\ 1)$, $b \sim U(0.1,\ 1)$. The major and minor axes have different ranges to avoid degenerated ellipses. The tilt angle is also chosen uniformly $\theta \sim U\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$.
2. Secondly, $N = 50$ sample points are generated on the canonical coordinate system: $(a\cos\phi, b\sin\phi) \in \mathbb{R}^2$, where $\phi \sim U(\phi_s,\ \phi_e)$, $\phi_s, \phi_e \sim U(-\pi, \pi)$, and satisfying that $\phi_e - \phi_s > 1$ in order to avoid the prediction of degenerated ellipses caused by datasets with small curvature. The random choice of the starting and ending angles combined with this constraint allows having different levels of occlusion in the training points of the ellipse.
3. Finally, 1% of normally distributed Gaussian noise is added and a specific percentage of the points are modified to convert them to outliers.
4. In addition to this, 1000 test points on the true ellipse are generated to carry out the computation of the RMSO error.

Fig. 4 presents four examples of synthetic ellipses whose training samples contain 5%, 10%, 15% and 20% of anomalous points respectively. Fig. 4a shows almost a degenerated ellipse, and most of the competing methods failed in the fit. Muñoz, Halir&Flusser and Szpak methods achieved to output an ellipse, but SAREfit was the only one able to produce an ellipse with very similar characteristics to the ground truth, with only a small displacement. Second and third examples (Figs. 4b-4 c) also have a high level of occlusion (around 50%), and the number of outliers was increased. In both cases, all methods had a better outcome, although most of them were affected by the presence of outliers. Again, the most accurate prediction was the one generated by SAREfit, followed by Rosin and Muñoz methods. At last, an example with 20% of outliers and less occlusion is shown in Fig. 4d. All the algorithms approximately fitted the ellipse with the exception of Liang algorithm, and our proposal is the only method which completely overwrites the edge of the ellipse.

Next, batches of 1000 different random initializations and with random occlusion were run varying the percentage of outliers between batches, and the results are summarized in Fig. 5 using the performance profiles described before. Note that the best method is the one that first reaches probability 1. Fig. 5a–f contains the outcomes of 5% and 10% of outliers. There are six methods that never attain to solve all the possible configurations posed in the batch: Taubin, Rosin, Prasad, Köning, Liang, and Sobhani. This means that they are very unstable if the input samples do not have a refined shape of an ellipse. There is more variability with the rest of the methods depending on the metric. In terms of Natural and Algebraic error, SAREfit is clearly the best method, solving always all the fits with less error than the competitors. Moreover, our proposal is also good using the RMSO measure as the hard training datasets ($\tau > 2$) are solved with less error, i.e. when the quantity of outliers is incremented, the results were more robust in favor of SAREfit. Muñoz and Szpak algorithms also perform well for the Natural and RMSO errors. Fig. 5g–l show the results of 15% and 20% of anomalous points. Here, Rosin method improved its results reaching almost 90% of the ellipse fits. On the other side, the positive differences observed with respect to Muñoz and Szpak methods is incremented and the gap based on RMSO error is clearly reduced. The case of Liang and Sobhani methods is

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

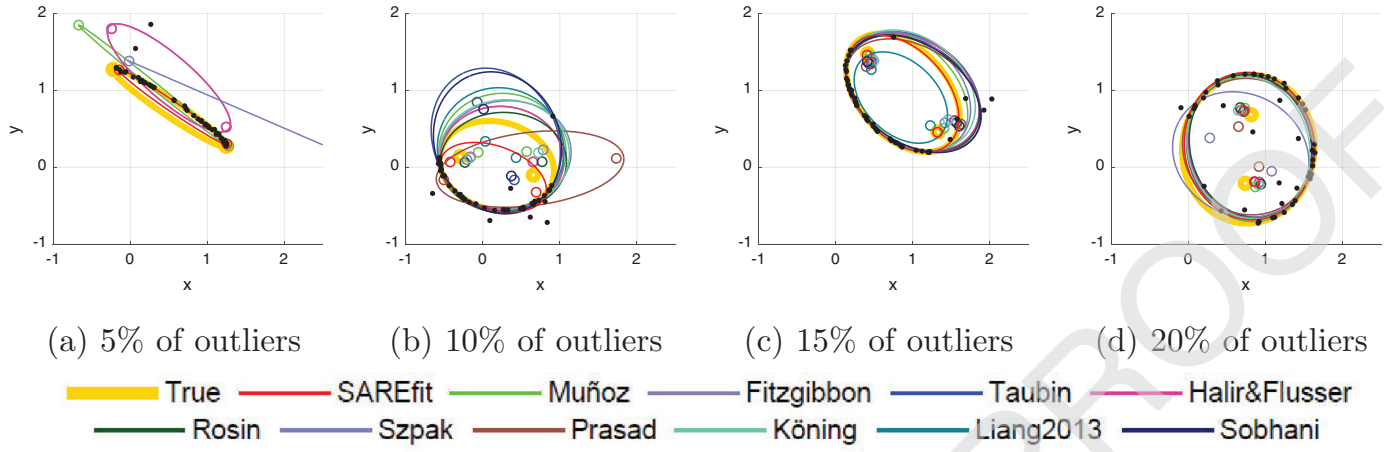


Fig. 4. Graphical comparison using synthetic data generation. Four different initializations and their solutions are shown. The black points are the training samples and the true ellipse is plotted as a thick solid curve, while the fitted ellipses are shown as a narrow solid curve.

very particular because one expects a better performance. However, these methods model the sparsity of the samples and they work well under controlled situations, such as the absence of occlusion and large numbers of samples. Unlike SAREfit, this makes both methods inappropriate for extreme configurations. Nevertheless, when the number of outliers is fixed and there is no occlusion (see Subsection 4.5) they improve although they are also affected by eccentricity.

Fig. 6 sums up the geometric errors for 10% of outliers. The mean and median values in the box plots are shown as a gray circle and a line, respectively. A small pattern can be found in each feature with respect to our proposal: SAREfit has located its average and median below the ones obtained by the competing methods, although there is a bit more dispersion across the fits. This means that most of the input samples are fitted better, although sometimes the differences are larger. The worse method is Szpak, showing a great dispersion caused by the generation of incorrect ellipses. SAREfit and Muñoz methods are the best ones focusing on the center, area and angle, followed by Fitzgibbon. Köning method stands out for its nice performance in center, semiaxes, and area but it has problems to fit adequately the angle of the ellipse, which might be caused by the outliers. Liang and Sobhani methods yield adequate values for the center and area, but the size and orientation of the fitted ellipse are imprecise, affecting all error measures.

The outcomes for 20% of outliers are presented in Fig. 7, where the differences between methods are more obvious. Fitzgibbon, Halir&Fluser, and SAREfit returned the lowest errors, especially for the center, major semiaxis, and area. Muñoz still worked fine in terms of angle and minor semiaxis but its median and mean value are larger for the other features, which are in consonance with the gap found in the performance profiles. Prasad method has the third best results focusing on angle orientation and minor semiaxis, but this is not enough to obtain a good fit, as it is observed with the other metrics. Liang and Sobhani methods are still behind most of the methods although their error range is small, which shows some level of stability. The main problem of most methods is the orientation of the fitted ellipse, which causes major differences in terms of performance.

In addition to the previous tests, an analysis of the upper limit of outliers was carried out to check how the competing methods deal with anomalous samples. Therefore, a collection of 1000 runs without occlusion was executed for different percentages of outliers, ranging from 0% to 50%. Note that higher values of outliers might be considered as extreme noise level that modifies the original elliptical shape of the data points. In the first three plots of Fig. 8, the median errors obtained for each measure are shown. It is worth noting that the higher level of outliers the more wrong fits are obtained by most of the methods, as shown in the last plot. These bad outputs were discarded in the computation of the median, making smoother the error curves from 30% onwards. Only SAREfit, Muñoz, Fitzgibbon and Halir&Flusser methods do not fail in the fit, which makes them very stable algorithms. Algebraic measure reveals that Taubin, Fitzgibbon, Sobhani and Liang methods are completely unstable. The break point of most of the algorithms is placed at 10% of outliers.

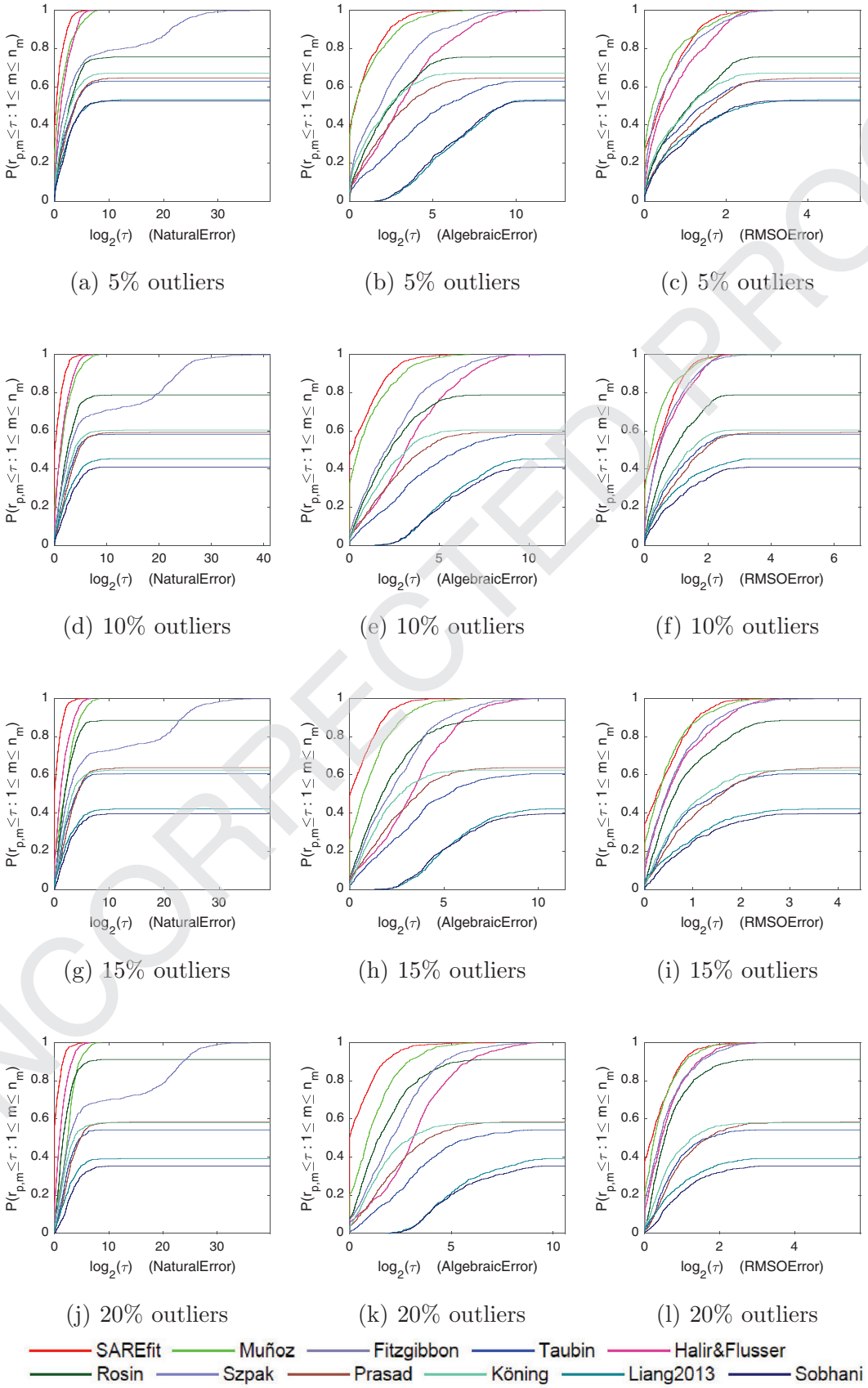
Table 2
CPU times and complexity comparison of the methods using synthetic data. Mean, median, standard deviation and interquartile range computed among 1000 runs.

Method	CPU time (s)				Complexity
	Mean	Median	Std	IQR	
SAREfit	0.0116	0.0108	0.0019	0.0018	$O(KN + N \log N)$
Muñoz	0.0066	0.0060	0.0017	0.0006	$O(N \log N)$
Fitzgibbon	0.0002	0.0002	0.0001	0.0001	$O(N)$
Taubin	0.0002	0.0002	0.0002	0.0000	$O(6N + 72) \sim O(N)$
Halir&Flusser	0.0001	0.0001	0.0001	0.0000	$O(N)$
Rosin	0.0002	0.0001	0.0002	0.0000	$O(N)$
Szpak	0.0189	0.0123	0.0231	0.0107	$O(200N) \sim O(N)$
Prasad	0.0001	0.0001	0.0002	0.0000	$O(30N) \sim O(N)$
Köning	0.0073	0.0070	0.0023	0.0010	$O(N^3)$
Liang	0.3901	0.3839	0.0306	0.0223	$O((2N + 1)^{3.5} + N^2)$
Sobhani	0.0026	0.0024	0.0005	0.0003	$O(N^2)$

4.4. Computational efficiency

The performance of each method in terms of computation time was first analyzed. The mean, median, standard deviation and interquartile range values of 1000 runs are presented in Table 2. In general terms, fitting an ellipse does not take a long time. Most of the methods compute a fit in less than one millisecond, with the exception of Szpak, Muñoz, Köning, Liang, and SAREfit. Liang and Szpak methods are the slowest ones, while their irregular performance does not justify this computational need. Taking into account that SAREfit employs Halir&Flusser method as a baseline carrying out 90 executions, 0.0116 seconds for the whole procedure is an acceptable execution time for any ellipse fitting method, being adequate for its use in real-time applications. Moreover, it has to be noted that the presence of a cache may improve the performance of the algorithm, by making the execution of the loops faster.

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406



(a) 5% outliers (b) 5% outliers (c) 5% outliers

(d) 10% outliers (e) 10% outliers (f) 10% outliers

(g) 15% outliers (h) 15% outliers (i) 15% outliers

(j) 20% outliers (k) 20% outliers (l) 20% outliers

Fig. 5. Performance profiles of synthetic experiments (the closer to the upper left corner, the better). Natural, Algebraic and RMSO errors are analyzed with 5%, 10%, 15% and 20% of outliers. X axis shows the factor of the best possible ratio in a logarithmic scale and Y axis represents the probability cumulative distribution.

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

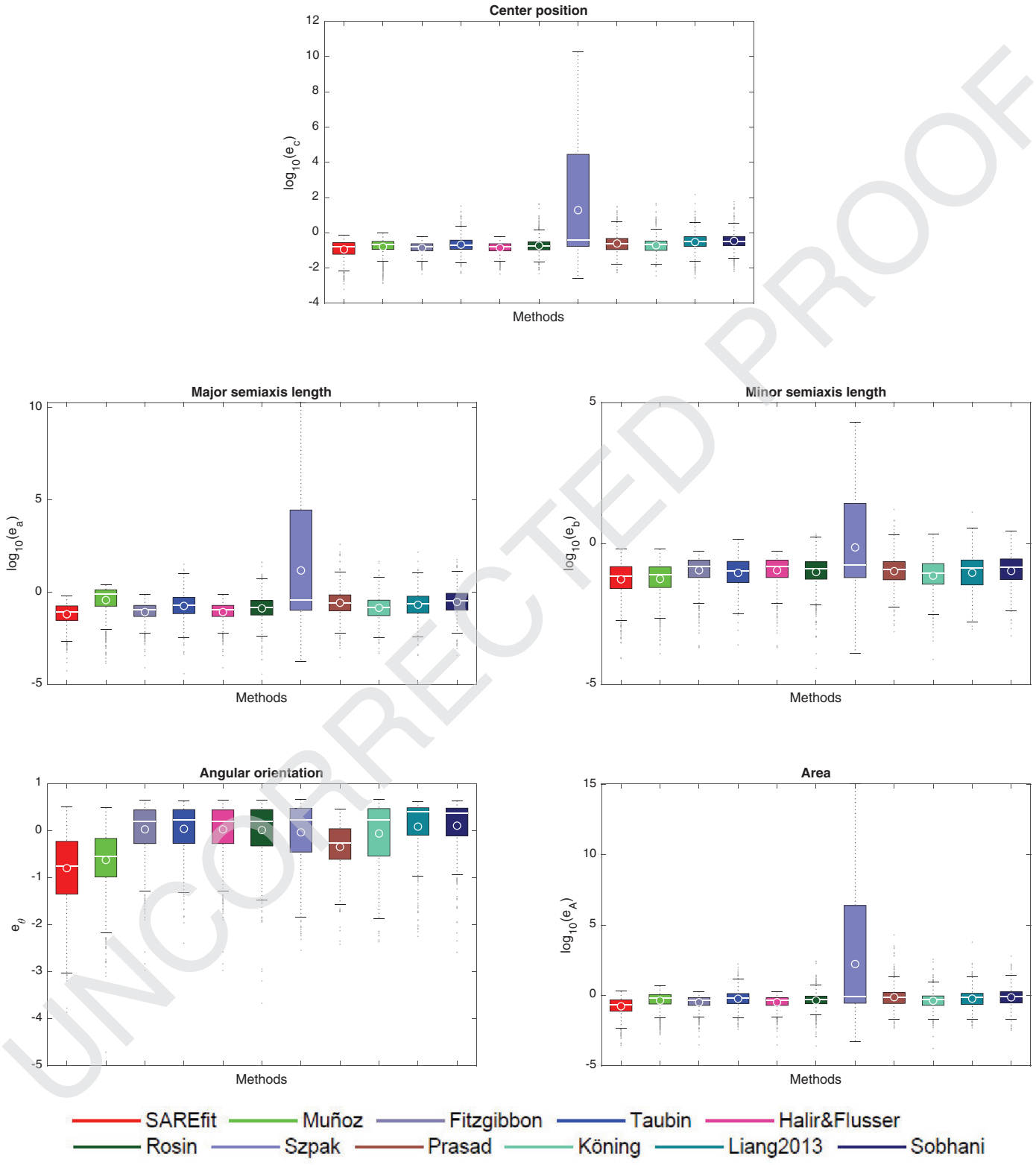


Fig. 6. Major and minor axes, center, angle and area box plots comparisons. Geometric errors are analyzed with 10% of outliers. X axis shows the factor of the best possible ratio in a logarithmic scale and Y axis represents the probability cumulative distribution.

From a theoretical point of view, SAREfit requires $O(K)$ iterations of its main loop, where the computational complexity of each iteration is the one of the baseline algorithm, which is linear with respect to the number of input samples [16], $O(N)$. Therefore, the complexity of our proposed algorithm is $O(KN + N\log N)$, where the last term comes from the computation of the spatial median. Compared with the competitors (see Table 2), our proposal has an acceptable complexity. The most recent works are highly dependent from the number of points, $O(N^2)$ and above, being $O(N\log N)$ more efficient than them with higher quality performance.

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

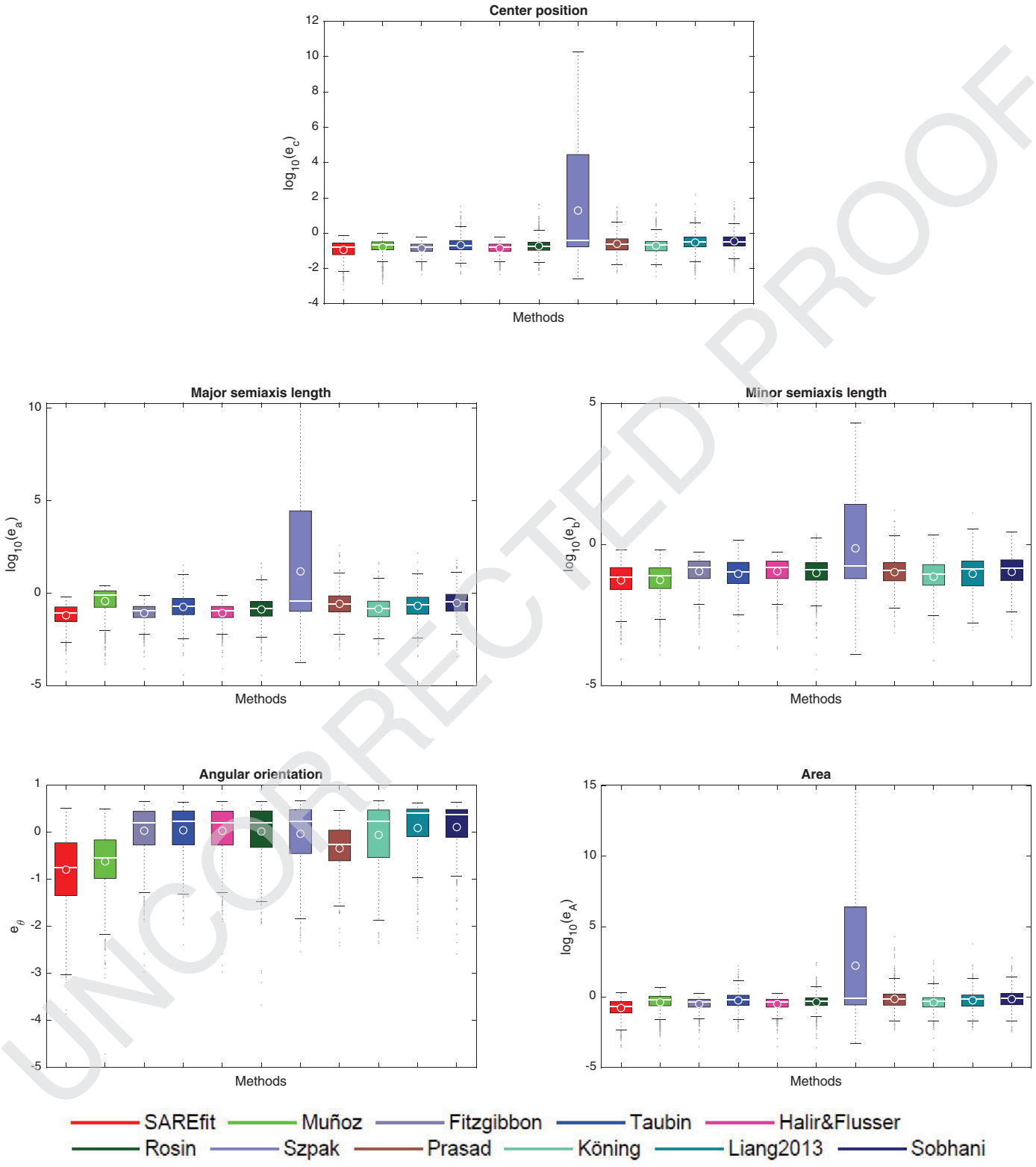


Fig. 7. Major and minor axes, center, angle and area box plots comparisons. Geometric errors are analyzed with 20% of outliers. X axis shows the factor of the best possible ratio in a logarithmic scale and Y axis represents the probability cumulative distribution.

4.5. Occlusion analysis

This Subsection intends to evaluate in detail the performance of each method with the presence of low, medium and high levels of occlusion. For this purpose, again a batch of 1000 runs with 0%, 25%, 50% and 75% of occlusion was executed and error measurements were computed. The procedure was as it follows: starting angle of the generated ellipse is selected from $\phi_s \sim U(-\pi, \pi)$. Then, the ending angle is computed as $\phi_e = \phi_s + occl \cdot 2\pi$, where *occl* represents the occlusion level in the range [0,1]. The fraction

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

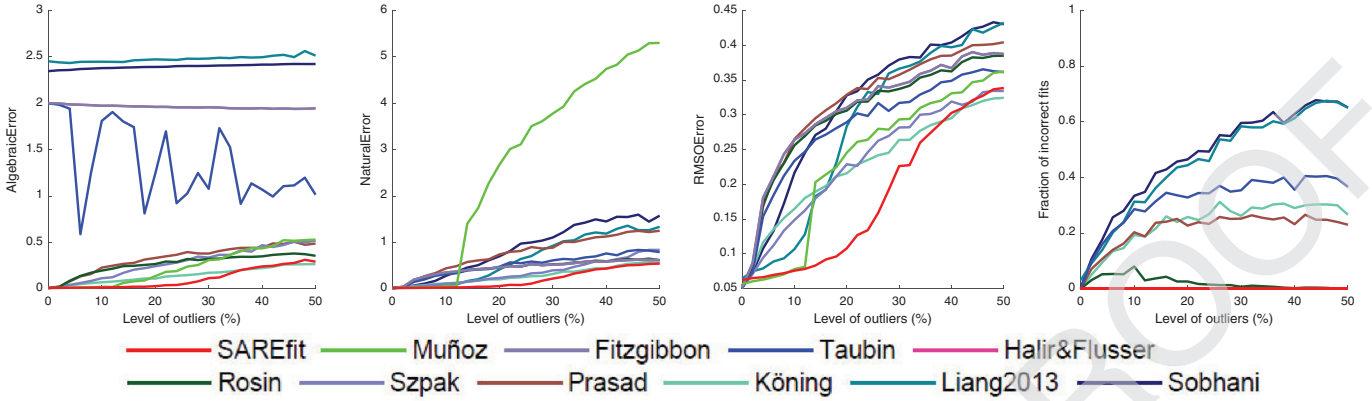


Fig. 8. Evolution of the performance of the methods varying the level of outliers (0-50%). Median values of the algebraic, natural and RMSO errors over 1000 runs are represented. Las t figure counts the number of not successful fits.

of outliers was chosen at random from the uniform distribution between 0% and 20%.

Fig. 9 presents the performance profiles of the Natural, Algebraic and RMSO errors The general tendency is the more occlusion is present in the training samples, the less is the quality of the fit for all methods. Nevertheless, in most cases, SAREfit method yields the best result. When the shape of the ellipse is complete or only a quarter is missing, our proposal is significantly better in terms of Algebraic and RMSO errors. Here, Liang and Sobhani methods perform much better. There are around 40% of ellipses that are fitted better than most of the methods, as the RMSO and Natural errors demonstrate. The comparisons with the Natural error are stretcher when the level of occlusion was increased. For 50% of occlusion, Szpak method become competitive for the natural error, and with Muñoz algorithm, the 1000 runs (probability 1) are solved with almost the same error as the obtained by SAREfit. When the tests are carried out with 75% of occlusion, Muñoz method is the best for the Algebraic error followed by our proposal, but being Szpak the second best in terms of orthogonal error and the first method for the Natural error.

An example of the outcomes with the presence of occlusion is depicted in Fig. 10. For 25% the fit of our proposal is almost perfect although Muñoz method fails. Also, Szpak method is near to the true ellipse, shown in yellow. Köning, Fitzgibbon, Prasad, Liang, and Sobhani do not fit an ellipse. With 50% of occlusion, the competitors are clearly affected by the presence of outliers. Finally, with 75% of occlusion, most of the methods go through the ellipse points but fail in length of the major and minor semiaxes. SAREfit yields the most similar ellipse to the true one.

4.6. Real data

This subsection presents real examples of four images: two of them from the Caltech 256 dataset [44], named as *Satellite dish* (image '169_0015') with size 448×336, and *Can* (image '195_0039') with size 247×350, and other two images of wheels captured by a normal camera, named as *Wheel 1* and *Wheel 2*, with sizes 2061×2891 and 2416×2024, respectively. Canny edge detector algorithm was used to randomly extract a set of 50 points from each image, using threshold parameter equal to 0.1 for the *Satellite dish*, and 0.9 for the other images. Resulting binary images were refined using the *imfill*, *imclearborder*, *im2bw* and *bwperim* Matlab morphological functions with default parameters.

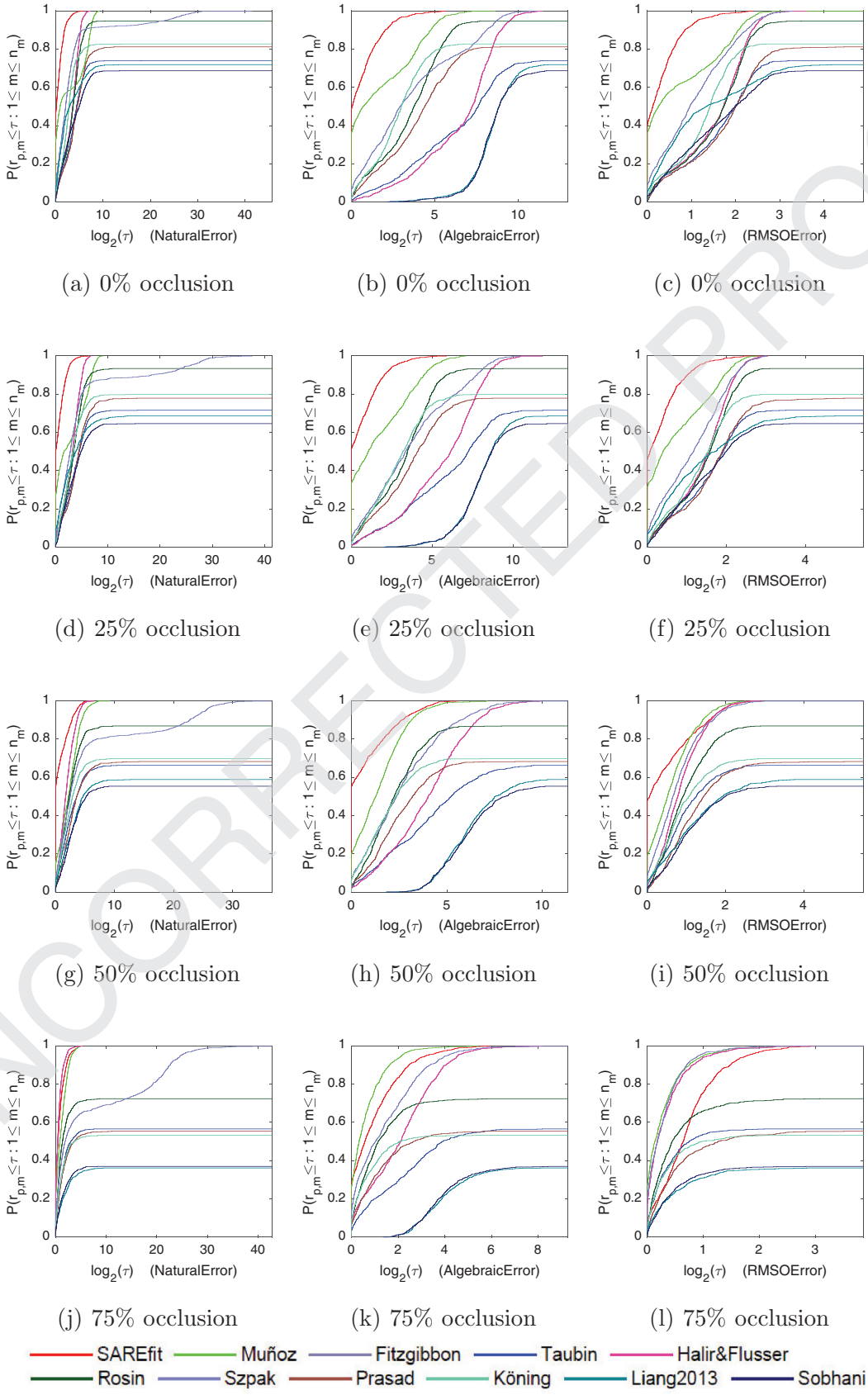
Fig. 11 shows the fits generated by the eleven tested methods. The *Satellite dish* (Fig. 11a) contains two extraneous points in the bottom right part of the image, as well as those located on the focus of the antenna, provoking that all methods failed in the fit except SAREfit and Szpak methods. The inferior part of the Szpak outcome falls outside the border of the satellite dish, being SAREfit more precise. In the case of the *Can* (Fig. 11b), there is a group of points located on the superior part which hampers the fit processing. Muñoz and our method are the only ones that were able to find the real solution. Finally, Fig. 11c–d are examples with some outliers outside and inside the ellipse shape respectively. Here, the predictions are more accurate, although again SAREfit and Muñoz algorithms achieved the best fit. The fits of the other methods are varied: Szpak and Liang are precise in the first wheel but not good in the second, Köning method has a small deviation, and the rest of the methods are affected to a greater or lesser extent.

Table 3
Quantitative comparison of the methods using real data according to the RMSO error. 1000 test points on the ellipse shape of the object were generated and used to compute the errors. Those methods which did not fit the ellipse are represented as a dash. Rank points of each example are shown in brackets, as well as the total count.

Method	Satellite dish	Can	Wheel 1	Wheel 2	Rank
SAREfit	**1.1180 (1)**	**1.0361 (1)**	**2.1877 (1)**	**2.0712 (1)**	**4**
Muñoz	4.4465 (4)	1.1322 (2)	2.1148 (2)	2.6731 (2)	10
Fitzgibbon	5.2327 (7)	2.1789 (3)	9.0742 (7)	5.2885 (3)	20
Taubin	–– (10.5)	2.4739 (7)	9.3524 (9)	5.9177 (7)	33.5
Halir&Flusser	5.2327 (8)	2.1789 (4)	9.0742 (6)	5.2885 (4)	22
Rosin	5.2089 (6)	2.4510 (6)	9.1103 (8)	5.7541 (5)	25
Szpak	1.8212 (3)	3.0478 (10)	3.3663 (4)	8.9530 (11)	28
Prasad	–– (10.5)	2.7081 (9)	9.4619 (10)	6.5871 (8)	37.5
Köning	5.1459 (5)	2.6124 (8)	5.4467 (5)	6.9363 (10)	28
Liang	1.3139 (2)	5.0095 (11)	2.3093 (3)	5.9070 (6)	22
Sobhani	5.5837 (9)	2.3151 (5)	11.2278 (11)	6.6144 (9)	34

In order to have a quantitative point of view of these fits, we computed the RMSO error with respect to a set of 1000 points generated on the shape of the real ellipse. For that purpose, we first used the *Ellipse Labeling Tool*[2] to select five points and we solved the general equation of an ellipse to assure its precision overlaying the ellipse on the real image. After that, we were able to generate any point of this ellipse. Detailed results and a ranking of methods are shown in Table 3. Results were ordered increasingly and their position determined the rank value. For all the images our method is the best and the error values are always very small, indicating that the fitted ellipse is very close to the real shape. The competing methods are rather unstable, as can be seen in the error values and the rank position. With 4 points, SAREfit is the winner, followed by Muñoz which is also robust achieving 10 points.

[2] https://sites.google.com/site/dilipprasad/Source-codes (accessed on June 10, 2019)



(a) 0% occlusion (b) 0% occlusion (c) 0% occlusion

(d) 25% occlusion (e) 25% occlusion (f) 25% occlusion

(g) 50% occlusion (h) 50% occlusion (i) 50% occlusion

(j) 75% occlusion (k) 75% occlusion (l) 75% occlusion

Fig. 9. Performance profiles of occlusion experiments (the closer to the upper left corner, the better). Natural, Algebraic and RMSO errors are analyzed with 0%, 25%, 50% and 75% of occlusion. X axis is shown in a logarithmic scale and Y axis represents the probability cumulative distribution.

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

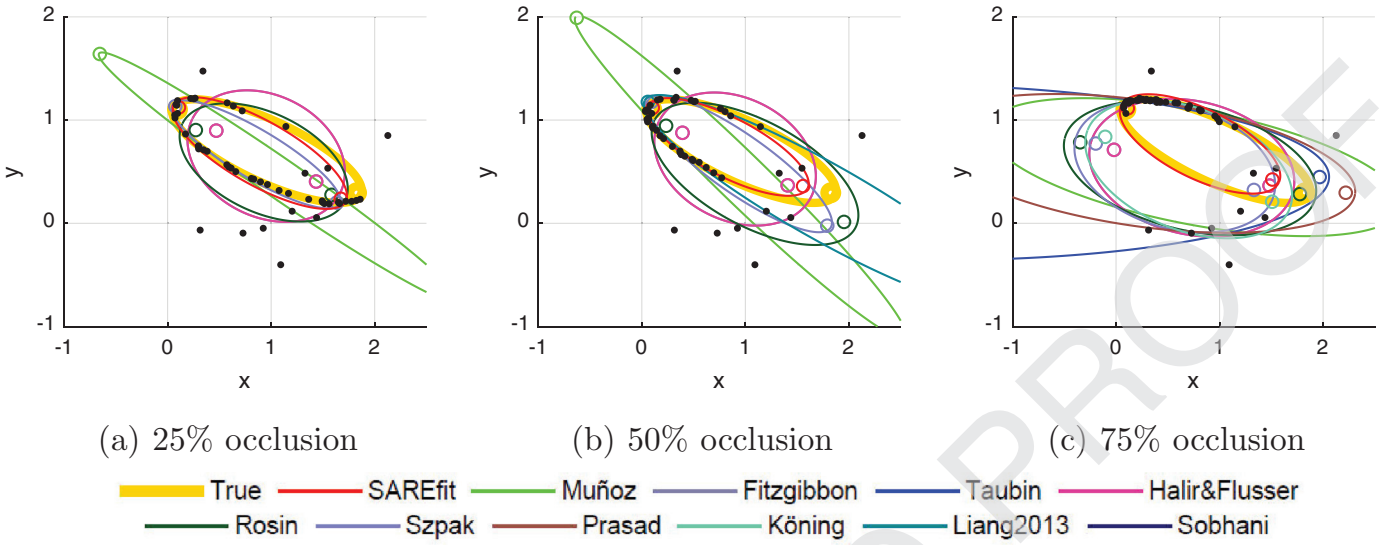


Fig. 10. Example of the outcomes with different levels of occlusion. The black points are the training samples and the true ellipse is plotted as a thick solid curve, while the fitted ellipses by each method are shown as a narrow solid curve.

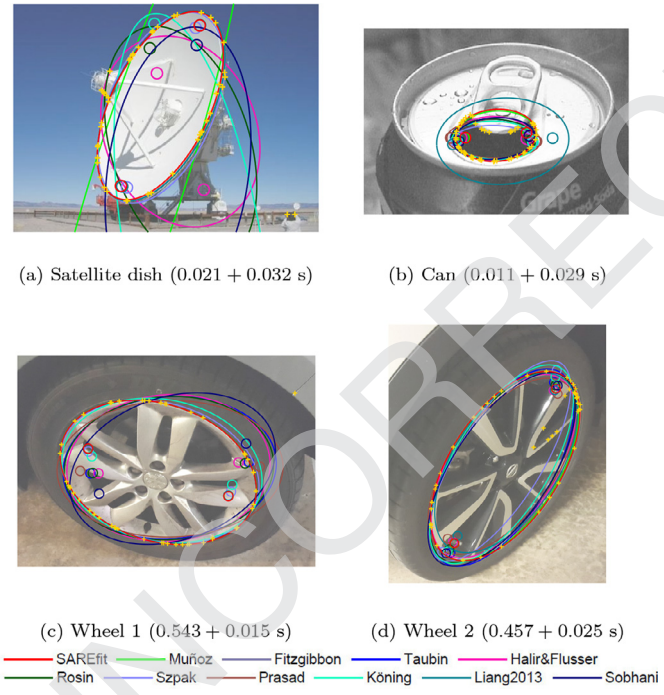


Fig. 11. Graphical comparison of the performance of the tested methods using real data. The yellow crosses are the training samples, while the fitted ellipses by each method are shown as a narrow solid curve. Captions include processing time divided in two terms: the first one represents the image loading and point extraction, and the second is the SAREfit execution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Finally, in the captions of the images, we have included the computation time needed to perform the ellipse fitting. These values are separated into two terms. The first represents the CPU time necessary to load the image and execute the Canny detection algorithm. This value is highly dependent on the size of the image. After that, the second term corresponds to the SAREfit execution time, whose values are quite similar to the CPU time analysis done before.

Fig. 12. Graphical comparison of the operation of the multiple ellipse fitting methods.

4.7. Multiple ellipse fitting

Multiple ellipse fitting is one of the main challenges of this field that has a variety of applications in biology, medicine, architecture, etc. Although SAREfit is developed only for single ellipse fitting. In this subsection, we address the case of multiple fits in the same cloud point.

For that purpose, a framework rather than design an algorithm for the problem could be used. Hui Li [45] has recently proposed a four-step methodology to find all possible ellipses based on three simple techniques: single ellipse fitting, anomaly detection, and clustering. Thus, its architecture allows to include any kind of single ellipse fitting algorithm. The source code was implemented in Mathematica and it was kindly provided by the author. Our proposal was connected with his code through MATLink library.[3]

Fig. 12 presents the results of both the original Li's method and the adapted one for a specific set of data samples. Both executions detect two ellipses and the final result depends on the quality of the fitting algorithm. In this example, SAREfit generates larger ellipses than the original method, which more closely follow the shape of the point cloud. The case of the Can example (Fig. 11b) can be addressed with the same methodology. First, the point extraction step should preserve points of all the ellipse shapes present in the image and then this procedure will cluster, detect and fit the corresponding ellipses.

An advantage of our proposal is that can be easily integrated into any other kind of method for multiple fits where the ellipse fitting procedure is a separate component of the whole process.

[3] http://matlink.org/

Please cite this article as: K. Thurnhofer-Hemsi, E. López-Rubio and E.B. Blázquez-Parra et al., Ellipse fitting by spatial averaging of random ensembles, Pattern Recognition, https://doi.org/10.1016/j.patcog.2020.107406

5. Conclusions

In this section, we extract some conclusions about the results of this research. An ensemble method that is based on a natural parametrization has been proposed for ellipse fitting. It selects the best fits among a set of fits obtained by random subsampling of the available training set. The spatial median is used to combine the best fits into a single ellipse fit which is provided as the output of the procedure.

Stability across different initializations is found after reaching 90 ellipse fits, of which only the best 10% are selected. This fact also assures performance improvement. According to the outcomes of the tests, it is confirmed that the best results are achieved for small values of subsampling factors, 15% of the input points are enough, reinforcing the idea of good performance under the presence of outliers. This is stated after performing experiments with different levels of outliers and occlusion and comparing the results for the Natural, Algebraic and RMSO errors, as well as individual Geometric errors. Although datasets with small curvature were dismissed to avoid degenerated ellipse as outcomes, the tests performed with an almost degenerated ellipse and 5% of outliers show that most competing methods failed in the fit, being SAREfit the only one to achieve an ellipse with similar characteristics to the true one. Unlike the competing methods, SAREfit solves all the fits with less error in terms of the Natural, Algebraic and RMSO errors. There is some variability depending on the metrics considered, but the combination of the spatial median with the post-processing refinement provides more precise natural and algebraic parameters. Regarding geometric parameters, most of the input sample sets are better fitted by SAREfit method than its competitors, especially focusing on the center, area and angle.

As a general conclusion, it can be said that SAREfit method outperforms the state-of-the-art of ellipse fitting methods, especially when outliers are present, measured in terms of three ellipse parametrizations and the closeness to the true ellipse through the orthogonal distances. Real data tests performed on several natural images also confirm such a statement. The accuracy of the estimation depends on the number of outliers that are present in the input set of points. This is determined by the ability of the procedure which provides the input set to our algorithm to overcome challenging effects such as illumination conditions and camera limitations. These data acquisition issues are outside the scope of our work, which deals with ellipse fitting only. A weakness of our procedure is that it relies on a base method so that gross estimation errors from the base method can affect the performance of our approach.

Our work is devoted to fit a single ellipse to a set of input points. In order to integrate our approach into a multiple ellipse detection system, a clustering subsystem should be added so as to provide our method with an input set of points that are expected to belong to a single ellipse. After our method produces an estimation of the ellipse parameters, the quality of the fit must be evaluated by a subsequent ellipse assessment subsystem, in order to ascertain whether the input set of points should be modified. This way, the segmentation of the overall set of points into ellipses would be refined over time. The advantage of integrating our approach into a multiple ellipse detection system is that its robustness allows obtaining a good fit of an ellipse even if the previous clustering subsystem wrongly introduces some points belonging to other ellipses into the input set of points for our algorithm.

Further works include the improvement of the ensemble technique. At this stage all base ellipse fits are equally likely to get into the ensemble and they are given the same importance. Therefore, a possible enhancement of the proposal would be the selection of promising base fits for their inclusion in the ensemble, and the usage of weights to give more importance to those base fits which are more likely to be accurate. Additional enhancements might be attained by, using a combination of the error criteria in order to have a comprehensive evaluation of the base fits and the ensemble fits. Hopefully this would allow solving more difficult scenarios with high occlusion and higher levels of noise. The integration of SAREfit into an ellipse detection system would make it suitable to be applied to several areas of engineering, medicine or astronomy where multiple ellipses exist in the same dataset so that clustering of the sample points must be done at the same time that the individual ellipses are fitted.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is partially supported by the Ministry of Economy and Competitiveness of Spain [grant numbers TIN2016-75097-P and PPIT.UMA.B1.2017]. It is also partially supported by the Ministry of Science, Innovation and Universities of Spain [grant number RTI2018-094645-B-I00], project name Automated detection with low-cost hardware of unusual activities in video sequences. It is also partially supported by the Autonomous Government of Andalusia (Spain) under project UMA18-FEDERJA-084, project name Detection of anomalous behavior agents by deep learning in low-cost video surveillance intelligent systems. All of them include funds from the European Regional Development Fund (ERDF). The authors thankfully acknowledge the computer resources, technical expertise and assistance provided by the SCBI (Supercomputing and Bioinformatics) center of the University of Málaga. They also gratefully acknowledge the support of NVIDIA Corporation with the donation of two Titan X GPUs used for this research. The authors acknowledge the funding from the Universidad de Málaga. Karl Thurnhofer-Hemsi is funded by a Ph.D. scholarship from the Spanish Ministry of Education, Culture and Sport under the FPU program [grant number FPU15/06512].

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.patcog.2020.107406.

References

[1] J. Muñoz-Pérez, O.D. de Cózar-Macías, E.B. Blázquez-Parra, I. Ladrón de Guevara-López, Multicriteria robust fitting of elliptical primitives, J. Math Imaging Vis. 49 (2014) 492–509.
[2] A. Nurunnabi, Y. Sadahiro, D.F. Laefer, Robust statistical approaches for circle fitting in laser scanning three-dimensional point cloud data, Pattern Recognit. 81 (2018) 417–431.
[3] A. Gontar, H. Tronnolone, B.J. Binder, M.J. Bottema, Characterising shape patterns using features derived from best-fitting ellipsoids, Pattern Recognit. 83 (2018) 365–374.
[4] E. López-Rubio, K. Thurnhofer-Hemsi, E.B. Blázquez-Parra, O.D. de Cózar-Macías, M.C. Ladrón-de Guevara-Muñoz, A fast robust geometric fitting method for parabolic curves, Pattern Recognit. 84 (2018) 301–316.
[5] A.F. Borges, Analysis of wave velocity anisotropy of rocks using ellipse fitting, Int. J. Rock Mech. Min. Sci. 96 (2017) 23–33.
[6] D. Mitchell, J.V. den Berg, Development of an ellipse fitting method with which to analyse selected area electron diffraction patterns, Ultramicroscopy 160 (2016) 140–145.
[7] M. Liao, Y. qian Zhao, X. hua Li, P. shan Dai, X. wen Xu, J. kai Zhang, B. ji Zou, Automatic segmentation for cell images based on bottleneck detection and ellipse fitting, Neurocomputing 173 (2016) 615–622.
[8] W. Gander, G. Golub, R. Strebel, Least-squares fitting of circles and ellipses, BIT 34 (4) (1994) 558–578.
[9] A. Fitzgibbon, M. Pilu, R. Fisher, Direct least squares fitting of ellipses, IEEE Trans. Pattern Anal. Mach.Intell. 21 (5) (1999) 476–480.

[10] S.J. Ahn, W. Rauh, H.-J. Warnecke, Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola, Pattern Recognit. 34 (12) (2001) 2283–2303.
[11] J. Yu, S.R. Kulkarni, H.V. Poor, Robust ellipse and spheroid fitting, Pattern Recognit Lett 33 (5) (2012) 492–499.
[12] K. Kanatani, P. Rangarajan, Hyper least squares fitting of circles and ellipses, Comput. Stat. 'I&' Data Anal. 55 (6) (2011) 2197–2208.
[13] J. Liang, P. Li, D. Zhou, H. So, D. Liu, C.-S. Leung, L. Sui, Robust ellipse fitting via alternating direction method of multipliers, Signal Process. 164 (2019) 30–40.
[14] 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. Pattern Anal. Mach. Intell. 13 (11) (1991) 1115–1138.
[15] Z. Szpak, W. Chojnacki, A. van den Hengel, Guaranteed ellipse fitting with a confidence region and an uncertainty measure for centre, axes, and orientation, J. Math. Imaging Vis. 52 (2) (2015) 173–199.
[16] R. Halir, J. Flusser, Numerically stable direct least squares fitting of ellipses, in: The Sixth International Conference in Central Europe on Computer Graphics and Visualization, 1998, pp. 125–132.
[17] P.L. Rosin, Further five-point fit ellipse fitting, Graph. Models Image Process. 61 (5) (1999) 245–259.
[18] D.K. Prasad, M.K. Leung, C. Quek, Ellifit: an unconstrained, non-iterative, least squares based geometric ellipse fitting method, Pattern Recognit. 46 (5) (2013) 1449–1465.
[19] R. Köning, G. Wimmer, V. Witkovský, Ellipse fitting by nonlinear constraints to demodulate quadrature homodyne interferometer signals and to determine the statistical uncertainty of the interferometric phase, Meas. Sci. Technol. 25 (115001) (2014) 0–11.
[20] J. Liang, M. Zhang, D. Liu, X. Zeng, O. Ojowu, K. Zhao, Z. Li, H. Liu, Robust ellipse fitting based on sparse combination of data points, IEEE Trans. Image Process. 22 (6) (2013) 2207–2218.
[21] E. Sobhani, M. Sadeghi, M. Babaie-Zadeh, C. Jutten, A robust ellipse fitting algorithm based on sparsity of outliers, in: 2017 25th European Signal Processing Conference (EUSIPCO), 2017, pp. 1195–1199.
[22] M. Amasyali, Improved space forest: a meta ensemble method, IEEE Trans. Cybern. 49 (3) (2019) 816–826.
[23] H. Pham, S. Olafsson, Bagged ensembles with tunable parameters, Comput. Intell. 35 (1) (2019) 184–203.
[24] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[25] T. Ho, The random subspace method for constructing decision forests, IEEE Trans. Pattern Anal. Mach. Intell. 20 (8) (1998) 832–844.
[26] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Mach. Learn. 63 (1) (2006) 3–42.
[27] R. Durrant, A. Kabán, Random projections as regularizers: learning a linear discriminant from fewer observations than dimensions, Mach. Learn. 99 (2) (2015) 257–286.
[28] J.J. Rodríguez, L.I. Kuncheva, Rotation forest: a new classifier ensemble method, IEEE Trans. Pattern Anal. Mach. Intell. 28 (10) (2006) 1619–1630.
[29] D.T. Bui, P. Tsangaratos, T.N. Phuong-Thao, T.D. Pham, B.T. Pham, Flash flood susceptibility modeling using an optimized fuzzy rule based feature selection technique and tree based ensemble methods, Sci. Total Environ. 668 (2019) 1038–1054.
[30] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140.
[31] Y. Freund, R.E. Schapire, Experiments with a new boosting algoritm, in: Machine Learning: Proceedings of the Thirteenth international Conference, volume 1, 1996, pp. 148–156.
[32] R.E. Schapire, The Boosting Approach to Machine Learning: An Overview, in: Nonlinear estimation and classification, Springer, 2003, pp. 149–171.
[33] J. Friedman, T. Hastie, R. Tibshirani, Additive logistic regression: a statistical view of boosting, Ann. Stat. 28 (2) (2000) 337–407.
[34] L. Song, H. Mi, Y. Lu, Q. Liu, Bagging-based system combination for domain adaption, in: Proceedings of the 13th Machine Translation Summit, 2011, pp. 293–299.
[35] D.T. Bui, T.-C. Ho, B. Pradhan, B.-T. Pham, V.-H. Nhu, I. Revhaug, Gis-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with adaboost, bagging, and multiboost ensemble frameworks, Environ. Earth Sci. 75 (14) (2016).
[36] P. Kadavi, C.-W. Lee, S. Lee, Application of ensemble-based machine learning models to landslide susceptibility mapping, Remote Sens. 10 (8) (2018).
[37] M. Lopes, Estimating the algorithmic variance of randomized ensembles via the bootstrap, Ann. Stat. 47 (2) (2019) 1088–1112.
[38] P.L. Rosin, Ellipse fitting by accumulating five-point fits, Pattern Recognit. Lett. 14 (8) (1993) 661–669.
[39] R.A. Maronna, R.D. Martin, Robust statistics: Theory and methods (with r), 2nd, John Wiley & Sons, 2019.
[40] R. Arratia, L. Gordon, Tutorial on large deviations for the binomial distribution, Bull. Math. Biol. 51 (1) (1989) 125–131.
[41] E.D. Dolan, J.J. Moré, Benchmarking optimization software with performance profiles, Math. Program. 91 (2) (2002) 201–213.
[42] Z. Zhang, Parameter estimation techniques: a tutorial with application to conic fitting, Image Vis. Comput. 15 (1) (1997) 59–76.
[43] O. Kurt, O. Arslan, A general accuracy measure for quality of elliptic sections fitting, Measurement 145 (2019) 640–647.
[44] G. Griffin, A. Holub, P. Perona, Caltech-256 object category dataset, CalTech Rep. (2007).
[45] H. Li, Multiple ellipse fitting of densely connected contours, Inf. Sci. 502 (2019) 330–345.

Karl Thurnhofer-Hemsi (born 1990) received his B.Sc. in Computer Engineering and his M.Sc. in Mathematics degrees from the University of Málaga, in 2014. He is a Ph.D. candidate at the Department of Computer Languages and Computer Science, University of Málaga. His interests are machine learning, pattern recognition, and image processing.

Ezequiel López-Rubio (born 1976) received his MSc and PhD (honors) degrees in Computer Engineering from the University of Málaga, in 1999 and 2002, respectively. He joined the University of Málaga in 2000, where he is currently a Professor of Artificial Intelligence. His interests are machine learning and pattern recognition.

Elidia Beatriz Blázquez-Parra (born 1974) is associate professor with the Department of Graphic Engineering, Design and Projects at the University of Málaga. She received her Ph.D. degree in Geodesy and Cartographic Engineering in 2003. Her current research activities include digital image analysis and visual pattern recognition.

M. Carmen Ladrón-de-Guevara-Muñoz (born 1986) received her Ph.D. (w/honors) degree from the University of Córdoba in 2018. She joined the department of Graphic Engineering as assistant professor at the University of Malaga in 2017. Her research interests currently include digital image analysis, pattern recognition, industrial heritage and sustainable technologies.

Óscar David de-Cózar-Macías is associate profesor at University of Málaga. He is Secretary of the Industrial Engineering School from 2001. He received his PhD degree in December 2010. His research interests are on digital image analysis, pattern recognition, cultural and industrial heritage and product design.