

# Fitting Parabolas in Noisy Images<sup>☆</sup>

Luis A. García-Escudero<sup>a</sup>, Agustín Mayo-Iscar<sup>a</sup>, Clara I. Sánchez-Gutiérrez<sup>b</sup>

<sup>a</sup>*Departamento de Estadística e Investigación Operativa, Universidad de Valladolid, Facultad de Ciencias, Paseo de Belén s/n, 47011 Valladolid, Spain.*

<sup>b</sup>*Diagnostic Image Analysis Group. Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA Nijmegen, The Netherlands*

---

## Abstract

A novel approach to fitting parabolas to scattered data is introduced by putting special emphasis on the robustness of the approach. The robust fit is achieved by not taking into account a proportion  $\alpha$  of the “most outlying” observations, allowing the procedure to trim them off. The most outlying observations are self-determined by the data. Procrustes analysis techniques and a particular type of “concentration” steps are the keystone of the proposed methodology. An application to a retinographic study is also presented.

*Keywords:* Parabola fitting, Robustness, Procrustes analysis, Retinography.

---

## 1. Introduction

The accurate fitting of conic sections to a given data set is an ongoing prerequisite in many applications, including image analysis. For instance, the automatic discovery of underlying conic sections in biomedical images is frequently needed. Moreover, the presence of noisy observations in these biomedical images is more often the rule than the exception, due, for instance,

---

<sup>☆</sup>Research partially supported by the Spanish Ministerio de Economía y Competitividad and FEDER, grant MTM2014-56235-C2-1-P, and by the Consejería de Educación de la Junta de Castilla y León, grant VA212U13.

*Email addresses:* [lagarcia@eio.uva.es](mailto:lagarcia@eio.uva.es) (Luis A. García-Escudero),  
[agustinm@eio.uva.es](mailto:agustinm@eio.uva.es) (Agustín Mayo-Iscar),  
[clara.sanchezgutierrez@radboudumc.nl](mailto:clara.sanchezgutierrez@radboudumc.nl) (Clara I. Sánchez-Gutiérrez)

to the technical limitations of the tools applied to obtain these images, or to the existence of unexpected “artifacts”. With this in mind, a method is presented here to robustly fit parabolas in scattered and noisy data sets. The key idea is to avoid trying to fit all the points in the data set, but only a proportion  $1 - \alpha$  of them. These retained points are supposed to include the “most informative” ones for fitting the parabola, and the existence of outliers among them is not expected. Outlying data frequently introduces serious problems for other (non-robust) methods.

By adopting a trimming approach, the whole data set does not have to be approximated by a single parabola, so other “secondary” or “less interesting” data point arrangements departing from that parabolic model are allowed in the data. If these outlying and secondary data point arrangements are no longer considered when determining the parabola fitting, then the robustness obviously arises in a very natural way. Moreover, these secondary data point arrangements may be further analyzed after they are properly detected and isolated by the method.

The idea of trimming data by taking into account only the dataset itself is not new in statistics. For instance, high-breakdown point methods such as the Least Trimmed Squares (LTS), the Least Median of Squares (LMS), the Minimum Volume Ellipsoid (MVE), and the Minimum Covariance Determinant (MCD) are frequently applied (see, e.g., Rousseeuw and Leroy 1987, Rousseeuw 1997 and Hubert et al. 2008).

The proposed trimming methodology, in this particular setting, may be stated as follows: Given a data set  $\{\mathbf{z}_i\}_{i=1}^n \subset \mathbb{R}^2$  and  $\mathcal{P}$  being the set of all the possible parabolas in  $\mathbb{R}^2$  defined as

$$\mathcal{P} = \{(x, y)' \in \mathbb{R}^2 : ax^2 + bxy + cy^2 + dx + ey + f = 0 \text{ with } b^2 - 4ac = 0\}, \quad (1)$$

we look for a subset of indices  $I \subset \{1, 2, \dots, n\}$  with  $\#I = [n(1 - \alpha)]$  and a parabola  $p \in \mathcal{P}$  such that

$$\sum_{i \in I} \|\mathbf{z}_i - \mathbf{z}_i^p\|^2 \quad (2)$$

is as small as possible, where  $\mathbf{z}^p$  stands for the closest point in the parabola  $p$  to the point  $\mathbf{z}$  (i.e.,  $\mathbf{z}^p = \arg \min_{\mathbf{x} \in p} \|\mathbf{z} - \mathbf{x}\|$ ).

This problem statement implies solving a double minimization problem on the set of all possible parabolas  $\mathcal{P}$  and on the set of all possible subsets of indexes  $I$  of size  $[n(1 - \alpha)]$ . Of course, this is not an easy task because we need to deal with the combinatorial space of subsets of a given size. In Section

2, we propose a feasible algorithm aimed at finding approximate solutions to this problem.

Instead of minimizing (2), we could have considered obtaining a “direct” robust fit of the coefficients defining the parabola in (1) through an LTS or LMS regression. This approach has previously been applied in Wang et al. (1997) to the case of fitting an ellipse in an optical-fiber quality control application. In fact, they propose an LMS fit of the coefficients  $a, b, c, d, e$  and  $f$  appearing in (1) by setting the constraint  $f = 1$ . Unfortunately, the “geometry” in this regression approach is not easy and typical assumptions for the regression model are far from being satisfied. Moreover, by setting  $f = 1$ , it becomes impossible to fit conics through the origin ( $f = 0$ ) and reasonable properties of invariance under translation, rotation and scale changes do not hold. Without trimming, several approaches have been introduced in the literature for fitting conics to scattered data based on minimizing

$$\sum_{i=1}^n Q(x_i, y_i)^2,$$

with  $Q(x_i, y_i) = ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + f$  and subject to different constraints on the coefficients  $a, b, c, d, e$  and  $f$ . In a seminal paper, Bookstein (1979) proposed a quadratic constraint on those coefficients that guarantees the invariance under shifts, rotations and scale changes.

In subsequent works, Sampson (1982), Taubin (1991) and Joseph (1994), among others, proposed some modifications to Bookstein’s approach, aiming to minimize (at least approximately) the sum of the so-called “geometric distances” given by

$$\sum_{i=1}^n \|\mathbf{z}_i - \mathbf{z}_i^p\|^2$$

(i.e., the *non-trimmed* version of our target function). Gnanadesikan (1977) had already included a justification for the appropriateness of the minimization of a sum of squared errors measured in the direction normal to the fitted conic. As we are using squared geometric distances in our target function, our approach is ideally aimed at dealing with independent normally distributed orthogonal errors with the same variance for the non-contaminating observations.

As mentioned above, robustness is a major goal for the methodology introduced here. Netanyahu et al. (1997) also paid special attention to robustness aspects, where the aim was to fit straight lines and circular arcs (but

not parabolas) to noisy images. We adopt here a different approach where *Procrustes analysis* (see, e.g., Seber 1984, Dryden and Mardia 1998) plays a key role. Notice that each parabola  $p \in \mathcal{P}$  can be obtained through location and scale changes and a rotation starting from the so-called “canonical” parabola. Taking that idea into account, the application of some “concentration” steps similarly to those applied in the fast-MCD algorithm (Rousseeuw and van Driessen 1999) or in the fast-LTS algorithm (Rousseeuw and Van Driessen 2006) yields the algorithm presented in this work.

The proposed algorithm is detailed and exemplified in Section 2. An application of the proposed methodology to a retinography study is given in Section 3. Finally, Section 4 makes concluding remarks and outlines a more general extension of the proposed approach.

## 2. Algorithm

A parabola in  $\mathbb{R}^2$  is defined as the geometrical place of the points satisfying (1). Let  $\mathcal{P}$  denote the set of all possible parabolas in  $\mathbb{R}^2$ . A very important parabola is the “canonical” parabola  $p_0$  defined as

$$p_0 = \{(x, y)' \in \mathbb{R}^2 : y = x^2\}. \quad (3)$$

It can be shown that each parabola  $p \in \mathcal{P}$  may be obtained as a composition of a translation, a scale change and a rotation starting from the canonical parabola  $p_0$  and this transformation is unique; i.e., there exists a set of parameters  $\theta = \{\mathbf{m}, \lambda, \omega\}$  such that  $p$  coincides with  $F_\theta(p_0)$  ( $p \equiv F_\theta(p_0)$ ) with

$$\begin{aligned} F_\theta : \mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix} \mapsto F_\theta(\mathbf{z}) &= \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \lambda \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \mathbf{m} + \lambda \mathbf{R} \mathbf{z}, \end{aligned}$$

where  $\mathbf{m} = (m_1, m_2)'$  is the translation vector parameter,  $\lambda$  is the scale change parameter and  $\mathbf{R}$  is the rotation matrix for the rotation angle  $\omega$ . Thus, wlog, we can denote each parabola  $p \in \mathcal{P}$  by  $p_\theta$  whenever  $p \equiv F_\theta(p_0)$ .

By using this notation, we present here an algorithm aimed at minimizing (2) on  $I \subset \{1, 2, \dots, n\}$  and on  $p \in \mathcal{P}$  for a fixed trimming size  $\alpha$ :

1. *Initialization of the iterative process:* Randomly select an initial set of parameters  $\theta^{(0)} = \{\mathbf{m}^{(0)}, \lambda^{(0)}, \omega^{(0)}\}$ . For instance, draw at random 4 observations in general position from the whole data set and use them to

obtain an initial parabola  $\hat{p}$ . Note that any 4 points in general position in  $\mathbb{R}^2$  determine 2 (conjugate) parabolas with complex coefficients. So we obtain 0, 1 or 2 parabolas with real coefficients. When a parabola with real coefficients is not found (one of the sampled points falls inside the triangle formed by the other three), we draw another 4-points random sample. When two parabolas with real coefficients are found, we can consider both of them as possible initializations or we can choose one at random. For  $\hat{p}$  being any parabola with real coefficients, we take  $\theta^{(0)}$  such that  $F_{\theta^{(0)}}(p_0) \equiv \hat{p}$ .

2. “*Concentration step*”: Assume that  $\theta^{(l-1)} = \{\mathbf{m}^{(l-1)}, \lambda^{(l-1)}, \omega^{(l-1)}\}$  is the set of parameters in the previous iteration:

- 2.1. *Trimming*: We look for the proportion  $1 - \alpha$  of observations closest to the associated parabola for this set  $\theta^{(l-1)}$  of parameters. To find these points, it is preferable to transform backwards the observations, i.e., to express them in terms of the canonical parabola  $p_0$ . Thus, we take  $\mathbf{z}_i^c = F_{\theta^{(l-1)}}^{-1}(\mathbf{z}_i)$  and search for  $\mathbf{z}_i^{c,p}$  being the closest point on the canonical parabola  $p_0$  to  $\mathbf{z}_i^c$ . The closest point to  $\mathbf{z}_i^c = (x_i, y_i)'$  has the form  $\mathbf{z}_i^{c,p} = (t, t^2)'$  where  $t$  is a real zero of the cubic polynomial  $2t^3 + (1 - 2y_i)t - x_i = 0$ . If this polynomial has more than one real zero, we consider the one with minimal value for  $\|\mathbf{z}_i^{c,p} - \mathbf{z}_i^c\|$ . Finally, take  $\mathbf{z}_i^p = F_{\theta^{(l-1)}}(\mathbf{z}_i^{c,p})$  and compute the distance of  $\mathbf{z}_i^p$  to  $\mathbf{z}_i$  through  $d_i = \|\mathbf{z}_i - \mathbf{z}_i^p\|$ . Sort the distances  $d_{(1)} \leq d_{(2)} \leq \dots \leq d_{(n)}$  and take

$$I^{(l)} = \{i \in \{1, 2, \dots, n\} \text{ such that } d_i \leq d_{[n(1-\alpha)]}\}.$$

Observations having indices outside  $I^{(l)}$  are those to be tentatively trimmed off at this stage of the iterative process.

- 2.2. *Update parameters*: In order to update  $\theta$ , we search for the best possible composition of a translation, a scale change and a rotation  $F_\theta(\mathbf{z}) = \mathbf{m} + \lambda \mathbf{R}\mathbf{z}$ , taking points  $\{F_\theta(\mathbf{z}_i^{c,p})\}_{i \in I^{(l)}}$  as close as possible to  $\{\mathbf{z}_i\}_{i \in I^{(l)}}$ . This task is done by applying Procrustes analysis:

*Procrustes analysis*: If  $I^{(l)} = \{i_1, \dots, i_{[n(1-\alpha)]}\}$  are the retained indices after the trimming stage, let  $\mathbf{Z} = (\mathbf{z}_{i_1}, \dots, \mathbf{z}_{i_{[n(1-\alpha)]}})'$  and  $\mathbf{Z}^{c,p} = (\mathbf{z}_{i_1}^{c,p}, \dots, \mathbf{z}_{i_{[n(1-\alpha)]}}^{c,p})'$  be two  $[n(1-\alpha)] \times 2$  matrices containing the points  $\{\mathbf{z}_i\}_{i \in I^{(l)}}$  and  $\{\mathbf{z}_i^{c,p}\}_{i \in I^{(l)}}$ , respectively, in their rows. The best possible  $\mathbf{m}$ ,  $\mathbf{R}$  and

$\lambda$  for “matching” those  $\mathbf{Z}$  and  $\mathbf{Z}^{c,p}$  data matrices are:

$$\begin{aligned}\widehat{\mathbf{m}} &= \text{ave}(\mathbf{Z}') - \text{ave}((\mathbf{Z}^{c,p})'), \\ \widehat{\mathbf{R}} &= (\widetilde{\mathbf{Z}^{c,p}}' \widetilde{\mathbf{Z}} \widetilde{\mathbf{Z}}' \widetilde{\mathbf{Z}^{c,p}})^{-1/2} (\widetilde{\mathbf{Z}^{c,p}}' \widetilde{\mathbf{Z}}), \text{ and,} \\ \widehat{\lambda} &= \text{trace}[(\widetilde{\mathbf{Z}^{c,p}}' \widetilde{\mathbf{Z}} \widetilde{\mathbf{Z}}' \widetilde{\mathbf{Z}^{c,p}})^{1/2}] / \text{trace}[\widetilde{\mathbf{Z}^{c,p}}' \widetilde{\mathbf{Z}^{c,p}}].\end{aligned}$$

The notation  $\text{ave}(\mathbf{A})$  stands for the column vector with the row-wise means of matrix  $\mathbf{A}$  and  $\widetilde{\mathbf{A}}$  for the mean-centered  $\mathbf{A}$  matrix. The optimal angle for the rotation is obtained as  $\widehat{\omega} = \arctan(\widehat{r}_{21}/\widehat{r}_{11})$  for

$$\widehat{\mathbf{R}} = \begin{pmatrix} \widehat{r}_{11} & \widehat{r}_{12} \\ \widehat{r}_{21} & \widehat{r}_{22} \end{pmatrix}.$$

If  $\det(\widehat{\mathbf{R}}) = -1$ , we should take  $\widehat{\mathbf{R}} = -\widehat{\mathbf{R}}$ .

This yields to an updated  $\theta^{(l)} = \{\widehat{\mathbf{m}}, \widehat{\lambda}, \widehat{\omega}\}$  set of parameters. More information about why the previous expressions are applied can be obtained from Section 5.6 in Seber (1984).

3. *Evaluate the objective function:* Perform  $L$  iterations of the process described in Step 2 and compute the associated value of the objective function

$$\mathcal{L}(\theta^{(L)}; \{\mathbf{z}_i\}_{i=1}^n) := \sum_{i \in I^{(L)}} \|\mathbf{z}_i - \mathbf{z}_i^p\|^2,$$

where  $\mathbf{z}_i^p$  are the projected points on the parabola defined by the final  $\theta^{(L)}$  set of parameters.  $L$  may be seen as a maximal number of iterations. However, this iterative process may be stopped if  $\|\text{vec}(\theta^{(l)}) - \text{vec}(\theta^{(l-1)})\| < \varepsilon$  ( $\text{vec}(\cdot)$  puts all the parameters into a single vector) and  $\varepsilon$  is a small enough constant to declare updating convergence.

4. *Initialize several times:* Draw  $S$  different random starting values (i.e., start from Step 1), repeat Steps 2-3, and, keep those  $\theta^{(L)}$  and  $I^{(L)}$  which have led to the smallest values for  $\mathcal{L}(\theta^{(L)}; \{\mathbf{z}_i\}_{i=1}^n)$ . These best possible solutions are fully iterated until complete convergence so the best one (i.e., the smallest value of the target function) can be chosen.

It may be shown that the iterative process described in Step 2 serves to monotonically decrease the value of the objective function. For each initialization, we end up in a local minimum of the target function (or close to

one of them). Therefore, the consideration of several random starts tries to search for the global minimum (or a close approximation to it).

Table 1 provides a summary of the notation used in the presentation of the algorithm.

Table 1: Summary of the notation in the “concentration” steps of the algorithm.

	Original space	Transforming toward canonical
Data points:	$\mathbf{z}_i$	$\mathbf{z}_i^c$
Projected pts. on parabola:	$\mathbf{z}_i^p$	$\mathbf{z}_i^{c,p}$

Figure 1 shows the evolution of the fitted parabolas through the iterative process in Step 2 for two different random initializations in a simulated data set. The data set was generated around the parabola  $y = 1 + x + x^2$  followed by a shift of  $\mathbf{m} = (-2, 2)'$  and a rotation of  $\omega = \pi/4$ . After generating 400 random points in the resulting parabola, we added to them some normally distributed error terms  $N(0, 0.15^2)$  in a direction normal to the parabola. Finally, 40 additional observations uniformly distributed in a rectangle, containing the first 400 (non-contaminated) data points, were considered. We can see how the number of updates  $L$  needed to get a solution close to the underlying parabola increases when the initialization is relatively not good (Figure 1(b) vs. Figure 1(a)). Moreover, we can see that there exist completely useless starting values as long as they eventually end up trapped in local minima. Therefore, the consideration of several random initializations is needed most of the times. Note that our chances of reaching the global minimum increase with higher  $L$  and  $S$  values, but the computational cost also increases. With this in mind, we recommend choosing a large value for  $S$ , but not too high a value for  $L$ . As commented in Step 4, we propose fully iterating only a few of the most “promising”  $\theta^{(L)}$  sets of parameters (i.e., those with the smallest values of the target function out of these  $S$  random initializations with just  $L$  updating steps).

The proposed methodology also needs to fix in advance the trimming level  $\alpha$ . The example in Figure 1 was run with a trimming size  $\alpha = 0.1$ , which allows the 40 added outlying observations to be completely removed. It can be seen that the estimation of the parabola does not necessarily change notably when choosing a slightly higher than needed trimming level whenever enough observations belonging to the “true” parabola are still detected. However, it is very important that the trimming level should allow us to remove outly-

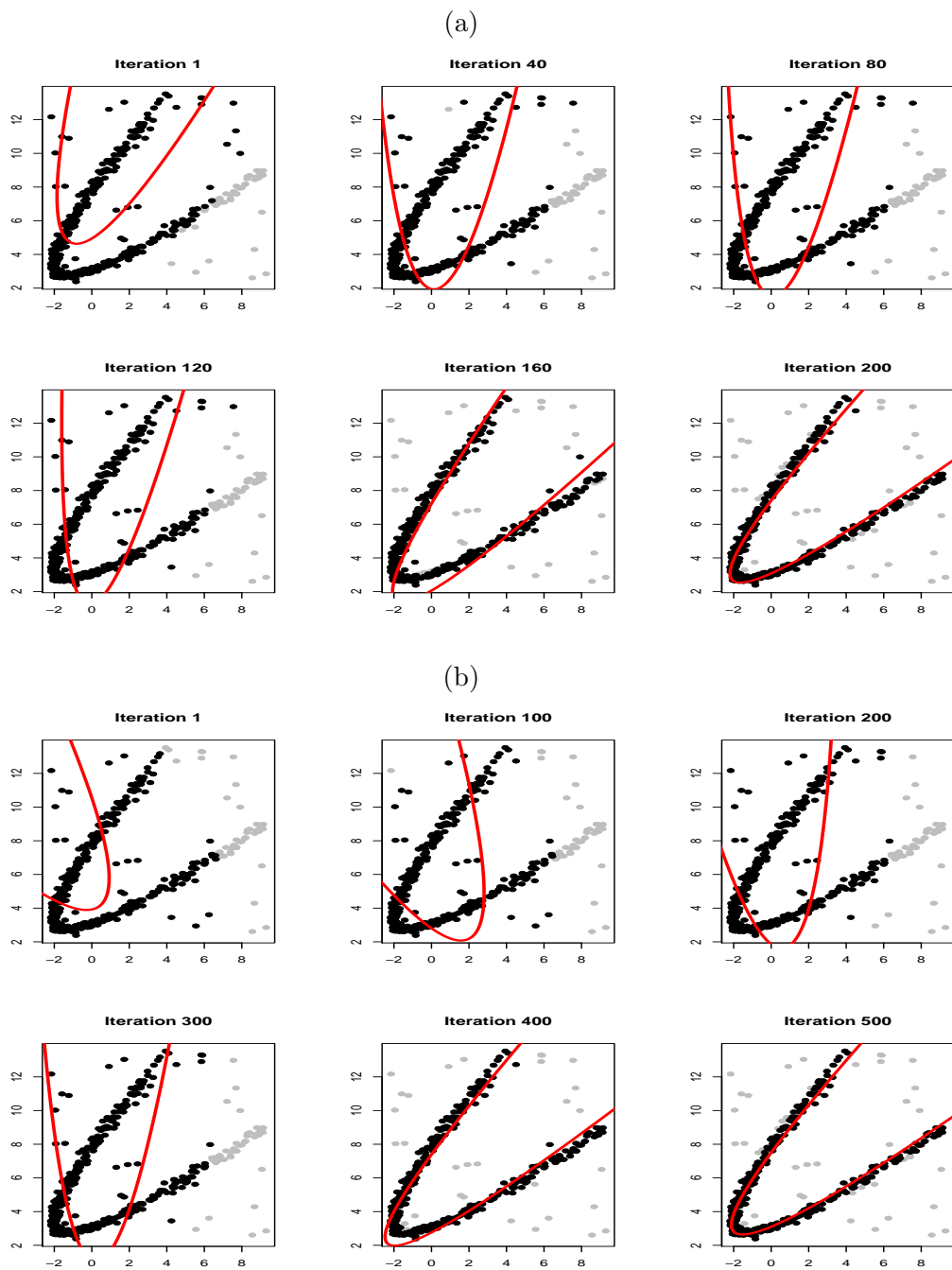


Figure 1: Evolution of the proposed algorithm with  $\alpha = 0.1$  for a simulated data set with two different initial “random” parabolas. Trimmed points are the gray ones.



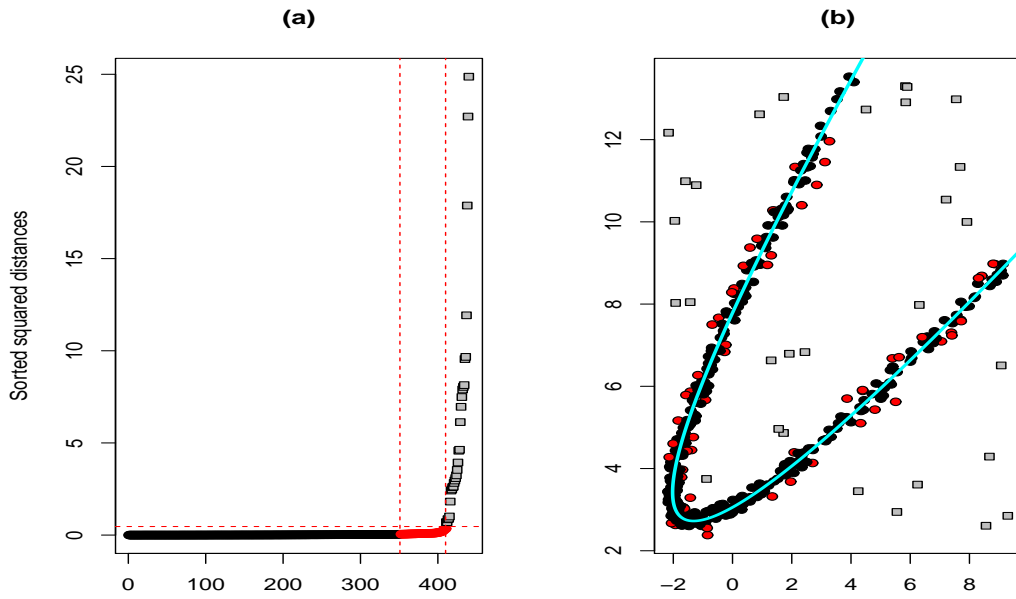
ing observations completely. Thus, we recommend choosing initially a rather high  $\alpha$  in a preventive way. Choosing  $\alpha$  slightly higher than its required value would result in the removal of some non-outlying observations; the latter could be easily restored by examining their associated geometric distances  $\|\mathbf{z}_i - \mathbf{z}_i^p\|$ . If these geometric distances are “similar” to other distances associated with non-trimmed data points, then these (wrongly trimmed) observations should be restored. For instance, the contamination level in the simulated data set shown in Figure 1 was  $40/440 \simeq 10\%$ . Figure 2 shows the results of applying the proposed methodology with  $\alpha = 0.2$  and  $\alpha = 0.35$  (two trimming levels higher than the trimming level  $\alpha = 0.1$  actually required). The left panels in Figure 2 show the squared geometric distances  $\|\mathbf{z}_i - \mathbf{z}_i^p\|^2$  sorted from lowest to highest. Therefore, the observations in the first positions of these panels correspond to the non-trimmed ones. These non-trimmed observations are represented by using solid-black points in the right panels. In Figure 2, we have restored those observations with squared geometric distances smaller than 10 times the largest geometric distance for the non-trimmed ones. These restored observations appear as circled-red points in the right panels of Figure 2. The squared-gray points are the observations that are finally declared as outliers. Note that the finally trimmed data points are not excessively dependent on the initial (larger than needed) trimming level.

Due to the application of the Procrustes analysis methods and the use of geometric distances, we directly achieve interesting properties of invariance under translations, scale changes and rotations.

### 3. Application to retinography

The vascular arch in retinal images plays an essential role in the screening of diabetic retinopathy (DR), an important ocular disease. In the retina, the major blood vessels arch above and below the area on the disc temporal side, known as the posterior pole. The curved shape described by the main vascular arcades can be approximated by a parabolic curve enclosing the posterior pole. When DR signs appear in the posterior pole, the risk of visual loss significantly increases. Therefore, the automatic detection of the retinal vascular arch could be used as a coordinate system on the retina to identify lesions that appears in the posterior pole, grading the severity of the disease (see, e.g., Li and Chutatape 2004). See also Sánchez et al. (2008,

Trimming level  $\alpha = 0.2$



Trimming level  $\alpha = 0.35$

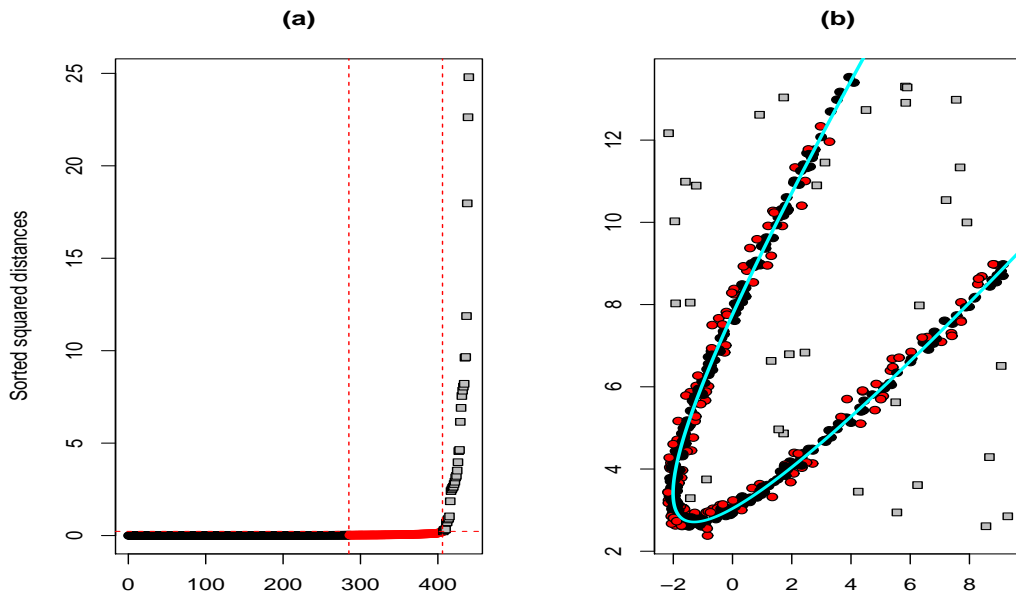


Figure 2: Results of applying  $\alpha = 0.2$  and  $\alpha = 0.35$  to the same data set as in Figure 1. The left panels show the sorted squared geometric distances. In the right panels, the solid-black points are the “initially” trimmed observations, the circled-red points are the “restored” ones and the squared-gray points are the “finally” trimmed ones.

2009) with more information about the type of retinal images analyzed in this example.

In order to achieve this automatic detection, pixels belonging to major blood vessels need to be extracted. First, the intensity variation in the background across the image is removed. This is accomplished by subtracting an estimate of these background variations by using mean filtering of the image green channel. We use the green component because it enhances the contrast of the vasculature. In order to highlight the major blood vessels, a two-dimensional Morlet wavelet transform (2D CWT) approach is applied to the normalized image (Antoine et al. 1993). This directional wavelet is convolved with the image varying its rotation angle from  $0^\circ$  to  $170^\circ$  at steps of  $10^\circ$ . The final filtered image is obtained by taking the maximum modulus of the 2D CWT over all angles. The wavelet parameters are chosen in order to provide stronger responses for pixels belonging to wider blood vessels. Finally, the filtered image needs to be thresholded to extract the pixels of interest from the major blood vessels. We apply a morphological double threshold operator to the image (Soille 1999). This operator thresholds the image for two ranges of gray values, one being included in the other. In contrast to the simple threshold approach, the double threshold operator obtains images less contaminated by artifacts and keeps the connectivity of the segmented features (Soille 1999). This segmentation is finally thinned to obtain its skeleton, which contains the pixels belonging to the major vessels as well as some other pixels belonging to adjacent vascular branches.

After this pre-processing stage, we apply the robust parabola fitting approach presented in this work with a high trimming level, say  $\alpha = 0.35$ , which certainly allows most of the outlying points and the secondary vascular configurations to be removed. The fitted parabola and the trimmed observations (shown as gray points) are depicted in Figure 3.

Notice how the proposed methodology provides a sensible approximation of the main vascular arcades by a parabola, in spite of the presence of other “secondary” blood vessels and artifacts. Similar results have been obtained with a large number of retinal images (see the “Supplementary Material” file for a few more examples of the types of fit obtained by applying the proposed methodology).

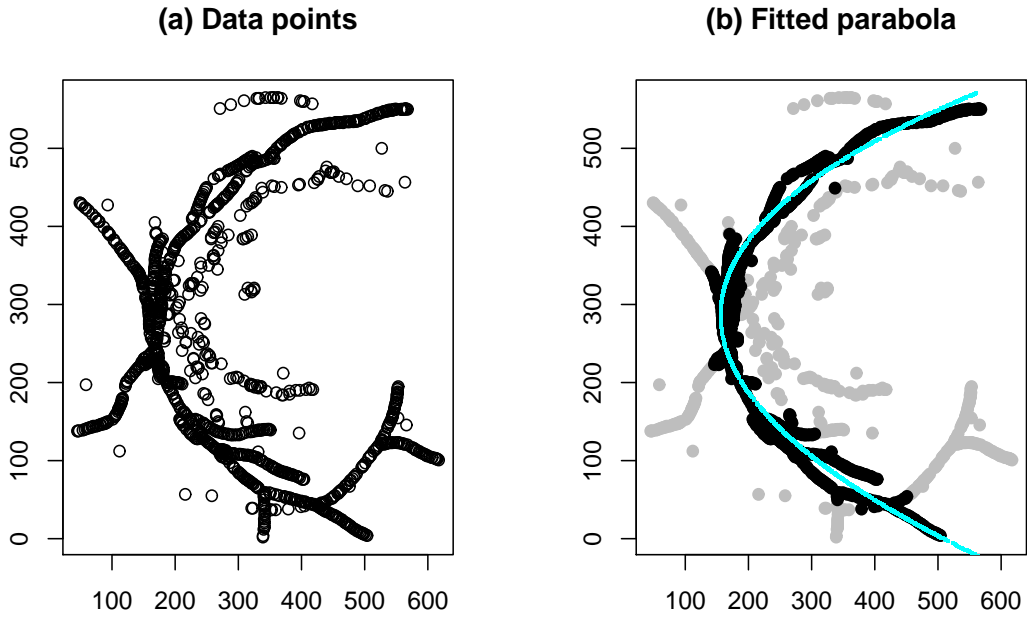


Figure 3: (a) A retinographic image and (b) the fitted parabola using the proposed methodology with  $\alpha = 0.35$ . Trimmed points are shown in gray.

#### 4. Discussion

A new methodology has been presented to fit parabolas to scattered data which is applicable in noisy (biomedical) images. The robustness of the proposed approach follows from its ability to trim a proportion  $\alpha$  of data. The dataset itself indicates the observations that should be trimmed. Procrustes analysis plays a key role in this approach.

The proposed methodology is being adapted (as part of ongoing work) with the aim of fitting other types of conic sections or objects, whenever these objects may be obtained through location and scale changes combined with rotations from a fixed “canonical object”. For that extension, we need a procedure to compute geometric distances to the given canonical object. If the canonical object can be summarized by a finite set of points in  $\mathbb{R}^2$ , then those distances can be approximated by searching for the closest point in this finite set. Trimming and Procrustes analysis are later applied in a similar fashion to that done in this work. The problem discussed here can be seen as a slightly more complex one, since the canonical parabola is an unbounded object.

## References

- [1] Antoine, J. P., Carette, P., Murenzi, R. and Piette, B. (1993) "Image analysis with two dimensional continuous wavelet transform," *Signal Processing*, **31**, 241-272.
- [2] Bookstein, F. (1979), "Fitting conic sections to scattered data," *Computer Vision, Graphics, and Image Processing*, **9**, 56-71.
- [3] Dryden, I.L. and Mardia, K.V. (1998). *Statistical Shape Analysis*. John Wiley, Chichester.
- [4] Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. Wiley, New York.
- [5] Hubert, M., Rousseeuw, P.J. and Van Aelst, S. (2008), "High-breakdown robust multivariate methods", *Statistical Science*, **23**, 92-119.
- [6] Joseph, S.H. (1994) "Unbiased least squares fitting of circular arcs", *CVGIP: Graphical Model and Image Processing*, **56**, 424-432.
- [7] Li, H. and Chutatape, O. (2004). "Automated feature extraction in color retinal images by a model based approach," *IEEE Transactions On Biomedical Engineering*, **51**, 246-254.
- [8] Netanyahu, N.S., Philomin, V., Rosenfeld, A. and Stromberg, A. (1997). "Robust detection of straight and circular road segments in noisy aerial images," *Pattern Recognition*, **30**, 1673-1686.
- [9] Rousseeuw, P.J. (1997), "Introduction to positive-breakdown methods," in *Handbook of Statistics, Vol. 15: Robust Inference*, edited by G.S. Maddala and C.R. Rao. North-Holland, 101-121.
- [10] Rousseeuw, P.J. and Leroy, A.M. (1987), *Robust Regression and Outlier Detection*, Wiley Interscience, New York (Series in Applied Probability and Statistics).
- [11] Rousseeuw, P.J. and Van Driessen, K. (1999), "A fast algorithm for the Minimum Covariance Determinant estimator," *Technometrics* , **41**, 212-223.

- [12] Rousseeuw, P.J. and Van Driessen, K. (2006), “Computing LTS regression for large data sets,” *Data Mining and Knowledge Discovery*, **12**, 2945.
- [13] Sánchez, C.I., Hornero, R., López, M.I., Aboy, M, Poza, J., and Abásolo, D (2008). “A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis,” *Medical Engineering & Physics*, **30**, 350-357.
- [14] Sánchez, C.I., García, M., Mayo, A., López, M.I. and Hornero, R. (2009). “Retinal image analysis based on mixture models to detect hard exudates,” *Medical Image Analysis*, **13**, 650-658.
- [15] Sampson, P.D. (1982), “Fitting conic sections to ‘very scattered’ data: An iterative refinement of the Bookstein algorithm,” *Computers Graphics Image Processing*, **18**, 97-108.
- [16] Seber, G.A.F. (1984), *Multivariate Observations*. John Wiley, New York.
- [17] Soille, P. (1999), *Morphological Image Analysis: Principles and Applications*. Springer-Verlag. New York.
- [18] Taubin, G. (1991). “Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 1115-1138.
- [19] Wang, C.M., Vecchia, D.F, Young, M. and Brilliant, A. (1997), “Robust regression applied to optical-fiber dimensional quality control”, *Technometrics*, **39**, 25-34.