Further Five Point Fit Ellipse Fitting

Paul L. Rosin
Department of Computer Science and Information Systems,
Brunel University,
Uxbridge,
Middlesex UB8 3PH
UK
email: Paul.Rosin@brunel.ac.uk

Abstract

The least squares method is the most commonly used technique for fitting an ellipse through a set of points. However, it has a low breakdown point, which means that it performs poorly in the presence of outliers. We describe various alternative methods for ellipse fitting which are more robust: the Theil-Sen, least median of squares, Hilbert curve, and minimum volume estimator approaches. Testing with synthetic data demonstrates that the least median of squares is the most suitable method in terms of accuracy and robustness.

1 Introduction

Ellipses commonly occur in man-made scenes, often being formed as the projection of circular objects onto the image plane. They provide a useful representation of parts of the image since 1/ they are more convenient to manipulate than the corresponding sequences of straight lines needed to represent the curve, and 2/ their detection is reasonably simple and reliable. Thus they are often used by computer vision systems for model matching [3, 5]. Over the years much attention has been paid to fitting ellipses to data samples, and many variations of the standard method for finding the least squares (LS) solution exist [4, 6]. However, computer vision often requires more robust methods that can tolerate large amounts of outliers since there is the likelihood that the data will be substantially corrupted by faulty feature extraction, segmentation errors, etc. While LS is optimal under Gaussian noise it is very sensitive to severe non-Gaussian outliers, and is therefore unsuitable for many vision applications.

Previously we described a robust method for fitting ellipses to curve data [11]. It uses the method of minimal subsets to accumulate ellipse hypotheses. A minimal subset is the smallest number of points necessary to uniquely define a geometric primitive (five for an ellipse). Many five-tuples of points are selected from the full data set. Those that define non-elliptic conics are rejected, and the remainder are considered as hypotheses of the ellipse that best fits all the data. Each of the intrinsic parameters of the ellipse (i.e. centre co-ordinates, major and minor axes, and orientation) is estimated as the median of the parameters of the hypothesised ellipses. In other words, if P_s^q is the qth parameter estimated from minimal subset s, then the final estimates are $\hat{P}^q = \text{med}_s P_s^q$ for $q = 1 \ldots 5$. This is in fact an application of the Theil-Sen estimator [19], which as already been used for fitting lines [8] and circular arcs [10]. Other applications of the minimal subset method to estimating geometric features are given in references [1, 15].

Unfortunately there are a number of inadequacies with the method just described, involving 1/ the treatment of circular parameters (i.e. orientation), 2/ statistical efficiency, and 3/ correlation between the five parameters. This paper presents several solutions to these problems and describes some variations on the theme of robust ellipse fitting.

2 Improved Ellipse Fitting

2.1 Approximate circular median

The Theil-Sen method described above estimates each parameter by the median of the set of hypothesised parameters produced from the minimum subsets. While this is adequate if the ellipses are described by their conic coefficients, there is a problem when the ellipses' intrinsic parameters are used.[1] Since the ellipse orientation is directional data then the median will produce incorrect estimates if the data is clustered around the ends of the range $[0, 2\pi]$.

The natural solution would be to use the circular median in place of the linear median [9, 18]. This is defined as the point P such that half of the data points lie on either side of the diameter PQ of the unit circle (on which the data lies), and that the majority of points are closer to P than Q. However, unless the values are binned (which is undesirable since the results will then depend on an arbitrary bin size) all pairs of points need to be considered, which is computationally expensive. Instead, as a compromise between efficiency and robustness, we use a combination of the circular mean and the circular median [9]. We rotate the data in order to centre its circular mean at the orientation midrange, π, a linear median is then performed, after which the rotation is undone. Since the median can be found in linear time, calculating the approximate circular median is also $O(n)$.

2.2 Repeated median

Although the Theil-Sen estimator is more robust than the LS method, for five dimensions its breakdown point[2] is only 12.9% since the fraction of outliers ϵ must satisfy $(1-\epsilon)^5 \geq 0.5$, so that $\epsilon = 1 - \left(\frac{1}{2}\right)^{\frac{1}{5}}$. Although the results published in Rosin [11] appear to show better robustness this is possibly due to many of the conics passing through the five-tuples being non-ellipses (i.e. hyperbolae or parabolae) and are therefore rejected.

[1] Experiments showed that it was preferable to use the intrinsic parameters of the ellipse rather than their corresponding conic coefficients [11]. Their values tend to lie in a smaller range than the coefficients, and are invariant to translation and rotation of the data. Therefore all the methods described in this paper use intrinsic parameters rather than conic coefficients.

[2] An estimator's breakdown point is defined as the percentage of outliers that may force the estimator to return a value arbitrarily larger than the correct result.

More robust estimators than the Theil-Sen are available, For instance, the repeated median (RM) [17] has a breakdown point of 50%, which is achieved by calculating

$$\hat{P}^q = \operatorname*{med}_i \operatorname*{med}_j \operatorname*{med}_k \operatorname*{med}_l \operatorname*{med}_m P^q_{i,j,k,l,m}$$

where $P^q_{i,j,k,l,m}$ is the qth parameter estimated from the five points $\{\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k, \mathbf{p}_l, \mathbf{p}_m\}$. For n points this recursive application of the median results in n^5 tuples being considered rather than the $\binom{n}{5}$ tuples used by the Theil-Sen estimator, and is therefore two orders of magnitude more costly in computation time. In any case, since both n^5 and $\binom{n}{5}$ are $O(n^5)$, in practise evaluating all the tuples is too costly, and fitting must be speeded up by only selecting a subset of the possible tuples. More details on the selection of these tuples as will be described later, however, they are not easily applicable to the repeated median estimator.

2.3 Least median of squares

Another robust estimator with a breakdown point of 50% is the Least Median of Squares (LMedS) estimator [16], which was applied to ellipse fitting by Roth and Levine [14]. For each of the minimal subsets s they calculated the errors ϵ_i from each point $\mathbf{p}_i$ to the ellipse defined by $\mathbf{p}_{i \in s}$. The LMedS ellipse has the parameters defined by the subset with the least median error, namely

$$\min_s \operatorname*{med}_i \epsilon_i.$$

Unlike the Theil-Sen and RM estimators this application of the LMedS uses the residuals of the full data set of each hypothesised ellipse. The ideal error measure would be the Euclidean distance between the point and the ellipse, but that is expensive to calculate, requiring solving a quartic equation and choosing the smallest of the four solutions. In practice there are many simpler approximations that are often used [12, 13]. The principle one (and both the simplest and most efficient) is the algebraic distance.

Although the LMedS method is significantly more robust than the Least Squares (LS) method it is less efficient as a statistical estimator (i.e. its asymptotic relative efficiency) as well as being generally less efficient computationally. This can be improved by "polishing" the fit returned by the best minimal subset. First the noise is robustly estimated using the median absolute deviation (mad) which finds the median of the errors and returns the median of the differences of the errors from that median

$$\text{mad} = \operatorname*{med}_{i=1}^{n} |\epsilon_i - \operatorname*{med}_{j=1}^{n} \epsilon_j|.$$

The median is modified by f which is a Gaussian normalisation and finite-sample correction factor [16]

$$f = 1.4826 \left(1 + \frac{5}{n-1}\right)$$

and can then enable outliers to be robustly detected by applying the following indicator function

$$w_i = \begin{cases} 1 & |\epsilon_i| < 3 \times f \,\text{mad} \\ 0 & \text{otherwise.} \end{cases}$$

The ellipse parameters are now reestimated by performing a standard LS fit, weighting the data points $\mathbf{p}_i$ by w_i to eliminate the effect of (robustly detected) outliers.

In addition to its robustness, a significant advantage of the LMedS approach just described is that all the parameters are estimated simultaneously. A weakness of the Theil-Sen and RM approaches is that each parameter is estimated independently, thereby ignoring any potential correlations between them.

We also note that the LMedS could be applied directly to the parameters rather than the residuals in a manner akin to the Theil-Sen estimator

$$\hat{P}^q = \min_x \operatorname*{med}_s \left(P_s^q - x\right)^2 .$$

2.4 Hilbert curve

A different approach to estimating the five ellipse parameters simultaneously is to map the 5D parameter space onto 1D. This allows a simple 1D estimation technique to be employed as before after which the result is mapped [2] back to the full 5D parameter space. We use the Hilbert curve which has the property that points close in 2D usually map into points close in 1D along the curve. Since the ellipse orientation is circular it is treated separately using the circular median. Therefore, in practice we analyse the remaining parameters using a 4D Hilbert curve quantised so that each axis is divided into 256 bins. The four parameters are all measured in the same units (pixels) obviating scaling problems. Since the Hilbert curve only covers a subset of $\mathcal{R}^4$ we must specify the desired range of each axis. In our experiments we have set all of them to be $[0, 1000]$.

2.5 MVE

Another way to estimate the parameters simultaneously rather than performing independent 1D medians is to use a multivariate generalisation of the median. Since this is computationally expensive we use instead another robust technique: the minimum volume ellipsoid estimator (MVE) [16]. The MVE is defined as the ellipsoid with the smallest volume that contains half of the data points. It is affine invariant and so it can be easily applied to multivariate data without need for normalising the dimensions, although that is not an issue here. The centre of the MVE is used as an estimate of the ellipse parameters. As usual, the ellipse orientation is estimated separately by an approximate circular median.

To speed up calculation of the MVE it is approximated by a random resampling algorithm similar to the minimal subset method we are using for ellipse fitting. From the best estimate, robust distances are calculated for each point relative to the MVE, and are used to compute a reweighted mean [16].

2.6 Sampling of tuples

All the estimators described in this paper use the minimal subsets generated from the different combinations of data points. For the line fitting application in [8] it was acceptable to use all the possible pairs of points as the complexity was only $O(n^2)$. Since ellipse fitting requires five-tuples taking all the combinations would result in $O(n^5)$ complexity. Unless only very small data sets are used this in not practically useful. However, reasonable results can still be obtained for must less computational cost if only some of the minimal subsets are computed. The simplest approach is just to randomly sample the data set with replacement to generate the number of minimal subsets which is estimated as necessary to achieve the desired likelihood of successful performance. For instance, for the LMedS just one minimal subset containing no outliers is sufficient for its success. This implies that if the proportion of outliers is ϵ, then the number of subsets N that must be chosen to include with $100 \times C\%$ confidence one composed entirely of inliers is [14, 16]

$$N = \frac{\ln(1-C)}{\ln(C-(1-\epsilon)^5)}.$$

Unfortunately, this analysis only considers type II noise – the outliers. But type I noise – consisting of low level noise often modelled by additive Gaussian noise – can also have considerable effect. In our earlier work on ellipse fitting [11] it was noted that even for near perfect test cases made up from uncorrupted synthetic data many of the minimal subsets generated ellipses providing poor estimates of the full data set. Just the effects of quantising the data to the pixel grid was sufficient to substantially distort these local ellipses, particularly if the points in the minimal subset were close together. Closely spaced points have a high "leverage", and therefore small variations in their position have a large effect on the ellipse through the minimal subset.

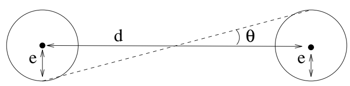


Figure 1: Orientation error of line

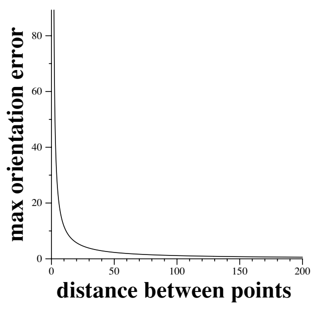


Figure 2: Orientation error as a function of distance between points

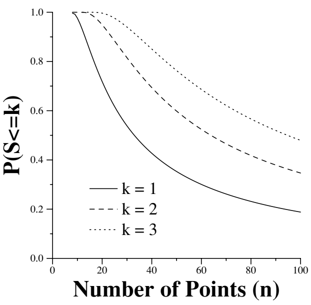
The instability of the ellipse fit when points are closely spaced can more easily be demonstrated by the example of the straight line passing through two points. If the error of the points is bounded by e and the distance between the points is d then from figure 1 the maximum orientation error can be seen to be approximately $\theta = \tan \frac{2e}{d}$. Plotting the error as a function of the distance between the two points (figure 2) shows that the error increases exponentially as the distance decreases.

Based on work carried out to analyse the distribution of spaces on lottery tickets we can calculate the probability of smallest spacing S of r sequentially ordered items selected from n as [7]

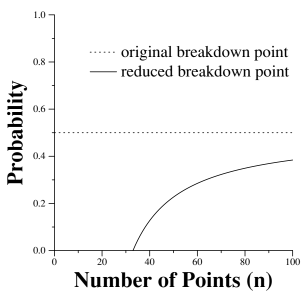
$$P(S \leq k) = 1 - \binom{n-(r-1)(k-1)}{r} / \binom{n}{r}.$$

Applying this to minimal subsets for ellipses figure 3a shows the probability of choosing five points. Naturally, as we increase the number of potential points that the subset is being drawn from the probability of two points in the subset being adjacent decreases. Nevertheless, it may appear surprising that even with 33 points there is still a 50% chance of a minimum subset of five points containing two directly adjacent members. Thus, random sampling will produce a number of subsets with closely spaced elements which could reduce their effectiveness, and therefore reduce the robustness of the estimator.

Let us take as an extreme example the situation in which minimal subsets containing adjacent points always produce erroneous ellipses. Then for a set of 33 points half the sampled tuples are likely to contain adjacent points so that even an estimator with a 50% breakdown point will be expected to have zero resistance to outliers. In general, an estimator with a breakdown point b will have its effective breakdown point reduced such that the fraction of outliers ϵ must satisfy $(1-\epsilon) \times P(S > 1) \geq b$. This is illustrated in figure 3b for an estimator with an original breakdown point of 0.5. Although in practise the effects of closely spaced points will not be as extreme as modelled here, they will have a definite lesser effect, depending on the spatial distance between the points as well as the amount of mislocalisation of the points.



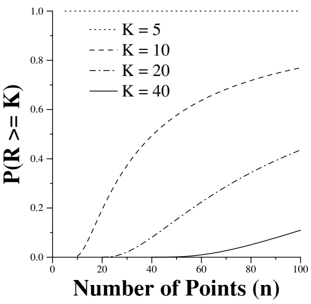
(a)



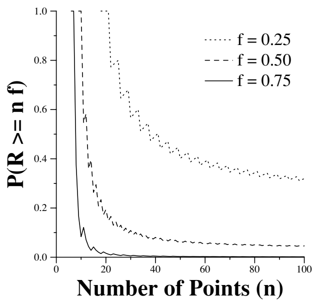
(b)

Figure 3: (a) Probability of choosing five points from n and getting a minimum spacing of one; (b) Reduced breakdown point of estimator

Not only are closely spaced points a problem, but points spaced too far apart can also lead to difficulties. In computer vision the data set for fitting the ellipse to is likely to be structured. For instance, the points may be sampled from a curve obtained by detecting edges in the image. This means that the outliers will not be randomly distributed among the data set. Instead, if the curve is poorly segmented then one end of the curve may be derived from the elliptical feature of interest, while the other end arises from a separate outlying object such as



(a)



(b)

Figure 4: (a) Probability of choosing five points from n and getting a overall range $R \geq K$; (b) Probability of range being larger than a fraction f of the number of points

background clutter. When outliers are clustered like this it is better to prevent the range of the minimal subsets being too large. Otherwise, at least one point will lie on the outlying section of the curve.

If the points $\mathbf{p}_j$ in a minimal subset are randomly selected then we can analyse the joint distribution of the spacings X_j between them [7]

$$P(X_j \geq k_j;\ j = 1 \dots r-1) = \binom{n - \sum_{j=1}^{r-1}(k_j - 1)}{r} \Big/ \binom{n}{r}.$$

Thus, the distribution of that total range that the subset covers, i.e. $R = 1 + \sum_{j=1}^{r-1} k_j$, is

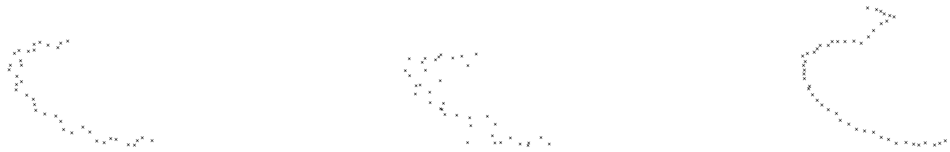
$$P(R \geq K) = \binom{n + r - K}{r} \Big/ \binom{n}{r},$$

and is plotted for various values of K in figure 4a. The likelihood of a minimal subset covering fixed a range increases as the number of available points increases. Of more interest is when the range of the minimal subset is taken as a fraction f of the full data set n. Replacing K by $\lceil nf \rceil$ gives the graph in figure 4b. The jagged outline is due to the non-linear $\lceil \bullet \rceil$ operation, necessary since there is only a discrete number of points. We can see from the graph that for small numbers of points there is a moderate probability that a large fraction will be covered by a randomly selected minimal subset. For example 19% of five-tuples taken from 20 points will extend to cover at least half the range of the data.

3 Experimental Results

The different ellipse fitting methods were tested on synthetic data to assess their performance. Four sets of test data were used, all based on 38 points sampled from an elliptic arc (axis lengths 333 and 250, subtended angle 200°) corrupted by Gaussian noise in the direction of the normals:

1. a sine wave was superimposed on the arc, generating some confusing small scale structure,



(a) (b) (c)

Figure 5: Examples of test data; (a) Ellipse with superimposed sine wave and Gaussian noise ($\sigma = 5$) (b) Ellipse with superimposed sine wave and Gaussian noise ($\sigma = 20$) (c) Ellipse with added clutter (20%) and Gaussian noise ($\sigma = 10$)

2. some points were additionally corrupted by extremely large amounts of Gaussian noise, producing outliers,
3. clutter was introduced (similar to structured outliers) by adding points along two noisy line segments, and
4. different lengths of the noisy arc are sampled (fewer points are taken from shorter lengths).

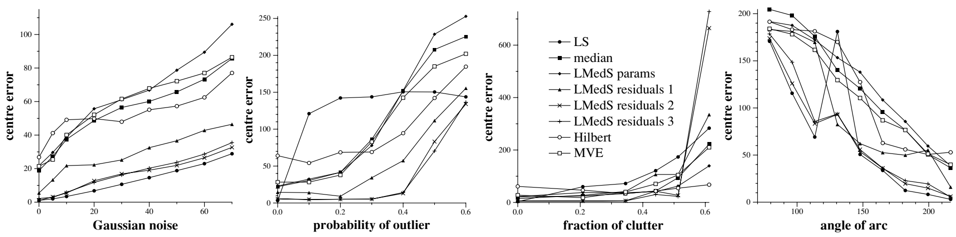
For each of the test types 500 samples were generated at each of a series of increasing levels of noise, probability of outliers, and amount of clutter, making up a total of 20,000 test cases. Some examples of the test data are shown in figure 5. An instance of the elliptic arc with the superimposed sine wave is given in figure 5a and figure 5b for two levels of noise, while figure 5c shows clutter added to the arc.

All the ellipse fitting methods were applied to the data, and the plots of the alpha-trimmed means of the error in the estimated ellipse centres are displayed in figure 6. The mean is trimmed since some incorrect ellipse fits produce extremely deviant parameter estimates that have a large influence on the mean. Some of the methods occasionally failed to fit an ellipse (fitting a hyperbola or parabola instead). In these cases the fit was ignored during the calculation of the mean. The Theil-Sen method is labelled as *median* and *LMedS params*, while the LMedS method applied to the residuals is labelled as *LMedS residuals 1, 2*, and *3* corresponding to the algebraic, weighted algebraic, and foci bisector distance approximations respectively. In figure 6a, despite the presence of the superimposed sine wave, the LS method outperforms all the other methods. However, this is explained by the fact that the noise is symmetric and uniformly distributed over all the data. We can also see that the LMedS method applied to the residuals gives better results than the remaining techniques.

Figure 6b shows how adding outliers ($\sigma = 10$ and 500 for types I and II noise) causes the LS method to break down. The LMedS residual methods work best now, particularly the weighted gradient and foci bisector distance approximations which consistently produce good results so long as there is a majority of inliers. The Theil-Sen, Hilbert curve, and MVE methods produce less accurate results, and also break down earlier. However, we note that these experiments show the Theil-Sen technique breaking down later than predicted.

The addition of clutter (figure 6c) also makes the LS estimator perform poorly, while the same two LMedS residual methods perform well again, only breaking down drastically when more than 50% of the data is made up of the clutter

We also analyse the effect of changing the angle of the arc drawn from the ellipse. The ellipse superimposed with the sinusoid was used, and the results are shown in figure 6d. As expected reducing the angle of arc increases the error. The spike in the graph for the LS fit for 131° arcs was surprising, but was confirmed by testing an additional 1000 random samples with a finer sampling of the angles. Many elongated ellipses tended to be fit, which may be an effect of the sinusoid.



(a) (b) (c) (d)

Figure 6: Alpha trimmed mean error of estimated centre location (α=0.1); (a) Superimposed sine wave and Gaussian noise (b) Types I and II noise (c) Clutter and Gaussian noise (d) Reducing arclength of ellipse

4 Conclusions

Various methods for fitting ellipses to data have been tested, although due to space limitations not all the results have been shown. The experiments suggest that in the presence of outliers the LMedS approach applied to either of the weighted algebraic or foci bisector distance approximations produces the best results. Both approximations result in robust and accurate fits. If the simpler but more biased algebraic distance approximation is used then the accuracy degrades significantly, and the robustness also suffers to a lesser extent. The LS method can produce excellent fits, but is not robust and therefore cannot be considered if the data may be contaminated by clutter or other outliers. The Theil-Sen method is markedly inferior to the LMedS residual method in terms of accuracy and robustness. Modifying the median operation of the Theil-Sen method and finding the LMedS estimates of the individual parameter values actually degrades the performance rather than improving it. Finally, both the Hilbert curve and MVE approaches were reasonably robust, but provided poor accuracy.

We discussed several issues concerning the sampling of points to form the minimal subsets. In order to avoid tuples with points separated by either too small or too large a gap a regular sampling of the points appeared advantageous. However, the experiments show that in fact no advantage for the LMedS method was gained by regular sampling, and that it actually caused the performance of the Theil-Sen method to deteriorate.

It should be noted that it is difficult to truly assess the performance of the fitting techniques. We have measured deviation in the parameter estimates from the underlying parameter set used to generate the contaminated data sets. However, many of the fits which produced low scores according to this criterion still represent the data quite adequately.

References

[1] R.C. Bolles and M.A. Fischler. A RANSAC-based approach to model fitting and its applications to finding cylinders in range data. In *7th Int. Joint Conf. on Artificial Intelligence*, pages 637–643, 1981.

[2] A.R. Butz. Alternative algorithm for Hilbert's space-filling curve. *IEEE Trans. Comp.*, 20:424–426, 1971.

[3] M. Dhome, J. T. Lapreste, G. Rives, and M. Richetin. Spatial localisation of modelled objects of revolution in monocular perspective vision. In *Proc. Europ. Conf. Computer Vision*, pages 475–485, 1990.

[4] A.W. Fitzgibbon and R.B. Fisher. A buyer's guide to conic fitting. In *British Machine Vision Conf.*, pages 513–522, 1995.

[5] D.A. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell. Invariant descriptors for 3-d object recognition and pose. *IEEE Trans. PAMI*, 13(10):971–991, 1991.

[6] W. Gander, G.H. Golub, and R. Strebel. Least squares fitting of circles and ellipses. *Bit*, 34:558–578, 1994.

[7] N. Henze. The distribution of spaces on lottery tickets. *Fibonacci Quarterly*, 33:426–431, 1995.

[8] B. Kamgar-Parsi, B. Kamgar-Parsi, and N.S. Netanyahu. A nonparametric method for fitting a straight line to a noisy image. *IEEE Trans. PAMI*, 11(9):998–1001, 1989.

[9] K.V. Mardia. *Statistics of Directional Data*. Academic Press, 1972.

[10] D.M. Mount and N.S. Netanyahu. Efficient randomised algorithms for robust circular arc estimation. Technical Report CS-TR-3198, University of Maryland, 1993.

[11] P.L. Rosin. Ellipse fitting by accumulating five-point fits. *Pattern Recognition Letters*, 14:661–669, 1993.

[12] P.L. Rosin. Analysing error of fit functions for ellipses. *Pattern Recognition Letters*, 17(14), 1996.

[13] P.L. Rosin. Assessing error of fit functions for ellipses. *CVGIP: Graphical Models and Image Processing*, 58:494–502, 1996.

[14] G. Roth and M.D. Levine. Extracting geometric primitives. *CVGIP: Image Understanding*, 58(1):1–22, 1993.

[15] G. Roth and M.D. Levine. Geometric primitive extraction using a genetic algorithm. *IEEE Trans. PAMI*, 16(9):901–905, 1994.

[16] P. Rousseeuw and A. Leroy. *Robust Regression and Outlier Detection*. Wiley, 1987.

[17] A.F. Siegel. Robust regression using repeated medians. *Biometrika*, 69:242–244, 1982.

[18] A. Stein and M. Werman. Robust statistics in shape fitting. In *Proc. Conf. Computer Vision Pattern Recognition*, pages 540–546, 1992.

[19] H. Theil. A rank-invariant method of linear and polynomial regression analysis. *Nederlandse Akademie Wetenchappen Series A.*, 53:386–392, 1950.