IMPROVED ELLIPSE FITTING BY CONSIDERING THE ECCENTRICITY OF DATA POINT SETS

Pankaj Kumar Jinhai Cai, Stan Miklavcic

Phenomics and Bioinformatics Research Centre and The Australian Centre for Plant Functional Genomics
School of Information Technology and Mathematical Sciences
University of South Australia
Mawson Lakes SA 5095

ABSTRACT

Ellipse and conic fitting is a highly researched and mature topic in image processing and computer vision. Surprisingly, however, none of the methods have thus far considered eccentricity of data point sets in the fitting of an ellipse. In this paper we show that irrespective of the method used to fit ellipses, the root mean square error (RMSE) of an algorithm increases with the eccentricity of the data point set. We propose a novel way of weighting data points based on their eccentricity to improve the results of ellipse fitting. Data points with higher weights are repeated and data points with insignificant weights are dropped. We empirically demonstrate that the proposed method improves the accuracy of ellipse fitting. Almost all methods of ellipse fitting irrespective of whether they minimize algebraic error or geometric error will benefit by the proposed method of pre-processing the data points.

Index Terms— Ellipse fitting, Eccentricity, Particle filter, Resampling

1. INTRODUCTION

Identifying and fitting ellipses to image data is an important process in computer vision and image processing with a wide range of applications. For example, ellipse or conic fitting was used to calibrate catadioptric cameras in [1] and [2], to calibrate pin hole camera for the geometry of single axis rotatory motion [3], in the segmentation of cells in microscopic images [4], segmentation of grains in [5] and in the study of galaxies in astrophysics in [6].

Previous approaches to conic fitting have all focused on minimizing a distance function to obtain the best fit to point data. An algebraic distance error was minimized in [7] and a numerically stable version of the same was presented in [8]. A geometric distance error measure was directly minimized in [9, 10] and by maximum likelihood estimation in [11, 12]. In contrast, Kanatani proposed hyperaccuracy methods for ellipse fitting in [13] and [14], while Yu, et al. in [15] proposed a new distance metric based on the intrinsic properties of ellipses. The new distance function had a clear geometric interpretation and was less computationally intensive than the geometric error. However, none of the above methods considered a non-uniform weighting of the error distance measure to account for the eccentricity of data points. Data points that are more distant from the ellipse center are considered more eccentric than data points which are closer. A mathematical definition of eccentricity is presented in Section 2. One method that considers weighting in the distance error measure calculation to make the fit more robust was adopted in [15]. Although somewhat related to the method proposed here, theirs involved a specific novel distance error measure based on the intrinsic properties of the ellipse. Our method is different from theirs as we don’t propose a new error measure; instead we propose a preprocessing method of point data to which an ellipse has to be fit. Our preprocessing is analogous to the resampling algorithm of particle filters where samples (data points) having more weight are repeated and samples with insignificant weights are dropped. After resampling of data points any ellipse fitting algorithm can be applied. We empirically show that the error of fitting improves after the data points have been processed with a resampling of data points.

The paper is organized as follows: In Section 2 we show how the error of ellipse fitting increases with the eccentricity of data points. The resampling based pre-processing of data points is explained in Section 3. Improved results of ellipse fitting are shown in Section 4. Finally, in Section 5 we conclude with a short summary and ideas for future work.

2. ERROR OF ELLIPSE FITTING DUE TO THE ECCENTRICITY OF DATA POINT SET

The eccentricity of an ellipse is defined by \( \Xi = \sqrt{\frac{a^2 - b^2}{a^2}} \), where \( a \) is the semi-major axis and \( b \) is the semi-minor axis of the ellipse and \( 0 \leq \Xi < 1 \). The value \( \Xi = 0 \) corresponds to a circle and the value \( \Xi = 1 \) corresponds to a straight line. To measure the eccentricity of data point sets to which an ellipse is to be fitted, we introduce the following pointwise...
eccentricity variable:
\[ \xi = \frac{d_a}{d_b + d_a}, \]  
where \(d_a\) is the orthogonal distance of a data point to the minor axis and \(d_b\) is the orthogonal distance to the major axis. This variable, whose values range between 0 and 1, takes on larger values for data points that are more distant from the minor axis than from the major axis.

We conducted experiments to show that the root mean square error (RMSE), of a fit increases with the average eccentricity of the data point set. In these experiments we generated data points from a parametric ellipse and added zero mean Gaussian noise to the points. Different ellipse fitting algorithms were applied to fit an ellipse to these data points. In our study we considered the following ellipse-fitting algorithms for which authors have provided the codes: CGIP-1979 [16], a basic and very initial approaches to conic fitting; PAMI-1999 [7] and WSCG-1998 [8], are methods which minimize algebraic errors; PAMI-1991 [17], minimizes geometric distance error; and ECCV-2012 [12], is a MLE based approach. These also form a good representation of the wide cross-section of ellipse fitting algorithms, which researchers can use. The RMSE values for data points in angular sectors as shown in Figure 1 are computed. The left and right

![Fig. 1](image)

regions between lines of the same colour are the sectors we refer to. Data points in different sectors have different average eccentricity and different average RMSE values. The average eccentricity of data points in the narrow sectors are higher than those in the broad sectors; in the former regions the data points are further from the ellipse centre. Different ellipse fitting methods give different average RMSE values. As such it is desirable to employ a method that gives a lower RMSE. However, apart from a lower RMSE, there may be other application specific criteria, such as computational efficiency, for choosing one method over another. In general, it is desirable to employ a method which gives a lower RMSE. In our experiments of computing RMSE for data points with different average eccentricity, we clearly see a trend of increasing RMSE with the increase in the average eccentricity of the data point set (Table 1). Furthermore, this trend is a common phenomena irrespective of the method used. This phenomena is visually demonstrated in Figure 2. Figure 2 shows magnified views of the results of ellipse fitting in the regions of different mean eccentricity; errors are higher for the more eccentric data points. The observations shown in Table 1 and Figure 2 are averages of 1000 iterations of each of the five different algorithms.

![Fig. 2](image)

### Table 1. Average RMSE per data point in the different sectors.

<table>
<thead>
<tr>
<th>Method</th>
<th>0.81</th>
<th>0.83</th>
<th>0.84</th>
<th>0.85</th>
<th>0.86</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGIP-1979</td>
<td>17.15</td>
<td>17.65</td>
<td>17.98</td>
<td>18.45</td>
<td>18.87</td>
</tr>
<tr>
<td>PAMI-1991</td>
<td>17.11</td>
<td>17.61</td>
<td>17.94</td>
<td>18.51</td>
<td>18.93</td>
</tr>
<tr>
<td>PAMI-1999</td>
<td>16.42</td>
<td>16.92</td>
<td>17.23</td>
<td>17.48</td>
<td>17.68</td>
</tr>
<tr>
<td>WSCG-1998</td>
<td>16.43</td>
<td>16.92</td>
<td>17.23</td>
<td>17.48</td>
<td>17.89</td>
</tr>
<tr>
<td>ECCV-2012</td>
<td>16.82</td>
<td>17.38</td>
<td>17.70</td>
<td>17.97</td>
<td>18.29</td>
</tr>
</tbody>
</table>

3. PROPOSED RESAMPLING OF DATA POINTS

Resampling is process used in particle filters to avoid the problem of particle degeneration [18]. In this process the par-
particles having greater weight are repeated and particles with an insignificant weights are dropped, with the overall number of particles being preserved. Algorithm 1 gives the pseudo code for this method. In the present application, we adopt the same concept to have multiple replicates of those data points having high pointwise eccentricity values. A weight is assigned to each data point \( r \) using the weight function \( w^r = e^{\xi r} \), where \( \xi = d/n \). The resampling algorithm described in Algorithm 1 is applied to the data points after the weights have been normalized to sum to unity \( w^r_n = w^r / \sum_{r=1}^{R} w^r \).

**Algorithm 1**

Given \( R \) particles/data points \( X^r_n \) and their weights \( w^r_n \), compute new set of particles/data points \( X^s_n \) and equal weights \( w^s_n \).

1. Initialize the cumulative sum of weights (CW) \( c_1 = w^r_n \).
   
   for \( r = 2 : R \) do
     
     construct CW: \( c_r = c_{r-1} + w^r_n \).
   
   end for

2. Start at the base of CW: \( r = 1 \).
3. Draw a starting point \( \mu_1 = \{0, R^{-1}\} \).
   
   for \( s = 1 : R \) do
     
     Move along CW \( \mu_s = \mu_1 + R^{-1}(s-1) \).
     
     while \( \mu_s > c_r \) do
       
       \( r = r + 1; \)
     
     end while
   
   end for

4. Reassign particles: \( X^s_n = X^r_n \).
5. Reassign weight: \( w^s_n = N^{-1} \).
6. Reassign parents: \( r^s = r \).

In our application of the resampling algorithm, the number of input data points and output data points can be changed. In fact better ellipse fitting results are obtained when the number of input data points exceeds the number of input data points. We adopted the strategy of keeping all data points and replicating those points having higher weights. For replication the data points were interpolated between current parent point and the previous neighbour of the parent. The number of replicates to be obtained came from the reassignments to each parent. To compute the weights of data points a knowledge of the major and minor axis is required. In our experiments using the synthesized data set referred to in Section 2, the major and minor axes are known from the original ellipse used to generate the data points. In the experiment with a real data set for the example shown in Section 4.2, an estimate of the major and minor axes is obtained by taking means of the ellipse parameters obtained by the five different methods used for evaluation. The estimation of the eccentric weight of data points is not sensitive to small errors in the estimate of minor and major axes.

### Table 2. Average RMSE per data points in the different sectors after resampling of the data points. The RMSE values are averages of 1000 iterations of each method.

<table>
<thead>
<tr>
<th>Ellipse fitting methods</th>
<th>Average Eccentricity of data points in sectors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.81  0.83  0.84  0.85  0.86</td>
</tr>
<tr>
<td>CGIP-1979</td>
<td>8.40  8.58  8.74  8.87  8.98</td>
</tr>
<tr>
<td>PAMI-1991</td>
<td>8.20  8.37  8.53  8.66  8.76</td>
</tr>
<tr>
<td>PAMI-1999</td>
<td>8.44  8.62  8.78  8.91  9.01</td>
</tr>
<tr>
<td>WSCG-1998</td>
<td>8.44  8.62  8.78  8.91  9.01</td>
</tr>
<tr>
<td>ECCV-2012</td>
<td>8.60  8.78  8.94  9.08  9.18</td>
</tr>
</tbody>
</table>

#### 4. RESULTS

To investigate the effectiveness of the method of resampling, we conducted experiments on both synthetic and real data. Resampling was applied as a precursor to the five different ellipse fitting methods: CGIP-1979 [16], PAMI-1999 [7], WSCG-1998 [8], PAMI-1991 [17], and ECCV-2012 [12]. We show below both a quantitative and a qualitative evaluation of the proposed method. For the quantitative evaluation, the synthetic data set is used, while for the qualitative evaluation, the real data set is used.

##### 4.1. Quantitative Evaluation

For the case of the synthetic data, the improvement in the fitting of an ellipse after resampling of data points is quite evident for the same noise level. Table 2 and Figure 3 show the RMSE of the fit and visualization of the fit, respectively, for the same experiment after resampling. Compared to the results of a direct application of the five methods without resampling, there is a notable improvement for all five methods (Table 1 and Figure 2). The reason for the improvement can be intuitively understood to be due to the introduction of more data points in the regions of high eccentricity. This tends initially to increase the error values. However, the subsequent minimization procedure tries to minimize this error leading to an overall improved fit.

##### 4.2. Qualitative Evaluation

The data set used for qualitative testing the proposed algorithm comes from an application aimed at a 3D reconstruction of plant root architecture. Tracking of root tips. Plant roots are grown in a transparent gellan gum medium and imaged every five degree rotation at regular time points. From the images root tips were detected according to the method of [19] and tracked as a function of rotation angle. Seventy two angularly displaced images were used for the root tip detection and tracking. The trajectory of a single root tip traces out an ellipse in the image plane. Ellipse fitting is then applied to the set of detected root tips in the 72 different positions. Figure 4 shows the results of this fitting. The small regions (a),
Fig. 3. Results of ellipse fitting of synthetic data set after resampling. The main figure shows the entire data set plus ellipse, while subplots show magnified views of small regions (a), (b), and (c). The results are to be compared with results in Figure 2.

(b), (c) of Figure 4 are magnified for better visualization. The left hand subplots are the results of fitting before resampling, while the right hand subplots show the results after resampling. There is significant improvement in the results of ellipse fitting after resampling of data points. An arrow marker is inserted in the images to highlight the improvement in the results, especially for the PAMI-1999 method.

5. SUMMARY AND FUTURE IDEAS

In this paper we propose a novel way of resampling and repeating data points for application to ellipse fitting problems based on an eccentricity function. Significant improvement in results is found for five different, well-established algorithms. The concept can be applied to other fitting problems where there is provision for assigning non-uniform weights to data points and then resampling according to their weights. We mention that MLE-based approaches are not suitable for high eccentricity ellipse data as the matrix generated during non-linear Levenberg-Marquardt optimization becomes singular. This is also the reason why the error of the MLE-based method, ECCV-1012, is sometimes greater than the errors obtained with the other four methods. In future work we shall extend the application of the proposal to RANSAC-based ellipse fitting methods, where the challenge will be to dynamically identify outliers and not assign them weights for the resampling algorithm.

Fig. 4. Results of ellipse fitting of a real data set after resampling. The main figure shows first the origin (root tip tracking) of a real data set, plus overall results of ellipse fitting. Subplots show magnified views of small regions (a), (b), and (c). Left hand subplots are results prior to resampling. Right hand subplots are results subsequent to resampling. An equal length arrow is included in both sets to highlight the obvious improvement.
6. REFERENCES


