Pattern Recognition Letters 31 (2010) 1860–1867

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

An iterative approach for fitting multiple connected ellipse structure to silhouette

Richard Yi Da Xu *, Michael Kemp

School of Computing and Mathematics, Charles Sturt University, Australia

ARTICLE INFO

Article history:
Available online 18 January 2010

Keywords:
Multiple connected ellipse fitting
Iterative Closest Point
Contour fitting

ABSTRACT

In many image processing applications, the structures conveyed in the image contour can often be described by a set of connected ellipses. Previous fitting methods to align the connected ellipse structure with a contour, in general, lack a continuous solution space. In addition, the solution obtained often satisfies only a partial number of ellipses, leaving others with poor fits. In this paper, we address these two problems by presenting an iterative framework for fitting a 2D silhouette contour to a pre-specified connected ellipses structure with a very coarse initial guess. Under the proposed framework, we first improve the initial guess by modeling the silhouette region as a set of disconnected ellipses using mixture of Gaussian densities or the heuristic approaches. Then, an iterative method is applied in a similar fashion to the Iterative Closest Point (ICP) (Alshawa, 2007; Li and Griffiths, 2000; Besl and Mckay, 1992) algorithm. Each iteration contains two parts: first part is to assign all the contour points to the individual unconnected ellipses, which we refer to as the *segmentation* step and the second part is the non-linear least square approach that minimizes both the sum of square distances between the contour points and ellipse's edge as well as minimizing the ellipse's vertex pair (s) distances, which we refer to as the *minimization* step. We illustrate the effectiveness of our methods through experimental results on several images as well as applying the algorithm to a mini database of human upper-body images.

1. Introduction

Object recognition and shape registration is an important image processing topic and its finding has been incorporated into many real-life applications. Over the last decade, various techniques have been developed for registering 2D and 3D models to images, from simple geometrical shape models such as a vehicle (Yiu et al., 2005) to more complex models such as human limbs tracking (Bernier et al., 2009; Fossati et al., 2008) and hand model fitting (Du and Charbon, 2007).

Many works have used the joint ellipse models to represent complex shapes (Fossati et al., 2008; Xu and Kemp, 2009; Jeune et al., 2004). The advantage of using a connected ellipse representation is that the ellipse is easily parameterized and the parameters convey information including the location, orientation and variation of data. The connectivity requirements of the adjacent ellipses further reduce the number of degree of freedoms, and hence constrain their movements to reflect real-world problems. For example, the upper and lower human limbs should always be joined. However, some of the recent works suffer from two drawbacks: firstly, they lack a continuous solution space to ensure only discrete "valid" ellipse parameters are achieved and secondly most of these methods require an accurate initialization.

A recent representative example which lacks in continuous solution space is found in (Fossati et al., 2008), where the authors use Generalized Expectation Maximization (GEM) to assign edge pixels to body parts and to find the body pose that maximizes the likelihood of the resultant assignments. However, in the Maximization (M)-step of the algorithm, a "better" pose parameter is selected from the predefined data sets to increase the likelihood probability. The use of only discrete number of classes can result in misalignment between the image and the model, especially when the number of predefined models is small. In this paper, we propose an effective and generalized approach to the fitting algorithm. Our algorithm does not use the predefined model parameter classes like (Fossati et al., 2008), but considers an infinite number of possible multiple ellipse representations under the connectivity requirement.

It is natural to present the problem of fitting as finding a solution for ellipse structure that minimizes the sum of squares of distances between the data points and the ellipse structure, in which a numerical method can be used to find the solution in a non-linear least square fashion. For example, in (Jeune et al., 2004), the authors fit multiple ellipsoids to the noisy 3D hand points using the geometrically constrained Levenberg–Marquardt (L–M) (Madsen et al., 2004) algorithm to reflect the fact that fingers can only swing around the joints within an angular threshold. However, when minimizing a single objective function for a complex structure containing many connected ellipses, the solution often aligns

* Corresponding author. Fax: +61 2 633 84649.
E-mail addresses: rxu@csu.edu.au (R.Y.D. Xu), mkemp@csu.edu.au (M. Kemp).

0167-8655/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2010.01.017

only a few ellipses in the structure and misaligns the others due to the vertex-joining constraints, which the details are shown in Section 3.1.

Another way to address our fitting problem is to simultaneously segment the contour points to each ellipse and at the same time, fitting each ellipse individually to only a subset of the points it is assigned to whilst maintaining the connectivity requirement. One of the possible solutions is to consider this problem as a mixture of densities one and hence could be solved using the Expectation–Maximization procedure (Bilmes, 1997). However, since a non-Gaussian probability function (*pdf*) is required to assign contour points with a higher probability when it is closer to an ellipse edge, then a close-from solution in the M-step cannot be guaranteed.

In our proposed work, we resort our fitting method to the well-known Iterative Closest Point (ICP) framework (Alshawa, 2007; Li and Griffiths, 2000; Besl and Mckay, 1992). In the original ICP algorithm (Besl and Mckay, 1992), given a set of data and model points, the algorithm estimates the parameters for a rigid transformation. The method starts from a pairing process, which identifies between each data point with the nearest point of the model one. Next, ICP estimates the provisional parameters of the rigid transformation of all the paired points using a least square solution. The estimated transformation is then applied to all the data points. In the next iteration, the pairing and transformation is executed again from the last position of the data cloud. The iterations repeats itself until a stopping criteria is met. The ICP framework has since been extended to triangles fitting (iterative closest triangle or ICT) (Li and Griffiths, 2000) and line fittings (iterative closet line or ICL) (Alshawa, 2007).

Our aim is different to ICP, because we not trying to solve for a global rigid transformation. Instead, each ellipse is allowed to have its own freedom of movements while maintaining the connectivity requirements. However, we borrowed the idea from the ICP algorithm, where an iterative approach is proposed. At each iteration step, the contour points are firstly segmented (or assigned) to individual ellipses. At the second step, a non-linear least square algorithm is applied to simultaneously minimize two terms. The first term is the sum of square Euclidean distances between the contour point and its assigned ellipse's edge; the second term is the sum of Euclidean distances of the vertices of all pre-specified joints.

The rest of this paper is organized as follows: In Section 2, we will describe our ellipse structure and also briefly describe the approach to the initial ellipse refinement using unconstrained Gaussian mixtures. In Section 3, we will describe the centerpiece of this paper where we present our iterative fitting algorithms in a detail manner. We also discussed the relationship of our approach with Maximum A Posterior (MAP) modeling. We quantitatively compared our proposed fitting method with the previous approach where minimization of a single objective function is used. We present our results and discussions subsequently in Section 4.

2. Initial ellipse model and refinement

2.1. Connected ellipses structure

In our work, the connected ellipses structure is modeled with a set of specified joints. These joint requirements are maintained throughout the fitting process. For example, in Fig. 1a, we show that a human upper-body is modeled using six ellipses and three joints. All joints connect ellipses at vertices. There are a few different ways to parameterize an ellipse. For the purpose of this paper however, we assume joints occur at vertices of the major axis (although the algorithm presented easily extends to joints on the minor axis). Therefore, each ellipse is modeled using the following five parameters: The first two are the pointer to the joint of the parent node, x_{maj}, y_{maj}. The third and fourth parameters are the position of the opposite vertex: x'_{maj}, y'_{maj}. The last parameter is the square of the ratio between the major and minor axes of the ellipse ρ.

Prior to the application of our algorithm, we need both the image silhouette and a rough initial estimate for the ellipse model. The method to obtain the very initial coarse ellipse model is application-dependent. Templates or heuristics are both reasonable choices. For example, for modeling a human pose, the PCA-based template matching (Fossati et al., 2008) can be used.

The image silhouette can be obtained in various ways: in many computer vision applications, the silhouette typically associates with the foreground object and can be obtained using the background subtraction (Stauffer and Grimson, 1999) or alike methods if the camera is fixed. In addition, there are many recent methods designed to handle noises after the background subtraction process, and hence able to produce a reasonable clean silhouette using various application-dependant measures. For example, in a region-based approach described in (Izadi and Saeedi, 2008), the author removed the noises caused by the shadow after background subtraction and obtained a good silhouette of a moving person. A good comparison paper on the subject of noise removal after background subtraction can be found in (Parks and Fels, 2008). Therefore, for the scope of this project, we assumed to have a noise-free silhouette and have left the noise handling in the pre-processing step.

In Fig. 1b, we illustrate an example used to explain our algorithm in this paper. The contour points are sampled at regular intervals and the initial configuration of the ellipse model is far from the accurate fitting.

2.2. Improved initial configuration

To improve the initial guess and generate a better "seeding" configuration for the later fitting, we use Expectation Maximization (EM) with a Gaussian Mixture Model (GMM) (Bilmes, 1997) as the clustering algorithm. One reason for choosing EM with a GMM is because the model parameters can be converted back and forth between the ellipse parameters.

The set of observed data sets $\{\boldsymbol{x}_i\}$ is the set of 2D location features of every pixel inside the enclosed contour. Latent variables $\{z_i\}$ indicate which density of the mixtures generated $\boldsymbol{x}_i$. The initial parameter vector $\boldsymbol{\pi}^{(0)} = \left\{\left(\boldsymbol{\mu}_i^{(0)}, \boldsymbol{\Sigma}_i^{(0)}, \alpha_i^{(0)}\right)\right\}$ is determined from the rough estimate of the ellipse set configurations, where each mean is computed from the corresponding ellipse center: $\boldsymbol{\mu}_i = [x_c, y_c]^T$ and each covariance matrix is computed from the ellipse shape by noticing that two unit eigenvectors are in the same direction as the major and minor ellipse axes. The eigenvalues are the square of the lengths of the corresponding semi-axes divided by a constant z, where the choice of z is somewhat arbitrary. In order for the fitted ellipses to approximately line up with the edge pixels later, we need to make the probability that a pixel lies outside of the ellipse to be approximately zero. By adjusting the z constant we can alter this probability. For our work, we chose $z = 3.29$. This makes the probability that a pixel lies outside of the ellipse to be 0.45%. Lastly, we set each initial weight $\alpha_i^{(0)}$ to be the ratio of the area of the ith ellipse to the sum of areas of all ellipses. After the E–M procedures completes, we obtain a set of GMM parameters, $\boldsymbol{\pi}^{(f)} = \left\{\left(\alpha_i^f, \boldsymbol{\mu}_i^f, \boldsymbol{\Sigma}_i^f\right)\right\}$. We then convert them back to a set of ellipse parameters. Note that in this stage, the ellipses are not connected. The resultant ellipse set is shown in Fig. 2.

Alternatively, in some application-dependant scenarios, such refinement can be substituted by a more heuristic approach. For example, in our previous work (Xu and Kemp, 2009), the smoothed

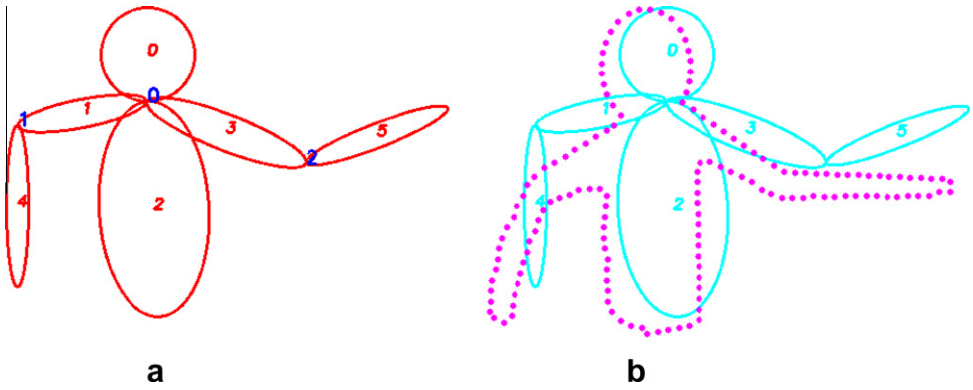

Fig. 1. (a), Six connected ellipses representation of a human upper-body, the blue number indexes the connecting joints, and (b) the model is superimposed to the rough initial contour points in pink. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

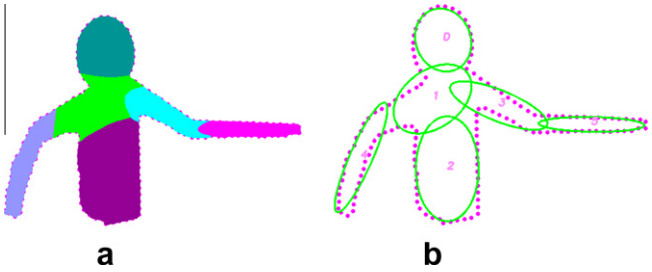

Fig. 2. Result of initial model refinement (a) shows the pixel association to each Gaussian mixture components after EM converges and (b) shows unconnected ellipse sets by converting from the final GMM parameters.

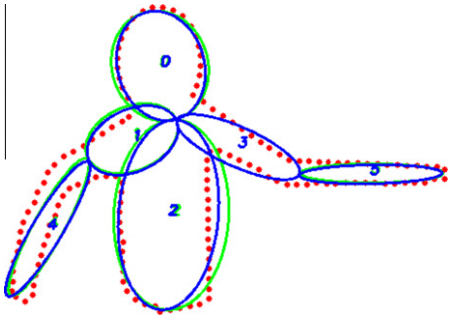

Fig. 3. Fitting as a result of using a single objective function with the entire ellipse structure.

contour curvature information is used to divide the human silhouette into segments where individual ellipses are then fitted for the model's initial setting.

3. Iterative multiple connected ellipse fitting framework

3.1. The drawback of a single objective function method

In the previous fitting method (Xu and Kemp, 2009), after obtaining the initial estimate of the unconnected ellipse set, the ellipses are "stretched" to have their vertices joined according to the original model. This is shown by the green ellipses in Fig. 3.[1] Then, a single objective function is defined as an expression for the sum of square of approximate closest Euclidean distances $fd()$ between a contour point $\boldsymbol{x}_i$ and the entire ellipse structure, parameterized by $\boldsymbol{\theta}$:

$$\arg\min_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{N_{point}} (fd(\boldsymbol{\theta}, \boldsymbol{x}_i))^2 \right\}$$

Each element in the vector $\boldsymbol{\theta}$ is a parameter of the ellipse structure. In the case of a human upper-body structure, depicted in Fig. 1, the structure has 22 degree of freedom (six ellipses and with three connections), hence $\boldsymbol{\theta}$ has 22 dimensions as the ellipses are not modeled individually. The numerical methods, for example, the Levenberg–Marquardt method (Rosin, 1993) is used to solve $\boldsymbol{\theta}$. While this method guarantees the connectivity requirement, the solution we found may satisfy only a partial number of ellipses, leaving others to be fitted less satisfactorily. In Fig. 3, the green ellipses are the initial $\boldsymbol{\theta}$ of the L–M algorithm. The blue ellipse structure is the final fitting result, i.e., the final $\boldsymbol{\theta}$ when L–M algorithm converges. We observe that only ellipse 2 and ellipse 0 were improved from the green structure. The other ellipses were virtually unchanged from before the L–M algorithm was performed.

In our proposed method, we aim to alter the fitting problem; such that each ellipse is fitted individually. Hence each contains five parameters. At the same time, only the contour points close to an ellipse are used to contribute to the fitting of that ellipse. Therefore, the connectivity constraints also needed to appear in the minimization equation as a penalty function. Similar to the ICP framework described in the introduction, where the *pairing* process and the *transformation* process needs to be applied recursively, in our work, as the minimization process changes the values of the ellipse structure parameters, so will it affect the segmentation of the contour points to the ellipses. Therefore, our proposed framework also contains the following two steps in an iterative manner: the *Segmentation* step and the *Minimization* step:

3.2. Segmentation step

During the *Segmentation* step, each point on the contour $\boldsymbol{x}_i$ is assigned to an ellipse that it has the closest distance to its edge. The shortest Euclidean distance of a point from an ellipse does not have a simple formula. We use the formula from (Jeune et al., 2004) for the distance function which is a compromise between a simple formula and a reasonable approximation to the real distance between a point $\boldsymbol{x}_i$ and an ellipse with parameters $\boldsymbol{\theta}_j = (x_{maj,j}, y_{maj,j}, x'_{maj,j}, y'_{maj,j}, \rho_j)$:

$$d_{eps}(\boldsymbol{x}_i, \boldsymbol{\theta}_j) = \begin{cases} d_1, & \text{if } \boldsymbol{x}_i \text{is inside of ellipse} \\ d_2, & \text{if } \boldsymbol{x}_i \text{is outside of ellipse} \end{cases} \quad (1)$$

For an ellipse centered at (0,0) with semi-axes lengths of a and b these distances are given by:

[1] For interpretation of color in Figs. 3,4,8,10, the reader is referred to the web version of this article.

$$d_2 = \sqrt{x_i^2 + y_i^2}\left(1 - \frac{1}{\sqrt{(x_i/a)^2 + (y_i/b)^2}}\right)$$
$$d_1 = r\left(\sqrt{(x_i/a)^2 + (y_i/b)^2} - 1\right) \quad (2)$$

where $r = \min(a,b)$ is the half-length of the minor axis

To obtain the distance for a general ellipse, we use an orthogonal transformation to center it at (0,0) and align the axes with the coordinate axes. At the end of the segmentation step each point $\boldsymbol{x}_i$ on the contour is mapped to the $e(\boldsymbol{x}_i)$th ellipse. The result of the contour point segmentation is shown in Fig. 4, where each color represents the ellipse assignments for the contour points.

3.3. Minimization step

3.3.1. The Two Individual Minimization Terms

If our goal is to minimize *only* the sum of distances between each individual ellipse and the contour points assigned to it, then we can express the sum of square distance between ellipse structure $\boldsymbol{\theta}$ and contour point set: $\boldsymbol{X} = \{\boldsymbol{x}_i\}$:

$$f(\boldsymbol{\theta}, \boldsymbol{X}) = \sum_{i=1}^{N} d^2_{eps}(\boldsymbol{x}_i, \boldsymbol{\theta}_{e(\boldsymbol{x}_i)}) \quad (3)$$

where $d^2_{eps}(.)$ is described in Eq. (1), and N is the number of contour points.

At the same time, if our goal is to *only* make the unconnected ellipse models in Fig. 3b connected (or closer to connection), then we need to define an objective function for the sum of the distances between the ellipse joints, which we refer to as $g(\boldsymbol{\theta})$. Note that g is independent of the contour points $\boldsymbol{x}_i$.

Section 2.1 indicates that we express the ellipse joints structure in a tree model. Therefore, there should be a unique joint between each pair of ellipses. Subsequently, we define:

$$\delta(p,q) = \begin{cases} 1, & \text{if ellipse } \boldsymbol{\theta}_p \text{ and } \boldsymbol{\theta}_q \text{ has a joint} \\ 0, & \text{if ellipse } \boldsymbol{\theta}_p \text{ and } \boldsymbol{\theta}_q \text{ does not have a joint} \end{cases} \quad (4)$$

Therefore, the sum of all Euclidean distances between the corresponding joints is:

$$g(\boldsymbol{\theta}) = \frac{1}{2}\sum_{p=1}^{M}\sum_{q=1}^{M} d^2_{jts}(p,q)\delta(p,q) \quad (5)$$

where $d_{jts}(p,q)$ is the Euclidian distance between the two closest major axis of ellipse p and q.

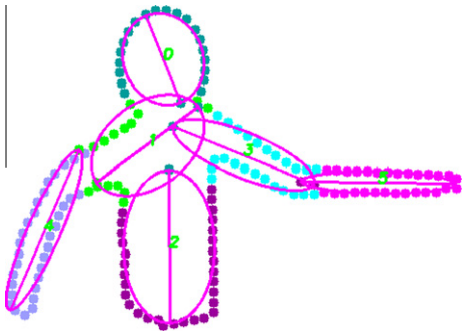

Fig. 4. Contour points to individual ellipse segmentation, different colors represent each of the ellipse assignments.

3.3.2. Combined minimization

Our ultimate aim is to obtain an optimal fitting for the joint ellipse structure which minimizes the sum of square distances between the ellipses with its respected segmented contour points while maintaining the defined joint structure. Therefore, we must minimize both f and g function simultaneously. More formally, in the *minimization* step, given a set of contour point $\{\boldsymbol{x}_i\}$, we would like to compute iteratively, $\boldsymbol{\theta}^{(k+1)}$, such that:

$$\boldsymbol{\theta}^{(k+1)} = \arg\min_{\boldsymbol{\theta}} \{w_{eps} f(\boldsymbol{\theta}, \boldsymbol{X}) + w_{jts} g(\boldsymbol{\theta})\}$$

The scalar weights, w_{eps} and w_{jts} are inserted into the above equation to control the balance between optimizing distances between the joints and optimizing the distances between the contour and the ellipses (it is elaborated in Section 3.4.1). Substituting Eqs. (3) and (5) in, we can rewrite the above equation to be:

$$\boldsymbol{\theta}^{(k+1)} = \arg\min_{\boldsymbol{\theta}} \left\{ w_{eps}\sum_{i=1}^{N} d^2_{eps}(\boldsymbol{x}_i, \boldsymbol{\theta}_{e(\boldsymbol{x}_i)}) + \frac{w_{jts}}{2}\sum_{p=1}^{M}\sum_{q=1}^{M} d^2_{jts}(\boldsymbol{\theta}_p, \boldsymbol{\theta}_q)\delta(p,q) \right\} \quad (6)$$

3.3.3. Solution using Levenberg–Marquardt (L–M) algorithm

We use Levenberg–Marquardt (L–M) algorithm (Madsen et al., 2004) to solve Eq. (6). The L–M algorithm iteratively finds the minimum value of a multivariate function of parameter vector $\boldsymbol{\theta}$ that is expressed as the sum of squares of non-linear functions $\xi_i(\boldsymbol{\theta})$, with an initial guess $\boldsymbol{\theta}^{(i)}$, i.e.:

$$\boldsymbol{\theta}^{(f)} = \arg\min_{\boldsymbol{\theta}} \left\{ \frac{1}{2}\sum_{i=1}^{N} (\xi_i(\boldsymbol{\theta}))^2 \right\}, \text{with initial } \boldsymbol{\theta} = \boldsymbol{\theta}^{(i)} \quad (7)$$

Therefore, we can express Eq. (6) in the form of a non-linear least square, with the initial estimate $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$, derived from the previous *Segmentation–Minimization* iteration:

$$\boldsymbol{\theta}^{(k+1)} = \arg\min_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{N} (F(\boldsymbol{x}_i, \boldsymbol{\theta}))^2 \right\}$$
$$\text{where} \quad F(\boldsymbol{x}_i, \boldsymbol{\theta}) = \sqrt{w_{eps} d^2_{eps}(\boldsymbol{x}_i, \boldsymbol{\theta}_{e(\boldsymbol{x}_i)}) + \frac{w_{jts}}{2N}\sum_{p=1}^{M}\sum_{q=1}^{M} d^2_{jts}(\boldsymbol{\theta}_p, \boldsymbol{\theta}_q)\delta(p,q)} \quad (8)$$

As both $d^2_{eps}(\boldsymbol{x}_i, \theta_{c_i})$ and $d^2_{jts}(\theta_p, \theta_q)$ functions are positive, taking square root in Eq. (8) is valid.

Note that the value of $g(\boldsymbol{\theta})$ is the same at each L–M iteration. Therefore, it can be computed just once before L–M starts at each recursion, which we divide it by N, the number of contour points as shown in Eq. (8). After L–M converges, we obtain a set of new ellipse parameters. We then begin a new iteration.

The results of the iterative steps are shown in Fig. 5, where we show the ellipse structure fitting results after the 1st, 10th and 30th recursions, respectively. It shows improvements notably in the left arm, i.e., ellipses 1 and ellipse 4 after the 30th recursion.

3.4. Measure for the Quality of the Fitting

In order to quantitatively measure the result of fitting, we define Quality of the Fitting (QoF) to be the root mean of square (RMS) of distances between each point $\boldsymbol{x}_i$ and an ellipse indexed by $e(\boldsymbol{x}_i,\boldsymbol{\theta})$ (using Eq. (1)) that enjoys the closest distance from the point $\boldsymbol{x}_i$.

$$\text{QoF}(\boldsymbol{\theta}, \{\boldsymbol{x}_i\}) = \sqrt{\frac{\sum_{i=1}^{N} d^2(e(\boldsymbol{x}_i, \boldsymbol{\theta}), \boldsymbol{x}_i)}{N}} \quad (9)$$

For the ellipse structure of the "upper-body" image and the associated contour points given in the previous section, QoF of

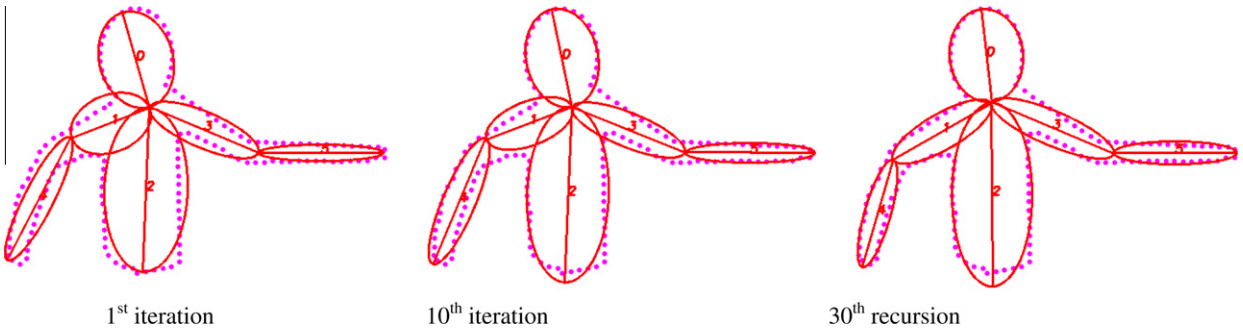

Fig. 5. The result of ellipse fitting at different recursions.

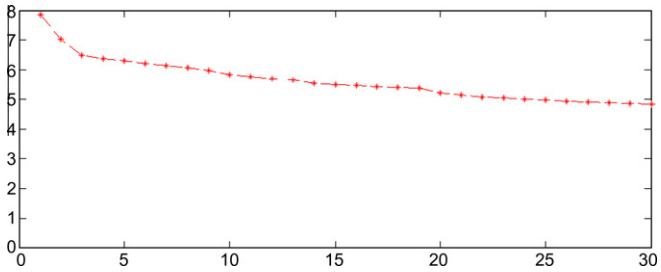

Fig. 6. The plot of Quality of the Fitting over the 30 iterations.

the rough initial structure is 31.54. After five iterations, the QoF is further reduced to 6.30. The plot of QoF for the "upper-body" image over 30 iterations is shown in Fig. 6.

3.4.1. The weights

The weights, w_{eps} and w_{jts} proposed in our framework indicate the balance of the L–M procedure between the two minimizations. To illustrate the implication of these weights, as shown in Fig. 7a, when we let $w_{eps} = 0$, the procedure is merely trying to close the distances between the corresponding joints. For this reason, as shown in Fig. 6a, the result of the iteration without the first term simply "dragged" the ellipse's joints together, resulting in a poor fit to the contour points. This is noticeable particularly in ellipse 2 and 4 in Fig. 7a.

Likewise, if we let $w_{jts} = 0$, then the procedure merely minimizes the distances between the points and the individual ellipses associated with those points. As the result in Fig. 7b shows, each ellipse fits well to the contour segments assigned to it (depicted by the different colors). However, the ellipses are drastically changed from their initial shapes and the joints are not connected.

3.5. Relationship to maximum a posterior (MAP)

The Eq. (6) is used to find a parameters set that minimizes the sum of the two square distance functions simultaneously. This can also be thought as to maximize the log of the posterior probability, where the posterior is expressed as:

$$p(\boldsymbol{\theta}|\boldsymbol{X}) \propto p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

The likelihood function is a joint probability and since we segmented contour point sets $\boldsymbol{X} = \{\boldsymbol{x}_i\}$ to each individual ellipse $\boldsymbol{\theta}_{e(\boldsymbol{x}_i)}$ in step 3.1, therefore, the likelihood function can be written as:

$$L(\boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta}) = \prod_i^N p(\boldsymbol{x}_i|\boldsymbol{\theta}_{e(\boldsymbol{x}_i)}) \tag{10}$$

When we consider it as a function of $\boldsymbol{x}_i$, we could let $p(\boldsymbol{x}_i|\boldsymbol{\theta}_{e(\boldsymbol{x}_i)}) \propto \exp(-d^2_{eps}(\boldsymbol{x}_i, \boldsymbol{\theta}_{e(\boldsymbol{x}_i)}))$. However, when considered as a likelihood function in terms of $\boldsymbol{\theta}$, $p(\boldsymbol{x}_i|\boldsymbol{\theta}_{e(\boldsymbol{x}_i)})$ needs to be normalized by a function $\psi(\boldsymbol{\theta}_{e(\boldsymbol{x}_i)})$, which can be derived through integration. Therefore:

$$L(\boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta}) = \prod_i^N \frac{1}{\psi(\boldsymbol{\theta}_{e(\boldsymbol{x}_i)})} \exp(-d^2_{eps}(\boldsymbol{x}_i, \boldsymbol{\theta}_{e(\boldsymbol{x}_i)})) \tag{11}$$

The prior function can be used to ensure the joints join and can be written as the joint probability:

$$p(\boldsymbol{\theta}) \propto \prod_{1 \leqslant p < q \leqslant M} \exp(-d^2_{jts}(\boldsymbol{\theta}_p, \boldsymbol{\theta}_q))^{\delta(p,q)} \tag{12}$$

This can be guaranteed to be a probability distribution function, if necessary, by constraining the parameters uniformly to a compact set. Therefore:

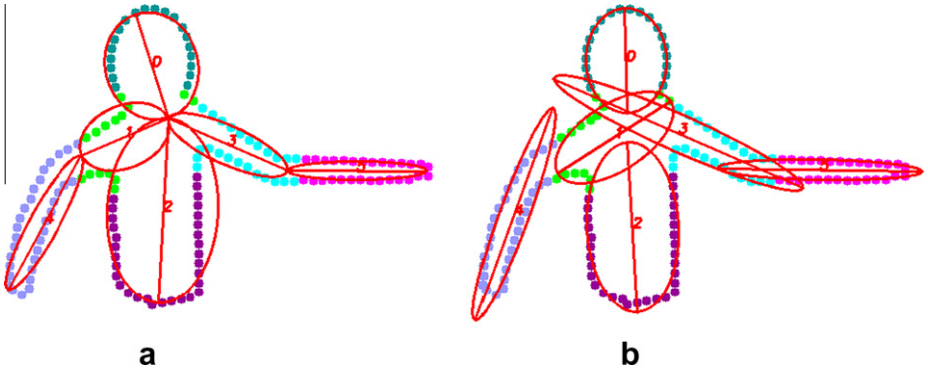

Fig. 7. Both extreme cases of weights, (a) $w_{eps} = 0$, and (b) $w_{jts} = 0$. The contour point segmentation is also shown in different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

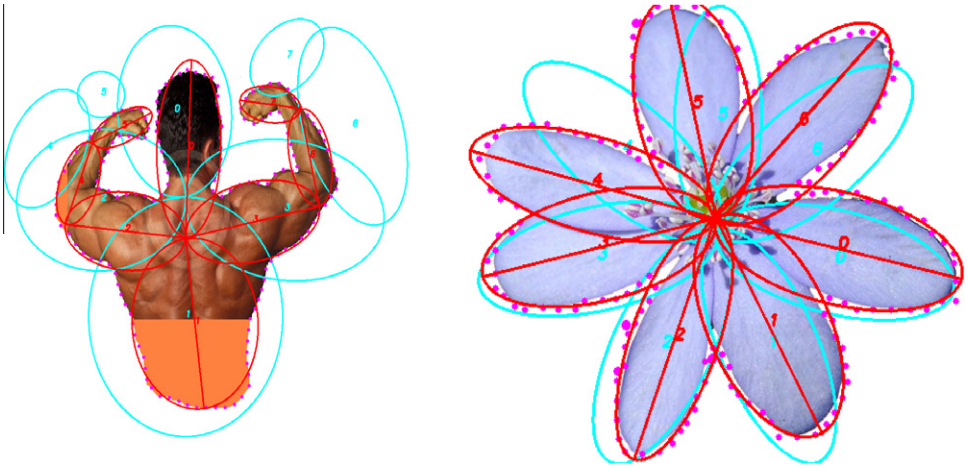

Fig. 8. Other results of multiple ellipse fittings.

Fig. 9. Images of human upper-body foreground blob.

$$\begin{aligned}&\arg\max_{\theta}\{\log(p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta}))\}\\&=\arg\min_{\theta}\left\{\sum_{i=1}^{N}\left\{\log(\psi(\boldsymbol{\theta}_{e(\boldsymbol{x}_i)}))+d^2_{eps}(\boldsymbol{x}_i,\boldsymbol{\theta}_{e(\boldsymbol{x}_i)})\right\}\right.\\&\left.+\frac{1}{2}\sum_{p=1}^{M}\sum_{q=1}^{M}d^2_{jts}(\boldsymbol{\theta}_p,\boldsymbol{\theta}_q)\delta(p,q)\right\}\end{aligned}$$

The above equation has a close resemblance of Eq. (6).

4. Results and discussion

4.1. Test images

In order to test the effectiveness of our proposed framework, we have also performed experiments using other images with their corresponding pre-specified ellipse structure for the purpose of rough initial guess. Results obtained are shown in Fig. 8, where the cyan color is initial ellipse joint models and the red is the final results after the 5th recursion. Note that although color images are used, we merely use them as foreground silhouettes.

The QoF for all three test images are shown below in Table 1. In the first column, we listed the QoF for the initial setting, in the second column, we listed the QoF when a single objective function after GMM refinement is used (Section 3.1). Finally, we listed the QoF using our proposed algorithm at the 1st, 5th and 10th iterations.

Table 1
The measure of QoF for the three test cases to compared with the single objective function fitting (Xu and Kemp, 2009).

	Rough initial setting	Single objective function	1st Iteration	5th Iteration	10th Iteration
Upper-body	31.54	8.83	7.85	6.30	5.84
Strong man	37.47	9.75	9.36	8.46	8.17
Flower	23.90	5.06	4.17	3.72	3.48

4.2. Application to human body part detection

In addition to the above images, we also tested our iterative algorithm against human upper-body images. In this scenario, the way which we obtained the initial unconnected ellipses (Fig. 2b) is different to the methods described in Section 2.2. Instead of using Gaussian Mixture Modeling as a refinement step, we obtain them using a curvature-based heuristic approach (Xu and Kemp, 2009). We first obtained the foreground blobs using the background subtraction algorithms (Stauffer and Grimson, 1999). The results are shown in Fig. 9.

We removed the noise by selecting the largest connected component containing the centroid. We then acquired the contour from the blob's bottom left to its bottom right. This is shown in Fig. 10. The green and blue points indicate, respectively the beginning and end of the contour. For background images containing significant noises, the true contour may be obtained using more advanced noise handling algorithms, such as (Izadi and Saeedi, 2008) and (Parks and Fels, 2008).

We then smoothed the silhouette so that we can determine just the global features: the head, hands and armpits. Next we reduced the smoothing to detect the local features of the neck and elbows. The individual ellipses are fitted to each segment, and we obtained a set of unconnected ellipses shown in Fig. 11. Please see Xu and Kemp (2009) for details:

Finally, we applied our iterative fitting algorithm described in this paper. The result is shown in Fig. 12.

Our final comment is that although the methods and data described in this paper are concentrating on 2D ellipse fitting, this is merely for the purpose of clear illustration. However, our method can easily extend into fitting multiple connected ellipsoids to an

Fig. 10. Result of human upper-body silhouettes pre-processing.

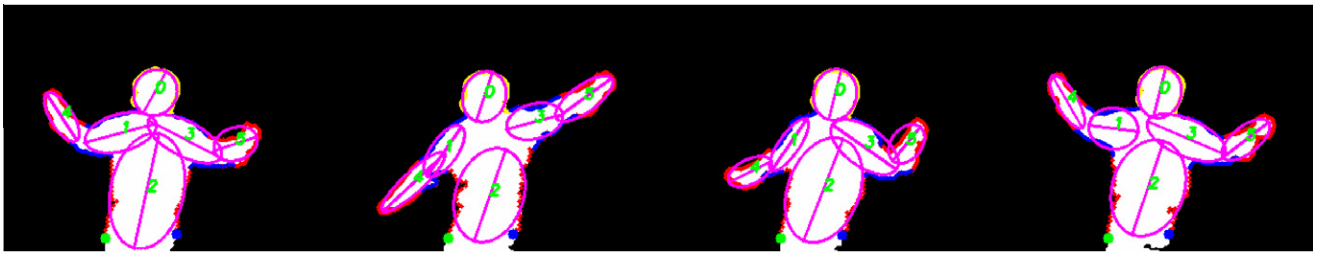

Fig. 11. Result of obtaining individually fitted unconnected ellipses using Xu and Kemp (2009).

Fig. 12. The fitting result using real human foreground data: six body segments are firstly obtained using curvature analysis (Xu and Kemp, 2009) (shown in different colors) before applying our algorithm. The results shows ellipse configuration after the 5th iteration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

enclosed volumetric data, and this is where our future work will be focusing on.

Acknowledgments

Our thanks to Charles Sturt University's Center for Research in Complex Systems for supporting our research; we also thank Mr. Anthony Wallis and Mrs. Oriana Zanon for their proofreading efforts.

References

Alshawa, M., 2007. lCL: Iterative closest line A novel point cloud registration algorithm based on linear features. Ekscentar 10, 53–59.

Bernier, O., Cheung-Mon-Chan, P., Bouguet, A., 2009. Fast nonparametric belief propagation for real-time stereo articulated body tracking. Computer Image and Vision Understanding 113, 29–47.

Besl, P.J., Mckay, N.D., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Machine Intell. 14, 239–256.

Bilmes, J.A., 1997. A gentle tutorial on the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden markov models.

Du, H., Charbon, E., 2007. 3D hand model fitting for virtual keyboard system. In: Eighth IEEE Workshop on Applications of Computer Vision, p. 31.

Fossati, A., Arnaud, E., Horaud, R., Fua, P., 2008. Tracking articulated bodies using generalized expectation maximization. In: CVPR Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment, Anchorage, AK, USA.

Izadi, M., Saeedi, P., 2008. Robust region-based background subtraction and shadow removing using color and gradient information. In: 19th Internat. Conf. Pattern Recognition, 2008. ICPR 2008, Tampa, FL.

Jeune, F.L., Deriche, R., Keriven, R., Fua, P., 2004. Tracking of hand's posture and gesture, CERTIS, ENPC, %3chttp://certis.enpc.fr/publications/papers/04certis02.pdf2004%3e.

Li, Q., Griffiths, J.G., 2000. Iterative closest geometric objects registration. Comput. Math. Appl. 40, 1171–1188.

Madsen, K., Nielsen, H.B., Tingleff, O., 2004. Methods for non-linear least squares problems: IMM, DTU.

Parks, D.H., Fels, S.S., 2008. Evaluation of background subtraction algorithms with post-processing. In: IEEE Conf. Advanced Video and Signal Based Surveillance.

Rosin, P.L., 1993. A note on the least squares fitting of ellipses. Pattern Recognition Lett. (PRL) 14, 799–800.

Stauffer, C., Grimson, W.E.L., 1999. Adaptive background mixture models for real-time tracking. In: IEEE CVPR, pp. 246–252.

Xu, R.Y.D., Kemp, M., 2009. Multiple curvature based approach to human upper body parts detection with connected ellipse model fine-tuning. In: IEEE Internat. Conf. on Image Processing (ICIP2009), Cairo Egypt.

Yiu, B.W.-S., Wong, K.-Y.K., Chin, F.Y.L., Chung, R.H.Y., 2005. Explicit contour model for vehicle tracking with automatic hypothesis validation. In: IEEE Internat. Conf. on Image Processing, 2005. ICIP 2005, pp. II-582–II-589.