Pattern Recognition 53 (2016) 259–275

Contents lists available at ScienceDirect

Pattern Recognition

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

Parameter-free modelling of 2D shapes with ellipses

Costas Panagiotakis [a,b,*], Antonis Argyros [b,1]

[a] Department of Business Administration, Technological Educational Institute of Crete, 72100 Agios Nikolaos, Crete, Greece
[b] Foundation for Research and Technology-Hellas (FORTH), Institute of Computer Science, 70013 Heraklion, Crete, Greece

ARTICLE INFO

Article history:
Received 24 May 2015
Received in revised form
27 October 2015
Accepted 7 November 2015
Available online 22 November 2015

Keywords:
Ellipses fitting
Shape analysis
Shape complexity
Expectation-Maximisation
Model selection
AIC
BIC

ABSTRACT

Our goal is to represent a given 2D shape with an automatically determined number of ellipses, so that the total area covered by the ellipses is equal to the area of the original shape without any assumption or prior knowledge about the object structure. To solve this interesting theoretical problem, first we employ the skeleton of the 2D shape which provides important information on the parameters of the ellipses that could approximate the original shape. For a given number of such ellipses, the hard Expectation-Maximisation (EM) algorithm is employed to maximise the shape coverage under the Equal Area constraint. Different models (i.e., solutions involving different numbers of ellipses) are evaluated based on the Akaike Information Criterion (AIC). This considers a novel, entropy-based shape complexity measure that balances the model complexity and the model approximation error. In order to minimise the AIC criterion, two variants are proposed and evaluated: (a) the augmentative method that gradually increases the number of considered ellipses starting from a single one and (b) the decremental method that decreases the number of ellipses starting from a large, automatically defined set. The obtained quantitative results on more than 4000 2D shapes included in standard as well as in custom datasets, quantify the performance of the proposed methods and illustrate that their solutions agree with human intuition.

1. Introduction

Object modelling and representation through a sufficient number of geometric primitives (prototypes) is a key problem in computer vision and pattern recognition with several applications including object detection and retrieval [1,2], tracking [3], motion analysis and action recognition [4,5]. The selection of the type of prototypes is application dependent. For example, when curves are modelled, straight line segments are usually preferred due to their simplicity. In grey scale or colour images, segmentation methods [6–9], mixtures of simpler models such as rigid templates and bag-of-features or deformable part models such as pictorial structures have been used for object representation and detection [2,10]. Ellipsoids and superquadrics [11,12] have been considered as volumetric primitives for 3D shape representation, since they are convenient part-level models that can further be deformed together to model articulated objects.

In the case of region modelling, a 2D binary image is given with foreground points representing the shape to be modelled. This image can be the result of any object detection or image segmentation method (e.g., [6,9]). The goal is then to estimate the parameters of k-prototypes that approximate the foreground shape according to some predefined optimality criterion. Ellipses are convenient such prototypes, since their parameters convey information including the location, orientation and variation of data [1]. Depending on the selection of the optimality criterion, two problem formulations can be identified:

- The problem of Maximum Coverage (MAX-α), where the goal is to maximise the coverage α of foreground points given a fixed number of ellipses k.
- The problem of Minimum Number of Ellipses (MIN-k), where the goal is to find the minimum number k of ellipses that achieve a certain coverage α.

Both formulations are defined under the Equal Area constraint which states that the sum of the areas of all ellipses should be equal to the area of the given shape.

Analogous problem formulations exist in the case of approximating curves with straight line segments. In that context, the aforementioned problems are optimally solved in polynomial time by graph-theoretical and dynamic programming methods [13]. For the case of approximating 2D shapes by ellipses, both formulations exhibit a high computational complexity. For the MAX-α formulation, the number k of ellipses needs to be known. Their

* Corresponding author.
E-mail addresses: cpanag@staff.teicrete.gr (C. Panagiotakis), argyros@ics.forth.gr (A. Argyros).
URLS: http://www.csd.uoc.gr/~cpanag (C. Panagiotakis), http://users.ics.forth.gr/~argyros (A. Argyros).
[1] Also with the Computer Science Department, University of Crete, Greece.

http://dx.doi.org/10.1016/j.patcog.2015.11.004
0031-3203/

parameters may be estimated by a hard Expectation Maximisation (EM) algorithm which, nevertheless, is not guaranteed to converge to an optimal solution. For the MIN-k formulation, optimality is not guaranteed again. Additionally, the parameter α needs to be decided a priori, based on some heuristic and/or application dependent manner. Additionally, large values of α result in a large number of ellipses that, also due to discretisation effects, model the shape with unnecessary detail.

We are interested in a parameter-free method that automatically computes a sufficient number of ellipses that achieve a good balance between model complexity and shape coverage under the Equal Area constraint. This theoretically interest computer vision problem has compact and easy-to-grasp description but a very high algorithmic complexity due to the large search space and to the unspecified number of the ellipses. Even if just one ellipse is used, there does not exist any trivial method to compute the optimal solution for MAX-α.

In this work, we define a new shape complexity measure, which is combined with AIC [14] to automatically approximate a 2D shape by a collection of ellipses without a priori knowledge or specification of k and α. To the best of our knowledge, we propose the first parameter-free method that automatically estimates an unknown number of ellipses that achieve a good balance between model complexity and shape coverage. Although not necessarily optimal, the resulting methodology suggests solutions that outperform existing methods in quantitative terms and also agree with human intuition.

As an example, Fig. 1 illustrates the solutions of our approach for two different human figures. In the left figure, five ellipses cover 92.5% of the shape, while in the left, twelve ellipses cover 91.5% of it. An increase of the number of ellipses will increase shape coverage, at the cost of an increase in modelling complexity that is incompatible to the complexity of the given shapes. Both solutions agree with human intuition. In both cases, the number of ellipses and their parameters are automatically determined by the proposed method.

2. Related work

Several methods have been proposed for deciding how a 2D shape can be represented as a set of ellipses. They can be classified into (a) boundary based methods that minimise the approximation error between the boundaries of ellipses and the scattered data points and (b) region based methods that minimise the approximation error between the regions of ellipses and the given shape.

The fitting of a single elliptic curve to a given set of 2D points is a well-studied problem. A robust method for direct least square fitting of a single ellipse that solves a generalised eigensystem has been proposed in [15]. The approaches for detecting ellipses that fit on edge pixels (data points) in an image can be grouped into those that employ (a) the Hough Transform, (b) Genetic Algorithms and (c) edge-following [16]. In order to detect ellipses effectively and quickly using Hough Transform, many variations have been suggested that overcome the algorithmic time and space complexity of the original Hough Transform formulation. For example, the use of a subset of the data points [17] and the decomposition of the five-dimensional parameter space into several sub-spaces of lower dimensions [18] accelerates the process. The Genetic Algorithm-based methods solve the problem by finding a suboptimal solution to an optimisation problem [19]. Finally, the edge-following algorithms consist of two main steps. First, they compute the chained edge points from the original edge image and then they perform grouping of these short segments and fitting to a set of ellipses [16].

Most of the aforementioned methods work well when the ellipses are partially occluded, but cannot be used to fit a set of ellipses in a region. Region based methods do not restrict themselves to the boundaries of shapes but rather take into account all shape points. Thus, they are more tolerant to noise. For a single ellipse, its parameters can be analytically determined based on its second order moments [4]. For more than one ellipses, the fitting problem is not as trivial. Several methods provide approximate solutions. There exists methods where the number of prototypes is given [20]. Other methods assume that aspects of the model structure is known a priori [21]. However, the problem of automatically determining the model structure is still an open one [21].

In [22], the head and the body parts of a human figure are identified as two ellipses. The K-means algorithm is used to minimise the fitting error between the figure points and the two ellipses. Then, individual ellipses are fit to each cluster using the direct least square fitting method [15]. The clusters are recalculated and the process is repeated until convergence. In [20], a method based on image moments and the bsp-tree data structure for structuring and tracking of objects in video sequences has been proposed. That algorithm uses image moments to build a hierarchical structure which approximates the object by a set of ellipses at different levels. The main limitation of this method is that the number of the ellipses should be a power of two. When the given object is represented using a tree structure with known connections and possible locations for each part are also given, then the best match of the model to an image can be nearly optimal and can be estimated in linear time [23].

The method proposed in [21] has some similarities to our work. In that work, a Gaussian Mixture Model (GMM) is estimated using the classical Expectation-Maximisation (EM) algorithm to estimate the ellipse decomposition of a silhouette image. However, this method assumes a known number of ellipses which are additionally organised in a known hierarchical tree structure. These two constraints greatly simplify the optimisation problem because they permit the optimisation of a subset of ellipses at a time as

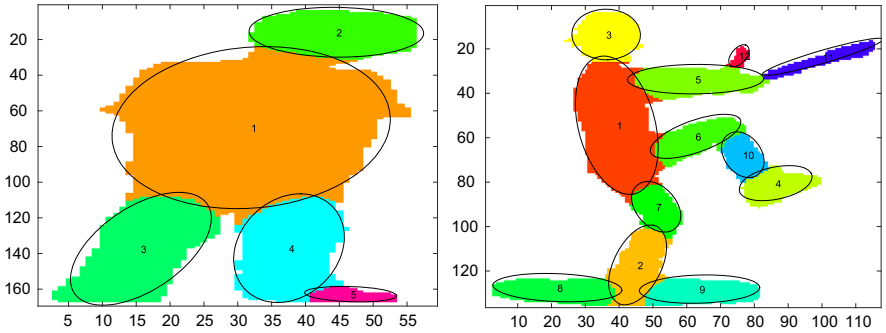

Fig. 1. Example outputs of the proposed method for two different human figures.

opposed to the optimisation of the full set of parameters of all ellipses whose number is unknown.

In contrast to the existing approaches, in this work the model structure as well as the number of ellipses is considered unknown. These are automatically computed as the method evolves based on a new entropy based shape complexity measure that balances model complexity and the fitting error between the model and the data. We propose two different search and optimisation strategies. The first one is called Augmentative Ellipse Fitting Algorithm (AEFA) and the second Decremental Ellipse Fitting Algorithm (DEFA). AEFA gradually increases the number of ellipses starting from a single one, while DEFA decreases the number of ellipses starting with a large (though automatically determined) number of ellipses.

In the literature, a lot of efforts target the definition of a proper measure of the complexity of a 2D shape. Such a measure is important to several applications in computer vision as well as in developing efficient shape classification algorithms [24,25]. In [24], shape complexity $H(p)$ is computed as the entropy to the histogram of curvature:

$$H(p) = -\sum_i p_i \log\ p_i. \tag{1}$$

According to this definition, the more symmetrical curves have lower shape complexity. This definition is scale, rotation and translation invariant. A disadvantage is that the local variation of curvature is ignored, since the global histogram of curvature is used. In [25], the shape complexity is computed by a weighted sum of three criteria: (a) entropy of the global distance distribution, (b) entropy of local angle distribution, and (c) shape randomness. Shape complexities vary in the range [0..2]. A drawback of this method is that the relative importance of the three criteria has to be determined by the user who has to set an appropriate weighting scheme. In contrast, in this work, we propose a new shape complexity measure that is efficiently combined with AIC [14] to aid towards the solution of the model selection problem in a parameter-free manner.

Besides its theoretical interest, the proposed methodology has interesting applications, e.g., to the problem of recovering automatically the unknown kinematic structure of an unmodelled articulated object based on a given set of views. This also explains why it is important not to pre-specify the number of primitives and why [21] does not fit our goals.

In summary, the major contributions of our work are the following:

- We propose a parameter-free method that approximates a given 2D shape with a number of ellipses, so that the total area covered by the ellipses is equal to the area of the original shape (Equal Area constraint).
- The proposed methodology automatically computes the number and the parameters of the ellipses that achieve a good balance between model complexity and shape coverage under the Equal Area constraint. No prior knowledge of the shape is required.
- Our approach capitalises on a novel definition of shape complexity that exploits the skeleton of a shape. In combination with the AIC, shape complexity defines an information-theoretic criterion for model selection that balances model complexity and the shape coverage error.

The rest of the paper is organised as follows. Section 3 sets the scene by formulating the problem. Sections 4 and 5 present the proposed shape complexity measure and the model selection problem, respectively. Two ellipse fitting variants are proposed (Sections 6 and 7) and their computational complexity is investigated in Section 8. The experimental results are provided in Section 9. Finally, Section 10 summarises our work by providing the main conclusions of our study.

3. Problem formulation

We assume a binary image I that represents a 2D shape. A pixel p of I belongs either to the foreground FG ($I(p)=1$) or to the background BG ($I(p)=0$). The area A of the 2D shape is given by

$$A = \sum_{p \in FG} I(p). \tag{2}$$

We also assume a set E of k ellipses E_i, each with individual area $|E_i|$. The binary image U_E is also defined so that $U_E(p)=1$ at points p that are inside any of the ellipses E_i and $U_E(p)=0$, otherwise. Then, we define the coverage $\alpha(E)$ of the 2D shape by the given set of ellipses E as:

$$\alpha(E) = \frac{1}{A}\sum_{p \in FG} I(p)U_E(p). \tag{3}$$

Essentially, $\alpha(E)$ is the percentage of the 2D shape points that are under some of the ellipses in E.

Let $|E|$ denote the sum of the areas of all ellipses $|E| = \sum_{i=1}^{k} |E_i|$. As reported in Section 1, the problem of Maximum Coverage (MAX-α) amounts to computing the parameters of a set E^* of k ellipses E_i (k is fixed), so that $\alpha(E^*)$ as defined in Eq. (3) is maximised, under the constraint that the sum of the areas of all ellipses is equal to the area of the 2D shape (Equal Area constraint).

Formally,

$$E^* = \arg\max_E \alpha(E), \quad \text{s.t. } |E| = A. \tag{4}$$

The problem of Minimum Number of Ellipses (MIN-k) amounts to computing the parameters of the minimum number k of ellipses E_i, so that the coverage $\alpha(E^*)$ is greater than or equal to a given threshold τ under the Equal Area constraint ($|E| = A$). Both MAX-α and MIN-k formulations define a fixed parameter (k or τ, respectively). In this work, we define a new, parameter-free version of the problem. The goal is now to compute the parameters of an automatically defined number k of ellipses E_i, so that the trade-off between shape coverage $\alpha(E)$ and model complexity is optimised. Intuitively, if a 2D shape is "simple", increasing the number of ellipses towards increasing the coverage, should be more penalised than the case of a complex shape. Section 4 defines shape complexity and Section 5 presents how this is used to drive the model selection process.

4. Shape complexity

We define a new shape complexity measure based on the skeleton S of the shape to be modelled. The skeleton or medial axis is the set of the centroids of circles that are tangent to shape boundary in at least two points, where all such circles are contained in the shape [26]. The medial axis together with the values of the radii of the maximally inscribed circles is called the Medial Axis Transform (MAT). MAT is a compact and lossless representation of the shape, as it can be used to reconstruct the original shape. As such, it can be used to assess shape similarity and assist in object recognition and retrieval, outperforming, in may cases, other shape descriptors in the presence of partial occlusion and articulation of parts [27]. Fig. 2 (left) illustrates the MAT representation. The colouring of the skeleton points denotes the radius of the maximal inscribed circle at that point.

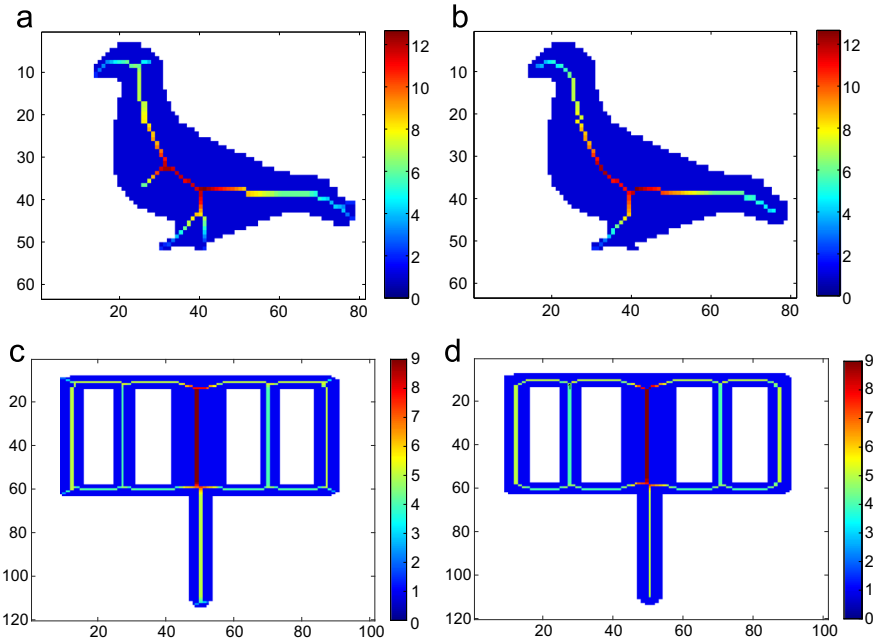

Fig. 2. The medial axis ((a) and (c)) and the result of thinning ((b) and (d)) of two shapes. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

Let $G=(V,W)$ be the graph comprising a set V of nodes together with a set W of edges, where V are the branching nodes and the end points of a skeleton. The edges W naturally segment the skeleton into parts $l_i \in S, i \in \{1,\ldots,|W|\}$, where $|W|$ denotes the number of edges. For each of these segments, the 16-bin histogram of the radii of the circles is computed. Shape complexity is then defined as a function of the entropy of these MAT-based histograms. More specifically, the proposed shape complexity C is defined as

$$C = -\sum_{i=1}^{|W|}\sum_{j=1}^{16} p_{ij} \log\ p_{ij} + \log|S|, \tag{5}$$

where p_{ij} is the j bin of the i-th histogram. The term $\log|S|$ represents global information of the skeleton as it is an upper bound of the entropy of the radii histogram computed over the whole skeleton. Since we employ 16-bin histograms, C is bounded by $4|W|+\log|S| = O(\log A)$.

Shape complexity as previously defined is translation and rotation invariant, while it slightly increases with scale due to the $\log|S|$ term. However, minor boundary deformations may create several branches on the extracted skeleton, thus affecting shape complexity. To alleviate this problem, instead of the medial axis, we employ the skeleton resulting out of shape thinning [28], since it provides a more compact representation keeping the most important skeleton branches. Additionally, prior to that, we employ a closing morphological filter [29] to the input shape image. Another option would be to keep the most important skeleton branches of MAT by applying some skeleton pruning method [30]. We refrain from doing this, as skeleton pruning algorithms are associated with a predefined threshold that controls the level of pruning. It should be noted that while model selection is based on the adopted shape complexity measure, the proposed method could, in principle, accommodate any other definition.

It has to be noted that the proposed shape complexity measure is well defined also for shapes with holes. Fig. 2 illustrates two examples of the medial axis ((a) and (c)) and the result of thinning ((b) and (d)) of a shape without holes ((a) and (b)) and a shape with holes ((c) and (d)). In the first example ((a) and (b)), for the medial axis, $G=(V,W)$ consists of 12 nodes and 11 segments and results in a shape complexity of 31.43. The skeleton resulting from thinning corresponds to a graph consisting of 4 nodes and 3 segments and a shape complexity of 16.74. In the second example ((c) and (d)), for the medial axis, $G=(V,W)$ consists of 18 nodes and 21 segments and results in a shape complexity of 33.18. The skeleton resulting from thinning corresponds to a graph consisting of 8 nodes and 12 segments and a shape complexity of 19.94.

5. Model selection

In order to solve the parameter-free ellipse estimation problem, we employ the Akaike Information Criterion (AIC) [14] on the defined shape complexity measure (Eq. (5)), as a means to automatically compute the suitable ellipse set. AIC is preferred since it is one of the most popular information-theoretic criteria for model selection dealing well with the trade-off between the accuracy of model fitting and modelling complexity. However, the proposed methodology can be easily modified to accommodate any other criterion like BIC, AICc, etc.

Given a statistical model of k parameters, the AIC is given by $AIC = 2k - 2\ln(L)$, where L is the maximised value of the likelihood function for the model. AIC is also connected with the mean squared prediction error (MSE) of the model by $N\ln(MSE)+2k+\varrho$, where N denotes the sample size (in lossless representation) and ϱ is a constant independent of the model used, therefore it can be ignored [31]. According to our problem formulation, the fitting performance $\alpha(E) = 1 - MSE$. We also propose to use the defined shape complexity C as an estimation of the size of a lossless representation of the shape. Therefore, the AIC-based model selection criterion in this work amounts to the minimisation of the quantity

$$AIC(E,C) = C\ln(1-\alpha(E)) + 2k. \tag{6}$$

It should be noted that the above quantity is invariant to shape rotation and translation. Scale changes slightly affect the model selection process.

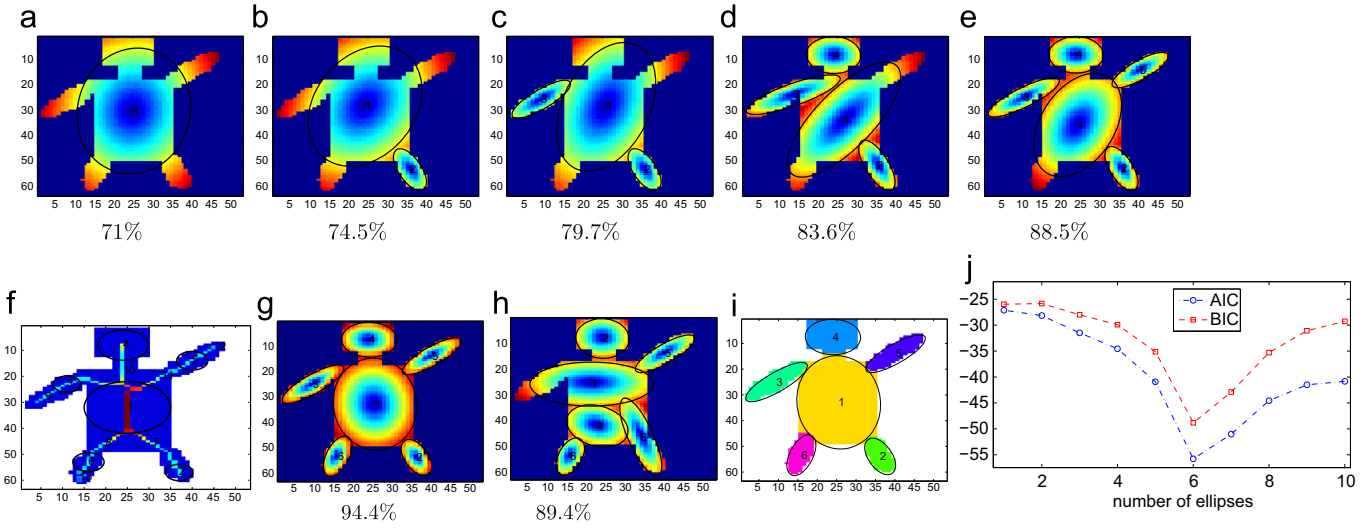

Fig. 3. (a)–(e) The solutions proposed by AEFA using one to five ellipses. (f) The six circles in *SCC* that initialise GMM-EM for $k=6$. (g) The solution of AEFA for $k=6$. (h) The solution in case that circles were selected only based on their size. (i) The association of pixels to the final solution of AEFA for $k=6$ ellipses. (j) The AIC and BIC criteria for different values of k. Captions show the estimated values of shape coverage α.

Another frequently used criterion for model selection is the Bayesian Information Criterion (BIC) [32] that is given by $BIC = 2k - 2\ln(L)$. However, BIC penalises model complexity more heavily yielding under-fitting. In our approach, AIC is preferred also because of its theoretical and performance advantages over BIC [31,33]. In the experiments we performed, both criteria have been tested, resulting in similar results for most of the cases.

6. Augmentative Ellipse Fitting Algorithm

The proposed Augmentative Ellipse Fitting Algorithm (AEFA) gradually increases the number of considered ellipses, starting from a single one. The input to AEFA is the binary image I representing the shape to be modelled by ellipses. Its goal is to compute a set of ellipses E^* with the lowest AIC (AIC^*), as this is defined in Eq. (6). The main steps of the AEFA algorithm are summarised in Algorithm 1.

Algorithm 1. Augmentative Ellipse Fitting (AEFA).

input: Binary image I.
output: Set of ellipses E^* with the lowest AIC^*

1 $[S,R] = ComputeShapeSkeleton(I)$
2 $C = ComputeShapeComplexity(S,R)$
3 $CC = InitializeEllipseHypotheses(S,R)$
4 $k=1$
5 $AIC^* = \infty$
6 **repeat**
7 $SCC = SelectHypotheses(k, CC)$
8 $E = GMM{-}EM(I, SCC, k)$
9 $AIC = ComputeAIC(I,E,C)$
10 $AIC_{min} = C \cdot \log(1-0.99) + 2 \cdot k$
11 **if** $AIC < AIC^*$ **then**
12 $AIC^* = AIC$
13 $E^* = E$
14 **end**
15 $k = k+1$
16 **until** $AIC^* < AIC_{min} \| k > |CC|$

Initially, the thinning skeleton points S of the 2D shape together with their associated radii R are computed (line 1). Then, the shape complexity C is computed as defined in Eq. (5) (line 2). The skeleton points are stored in S in descending order of R, that is $R(1) \geq R(2) \geq \cdots \geq R(|S|)$. In line 3, based on S and R, we compute a set of candidate circles CC that will constitute initial hypotheses of ellipses representing the 2D shape. The cardinality $|CC|$ of CC provides an upper bound on the number of ellipses that will eventually be used to model the 2D shape. The details of this algorithmic step are presented in Section 6.1.

Then, the following loop is executed (lines 6–16). In the k-iteration of the loop, first (line 7), k circles are selected from CC and stored in SCC (Section 6.3). These k circles initialise a hard Expectation-Maximisation algorithm (line 8) that is responsible for estimating the parameters of k ellipses that best (in the MAX-α problem formulation sense) represent the given 2D shape (Section 6.2). The AIC of the achieved solution is then computed based on Eq. (6). Also, AIC_{min} is computed (line 10) as a lower bound of the achievable AIC given the model cardinality k and the shape complexity C. The best fit ellipses set (E^*) and the lowest AIC (AIC^*) are updated (lines 11–14). The whole process terminates when the AIC of the current best fit ellipses is already very low[2] or the number of ellipses k exceeds $|CC|$ (line 16). The following sections describe in more detail the algorithmic steps of AEFA.

6.1. Initialising ellipse hypotheses

CC holds a set of non-overlapping or partially overlapping[3] circles whose centroids lie on the shape skeleton S. Their radii is defined by the corresponding radii R of skeleton points. Circles are considered for inclusion in CC in decreasing order with respect to their radius. Initially, $CC = \varnothing$. Then, in each iteration step i, the circle of centroid $K(i)$ and radius $R(i)$ is checked regarding whether it is overlapping with any of the already selected circles. If this is not the case, then this circle is added to CC. In order to reduce the size of CC and the complexity of the next steps of AEFA, circles with radius lower than 3% of the maximum radius are ignored. Thus, the initial ellipse hypotheses are non-overlapping circles whose centroids lies on the skeleton of the 2D shape and radii are determined by the distance of these centroids from the shape. This

[2] We assume that $\alpha(E) < 0.99$, since the ellipse prototype cannot exactly fit on the image grid. In order to speed up the method even lower values can be used.
[3] Partially overlapping is allowed only if the centroids lie on the end points of shape skeleton S.

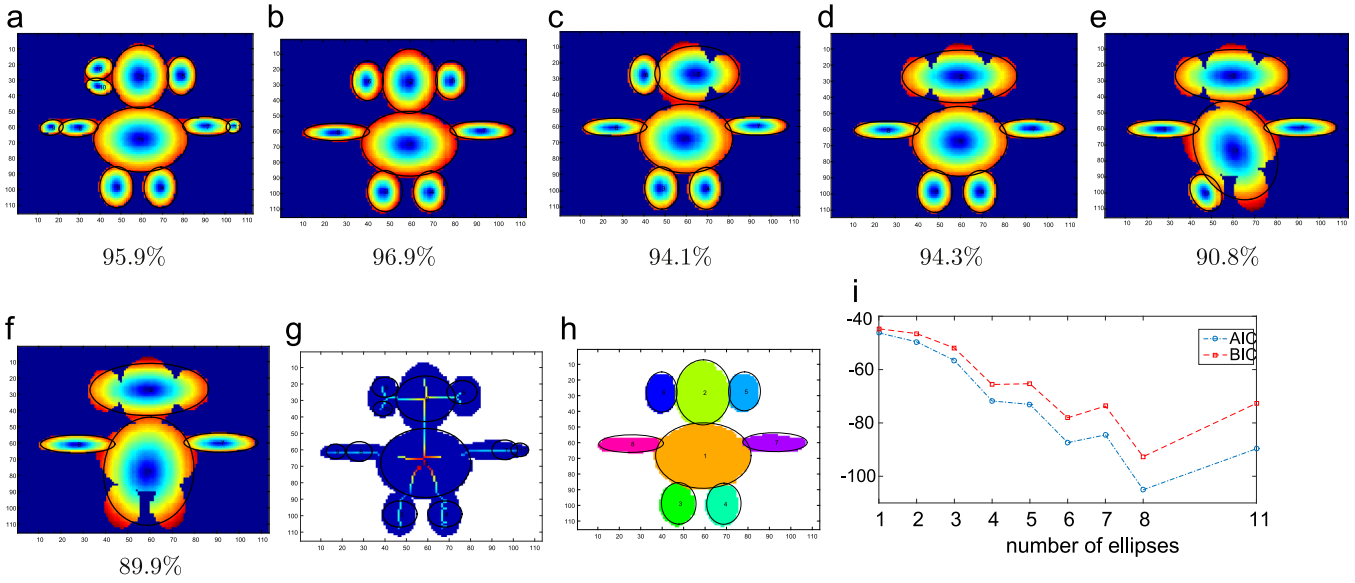

Fig. 4. (a)–(f) The intermediate solutions proposed by DEFA using 11, 8, 7, 6, 5 and 4 ellipses. Captions show the estimated values of shape coverage α. (g) The skeleton of the 2D shape. (h) The association of pixels to $k=8$ ellipses which is the final solution estimated by DEFA. (i) the AIC and BIC criteria for different values of k.

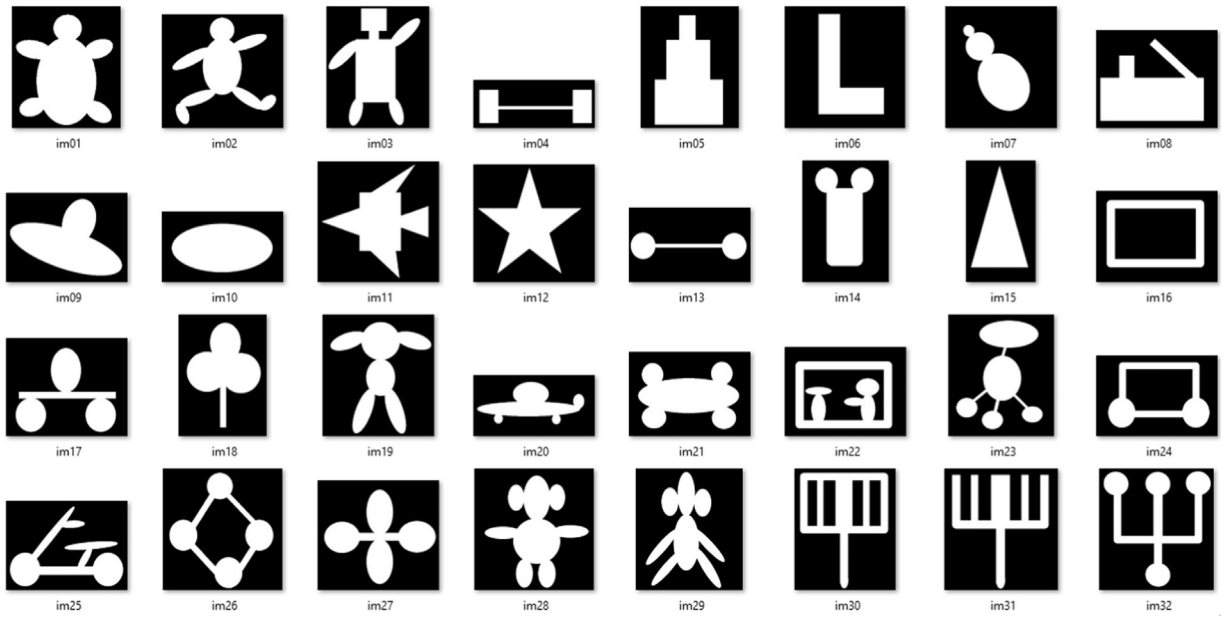

Fig. 5. The 32 shapes of the SISHA dataset.

computationally cheap method achieves a high coverage of the shape with a small number of initial ellipse hypotheses providing a upper bound for the maximum number of fitting ellipses.

6.2. Evolving ellipse hypotheses based on GMM-EM

The GMM-EM algorithm is responsible for computing the parameters of a fixed number k of ellipses E that achieve the best coverage $\alpha(E)$ of the given 2D shape. Essentially, this boils down to solving the MAX-α formulation of the modelling problem for k ellipses. This is achieved by solving a hard Expectation-Maximisation algorithm under the Equal Area constraint.

The clustering of foreground image points that is required for the hard Expectation-Maximisation algorithm is performed based on the 2D Gaussian function

$$P_i(p) = A_i \cdot e^{-(p-c_i)^T M_i (p-c_i)}, \tag{7}$$

that defines the probability of a pixel $p \in FG$ to belong to an ellipse E_i. In Eq. (7), c_i is the origin of E_i and M_i is a positive-definite 2×2 matrix representing the eccentricity and the orientation of E_i. We set each Gaussian amplitude equal to one ($A_i=1$), so that the values of $P_i(p)$ on the ellipse boundary are the same for all ellipses. This means that the probability of a point to belong to an ellipse E_i is independent of the size (area) of the ellipse and only depends on its position, orientation and eccentricity.

Due to the fact that $\forall i$, $A_i=1$, the natural logarithm of $P_i(p)$ is proportional to the ratio of distances $F(p,E_i)$ between the c_i and the points p, p', respectively, where p' is the intersection of the line

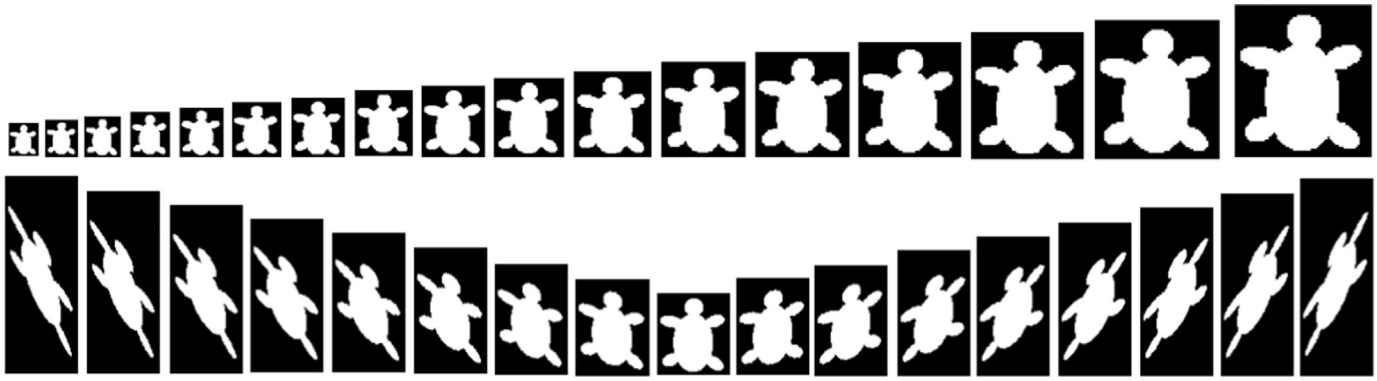

Fig. 6. The 17 images of SISHA-SCALE (first row) and SISHA-SHEAR (second row) datasets for the first shape of the SISHA dataset.

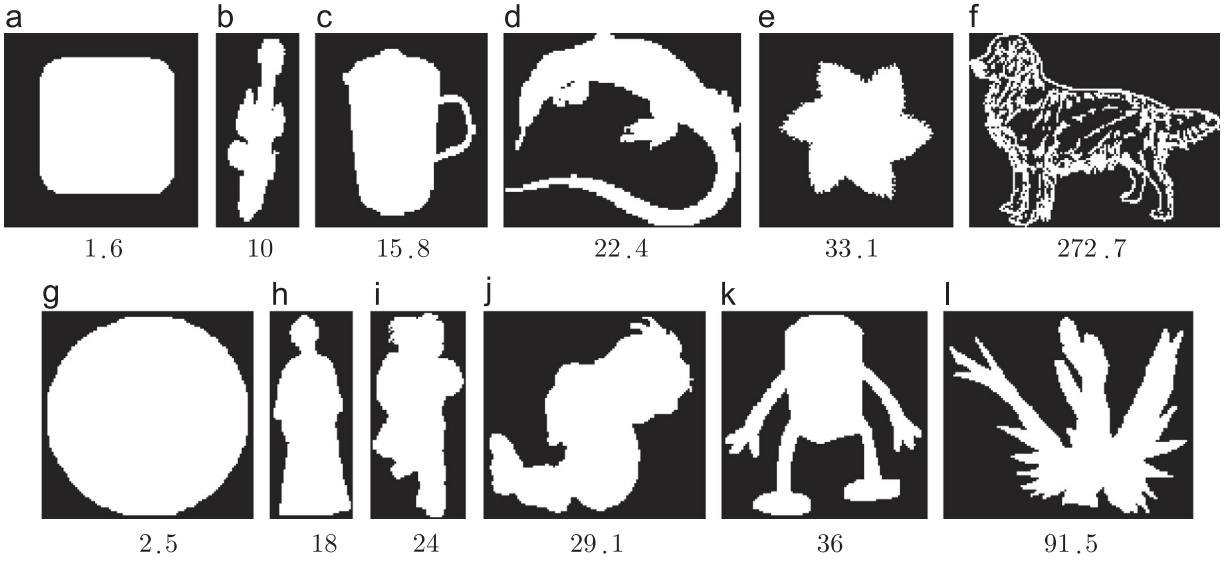

Fig. 7. Results on estimating shape complexity. Rows: MPEG-7, LEMS datasets. Columns: increasing estimated shape complexity (from 0% to 100%) in the corresponding dataset, with a step of 20%. Captions show the value of the estimated shape complexity: (a) 1.6; (b) 10; (c) 15.8; (d) 22.4; (e) 33.1; (f) 272.7; (g) 2.5; (h) 18; (i) 24; (j) 29.1; (k) 36; and (l) 91.5.

pc_i and the E_i boundary:

$$F(p,E_i)=\frac{\|p-c_i\|}{\|p'-c_i\|}. \tag{8}$$

It holds that the higher the $F(p,E_i)$, the lower the probability of p to belong to E_i. If $F(p,E_i)=1$, then p lies on the boundary of E_i. Thus, the foreground points ($p\in FG$) are clustered to k groups $G_i, i\in\{1,\ldots,k\}$ according to the following rule:

$$G_i=\{p\in FG: F(p,E_i)\le F(p,E_j)\ \forall j\neq i\}. \tag{9}$$

Then, according to the hard Expectation-Maximisation algorithm, the E_i is directly updated by the second order moments of G_i under the constraint that $\pi a_i b_i=|G_i|$ where a_i and b_i denote the semi-major and semi-minor axes of E_i, respectively.

6.3. GMM-EM initialisation

One important issue to the performance of the overall method is the proper initialisation of the GMM-EM algorithm. In case of random initialisation, GMM-EM has more chances of converging to a suboptimal local minimum of the problem. Fortunately, the shape information encapsulated by the shape skeleton can be exploited towards an informed decision on the GMM-EM initialisation process.

More specifically, the goal is to guess a subset of k of the circles in CC that yields the best possible solution with a single execution of the GMM-EM. It is to be noted that the exhaustive consideration of all possible subsets of circles requires $\binom{|CC|}{k}$ executions of GMM-EM which is computationally intractable even for very small values of k.

To avoid this computational burden, we adopt the following strategy for selecting the best k circles of CC that constitute the SCC hypotheses set which initialises the GMM-EM. While in iteration k, SCC already contains $k-1$ circles selected during iteration $k-1$. Then, the goal is to select the "best" possible circle in CC so as to form a set of k circles. The circle to be introduced should be the one that maximises the expected coverage of the 2D shape. To decide this, we take into account the distance of each 2D shape point from the $k-1$ ellipses already identified at the $k-1$ step. This is defined as $F(p)=\min_{i\in\{1,\ldots,k-1\}}F(p,E_i)$. Intuitively, a good circle to introduce is the one that covers the pixels that are further away from the already identified ellipses. To achieve this, from the set of circles CC-SCC, we select the one under which the sum of $F(p)$ is minimised.

Algorithm 2. Decremental Ellipse Fitting (DEFA).

input: Binary image I.
output: Set of ellipses E^* with the lowest AIC^*

1 $[S,R]=ComputeShapeSkeleton(I)$
2 $C=ComputeShapeComplexity(S,R)$
3 $CC=InitializeEllipseHypotheses(S,R)$

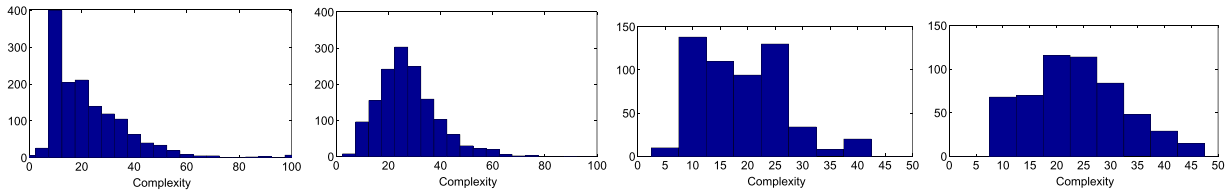

Fig. 8. Histograms of shape complexity C for the four employed datasets: MPEG-7 (top-left), LEMS (top-right), SISHA-SCALE (bottom-left) and SISHA-SHEAR (bottom-right).

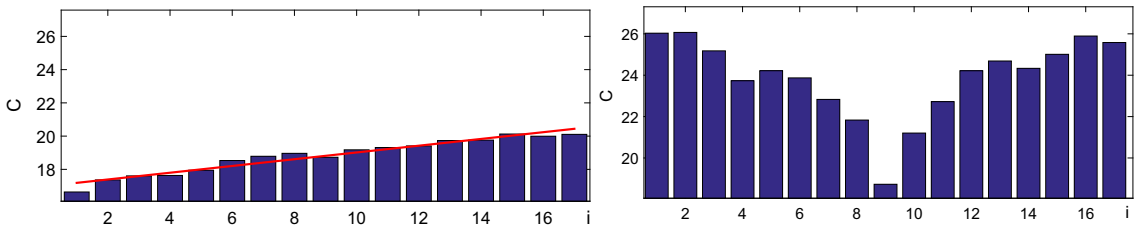

Fig. 9. The average C for each of the transformations across all shapes on SISHA-SCALE (left) and SISHA-SHEAR (right).

Table 1
The average AIC of AEFA, DEFA and EMAR, the $Pr(AEFA/AIC)$, $Pr(DEFA/AIC)$ and the $Pr(EMAR/AIC)$ computed on all images of the four used datasets.

Dataset	AIC(AEFA)	AIC(DEFA)	AIC(EMAR)	Pr(AEFA/AIC) (%)	Pr(DEFA/AIC) (%)	Pr(EMAR/AIC) (%)
MPEG-7	−44.74	**−45.13**	−39.95	29.57	**43.79**	9.71
LEMS	−63.50	**−63.93**	−54.72	34.88	**53.21**	1.37
SISHA-SCALE	−43.60	**−44.60**	−35.57	31.07	**41.54**	7.72
SISHA-SHEAR	−53.42	**−54.33**	−46.06	36.95	**43.93**	10.48

4 $k = |CC|$
5 $E = GMM\text{-}EM(I, CC, k)$
6 $AIC^* = ComputeAIC(I, E, C)$
7 **repeat**
8 **for** $e_1, e_2 \in E$ **do**
9 $\Delta AIC(e_1, e_2) = \infty$
10 **if** e_1 *and* e_2 *are adjacent* **then**
11 $\Delta AIC(e_1, e_2) = mergeGain(I, C, E, e_1, e_2)$
12 **end**
13 **end**
14 $[e_1^*, e_2^*] = \text{argmin}_{e_1, e_2 \in E}(\Delta AIC(e_1, e_2))$
15 **repeat**
16 $E = merge(E, e_1^*, e_2^*)$
17 $\Delta AIC(e_1, e) = \infty, \ \forall e \in E$
18 $\Delta AIC(e_2, e) = \infty, \ \forall e \in E$
19 $[e_1^*, e_2^*] = \text{argmin}_{e_1, e_2 \in E}(\Delta AIC(e_1, e_2))$
20 $k = k - 1$
21 **until** $\Delta AIC(e_1^*, e_2^*) > 0$
22 $E = GMM{-}EM(I, E, k)$
23 $AIC = ComputeAIC(I, E, C)$
24 **if** $AIC < AIC^*$ **then**
25 $AIC^* = AIC$
26 $E^* = E$
27 **end**
28 **until** $k = = 1$

The importance of this process is illustrated in the example of Fig. 3. The colourmap of Fig. 3(a)–(e), (g) and (h) corresponds to $F(p)$, that is, the distance of foreground pixels from the ellipses introduced so far (cold and warm colours denote small and large distances, respectively). In the evolution of the proposed method (Fig. 3(a)–(e) and (g)) it is shown that the next circle selected in SCC is located in an uncovered region so as to maximise the expected coverage. In Fig. 3(g) and (i) the final solution and the clustering of pixels proposed by AEFA are illustrated, respectively. This solution achieves $\alpha = 94.4\%$. Fig. 3(h) presents the corresponding result using the six largest circles of CC. This result is clearly suboptimal yielding $\alpha = 89.4\%$. Fig. 3(j) shows the AIC and BIC criteria for different k. A clear minimum at $k = 6$ is identified. It can also be verified that the minimisation of AIC or BIC yields the same solution.

7. Decremental Ellipse Fitting Algorithm

The Decremental Ellipse Fitting Algorithm (DEFA) decreases the number of considered ellipses starting from all ($|CC|$) of them. Similarly to AEFA, the input to DEFA is the binary image I representing the shape to be modelled by ellipses. Its goal is to compute a set of ellipses E^* with the lowest AIC (AIC^*), as this is defined in Eq. (6). The main steps of the DEFA are summarised in Algorithm 2.

In the first three lines of DEFA, that are the same as AEFA, the computation of a set of candidate circles CC is performed. These constitute the initial ellipse hypotheses representing the 2D shape. The cardinality $|CC|$ of CC provides an upper bound on the number of ellipses that can be used to initiate DEFA (lines 4–6).

Then, the main loop of DEFA is executed (lines 7–28). In each iteration, a pair of adjacent ellipses is selected in order to be merged and to possibly decrease the AIC. Since there is no lower bound of AIC as the number of ellipses decreases, this process continues until the set of considered ellipses is reduced to just one.

First (lines 8–13), the difference of the AIC that corresponds to the merging of adjacent ellipses $e_1, e_2 \in E$ and the current AIC is stored in ΔAIC matrix. It holds that the lower $\Delta AIC(e_1, e_2)$, the more suitable the pair for merging. Therefore, the multiple merging process starts from the pair of adjacent ellipses $e_1, e_2 \in E$ with

Table 2
The average coverage α of AEFA, DEFA and EMAR, the $Pr(AEFA/\alpha)$, $Pr(DEFA/\alpha)$ and the $Pr(EMAR/\alpha)$ computed on all images of the four used datasets.

Dataset	α(AEFA) (%)	α(DEFA) (%)	α(EMAR) (%)	Pr(AEFA/α) (%)	Pr(DEFA/α) (%)	Pr(EMAR/α) (%)
MPEG-7	89.89	**90.10**	86.58	31.86	**36.43**	7.43
LEMS	92.97	**93.05**	88.91	34.40	**40.08**	0.82
SISHA-SCALE	93.57	**93.71**	90.64	21.88	**33.09**	8.82
SISHA-SHEAR	92.89	**93.00**	90.31	31.80	**39.71**	9.56

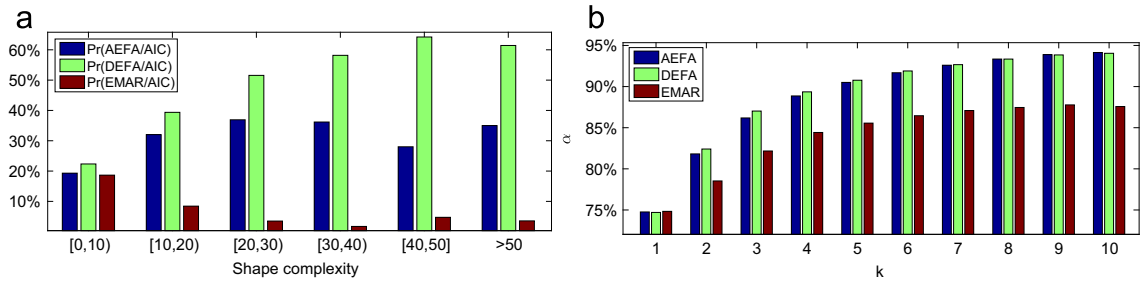

Fig. 10. (a) $Pr(AEFA/AIC)$, $Pr(DEFA/AIC)$ and $Pr(EMAR/AIC)$ for different ranges of complexity computed over all 3950 images. (b) The average values of coverage α as a function of the number k of ellipses over all 3950 images.

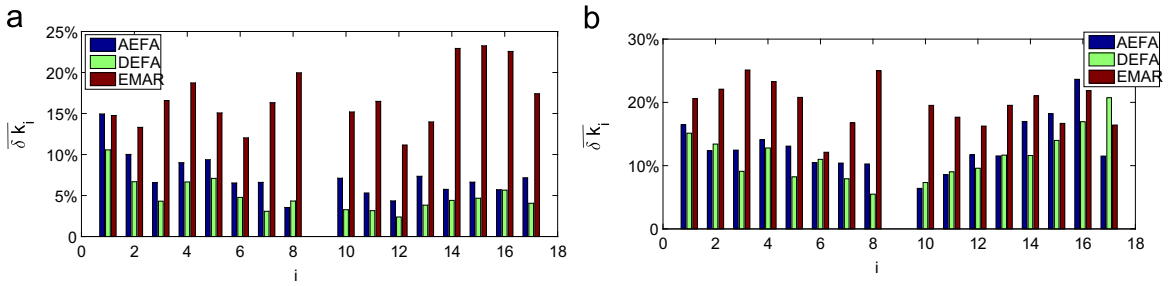

Fig. 11. $\overline{\delta k_i}$ on the SISHA-SCALE (top) and SISHA-SHEAR (bottom) datasets.

the minimum ΔAIC and stops when ΔAIC is positive (line 21). Ellipse merging (line 16) merges e_1, e_2 to one ellipse by running the GMM-EM on the union of e_1, e_2, without affecting the rest of the ellipses in E. In line 22, the resulting (reduced) ellipse set E initialises the hard Expectation-Maximisation algorithm (Section 6.2) that updates the ellipse parameters. This process converges quite faster than the corresponding step in AEFA. This is because for a given k, AEFA starts from circles and evolve them to ellipses, while DEFA starts from the ellipses as they have already evolved to fit the 2D shape, so less GMM-EM iterations are required until convergence.

The adopted scheme may result in multiple ellipse merges at a time. This multiple ellipse merging strategy does not affect significantly the quality of the obtained results if each ellipse $e \in E$ participates in at most once pairs to be merged. This is implemented by lines (lines 17 and 18). According to our experiments in 3950 images (see Section 9), the use of multiple merges decreases the total computational cost about 30%, while the average AIC was only 0.086% lower than the corresponding AIC without the multiple merging step. In addition, due to the fast convergence of DEFA, DEFA is about tree times faster on average than AEFA.

Fig. 4 illustrates an example run of DEFA. The colourmap of Fig. 4(a)–(f) corresponds to $F(p)$ (see Section 6.2), that is, the distance of foreground pixels from the ellipses introduced so far (cold and warm colours denote small and large distances, respectively). As shown in Fig. 4(a)–(f), the ellipses that are located in the most over-segmented regions are selected so as to maximise the expected coverage. Fig. 4(b) and (h) shows the final solution and the clustering of pixels (α=96.9%), respectively. Fig. 4(i) shows the AIC and BIC criteria for different values of k. A clear minimum at k=8 and a local minimum at k=6 are identified, while the values of k=9 and k=10 are not computed due to the multiple merging step that merges three pairs of ellipses in the first iteration of the main loop. As it can be verified, the minimisation of AIC or BIC yields the same solution.

8. Computational complexity

The computational complexity of AEFA and DEFA depends on the number of foreground pixels ($n = |FG|$), the number of candidate circles ($c = |CC|$) and the number of iterations that the main loop of AEFA is executed k_{max} with $k_{max} \leq c$. It holds that c and k_{max} are independent of image size (n) and only depend on the shape complexity.

The computation of shape complexity and of the shape skeleton is $O(n)$. The computational complexity of the process of forming ellipse hypotheses is also $O(n)$. Then, the main loop of AEFA is executed k_{max} times. In each k-iteration of this loop, GMM-EM and hypothesis selection are executed. Similarly with k-means, the complexity of GMM-EM is $O(kn)$. The complexity of *SelectHypotheses* is also $O(kn)$. Conclusively, the most computationally intensive part of the proposed method is the main loop of AEFA with $O(k \cdot n)$ complexity. The total computational cost of AEFA is $O(k_{max}cn) = O(c^2 n)$.

In DEFA, the main loop is executed c times. In each k-iteration of this loop, GMM-EM is executed with $O((c-k)n)$ computational cost. The computational complexity of the selection of multiple merging set (lines 8–21) is ($O(c^2)$). Therefore, the computational complexity of DEFA is $O(c^2 n + c^3) = O(c^2 n)$.

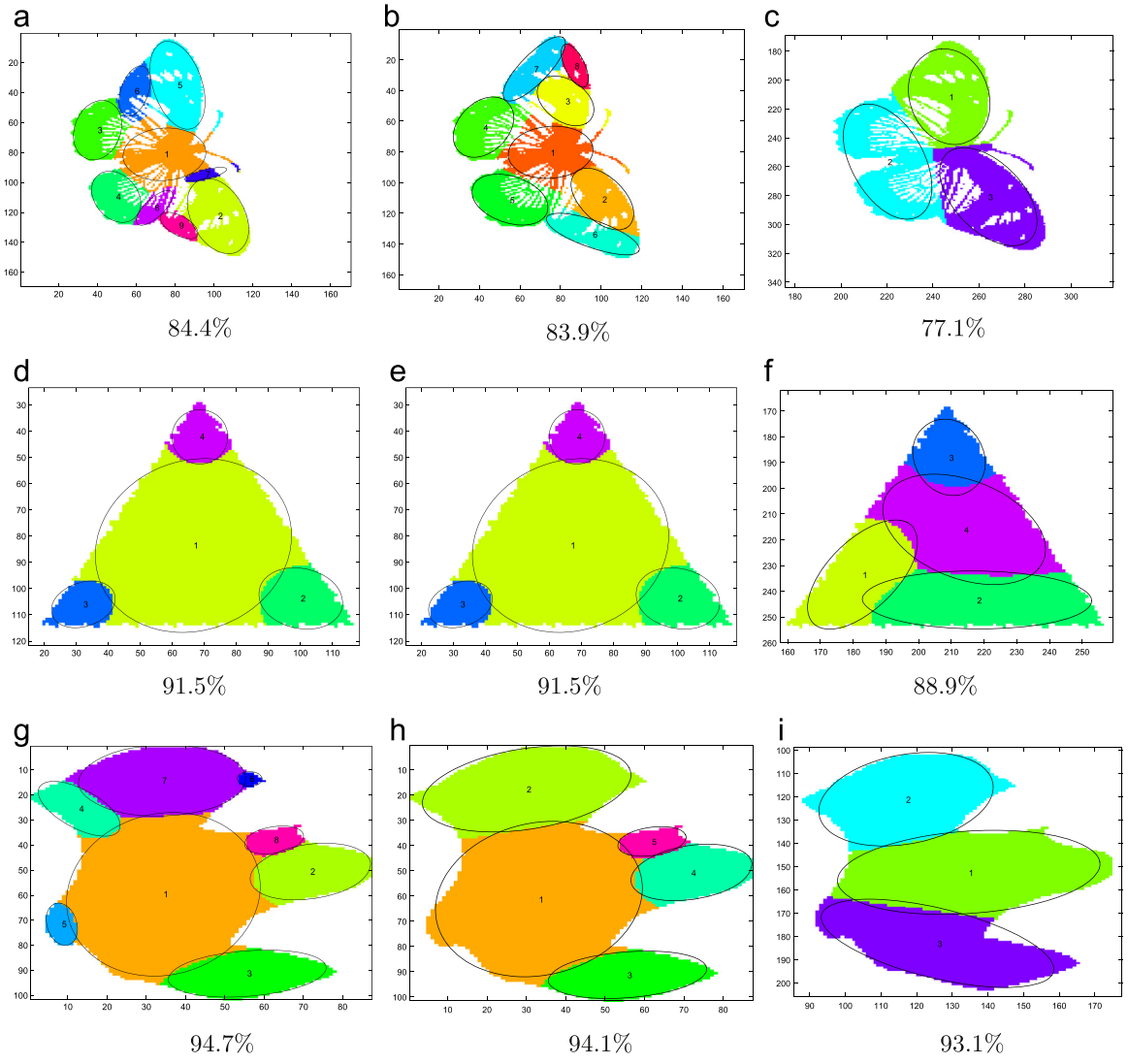

Fig. 12. Results of AEFA (first column), DEFA (second column) and EMAR (third column) on selected shapes from MPEG7 dataset. Rows correspond to the 10%, 50%, and 90% percentile of the value α of AEFA. Captions show the estimated values of shape coverage α: (a) 84.4%; (b) 83.9%; (c) 77.1%; (d) 91.5%; (e) 91.5%; (f) 88.9%; (g) 94.7%; (h) 94.1%; and (i) 93.1%.

9. Experimental evaluation

The evaluation of the proposed approaches was based on two standard databases as well as on the SISHA (SImple SHApe) dataset that was developed for the purposes of this research. More specifically, we employ:

- MPEG-7 [34], which consists of 1400 binary shapes organised in 70 categories with 20 shapes per category. This dataset has been extensively used on shape retrieval tasks [35].
- A subset of LEMS [36], i.e., 1462 shapes that come from the following categories of the original database: Buildings, Containers, Fish, Fruit and vegetables, Misc Animal, People, Robots, Toddlers and Turtles.

SISHA contains the 32 shapes shown in Fig. 5 and has been employed to measure the performance of the proposed methods under scaling and shearing transformations. For each 2D shape of SISHA, 16 scale and 16 shear transformations are applied resulting in the SISHA-SCALE and the SISHA-SHEAR datasets. In those, we include the original images, so each of them consists of 544 ($32 \times 16+32$) binary shapes.

The 16 scale transformations change both image dimensions by a factor of $sc_i = 1.1^{i-9}$, $i \in \{1,\ldots,17\}-\{9\}$. Thus, it holds that the first 8 transforms correspond to a shrinking of the original image, while the last 8 correspond to an enlargement. The ratio of the number of pixels between the smallest image ($sc_1 = 0.466$) and the largest image ($sc_{17} = 2.14$) is high enough ($sc_{17}/sc_1)^2 = 21.11$ to represent a very broad range of scales. For the 16 shear transformations, we employed only vertical shear, since the effects of horizontal shear are the same when applied on a rotated image. The vertical shear transform is given by $\begin{pmatrix} 1 & 0 \\ sh_i & 1 \end{pmatrix}$, where $sh_i = 0.2 \cdot (i-9)$, $i \in \{1,\ldots,17\}-\{9\}$ controls the shearing effect. Fig. 6 illustrates SISHA-SCALE and SISHA-SHEAR for the first image of the SISHA dataset.

The employed datasets together with the results obtained with the proposed methods are available online.[4]

Estimation of shape complexity: We computed the shape complexity C for all shapes in the MPEG-7 and LEMS datasets. In Fig. 7, each row corresponds to each of these datasets. Columns correspond to 0%, 20%, 40%, 60%, 80% and 100% percentile of shape

[4] https://sites.google.com/site/costaspanagiotakis/research/EFA

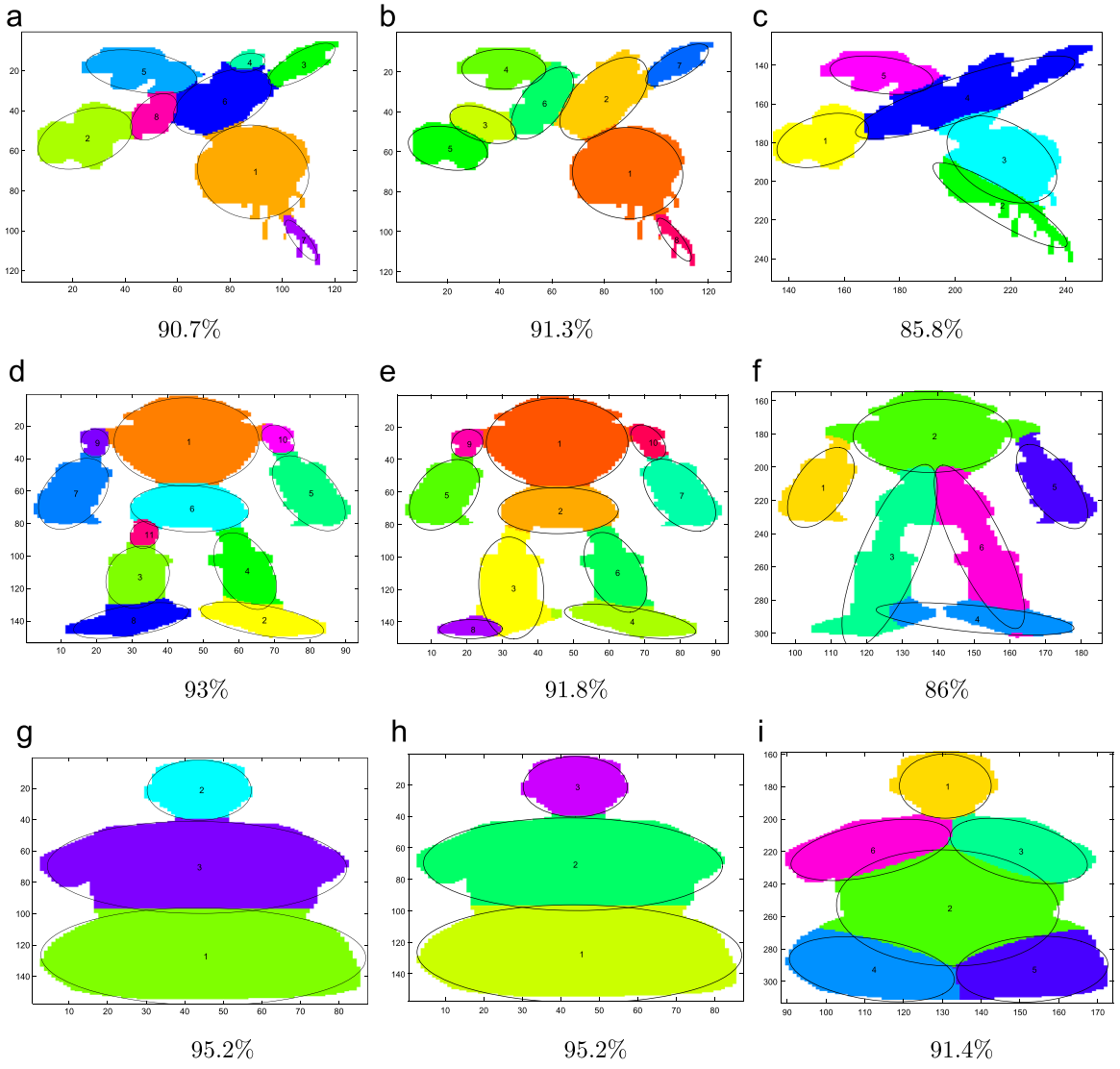

Fig. 13. Results of AEFA (first column), DEFA (second column) and EMAR (third column) on selected shapes from LEMS dataset. Rows correspond to the 10%, 50%, and 90% percentile of the value α of AEFA. Captions show the estimated values of shape coverage α: (a) 90.7%; (b) 91.3%; (c) 85.8%; (d) 93%; (e) 91.8%; (f) 86%; (g) 95.2%; (h) 95.2%; and (i) 91.4%.

complexity. That is, shapes in each row should vary from the simplest to the most complex in the corresponding dataset, according to the adopted shape complexity measure. It can be verified that this is actually the case.

We also computed the histograms of 2D shape complexity C for the four datasets (see Fig. 8). The most complex shapes come from MPEG-7 datasets. However, the average shape complexity of LEMS is higher than that of MPEG-7. In addition, SISHA-SHEAR appears more complex than SISHA-SCALE. This is attributed to the fact that the shearing transformation turns rectangles into slanted parallelograms.

This is also quantified by the results of the following experiment. For the SISHA-SCALE and SISHA-SHEAR, we computed the variance of shape complexity for each of the shapes across all transformations. Fig. 9 illustrates the average shape complexity for each of the transformations across all shapes on SISHA-SCALE (left) and SISHA-SHEAR (right). For SISHA-SCALE, the mean variance for all shapes is 3.7, while for SISHA-SHEAR this is 15.5. Thus, shape complexity does not vary significantly with scale. Moreover, it can be well approximated by a straight line. Since the x-axis of this figure corresponds to the logarithm of image scale, this experiment verifies the theoretical result of Section 4 that the shape complexity increases logarithmically with scale. With respect to shear, complexity increases as we move further from the original shape (middle) and shear transformations become stronger.

Comparison with EMAR: We compared the proposed AEFA and DEFA methods with the method presented in [21] that employs the EM algorithm [21] under random initialisation of GMMs (EMAR). We compare with it because, to the best of our knowledge, this is the closest method we can compare with. As originally proposed, EMAR operates for a fixed number of ellipses k. To perform a fair and interesting comparison, similarly to AEFA, we let EMAR successively increase the number of considered ellipses, yielding the solution that maximises the AIC criterion as defined in Section 4. The size of the ellipses is determined by the Equal Area constraint.

We compared the performance of AEFA, DEFA and EMAR on the basis of α and on the AIC that try to minimise. For a given dataset, we also compute $Pr(m/AIC)$, where m is a method in $\{AEFA, DEFA, EMAR\}$. $Pr(m/AIC)$ is used to measure performance the method m that is defined as the balance between model complexity and shape coverage. $Pr(m/AIC)$ is quantified as the percentage of images of the datasets where the method m clearly

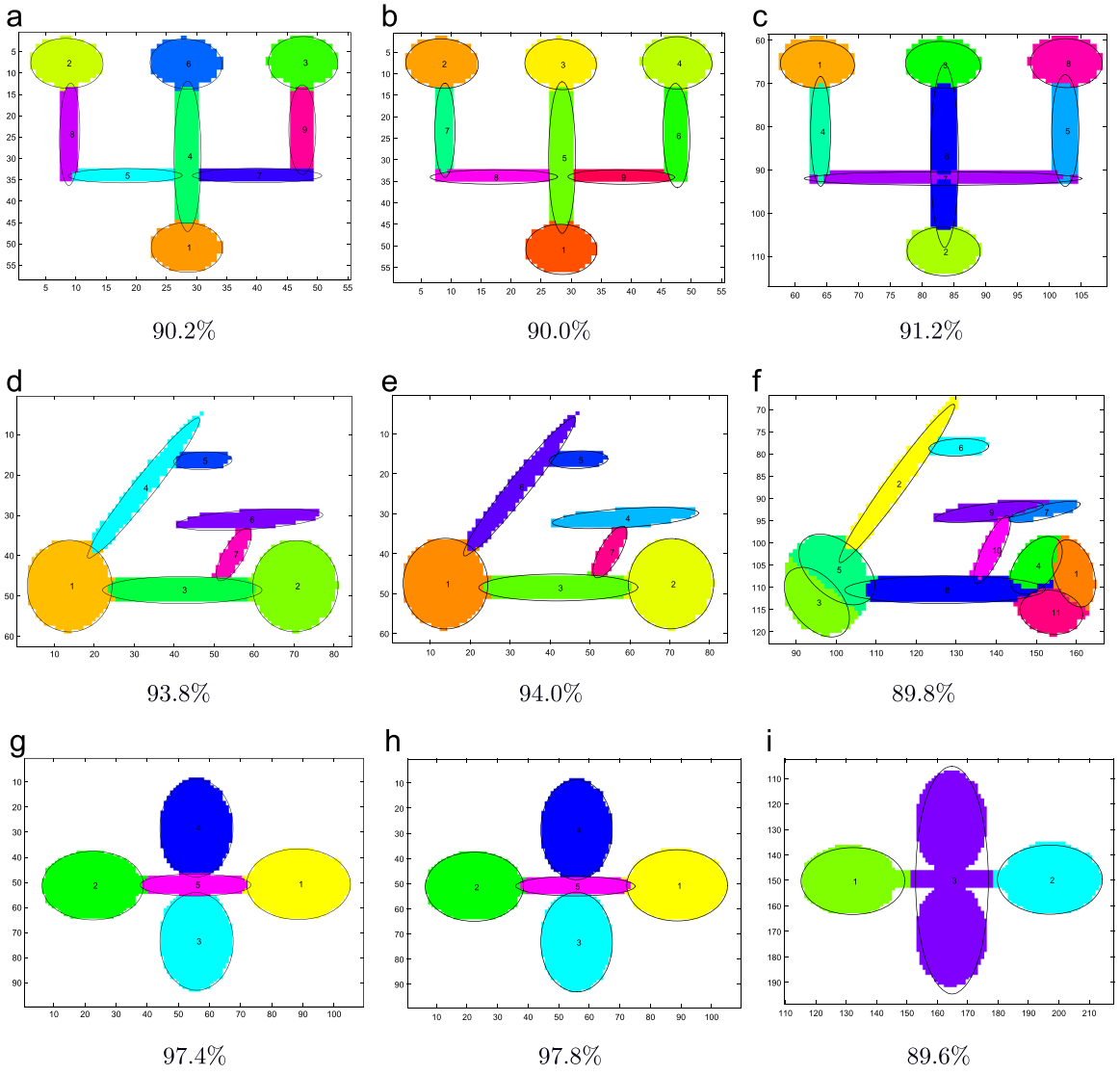

Fig. 14. Results of AEFA (first column), DEFA (second column) and EMAR (third column) on selected shapes from SISHA-SCALE dataset. Rows correspond to the 10%, 50%, and 90% percentile of the value α of AEFA. Captions show the estimated values of shape coverage α: (a) 90.2%; (b) 90.0%; (c) 91.2%; (d) 93.8%; (e) 94.0%; (f) 89.8%; (g) 97.4%; (h) 97.8%; and (i) 89.6%.

outperforms the two others under the AIC.This also means that the value $100\%-\sum_m Pr(m/AIC)$ gives the percentage of images for which there is no clear winner method.

Table 1 shows the average AIC of AEFA, DEFA and EMAR, the $Pr(AEFA/AIC)$, $Pr(DEFA/AIC)$ and the $Pr(EMAR/AIC)$ computed in the four employed datasets. The corresponding statistics for the shape coverage α are presented in Table 2. The results with respect to coverage α and AIC are similar. AEFA and DEFA clearly outperform EMAR under any dataset and metric. DEFA slightly outperforms AEFA under any dataset and metric. When AIC is taken into account, DEFA outperforms AEFA and EMAR in 47% of the 3950 shapes. When α is taken into account, DEFA outperforms AEFA and EMAR in 37% of the 3950 shapes.

Fig. 10(a) illustrates the $Pr(AEFA/AIC)$, $Pr(DEFA/AIC)$ and the $Pr(EMAR/AIC)$ for different intervals of complexity computed over the 3950 images of all datasets. The plots show that the higher the complexity of a given shape, the more AEFA outperforms EMAR and the more DEFA outperforms AEFA and EMAR. This is due to the fact that as the shape complexity C increases, more ellipses are required to represent the shape, therefore, the proposed framework and especially DEFA, that starts from high number of ellipses, perform much better compared to the random initialisation of EMAR.

Fig. 10 (b) depicts the average values of α as a function of number of ellipses (k) over all 3950 images. It holds that when $k=1$ all the methods perform similarly. The higher the number of ellipses k, the larger the difference between the performances of the proposed methods and EMAR, since the initial ellipses of AEFA and the merging pairs of DEFA are selected so that the expected coverage is maximised.

Robustness to transformations: The robustness and stability of the AEFA, DEFA and EMAR on scaling and shearing transformations are measured on SISHA-SCALE and SISHA-SHEAR datasets. Let $k_{i,j}$ denote the number of ellipses, that an algorithm yields for the shape $j \in \{1,\ldots,32\}$ of SISHA-SCALE dataset under the factor sc_i. The following relative change ($\delta k_{i,j}$) between the $k_{i,j}$ and the $k_{i,9}$ measures the effect of the scaling sc_i on the model selection.

$$\delta k_{i,j} = \frac{|k_{i,j} - k_{i,9}|}{k_{i,9}}. \tag{10}$$

The same metric is applied to SISHA-SHEAR, where the effect of sh_i is measured. Fig. 11 illustrates the average value of the relative

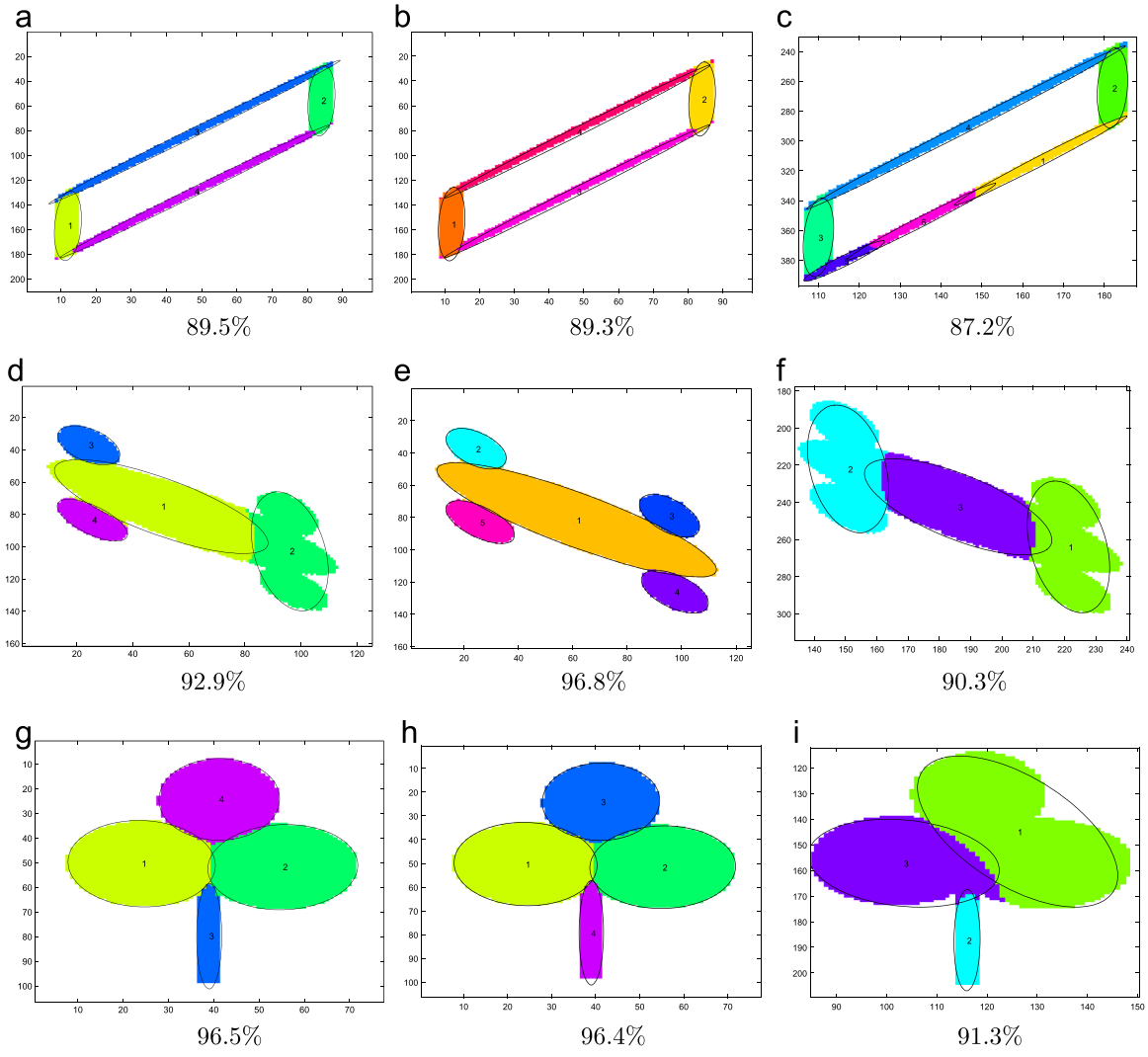

Fig. 15. Results of AEFA (first column), DEFA (second column) and EMAR (third column) on selected shapes from SISHA-SHEAR dataset. Rows correspond to the 10%, 50%, and 90% percentile of the value α of AEFA. Captions show the estimated values of α: (a) 89.5%; (b) 89.3%; (c) 87.2%; (d) 92.9%; (e) 96.8%; (f) 90.3%; (g) 96.5%; (h) 96.4%; and (i) 91.3%.

changes $\overline{\delta k_i}$ over the 32 images of SISHA-SCALE and SISHA-SHEAR datasets.

Under any case of scale and in most cases of shear transformations, AEFA and DEFA perform better than EMAR. The average effect of scaling and shearing on AEFA is 7% and 13% on model selection, respectively. The corresponding average effect of scaling and shearing on DEFA is 5% and 11.5% on model selection, respectively. The corresponding average effect on EMAR is 17% and 20%.

Shape coverage: The goal of another experiment was to examine in more detail the behaviour of AEFA with respect to the coverage parameter α and to compare all the methods under the same instances. In each of the four datasets, we sorted the 2D shapes in increasing order of the coverage α that was achieved by AEFA. We then identified the shape at the $x\%$ percentile of the sorted list with $x \in \{10\%, 50\%, 90\%\}$. This results in three images of increasing α per dataset. In Figs. 12–15 we show these three samples (rows) for the AEFA (first column) and the corresponding results for DEFA (second column) and EMAR (third column) under the four datasets. One expects complicated shapes consisting of parts that are not accurately modelled by ellipses to have lower α and to appear in the top rows. Less complicated shapes with mostly elliptical parts should have higher α and appear in the bottom rows. This is indeed the case. Still, coverage is quite high for all shapes. The results proposed by AEFA and DEFA are quite similar and usually clearly outperform EMAR. From a qualitative point of view, the solutions proposed by AEFA and DEFA agree with human intuition.

Robustness to noise: This experiment studies the behaviour of the proposed method with respect to noise. Various noise models have been proposed in the literature [37]. In this work, two datasets (SISHA-NOISE1 and SISHA-NOISE2) have been employed to measure the performance of the proposed methods under two different types of noise.

Each dataset contains the 32 original images of the SISHA dataset and five additional versions of each image, contaminated with a different noise level. So, each of the SISHA-NOISE1 and SISHA-NOISE2 datasets consists of 192 $(32+5 \cdot 32)$ images.

The employed noise process operates as follows. In an image, we randomly pick a foreground pixel. We define a circular neighbourhood of radius r around this point and, within this neighbourhood, we flip the label (foreground/background) of each pixel with some probability ($p=0{,}80$). This process is repeated for several randomly selected points and it terminates when the percentage of label changes is equal to the given noise level (e.g. 2%, 4%,…, 10%). The radius of each circular neighbourhood is uniformly selected in the interval $(0, R_{max})$, where R_{max} is the 5% and 20% of the square root of the object area on SISHA-NOISE1 and

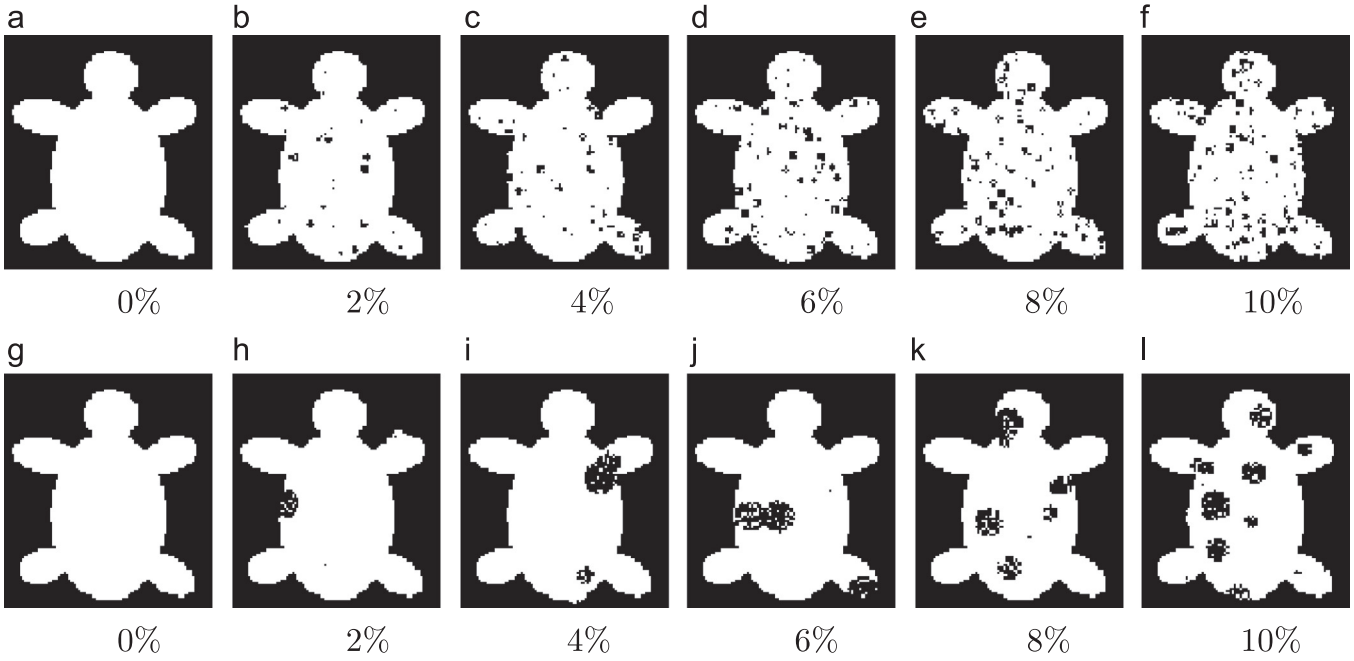

Fig. 16. The first shape of the SISHA dataset contaminated with different noise levels in the SISHA-NOISE1 (first row) and SISHA-NOISE2 (second row) datasets. Captions show the noise level: (a) 0%; (b) 2%; (c) 4%; (d) 6%; (e) 8%; (f) 10%; (g) 0%; (h) 2%; (i) 4%; (j) 6%; (k) 8%; and (l) 10%.

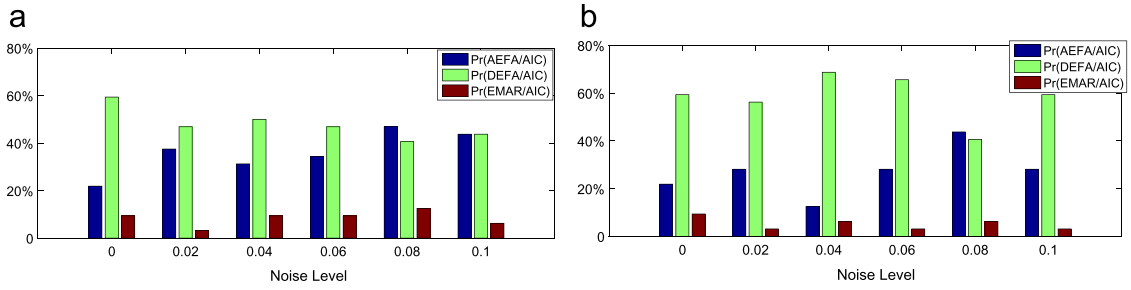

Fig. 17. *Pr*(*AEFA*/*AIC*), *Pr*(*DEFA*/*AIC*) and *Pr*(*EMAR*/*AIC*) under different noise levels of (a) SISHA-NOISE1 and (b) SISHA-NOISE2.

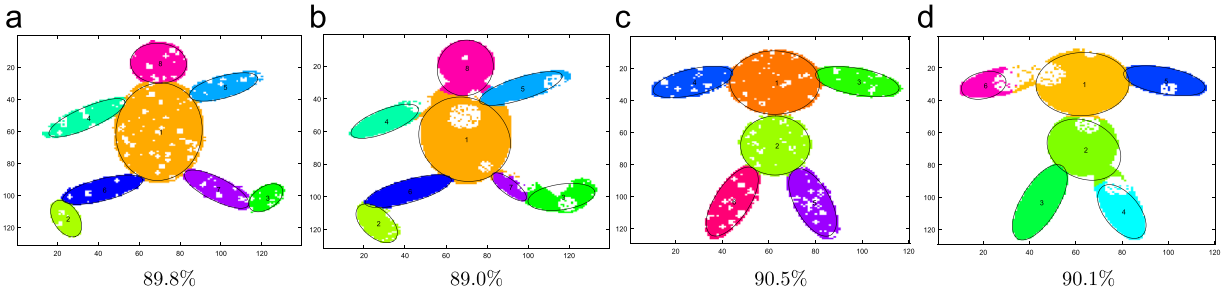

Fig. 18. Representative results from employing AEFA ((a) and (b)) and DEFA ((c) and (d)) in images of the SISHA-NOISE1 ((a) and (c)) and SISHA-NOISE2 ((b) and (d)) datasets. Captions show the estimated values of shape coverage α.

SISHA-NOISE2, respectively. Thus, on SISHA-NOISE1 we get a large number of "small spots" (e.g. similar to salt and pepper noise), while on SISHA-NOISE2 we get a small number of "large spots", that affect more significantly the shape characteristics. Fig. 16 depicts the six images of SISHA-NOISE1 (first row) and SISHA-NOISE2 (second row) datasets for the first shape of the SISHA dataset.

We compared the performance of AEFA, DEFA and EMAR by measuring the *Pr*(*AEFA*/*AIC*), *Pr*(*DEFA*/*AIC*) and *Pr*(*EMAR*/*AIC*) in the SISHA-NOISE1 and SISHA-NOISE2 datasets (see Fig. 17). The plots show that the proposed methods clearly outperform the EMAR method under any dataset and noise level. Additionally, in most cases DEFA outperforms AEFA. This is more pronounced in SISHA-NOISE2 where the employed noise model increases the complexity of the shapes. More specifically, DEFA clearly outperforms the rest methods in 48% and 58% of the images of SISHA-NOISE1 and SISHA-NOISE2, respectively. AEFA outperforms the rest methods in 36% and 27% of the images of SISHA-NOISE1 and SISHA-NOISE2, respectively. Fig. 18 shows some representative results of AEFA (left) and DEFA (right) method from SISHA-NOISE1 and SISHA-NOISE2 datasets.

Qualitative results: Shape examples that show the capabilities but also some limitations of AEFA and DEFA are shown in Figs. 19 and 20. When the shape parts are close to ellipses, then the

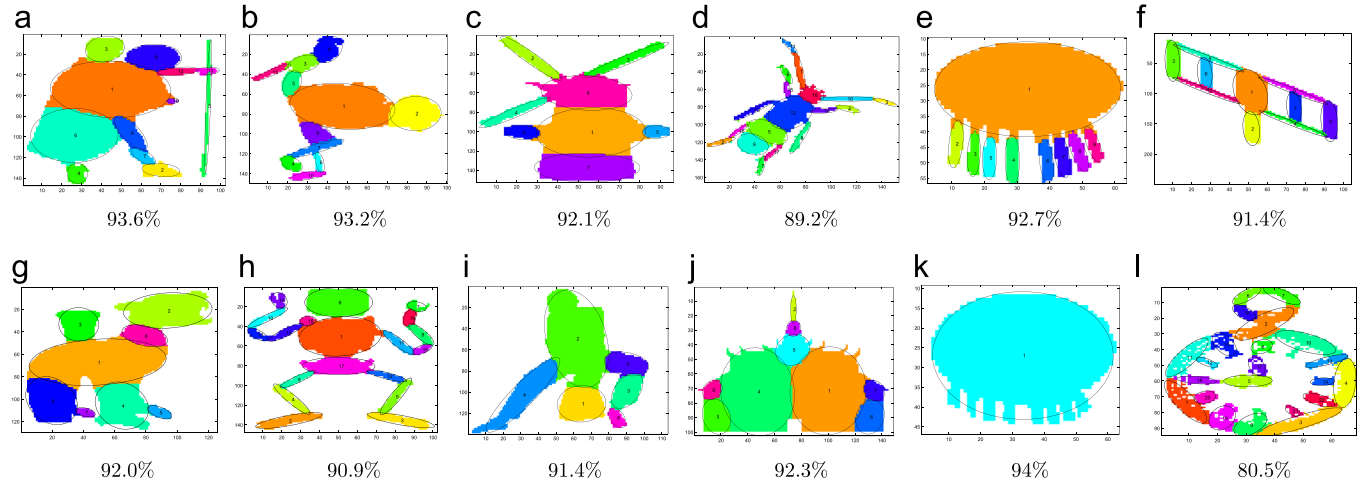

Fig. 19. Representative success (top) and failure (bottom) examples of AEFA method. Captions show the estimated values of shape coverage α: (a) 93.6%; (b) 93.2%; (c) 92.1%; (d) 89.2%; (e) 92.7%; (f) 91.4%; (g) 92.0%; (h) 90.9%; (i) 91.4%; (j) 92.3%; (k) 94%; and (l) 80.5%.

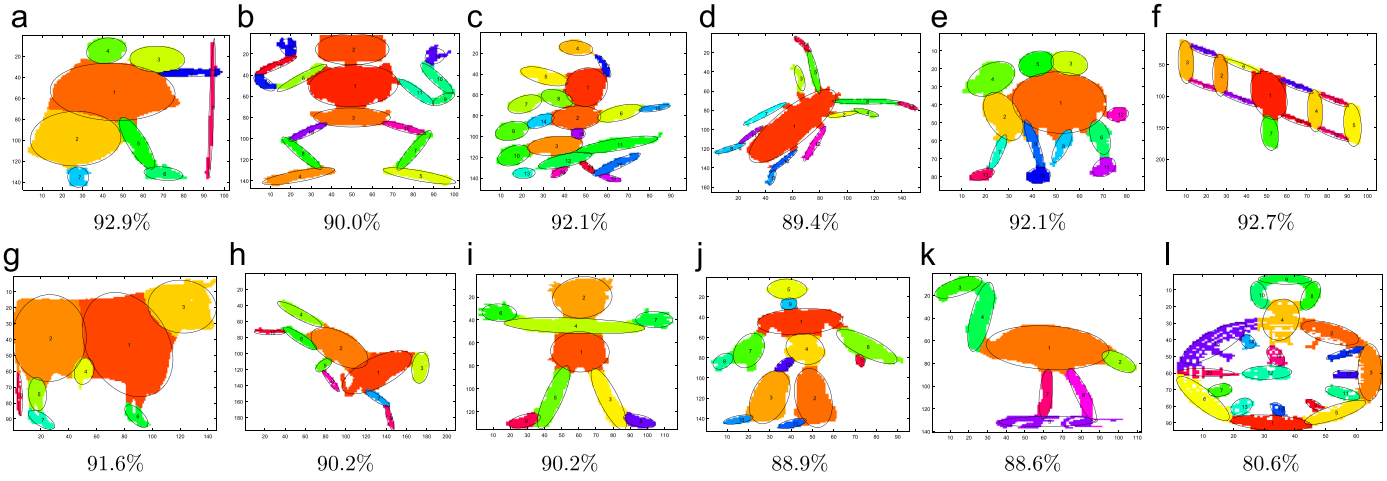

Fig. 20. Representative success (top) and failure (bottom) examples of DEFA method. Captions show the estimated values of shape coverage α: (a) 92.9%; (b) 90.0%; (c) 92.1%; (d) 89.4%; (e) 92.1%; (f) 92.7%; (g) 91.6%; (h) 90.2%; (i) 90.2%; (j) 88.9%; (k) 88.6%; and (l) 80.6%.

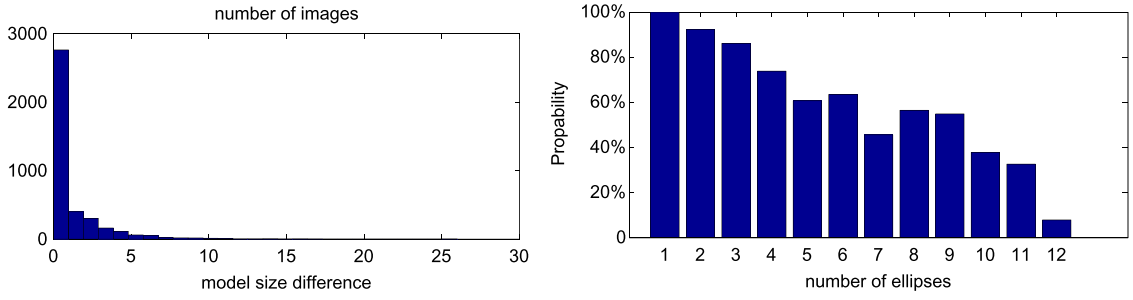

Fig. 21. Left: The histogram of the difference between the model size when AIC and BIC are preferred. Right: The probability of agreement of AIC and BIC as a function of k that AIC proposes computed over all images of all datasets.

solutions of AEFA and DEFA are near optimal even if the shape structure is complex (see Figs. 19 and 20, top rows). However, due to the minimisation of AIC, the large and small shape parts that are not well fitted by ellipses, may be over-segmented and merged, respectively (Figs. 19 and 20, bottom row). The existence of shape holes increases the number of ellipses that should be used to model the shape (see Fig. 19(f) and (l) and Fig. 20(f) and (l)). Fig. 19 (a), (d), (f), (h), and (l) and Fig. 20(f), (a), (d), (b) and (f) depict results of AEFA and DEFA on the same shapes. DEFA slightly outperforms AEFA according to the AIC criterion and provides result that are more compatible to human intuition.

AIC vs BIC: A final experiment compares model selection based on AIC versus BIC. For each image of the 3950 images of the datasets, the difference between the number of ellipses that AIC or BIC propose has been computed. This is almost always positive, since BIC penalises model size more heavily than AIC. This also means that the coverage achieved with AIC is higher than the one achieved with BIC. Fig. 21 (left) shows the histogram of this

difference. In 70% of instances, BIC and AIC agree on model selection. The difference is at most 2 in 88% of the cases, which means that usually AIC and BIC agree on model selection. Fig. 21 (right) illustrates the probability of agreement of AIC and BIC as a function of the number of ellipses that AIC proposes. When this is high, AIC and BIC agree on model selection. For more complex cases, agreement is less probable.

10. Conclusions

We proposed a parameter-free methodology for estimating automatically the number and the parameters of ellipses that approximate effectively a given 2D shape under the Equal Area constraint. Our approach capitalises on a novel definition of shape complexity that exploits the skeleton of a shape. In combination with the AIC, shape complexity defines an information-theoretic criterion for model selection that balances model complexity and the shape coverage error. For a given model, ellipse fitting is performed with a hard EM algorithm that is carefully initialised to improve shape coverage based on an augmentative (AEFA) or a decremental (DEFA) approach. Experiments on more than 4000 2D shapes assess the effectiveness of AEFA and DEFA on a variety of shapes, shape transformations, noise models and noise contamination levels. The obtained results demonstrate that DEFA outperforms AEFA especially for middle and high complexity shapes. While there is no guarantee for global optimality, AEFA and DEFA clearly outperform the competitive EMAR method in all datasets and metrics. From a qualitative point of view, the solutions proposed by AEFA and DEFA seem to agree with human intuition. Ongoing work targets the exploitation of the compact representation resulting from the proposed approach on the problem of recovering automatically the unknown kinematic structure of an unmodelled articulated object based on several, temporally ordered views of it. Additionally, we consider extensions of DEFA/AEFA towards handling shape primitives other than ellipses.

Conflict of interest

None declared.

Acknowledgements

This work was partially supported by the EU FP7-ICT-2011-9-601165 project WEARHAP.

References

[1] N.H. Trinh, B.B. Kimia, Skeleton search: category-specific object recognition and segmentation using a skeletal shape model, Int. J. Comput. Vis. 94 (2) (2011) 215–240.
[2] A. Toshev, B. Taskar, K. Daniilidis, Shape-based object detection via boundary structure segmentation, Int. J. Comput. Vis. 99 (2) (2012) 123–146.
[3] N. Kyriazis, A. Argyros, Scalable 3D tracking of multiple interacting objects, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3430–3437.
[4] C. Panagiotakis, E. Ramasso, G. Tziritas, M. Rombaut, D. Pellerin, Shape-based individual/group detection for sport videos categorization, Int. J. Pattern Recogni. Artif. Intell. 22 (06) (2008) 1187–1213.
[5] X. Ben, W. Meng, R. Yan, Dual-ellipse fitting approach for robust gait periodicity detection, Neurocomputing 79 (2012) 173–178.
[6] C. Panagiotakis, H. Papadakis, E. Grinias, N. Komodakis, P. Fragopoulou, G. Tziritas, Interactive image segmentation based on synthetic graph coordinates, Pattern Recognit. 46 (11) (2013) 2940–2952.
[7] X. Yang, X. Gao, D. Tao, X. Li, Improving level set method for fast auroral oval segmentation, IEEE Trans. Image Process. 23 (7) (2014) 2854–2865.
[8] X. Yang, X. Gao, J. Li, B. Han, A shape-initialized and intensity-adaptive level set method for auroral oval segmentation, Inf. Sci. 277 (2014) 794–807.
[9] X. Yang, X. Gao, D. Tao, X. Li, J. Li, An efficient mrf embedded level set method for image segmentation, IEEE Trans. Image Process. 24 (1) (2015) 9–21.
[10] P.F. Felzenszwalb, R.B. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part-based models, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1627–1645.
[11] A. Leonardis, A. Jaklic, F. Solina, Superquadrics for segmenting and modeling range data, IEEE Trans. Pattern Anal. Mach. Intell. 19 (11) (1997) 1289–1295.
[12] C. Goldfeder, P.K. Allen, C. Lackner, R. Pelossof, Grasp planning via decomposition trees, in: 2007 IEEE International Conference on Robotics and Automation, Roma: IEEE, 2007, pp. 4679–4684.
[13] C. Panagiotakis, G. Tziritas, Any dimension polygonal approximation based on equal errors principle, Pattern Recognit. Lett. 28 (5) (2007) 582–591.
[14] H. Akaike, A new look at the statistical model identification, IEEE Trans. Autom. Control 19 (6) (1974) 716–723.
[15] A. Fitzgibbon, M. Pilu, R.B. Fisher, Direct least square fitting of ellipses, IEEE Trans. Pattern Anal. Mach. Intell. 21 (5) (1999) 476–480.
[16] F. Mai, Y. Hung, H. Zhong, W. Sze, A hierarchical approach for fast and robust ellipse extraction, Pattern Recognit. 41 (8) (2008) 2512–2524.
[17] N. Kiryati, Y. Eldar, A.M. Bruckstein, A probabilistic Hough transform, Pattern Recognit. 24 (4) (1991) 303–316.
[18] S.-C. Zhang, Z.-Q. Liu, A robust, real-time ellipse detector, Pattern Recognit. 38 (2) (2005) 273–287.
[19] J. Yao, N. Kharma, P. Grogono, A multi-population genetic algorithm for robust and fast ellipse detection, Pattern Anal. Appl. 8 (1–2) (2005) 149–162.
[20] L. Rocha, L. Velho, P.C.P. Carvalho, Image moments-based structuring and tracking of objects, in: Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing, 2002, IEEE, 2002, pp. 99–105.
[21] R.Y. Da Xu, M. Kemp, Fitting multiple connected ellipses to an image silhouette hierarchically, IEEE Trans. Image Process. 19 (7) (2010) 1673–1682.
[22] N. Grammalidis, M.G. Strintzis, Head detection and tracking by 2-D and 3-D ellipsoid fitting, in: Proceedings of the Computer Graphics International, 2000, Geneva: IEEE, 2000, pp. 221–226.
[23] P.F. Felzenszwalb, D.P. Huttenlocher, Pictorial structures for object recognition, Int. J. Comput. Vis. 61 (1) (2005) 55–79.
[24] D.L. Page, A. Koschan, S.R. Sukumar, B. Roui-Abidi, M.A. Abidi, Shape analysis algorithm based on information theory, in: International Conference on Image Processing, vol. 1, IEEE, 2003, p. 229–232.
[25] Y. Chen, H. Sundaram, Estimating complexity of 2D shapes, in: IEEE Workshop on Multimedia Signal Processing, Shanghai: IEEE, 2005, pp. 1–4.
[26] H. Blum, et al., A transformation for extracting new descriptors of shape, in: Models for the Perception of Speech and Visual Form 19, vol. 5, 1967, pp. 362–380.
[27] X. Bai, L.J. Latecki, Path similarity skeleton graph matching, IEEE Trans. Pattern Anal. Mach. Intell. 30 (7) (2008) 1282–1292.
[28] L. Lam, S.-W. Lee, C.Y. Suen, Thinning methodologies—a comprehensive survey, IEEE Trans. Pattern Anal. Mach. Intell. 14 (9) (1992) 869–885.
[29] A.K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Inc., 1989.
[30] X. Bai, L.J. Latecki, W.-Y. Liu, Skeleton pruning by contour partitioning with discrete curve evolution, IEEE Trans. Pattern Anal. Mach. Intell. 29 (3) (2007) 449–462.
[31] K.P. Burnham, D.R. Anderson, Model Selection and Multimodel Inference: A Practical Information-theoretic Approach, Springer, 2002.
[32] G. Schwarz, et al., Estimating the dimension of a model, Ann. Stat. 6 (2) (1978) 461–464.
[33] K.P. Burnham, D.R. Anderson, Multimodel inference understanding AIC and BIC in model selection, Sociol. Methods Res. 33 (2) (2004) 261–304.
[34] L.J. Latecki, R. Lakamper, T. Eckhardt, Shape descriptors for non-rigid shapes with a single closed contour, in: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, Hilton Head Island, SC: IEEE, 2000, pp. 424–429.
[35] X. Bai, X. Yang, L.J. Latecki, W. Liu, Z. Tu, Learning context-sensitive shape similarity by graph transduction, IEEE Trans. Pattern Anal. Mach. Intell. 32 (5) (2010) 861–874.
[36] B. Kimia, A Large Binary Image Database, LEMS Vision Group at Brown University ⟨http://www.lems.brown.edu/~dmc/⟩, 2002.
[37] T. Liu, D. Tao, Classification with Noisy Labels by Importance Reweighting, arXiv preprint arXiv:1411.7718.

Costas Panagiotakis received the B.A., M.Sc., and Ph.D. degrees from the Department of Computer Science, University of Crete, Greece, in 2001, 2003, and 2007, respectively. He is an Associate Professor with the Department of Business Administrator (Agios Nikolaos), TEI of Crete, Greece. He is the author of one book and more than 60 articles in international journals and conferences. His interests include image analysis, pattern recognition, multimedia and signal processing.

Antonis Argyros is a Professor of Computer Science at the Computer Science Department (CSD), University of Crete (UoC) and a Researcher at the Institute of Computer Science (ICS), Foundation for Research and Technology-Hellas (FORTH) in Heraklion, Crete, Greece. His research interests fall in the areas of computer vision with emphasis on tracking, human gesture and posture recognition, 3D reconstruction and omnidirectional vision. He is also interested in applications of computer vision in the fields of robotics and smart environments. In these topics, he has published more than 120 papers in peer reviewed scientific journals, conferences and workshops.