PROBABILISTIC SHAPE MODELS: THE ROLE OF THE PARTITION FUNCTION

Arnaldo J. Abrantes[1] *Jorge S. Marques*[2]

[1]Instituto Superior de Engenharia de Lisboa and INESC, Portugal
[2]Instituto Superior Técnico and INESC, Portugal

ABSTRACT

Deformable models have been intensively investigated during the last decade. Several well known algorithms, proposed in other contexts can also be included in this class (e.g., Kohonen maps, elastic nets and fuzzy c-means). In all these methods the model parameters are obtained in a deterministic framework by the minimization of an energy function. This paper proposes a novel class of probabilistic shape models related to the unified framework presented in [1]. Shape modelling is addressed as a MAP estimation problem, by assuming that the image features are random variables with Gibbs-Boltzmann distribution, and provides extensions for several well known algorithms. The main difference between the proposed algorithms and the original ones lies on the partition function which depends on the model parameters and influences the shape estimates. For example, it is shown that in snakes the partition function generates short-range repulsive forces between the model units which prevent their collapse when they are attracted by common data.

1. INTRODUCTION

Deformable models have been intensively investigated as a set of adaptive tools for shape representation and tracking. Inspired in physics, these models have the ability to dynamically modify their shape to fit the object boundary. Some well known examples are: Snakes [2], Balloons [3], B-snakes [4] and eigen shape models [5]. We have recently proposed in [1] a unified framework for the study of deformable models, based on a common energy function, which includes several well known algorithms (e.g., snakes, elastic nets [6], Kohonen maps [7], fuzzy c-means [8] and hard c-means [9]) and allows a joint study of these methods. As an alternative to the physic interpretation, shape estimation can be formulated as a statistical inference problem. Previous attempts to link these two perspectives are described in [10][11] which discuss elastic nets and probabilistic snakes as Bayesian estimation methods.

This paper addresses shape modeling in a Bayesian context, assuming that image features are random variables with shape dependent probability distribution. The estimation of the unknown parameters is addressed as a MAP estimation problem. Two basic questions concern the choice of the probability models and its relationship with previous algorithms. Inspired in statistical mechanics we assume that the 2D features detected in the image (e.g., edge points) have a Gibbs-Boltzmann distribution $p(\mathcal{P}|v) = e^{-\beta E(\mathcal{P};v)}/Z(v)$ where E is the energy function proposed in [1], $\mathcal{P}$ is the set of observed features and v is a vector with the unknown (shape) parameters. Z is a normalization term denoted as partition function. By an appropriate choice of the weighting functions we may obtain the probability distribution for snakes, elastic nets, Kohonen map or fuzzy c-means [1]. Since the partition function depends on the unknown parameters, the MAP estimate obtained from the Gibbs-Boltzmann distribution is, in general, different from the shape estimate which minimizes the energy E used in classic active contour algorithms. The difference lies on the structure of the partition function. Therefore, the partition functions of several data representation algorithms are also discussed in the paper.

2. CONSTRAINED CLUSTERING

It was proposed in [1] a unified framework for a class of constrained clustering algorithms based on the minimization of a fuzzy energy function. This framework includes several data representation methods, (e.g., snakes, Kohonen maps, elastic nets, c-means and fuzzy c-means), derived in the context of neural networks, constrained clustering and deformable models. The usefulness of this approach is twofold. It allows a unified understanding of different algorithms and it provides design tools for the development of new techniques by an appropriate choice of a set of weighting functions. The unified framework is briefly summarized in this section (details can be found in [1]).

Let $\mathcal{P}$ be a set of M edge points detected in the image and let $v = [v_1, \cdots, v_N]^T$, $v_k \in R^2$, be a sequence of model units (v will be denoted as shape model in this paper). We wish to estimate v, in order to minimize an energy function

$$E = E_r + E_d \tag{1}$$

containing a regularization term $E_r = \frac{1}{2}tr\{v^T A v\}$, which measures the smoothness of the contour, and a data representation term

$$E_d = \sum_{p\in\mathcal{P}} \sum_{k=1}^{N} w_k(p)\|v_k - p\|^2 \tag{2}$$

where $w_k(p), k = 1, \cdots, N$, is a set of weighting functions which depend on the squared Euclidean distances $d_j = \|p - v_j\|^2$ from pattern p to all the model units (see figure 1).

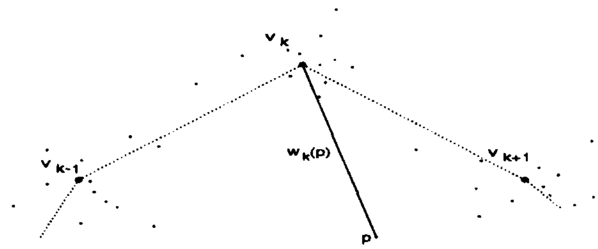

Figure 1. Weighting functions. (p - edge point; v_{k-1}, v_k, v_{k+1} model units)

The minimization of (1) is often performed by a gradient algorithm

$$v^{t+1} = v^t + \alpha(F_r + F_d) \tag{3}$$

where $F_r = -\frac{\partial E_r}{\partial v} = -Av$ and $F_d = -[\frac{\partial E_d}{\partial v_1}, \cdots, \frac{\partial E_d}{\partial v_N}]^T$ with [1]

$$\frac{\partial E_d}{\partial v_k} = \sum_{p \in \mathcal{P}} \vartheta_k(p)(v_k - p) \tag{4}$$

Function $\vartheta_k(p)$ is derived from the weighting functions $w_1(p), \cdots, w_N(p)$ and measures the influence of pattern p on the k-th model unit. The rows of matrices F_d and F_r are usually interpreted as data and regularization forces applied to each model unit. The data forces pull the model units towards the image patterns while regularization forces try to keep the contour shape smooth. Since all data points attract the model units, each unit is attracted towards the centroid of the data points near the unit (see [1] for details).

To define a deformable model belonging to this class, the user must specify a regularization matrix A and the set of weighting functions $w_k(p)$. By an appropriate choice of these functions, the algorithm becomes equivalent to snakes, Kohonen map, elastic net, hard or fuzzy c-means. Expressions for the weighting functions are given in [1].

3. MAP SHAPE MODELS

This section addresses shape estimation in a Bayesian framework. This requires two probabilistic models: a shape model (prior) and a data model (observation model). These models can either be estimated from a set of known images or chosen by the user, using the available knowledge on the shape properties and the image formation process. The second approach is adopted in this section. The shape and observation models used in this paper are based on the regularization and data energies defined in section 2.

Let the object shape, v, be a random variable with known probability density function $p(v) = Ce^{-\beta E_r}$ and let the image features be M independent random variables with Gibbs-Boltzmann distribution

$$p(\mathcal{P}|v) = \frac{e^{-\beta E_d(v)}}{Z(v)} \tag{5}$$

where

$$Z(v) = \int_{R^{2M}} e^{-\beta E_d(v)} \tag{6}$$

is the partition function (the notation stresses that Z depends on the model shape). The maximum a posteriori estimate of the model shape given a set of observed features, $\mathcal{P}$ is [12]

$$\hat{v} = \arg\max_v \log p(\mathcal{P}|v) + \log p(v) \tag{7}$$

Therefore,

$$\hat{v} = \arg\max_v -(\beta E_r + \beta E_d(\mathcal{P}; v) + \log Z(v)) \tag{8}$$

i.e., the estimated shape minimizes an equivalent energy

$$E^{eq} = \beta E + \log Z(v) \tag{9}$$

The MAP algorithm defined in (5-8) is not equivalent to the minimization of E described in section 2 due to $\log Z(v)$ term which depends on the model shape v. Using (2,6), the log partition function can be written as follows

$$logZ(v) = M \log \int_{R^2} \prod_{k=1}^{N} e^{-\beta w_k(p)\|v_k - p\|^2} dp \tag{10}$$

Equations (5-9) define a class of probabilistic shape models based on the fuzzy data energy (2). These models will be denoted as MAP Shape models and extend previous works. Durbin et al. proposed in [10] a probabilistic interpretation of the elastic net algorithm when $\beta = 1$. In this case, the partition function is shape independent and the MAP estimate also minimizes E). Gibbsian models are also used in probabilistic snakes described in [11]. However, the partition function is assumed to be constant (shape independent). This assumption is approximately valid if the model units are well separated but it is wrong if the distance between two or more units becomes smaller than 3σ. The partition function influences the shape estimates obtained from (8). This influence is addressed in sections 4 and 5.

4. TWO-UNIT MODELS

The difference between the energy functions minimized in constrained clustering and in the MAP methods is $logZ(v)$. In general, this term depends on all the shape units $v_1, \cdots, v_N$ and it is not easy to evaluate it analytically from (10), except in special cases (e.g. elastic nets with $\beta = 1$).

To understand the role of the log partition function we shall first consider models with two units: v_1, v_2. In this case $\log Z$ is a function of $v_1 - v_2$ and can be numerically computed. Figure 2 shows $\log Z$ in MAP snakes, elastic nets, Kohonen maps and fuzzy c-means, considering a model with two units. Different behaviors are observed. In snakes (Fig. 2a), the log partition function increases when the distance between the units becomes small, achieving a maximum at $v_1 - v_2 = 0$. The gradient of $logZ$ generates repulsive forces which pull the units apart. This can be interpreted as short range forces between units, similar to the Lennard-Jones interaction described by Szeliski et al. [13] , inspired in inter-molecular dynamics. A deep valley is observed in the MAP Kohonen map (see Fig. 2b). In this case attractive forces are applied to the model units. Finally fuzzy c-means exhibit a hybrid behavior (short range attraction

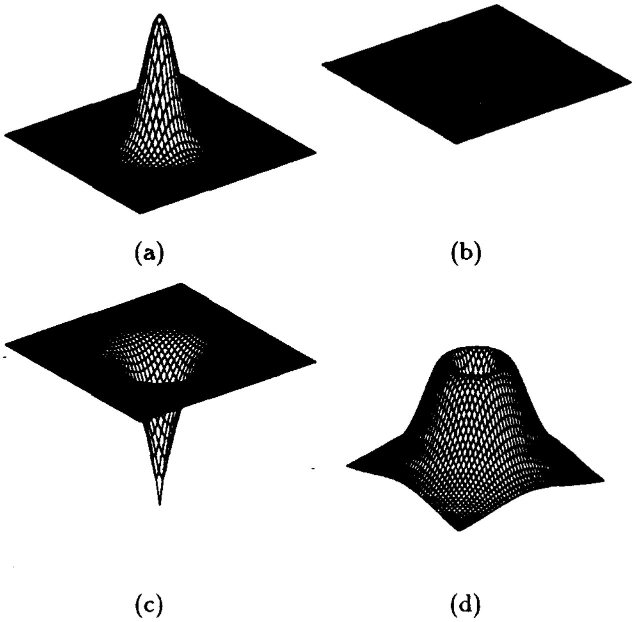

Figure 2. Log partition function of a pair of units in: a) snakes; b) elastic nets ($\beta = 1$); c) Kohonen maps (zero neighborhood radius); d) fuzzy c-means

forces, medium range repulsive forces) and elastic nets have a null force field if $\beta = 1$.

The force fields associated with these four algorithms are shown in Figure 3. During the optimization by the gradient algorithm, this force field is added to the data forces and regularization forces described in section 2 and influences the shape estimates. For example, the repulsive forces generated in snakes prevent the model units from collapsing. This can be seen in the next example.

Figure 4 illustrates the forces applied to a pair of units in the classic and MAP snake algorithm. In MAP snakes, two kinds of forces are displayed: the attraction forces towards the image features (gradient of the energy function) and the repulsive forces due to the partition function. In the deterministic algorithm, only the first type of forces is present. This shows the role of the partition function in keeping the model units apart, preventing their collapse.

5. MULTI-UNIT MODELS

It is harder to characterize the influence of the log partition function in multi-unit models. If the model units are well separated except a finite number of pairs, the previous analysis is valid and the log partition forces applied to each of the close units are identical to the ones computed in the previous section. If the number of close units in a group is larger than two, other interactions have to be considered.

In MAP snakes, it is possible to derive an analytical expression for the log partition forces, expressed as a sum of interactions between pairs, triples, ...,n-tuples of close units (this will be presented in detail in [14]).

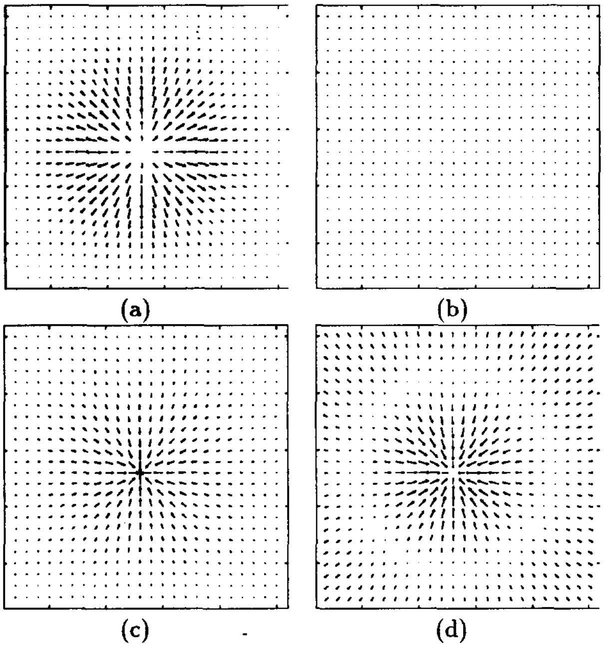

Figure 3. Force field of log partition function for a pair of units in: a) snakes; b) elastic nets; c) Kohonen maps; d) fuzzy c-means

To provide some insight about the influence of the log-partition function in real images preliminary tests were carried out to assess the performance of the deterministic and MAP snake algorithm. The edge patterns used in these tests were computed using the Sobel algorithm and the shape model was initialized outside the object and far from the object boundary. The optimization of the energy function was performed by the gradient algorithm.

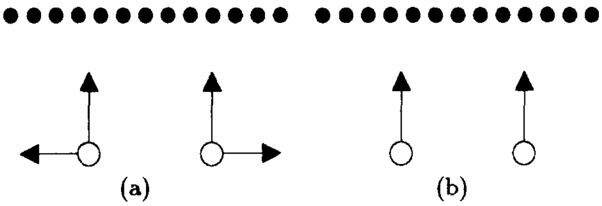

Figure 4. Forces applied to a pair of units: a) MAP and b) classic snakes; Black dots represent image features and white dots are the model units.

Figure 5 shows a comparison between snakes and MAP snakes. Figure 5a contains the original image and the initial model configuration (40 units were used in this example). Figures 5b,c show the edges detected by the Sobel algorithm in the upper left region of the knife image and the potential function in this region, obtained by convolving the edge image with a Gaussian lowpass filter [3]. Finally, Figures 5d,e show a detail of the shape estimates obtained by snakes and MAP snakes. The influence of the log partition function is clear since it avoids the concentration (collapse) of model units in the deep potential valleys (no redistribution of the model units was performed in these examples).

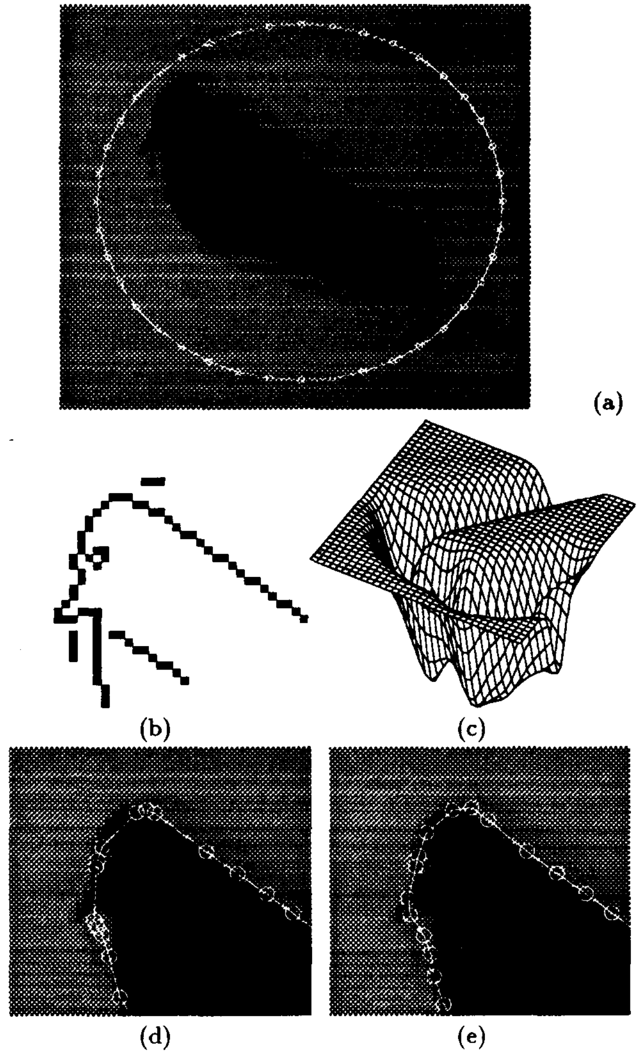

Figure 5. Object boundary extraction. a) initial model configuration; b) edge points; c) potential field; final configurations using d) Classic Snakes and e) MAP Snakes

6. CONCLUSION

This paper presented a class of MAP shape models based on the constrained clustering framework described in [1]. Shape is estimated by a MAP method, assuming a Gibbsian observation model with a fuzzy energy function. The fuzzy energy is defined by a set of weighting functions which control the algorithm behavior. The new class of algorithms, presented in the paper, provides probabilistic extensions for a set of well known techniques (snakes, elastic nets, Kohonen maps, c-means and and fuzzy c-means), allowing a joint characterization of these methods.

The MAP estimates derived from the Gibbs-Boltzmann distribution are not equivalent to the minimization of the energy E used in deterministic algorithms. An additional term (the log partition function) is added to the cost function. The influence of this term is discussed in the paper. First, simple models with two units are considered. It is concluded that the log partition function generates additional short range interaction forces between close units. These forces can be repulsive as in MAP snakes, attractive as in MAP Kohonen maps or hybrid as in MAP fuzzy c-means and resemble the Lennard-Jones interaction used by Szeliski et al. [13] , inspired in inter-molecular dynamics.

The role of the log partition function in larger models is also discussed. Short range forces are generated if two or more units become close. The amplitude of these forces depends on the choice of the weighting functions and on the geometry of the cluster. Experimental results are provided to assess the effect of these forces in MAP snakes with real images. It is concluded that they improve the shape estimates by avoiding the concentration or even collapse of model units in deep potential valleys.

The log partition function used in this paper was computed by numerical integration in R^2 using (2). This is a time consuming operation which is repeated in each iteration of the optimization process. To overcome this difficulty, an analytic expression of the log partition function for the MAP snake algorithm is presented in a forthcoming paper [14].

7. ACKNOWLEDGMENT

This work was partially supported by JNICT under contract TPR 2562/95.

REFERENCES

[1] A. Abrantes, J. Marques, A Class of Constrained Clustering Algorithms for Object Boundary Extraction, *IEEE Trans. Image Processing*, pp. 1507-1521, November, 1996.

[2] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active Contour Models, *Int. J. Computer Vision*, vol. 1, pp. 259-268,1987.

[3] L. Cohen, On Active Contour Models and Ballons, CVGIP: Image Understanding, vol. 53, pp. 211-218, 1991

[4] S. Menet, P. Saint-Marq, G. Medioni, B-snakes: implementations and Applications to Stereo, *Proc. DARPA Image Understanding Workshop*, 720-726, 1990.

[5] T. Cootes, C. Taylor, D. Cooper, J. Graham, Active Shape Models. Their Training and Application, *Comp. Vision and Image Understanding*, Vol. 61, n 1, pp. 38-59, Jan., 1995.

[6] R. Durbin, D. Willshaw, An Analogue Approach to the Traveling Salesman Problem Using an Elastic Net Method, *Nature*, Vol. 326, pp. 689-691, April, 1987.

[7] T. Kohonen, The Self-Organizing Map, *Proceedings of the IEEE*, Vol. 78, N. 9, pp. 1464-1480, September, 1990.

[8] J. Bezdek, A Convergence Theorem for the Fuzzy Isodata Clustering Algorithms, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, N. 1, January, 1980.

[9] McQueen, Some Methods for Classification and Analysis of Multivariate Observations, *Proc. 5th Berkeley Symp. Math. Stat. and Prob.*, Vol. I, Univ. California Press, pp. 281-286, 1967

[10] R. Durbin, R. Szeliski, A. Yuille, An Analysis of the Elastic Net Approach to the Travelling Salesman Problem, *Neural Comp.*, 1, pp. 348-358, 1989.

[11] D. Terzopoulos, R. Szeliski, Tracking with Kalman Snakes, in *Active Vision*, A. Blake, A. Yuille (eds.), pp. 3-20, MIT Press, 1992.

[12] H. Van Trees, Detection Estimation and Modulation Theory, Part I, Wiley, 1968.

[13] R. Szeliski, D. Tonnensen, D. Terzopoulos, Modelling Surfaces of Arbitary Topology with Dynamic Particles, *IEEE Conf. On Computer Vision and Pattern Recognition*, New York, 1993.

[14] J. Marques, A. Abrantes, Probabilistic Snakes, submitted for publication.