A New Energy Model for the Hidden Markov Random Fields

Jérémie Sublime[1,2], Antoine Cornuéjols[1], and Younès Bennani[2]

[1] AgroParisTech, INRA - UMR 518 MIA, F-75005 Paris, France
{jeremie.sublime,antoine.cornuejols}@agroparistech.fr
[2] Université Paris 13 - Sorbonne Paris Cité,
Laboratoire d'Informatique de Paris-Nord - CNRS UMR 7030,
99 Avenue Jean Baptiste Clément, 93430 Villetaneuse, France
Younes.bennani@lipn.univ-paris13.fr

Abstract. In this article we propose a modification to the HMRF-EM framework applied to image segmentation. To do so, we introduce a new model for the neighborhood energy function of the Hidden Markov Random Fields model based on the Hidden Markov Model formalism. With this new energy model, we aim at (1) avoiding the use of a key parameter chosen empirically on which the results of the current models are heavily relying, (2) proposing an information rich modelisation of neighborhood relationships.

Key words: Markov Random Fields, Hidden Markov Models, Image Segmentation

1 Introduction

A *Markov Random Fields network* is a graphical probabilistic model aiming at taking into consideration the neighborhood interactions between the data in addition to the observed a priori knowledge. Such model allows considering the explicit dependencies between the data and to weight their influence. In the case of image segmentation, these dependencies are the links between two neighbor pixels or segments (patches of pixels). Markov Random Fields networks rely on this notion of neighborhood, and are represented as non-oriented graphs the vertices of which are the data and the edges are the links between them. This additional information on the data has been shown to significantly improve the global results of the image segmentation process [1].

The *Hidden Markov Model* is also a probabilistic model in which a set of random variables $S = \{s_1, ..., s_N\}$, $s_i \in 1..K$ are linked to each other by neighborhood dependencies and are emitting observable data $X = \{x_1, ..., x_N\}$ where the x_i are the vectors containing the attributes of each observation. The goal is then to determine the optimal configuration for S, i.e. finding the values of the s_i in order to get the segmentation.

The *hidden Markov random field model* is the application of the hidden Markov model to the specific dependency structure of the Markov random fields.

This model is quite common in image segmentation [2]. An example of a typical structure for the HMRF model is shown in Figure (1).

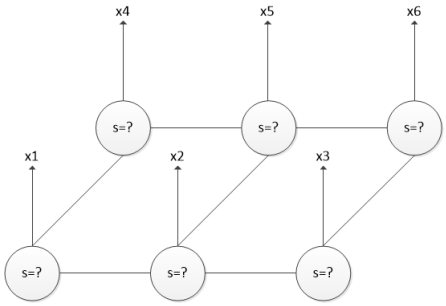

Fig. 1. An example of graph modelling the HMRF model.

The main goal of the hidden Markov random field model applied to image segmentation is to find the right label for each pixel or segment of a picture in order to get a segmentation in homogeneous and meaningful areas. Doing so requires to infer the $s_i \in 1..K$, which is often achieved by using the maximum a posteriori criterion (MAP) to determine S such as:

$$S = \arg\max_{S}(P(X|S,\Theta)P(S)) \tag{1}$$

$$P(X|S,\Theta) = \prod_{t} P(x_t|s_t,\theta_{s_t}) \tag{2}$$

As it is often the case in image segmentation, we will consider that $P(x_t|s_t,\theta_{s_t})$ follows a Gaussian distribution of parameters $\theta_{s_t} = (\mu_{s_t}, \Sigma_{s_t})$.

The most common way to maximize $P(X|S,\Theta)P(S)$ is to use the two-step HMRF-EM couple [4]:

- Research of the prototype parameters Θ and maximization of $P(X|S,\Theta)$ using the Expectation Maximization algorithm (EM) [5].
- Local optimization of $P(X|S,\Theta)P(S)$ using the Iterated Conditional Modes Algorithm (ICM) [6].

The principle of the ICM algorithm is to iteratively perform a local optimization of $P(x|s,\theta)P(s)$. Such optimization is often done by minimizing an energy function with a form deriving from the logarithm of $P(x|s,\theta)P(s)$ is shown in equation (3).

$$U(s,x) = U_{obs}(s,x,\theta_s) + \beta \sum_{v \in V_x} U_{neighbor}(s_v, s) \tag{3}$$

U_{obs} and $U_{neighbor}$ are potential energy functions: U_{obs} is the potential energy derived from the observed data, also known as the data term, and $U_{neighbor}$

the potential energy derived from the neighborhood configuration, sometimes refered as the smoothness term. V_x is the neighbor list of the observed data. β is an empirical constant used to modulate the weight of the neighborhood.

In its current versions, the HMRF-EM algorithm heavily relies on the choice of β to determine the quality of the results. Furthermore, while the energy function U_{obs} usually is the logarithmic form of $P(x|s,\theta_s)$ the emission law, the exact form of $P(s)$ is unknown. For this reason, the neighborhood energy $U_{neighbor}$ is chosen as simple as possible: another gaussian distribution for complex models, and in most cases something even simpler such as the Kronecker delta function [4, 7, 8].

Our goal in this paper is to propose a method to approach the local value of $P(s)$ in order to provide a more accurate and information rich neighborhood energy function, and avoid using the empirical constant β.

2 Proposed Modifications to the HMRF-EM Algorithm

In this section we consider a set of linked random variables $S = \{s_1...,s_N\}$, $s_i \in 1..K$ that are emitting observable data $X = \{x_1,...,x_N\}$, $x_i \in \mathbb{R}^d$. We assume the emission law of the x_i by the s_i to be a Gaussian distribution and we chose to consider a first order neighborhood for our Markov Random Fields. By convention, μ_s shall refer to the mean vector of a considered state s and Σ_s to its covariance matrix. Then, for each observation x associated to a state s we have the following data term as the observable energy:

$$U_{obs}(s,x) = \frac{1}{2}(x-\mu_s)^T \Sigma_s^{-1}(x-\mu_s) + log(\sqrt{|\Sigma_s|(2\pi)^d}) \tag{4}$$

2.1 Proposed Energy Model

In the hidden Markov model formalism, the neighborhood relations are often described by a transition matrix $A = \{a_{i,j}\}_{K\times K}$ where each $a_{i,j}$ is the transition probability from one state to another between two neighboring data. We propose to keep this formalism and to adapt it to the HMRF model. Then, for $P(x|s,\theta_s)P(s)$, we get the following expression:

$$P(x|s,\theta_s)P(s) = P(x|s,\theta_s) \times \prod_{v\in V_x} P(s|s_v) = \frac{1}{Z} \times \mathcal{N}(\mu_s,\Sigma_s) \times \prod_{v\in V_x} a_{s_v,s}^{\frac{1}{|V_x|}} \tag{5}$$

By considering the logarithm of equation (5) and the observable energy from equation (4), we get the following complete energy function to be minimized:

$$U_{HMRF}(s,x) = \frac{1}{2}(x-\mu_s)^T \Sigma_s^{-1}(x-\mu_s) + log(\sqrt{|\Sigma_s|(2\pi)^d}) - \sum_{v\in V_x} \frac{log(a_{s_v,s})}{|V_x|} \tag{6}$$

Since the $a_{i,j}$ are values between 0 and 1, our neighborhood energy is a positive penalty function. This penalty function can take as many as K^2 distinct values

instead of only two values (0 or β) for the regular neighborhood energy functions. We note that this fitness function is as difficult to optimize as those from other models and that there is no warranty that the global optimum can be found.

It is important to highlight that in addition of providing a more flexible energy model for the HMRF-EM framework, the transition probability matrix also provides semantic information on the clusters: The diagonal of the matrix gives some insight on the clusters compactness, non-diagonal values provide information on cluster neighborhood affinities, but also incompatibilities for values near 0, or the inclusion of a cluster inside another one for values above $0,5$.

2.2 Proposed Algorithm

The main difficulty in our approach is that we have to evaluate the transition matrix A in order to run our HMRF-EM algorithm. Without some external knowledge on the neighborhoods of the different clusters, there is no way to get the exact values for the transition probabilities. It is however possible to approach them by evaluating the a posteriori transition probabilities after the EM algorithm and after each step of the ICM algorithm. This method is described in Algorithm (1).

We note $\delta_{n,m}$ the Kronecker delta function the value of which is 1 if n and m are identical and 0 otherwise. We note C the set of all the oriented binary cliques of the considered dataset (since the Markov Random Field Model is usually non-oriented, all links are represented twice in C to cover both directions of each dependency). The a posteriori transition probabilities $a_{i,j}$ are evaluated by browsing all the data and counting the number of transitions from variables whose state is i to variables in the state j, divided by the number of all links starting from variables whose state is i:

$$a_{i,j} = \frac{Card(c \in C | c = (i,j))}{Card(c \in C | c = (i,\#))} \tag{7}$$

Algorithm 1: EM+ICM with transition matrix update

```
Initialize S and Θ with the EM algorithm
Initialize A using Equation (7)
while the algorithm has not converged do
    for each x ∈ X do
        Minimize U_HMRF(s, x) as defined in Equation (6)
    end
    Update A from the new distribution S using Equation (7)
end
```

The idea behind this computation is that the ICM algorithm is already making an approximation by considering that the states evaluated in its previous iterations can be used to evaluate the potential neighborhood energy in the current iteration. We are using this same approximation to approach the probability of a given transition between two states.

In our model each transition probability is computed with the goal of detecting statistically unlikely labels among the data depending on their neighborhood. While the classical energy formula penalty may give erratic results depending on both the picture and the penalty value β, our energy model proposes a more flexible penalty which is the result of a less coarse, but still highly improvable vision of how to model neighborhood relations in a Markov random field model.

The complexity of the regular HMRF-EM version is $O(i \times k \times d \times n \times |\bar{V}|)$. Since we have to browse the data twice instead of only once in order to update A with our proposed algorithm, we can expect that the optimization process will be up to twice slower. The complexity for our algorithm is $O(i \times (k \times d \times n \times |\bar{V}| + n \times |\bar{V}|))$, which is equivalent to the complexity of the original ICM algorithm.

3 Experiments

In our first experiment, we wanted to compare the performance of our energy model with the other models in term of raw segmentation results and cluster characteristics. To do so, we have been testing the HMRF-EM frameworks with three different energy models: An ICM algorithm with a basic energy formula based on the Potts model [7], see Equation (8), an HMRF-EM algorithm with one of the most commonly used energy formula, see Equation (9) [4, 8], and the HMRF-EM framework with our Energy model, see Equation (6).

$$U_{POTTS}(s,x) = (1 - \delta_{s_x, s_x^{t-1}}) + \beta \sum_{v \in V_x} (1 - \delta_{s_v, s}) \tag{8}$$

$$U_{MRF}(s,x) = (x - \mu_s)^T \Sigma_s^{-1} (x - \mu_s) + log(\sqrt{|\Sigma_s|(2\pi)^d}) + \beta \sum_{v \in V_x} (1 - \delta_{s_v, s}) \tag{9}$$

Fig. 2. From left to right : Original picture 481×321 pixels, 3-clusters segmentation using energy formula (9) and $\beta = 0.66$, 3-clusters segmentation using our energy formula, 3-clusters segmentation using energy formula (8) and $\beta = 1.0$.

The results in Figure (2) and Table (1) show that our model achieves decent segmentation results that are similar to what can be done with other energy models using a nearly optimal value for β.

Furthermore, our method seem to be effective to agglomerate pixels in homogeneous areas while conserving a low internal variance for the clusters which tends to prove that the resulting areas are meaningful.

Table 1. Purity of the expected clusters and average internal variance (3 clusters)

	Cluster Purity	Internal Variance
U_{POTTS}, $\beta = 2/3$	73.42%	6.10
U_{POTTS}, $\beta = 1$	73.38%	6.13
U_{MRF}, $\beta = 2/3$	69.37%	5.83
U_{MRF}, $\beta = 2/3$	69.36%	5.84
Our U_{HMRF}	73.46%	5.86

Similar results have been found on another experiment on a satellite image, see Figure (3) (2014 Cnes/Spot Image, DigitalGlobe, Google), thus confirming that our approach is competitive when compared to what has already been done in this area.

Fig. 3. From left to right, and from top to bottom : The original image 1131×575 pixels, the result of our algorithm, the result using energy equation (8) and $\beta = 1$, the result using energy equation (9) and $\beta = 1$.

As can be seen on a color and full scale version of Figure (3), our energy model achieves a decent segmentation with a fair amount of easily visible elements such as roads, rocky areas and some buildings. On the other hand the classical HMRF-EM algorithm using Equation (9) fails to aggregate neighbor pixels from common elements despite a relatively high value for β (the same phenomenon can be observed in Figure (2)), while the algorithm based on the Potts model using Equation (8) tends to give too coarse results. This second experiment emphasizes again that finding a right value for β is a difficult task.

Our third experiment was on a satellite image from a more difficult dataset. In this dataset from [9], a very high resolution image from the city of Strasbourg (France) is described in 187.058 unlabeled segments having 27 numerical attributes either geometrical or radiometrical. Another particularity of this dataset is that the segments have irregular neighborhoods: each segment can have one to fifteen neighbors. The neighborhood irregular configuration of this

dataset makes it impractical to find an optimal value for the constant β and the segmentation results using conventional HMRF-EM energy functions (8) and (9) were quite poor.

An extract of a 9-clusters segmentation on this dataset is available in Figure (4). The result shows that various elements of the city have been correctly detected such as sport areas, the river, and some buildings. On a full scale version of this image, we can see that some streets and houses have also been correctly labeled.

Furthermore, this experiment has also confirmed that our matrix representation for the neighborhood made sense, and that at a segment level it was possible to interpret some of the matrix elements (whereas it was difficult at a pixel level). For instance, on this dataset the matrix featured a low transition probability between building areas and water areas which is consistent with the satellite image. We also had an important transition probability from tree areas to grass areas (≈ 0.65) which is consistent with the fact that tree areas often come by patches instead of compact blocks and are often surrounded by grass.

Fig. 4. On the left: One of the original source image, ©CNES2012, Distribution Astrium Services / Spot Image S.A., France, All rights reserved. On the right: the resulting 9-clusters segmentation for the area.

In a last experiment, we have compared the computation times of the 3 energy models introduced in this article for 3 pictures of different sizes, see Table (2).

The results have confirmed what we had already hinted while describing the complexity of our modified HMRF-EM algorithm. Our version of the algorithm is slower, and it was to be expected because of the time needed to update the transition matrix after each iteration. Nevertheless, it is important to emphasize that it is still faster to run our version of HMRF-EM rather than trying several β values, or running an optimization algorithm for β, with another energy formula.

Table 2. Computation times in ms, Intel Core I5-3210M, 2.5GHz.

	$404 \times 476px$ 4 clusters	$380 \times 270px$ 4 clusters	$1131 \times 575px$ 4 clusters	$1131 \times 575px$ 8 clusters
U_{Potts} (8)	4451ms	2453ms	13370ms	26920ms
U_{MRF} (9)	4933ms	2717ms	14771ms	29433ms
Our U_{HMRF} (6)	6632ms	3418ms	20810ms	42736ms

4 Conclusion and Future Work

We have proposed an improvement to the energy model of the Hidden Markov Random Field Model applied to image segmentation. In our model, the neighborhood energy formula is based on an approached transition matrix rather than an empirical penalty parameter. Our preliminary experiments have shown our model to give competitive results compared with other models based on more classical energy functions who rely on finding an optimal penalty value for the neighborhood energy.

Furthermore, while our contribution does not bring any significant improvement on the quality of the results for image segmentation, our neighborhood energy model using a transition matrix gives the opportunity to have a semantic rich representation of the interactions between the clusters.

In our future works, we will focus on using the information collected in these transition matrices with the goal of proposing collaborative frameworks such as the collaborative segmentation of similar pictures using several HMRF-based algorithms sharing their prototypes and transition matrices, or the segmentation of a sequence of pictures using an HMRF-HMM hybrid framework.

References

1. Hernandez-Gracidas, C., Sucar, L.E.: Markov Random Fields and Spatial Information to Improve Automatic Image Annotation , Advances in Image and Video Technology, Lecture Notes in Computer Science, vol. 4872, p. 879-892 (2007)
2. Zhang, L., Ji, Q.: Image segmentation with a unified graphical model, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(8): p. 1406-1425 (2010)
3. Roth, S. and Black, M. J.: Fields of experts, Markov Random Fields for Vision and Image Processing, MIT Press, p. 297-310 (2011)
4. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm, IEEE Transactions on Medical Imaging, 20(1): p. 45-57 (2001)
5. Dempster, A.P., Laird, N.M., Rubin, D.: Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society. Series B (Methodological), vol. 39, no 1, p. 1-38 (1977)
6. Besag, J.: On the Statistical Analysis of Dirty Pictures. Journal of the Royal Statistical Society. Series B (Methodological), 48(3): p. 259-302 (1986)
7. Weinman, J.: A Brief Introduction to ICM, `http://www.cs.grinnell.edu/~weinman/courses/CSC213/2008F/labs/11-threads-icm.pdf` (2008)
8. Kato, Z., Berthod, M., Zerubia, J.: A Hierarchical Markov Random Field Model and Multi-temperature Annealing for parallel Image Classification, Graphical Models and Image Processing, Vol. 58, No 1, January, p. 18-37 (1996)
9. Rougier, S., Puissant, A.: Improvements of urban vegetation segmentation and classification using multi-temporal Pleiades images, 5th International Conference on Geographic Object-Based Image Analysis, Thessaloniki, Greece, p. 6 (2014)

Acknowledgements. This work has been supported by the ANR Project COCLICO, ANR-12-MONU-0001.