

(Hyper)-Graphical Models in Biomedical Image Analysis

Nikos Paragios¹, Enzo Ferrante³, Ben Glocker³, Nikos Komodakis², Sarah Parisot³, Evangelia I. Zacharaki¹

¹*CentraleSupélec, Inria, Université Paris-Saclay, France*

²*Université Paris-Est, École des Ponts ParisTech, France*

³*Imperial College of London, London, UK*

Abstract

Computational vision, visual computing and biomedical image analysis have made tremendous progress over the past two decades. This is mostly due the development of efficient learning and inference algorithms which allow better and richer modeling of image and visual understanding tasks. Hyper-Graph representations are among the most prominent tools to address such perception through the casting of perception as a graph optimization problem. In this paper, we briefly introduce the importance of such representations, discuss their strength and limitations, provide appropriate strategies for their inference and present their application to address a variety of problems in biomedical image analysis.

Keywords: (Hyper)Graphs, Random Fields, Message Passing, Graph Cuts, Linear Programming, Image Segmentation, Shape & Volume Registration

1. Introduction

The analysis of medical images aims at extraction of hidden, clinically relevant information, and has a wide range of applications including the detection or automatic outlining of internal structures (image segmentation), the summarization of knowledge from group populations (spatial normalization), or the association of imaging profile with the clinical state of the patient. Inverse modeling is often the paradigm used to interpret medical images. In simple words given an interpretation objective, the first step consists on describing the space of solutions through a parametric mathematical model. The parameters of this model are then somehow associated with the measurements through an objective specific cost function that also imposes some smoothness constraints to make the problem mathematically and computationally tractable. Optimal solution to the problem consists in finding the set of parameters able to produce the lowest cost of the energy function and is often determined through an approximate inference algorithm since exact solution is frequently not attainable.

Numerous mathematical paradigms have been considered in the field of medical imaging to implement the aforementioned strategy. Mathematical richness, computational complexity, inference

mechanisms and modularity and scalability are often their classification criteria. Ideally one would like to adopt models of limited complexity (it facilitates inference) but able to encode rich priors, as well as cost functions that can be exactly minimized (optimal solution) with reasonable computational complexity. Graphical models [1] is an elegant, powerful and computationally efficient method to such an objective. In the recent years, the field has witnessed an enormous progress due to the development of efficient optimization/inference methods coupled with machine learning algorithms and the availability of large scale training data. While probabilistic graphical models have a variety of useful variants, here we will focus on a Markov Random Fields (MRF) formulation, where inference is often expressed as a (undirected) graph optimization problem acting on a predefined graph structure (fixed number of nodes and connectivity) associated with a discrete number of variables.

A wide variety of tasks in medical image analysis can be formulated as discrete labeling problems. In very simple terms, a discrete optimization problem can be stated as follows: we are given a discrete set of variables \mathcal{V} , all of which are vertices in a graph \mathcal{G} . The edges of this graph (denoted by \mathcal{E}) encode the variables' relationships. We are also given as input a discrete set of labels \mathcal{L} . We must then assign one label from \mathcal{L} to each variable in \mathcal{V} . However, each time we choose to assign a label, say, x_{p_1} to a variable p_1 , we are forced to pay a price according to the so-called *singleton* potential function $g_p(x_p)$, while each time we choose to assign a pair of labels, say, x_{p_1} and x_{p_2} to two interrelated variables p_1 and p_2 (two nodes that are connected by an edge in the graph \mathcal{G}), we are also forced to pay another price, which is now determined by the so called *pairwise* potential function $f_{p_1 p_2}(x_{p_1}, x_{p_2})$. Both the singleton and pairwise potential functions are problem specific and are thus assumed to be provided as input. For example, if the aim is image segmentation, in order to assign a class label to every voxel of the image, an appearance model is learned (from training examples) for every class and encoded in the form of a probability distribution $\pi(x)$. This probability distribution can be naturally incorporated in the MRF model by setting the unary potentials of the segmentation grid for every label to the negative log-probability of the respective class:

$$g_p(x_p) = -\log(\pi(x_p)).$$

Our goal is then to choose a labeling which will allow us to pay the smallest total price. In other words, based on what we have mentioned above, we want to choose a labeling that minimizes the sum of all the MRF potentials, or equivalently the MRF energy. This amounts to solving the following optimization problem:

$$\arg \min_{\{x_p\}} \mathcal{P}(g, f) = \sum_{p \in \mathcal{V}} g_p(x_p) + \sum_{(p_1, p_2) \in \mathcal{E}} f_{p_1 p_2}(x_{p_1}, x_{p_2}). \quad (1)$$

The use of such a model can describe a number of challenging problems in medical image analysis. However these simplistic models can only account for simple interactions between variables, a rather constrained scenario for high-level medical imaging perception tasks. One can augment the

expressive power of this model through higher order interactions between variables, or a number of cliques $\{C_i, i \in [1, n] = \{\{p_{i1}, \dots, p_{i|C_i|}\}\}$ of order $|C_i|$ that will augment the definition of \mathcal{V} and will introduce hyper-vertices:

$$\arg \min_{\{x_p\}} \mathcal{P}(g, f) = \sum_{p \in \mathcal{V}} g_p(x_p) + \sum_{(p_1, p_2) \in \mathcal{E}} f_{p_1 p_2}(x_{p_1}, x_{p_2}) + \sum_{C_i \in \mathcal{E}} f_{p_1 \dots p_n}(x_{p_{i1}}, \dots, x_{p_{i|C_i|}}). \quad (2)$$

where $f_{p_1 \dots p_n}$ is the price to pay for associating the labels $(x_{p_{i1}}, \dots, x_{p_{i|C_i|}})$ to the nodes $(p_1 \dots p_{i|C_i|})$. Parameter inference, addressed by minimizing the problem above, is the most critical aspect in computational medicine and efficient optimization algorithms are to be evaluated both in terms of computational complexity as well as of inference performance. State of the art methods include deterministic and non-deterministic annealing, genetic algorithms, max-flow/min-cut techniques and relaxation. These methods offer certain strengths while exhibiting certain limitations, mostly related to the amount of interactions which can be tolerated among neighborhood nodes. In the area of medical imaging where domain knowledge is quite strong, one would expect that such interactions should be enforced at the largest scale possible.

The remainder of this paper reviews briefly our contributions in the field. Section 2 presents the work done in the area of inference algorithms while section 3 discuss their use to address the fundamental problems of biomedical image analysis, while the last section concludes the paper and presents perspectives and future directions of our work.

2. Inference on Graphical Models

Inference over graphs has been a well studied problem in a number of fields such as networks, operational research, computational biology and computer vision. Graph-based representations were introduced in computer vision at mid-eighties [2] through Markov Random Fields as a novel mathematical modeling framework constrained though from the lack of efficient inference methods as well as processing power - and became again popular during the past two decades thanks to the development of efficient optimization algorithms [3, 4].

On one hand, local iterative optimization methods or annealing-like methods were the first attempts to perform inference over graphs in computer vision and image processing. Computational efficiency was the main strength of local iterative methods with bottleneck being their inability to converge to a good optimum. Annealing methods - despite their theoretical guarantees with respect to the attained solution - were computationally inefficient and practically unusable. The principle of max-flow/min-cut theorem was first introduced in vision at the late eighties [5] and then re-introduced efficiently [6] along with the re-discovery of message passing methods [7] despite the absence of theoretical guarantees - such as belief propagation networks [8]. Such developments

had as a direct consequence the establishment of graphical models as one of the main stream computational vision mathematical formalism over the past decade. In the recent years, a number of extremely efficient optimization methods were re-introduced (e.g. Tree-reweighted Message Passing [9], the Fast Primal-Dual Method [10] or the Extended Roof Duality [11]) resulting on a great variety of optimization algorithms addressing the expectations of the field both in terms of optimality properties of the attained solution as well as in terms of computational complexity for low-rank (pair-wise) graphical models.

On the other hand, the case of higher order models with arbitrary interactions between nodes (both in terms of connectivity as well as in terms of potentials) is far from being considered that has reached a mature stage. Significant progress has been made when considering higher order interactions of simple nature such as generalized Potts model [12] or higher order models of limited complexity through their mapping to a pair-wise one and then use existing optimization methods to perform inference [13]. In the case of real-valued variables, an alternative consists of using message passing methods like belief propagation networks [14] that have linear complexity with respect to the order of clique. Dual decomposition [15] is another alternative that provides more freedom with respect to the class of potentials that could be handled while maintaining a competitive advantage in terms of computational complexity. In [16] a more efficient variant for mapping higher order problems to pair-wise ones using principles introduced in [10] was presented with better convergence and optimality guarantees. Furthermore, in [17] an efficient method to infer solutions for higher order graphical models with submodular potentials was presented while in [18] a submodular relaxation approach was proposed for pair-wise and higher order graphical models inference. More recently [19] tackled higher order inference for graphical models which can be expressed as the minimization of partially separable function of discrete variables, while in [20] the principle of active constraints adaptively learnt over multiple iterations was adopted. Further aspects on inference, and specifically on the most commonly used optimization principles in the context of graphical models, are discussed in [21].

3. Discrete Biomedical Image Analysis

Graph-based representations have attracted the interest of the biomedical image analysis community immediately after their re-appearance in the field of vision. Discrete labeling problems like semantic image segmentation were the first to be considered. Let us first refer to some seminal works like for example the one presented in [22] where an interactive segmentation method was introduced exploiting graph cuts as optimization technique. This concept was generalized in [23] to deal with the multi-class semantic segmentation problem. The same problem was addressed through an alternative but yet powerful optimization method, the random walker algorithm in [24]. Recently tree-structured graphs were preferred for medical image segmentation targeting efficiency in infer-

ence and training [25]. Let us now briefly review our own contributions in this area where the order of cliques (pair-wise versus higher order) is used as classification criterion.

3.1. Pair-Wise Graphical Models in Biomedical Imaging

Knowledge-based image segmentation is a fundamental problem in medical imaging. The central idea is to build a statistical model learned from a set of training examples constraining the space of solutions. This is a challenging problem in the context of graphical models since prior knowledge requires rich modeling and often global constraints. In [26] prior knowledge is expressed through a point distribution representation mapped to a pair-wise probabilistic graphical model. In such a context global shape priors could be built through the concatenation of local constraints expressed differently depending on the context [27], resulting on a powerful and flexible model both in terms of learning as well as in terms of inference using generic graph-based optimization methods. A different approach was adopted in [28, 29] where the prior knowledge is expressed globally through a multi-class statistical atlas and encoded as constraint throughout the random walker algorithm [24]. Furthermore, in order to define the optimal tradeoff between data, prior and smoothness constraints a mathematically elegant and principled learning framework was introduced in [29]. The work presented in [30] bears concept similarities with the one in [28]. The central idea is to extend the prior model towards accounting for multiple hypotheses. To this end, brain tumors were clustered geometrically to a common reference space towards creating a multi-hypotheses model where each class has its own statistical atlas of appearance. During segmentation, image likelihoods were combined with hypotheses selection and associated geometric position constraints toward automatic detection, segmentation and characterization of brain tumors.

The problem of image registration has been also considered within a pair-wise graphical model. In [31] a decomposed pair-wise fully connected chain model was introduced to address the problem of linear registration. The approximate impact of the different parameters was estimated for different combination of pairs of variables that have been jointly optimized. The problem of deformable registration was considered in [32] and was formulated through the deformation of a (super-imposed to the domain) grid according to a discrete set of displacements leading to an approach that is metric-free, modular, scalable and computationally efficient. Labels were assigned to displacements and therefore used to estimate the data similarity term, while the connectivity of the graph was used to impose smoothness on the deformation field. In such a model two aspects were critical; (i) the quantization of the search space, and (ii) the selection of the similarity metric. Insights as it concerns adaptive quantization of the search space determined according to the directional uncertainties of the obtained solution were presented in [33]. In order to address the selection of the similarity function the approach of [32] has been further extended with an online metric learning paradigm in [34], where mutual saliency was used to determine the impact of different feature spaces.

In order to deal with the ill-posedness nature of the deformable registration problem, regularization constraints have been further enhanced with learned deformation priors in [35]. The central idea was to determine domain-specific graph connectivity learned from the data. Integration of geometric landmark correspondences has been also considered through a simultaneous coupling of iconic and geometric methods. In [36] a decomposed/inter-connected graphical model was presented seeking to determine simultaneously a dense deformation model to create iconic correspondences and direct correspondences between candidate landmark points. Last but not least, the problem of slice-to-volume registration was considered in [37] through an over-parameterized pair-wise graph encoding simultaneously the plane selection and the in-plane deformations in a metric-free framework.

Despite their strengths, pair-wise graphical models inherit limited expressive power since variable interactions can only be encoded through links between two variables. In a number of problems, the use of more expressive priors is a necessity both from modeling view point as well as from accuracy one.

3.2. Higher-order Graphical Models in Biomedical Imaging

Graph matching is an interesting problem in biomedical imaging since it is often used to establish meaningful correspondences between anatomical landmarks. In [38] a method for landmark based segmentation was presented exploiting third order potentials. On top of local similarity constraints, pose invariant geometric constraints were introduced involving triplets of nodes that were able to explicitly account for global pose variations. The aforementioned principle has been integrated to a conventional edge/region-based segmentation paradigm [39] through an interconnected graph seeking to simultaneously determine correspondences between a sparse probabilistic model and detected landmarks, as the one in [38], and consistent voxel-based modeling in the volume domain. In order to address computational complexity issues, the problem of compacting the prior model (reduce the shape related higher order cliques) while maintaining its expressive power, has been studied in [40] where training examples were used to determine the optimal subset of higher order constraints through parallel training of models endowed with a coordination mechanism imposing convergence to a globally optimal solution. The use of such static models, when considering more complex articulated anatomical structures, becomes inefficient. In [41] a higher order graph has been considered to encode interactions between different vertebrae of the spine. Constraints on the vertebra relative positions were modeled and learned in R^6 towards facilitating inference in new patients. In order to reduce the complexity of the model (conjunction of increased dimensionality of the label space and multiple higher order cliques), a manifold learning approach was considered in [42] where both the higher order cliques and the corresponding label space have been mapped to a low dimensional manifold on which inference was then performed.

Higher order graphs have been also used in the context of linear and deformable registration. An approach that was able to determine a linear mapping through a graph-based local deformation

model endowed with higher order planarity constraints was presented in [43]. The same principle was used in [44] to deal with the task of image-to-volume registration through a decoupled inter-connected graphical model seeking to simultaneously determine the optimal plane through a higher order model and the associate in plane deformation through a pair-wise graph as suggested in [32]. The problem of coupled segmentation/atlas-based registration within the context of brain tumor delineation was considered within an adaptive graphical model in [45]. The central idea was to consider hyper-cliques modeling the interacting segmentation and deformation labels and adaptively resample the associating grids while explicitly accounting for the uncertainties, as suggested in [32], of the obtained solution. Deformable registration within a population alignment setting was also addressed through a higher order graphical model in [46]. In order to address this objective, multiple interacting graph-optimization problems seeking to simultaneously deform all volumes were simultaneously optimized such that: (i) visual correspondences between all possible pairs of volumes are established in the common space through smooth deformations, and (ii) individual distribution of voxels in the common space are as compact as possible (through congealing like criteria). Last, but not least the problem of population registration was coupled with multi-atlas segmentation in [47]. Such model was built upon the work presented in [46] involving the simultaneous deformation of all atlases and coupled with a multi-class segmentation map seeking to penalize disagreement between segmentation labels predicted from the deformed atlases.

4. Discussion and Perspectives

In this paper we have briefly introduced graphical models in biomedical imaging. After a brief introduction motivating the interest on such representations, we discuss existing state of the art inference algorithms and conclude with a brief representative snapshot of our work carried out during the past decade towards addressing some of the pillar problems of biomedical imaging. To summarize, during the past decade the mainstream effort was dedicated to the development of efficient inference methods for low-rank (pair-wise) user-defined application-driven graphical models involving limited interactions between the nodes of the graphical models. Such models have been very powerful in addressing a number of low and mid-level computational perception problems but failed to cope with high-level tasks. One can claim that this was due to:

- constraints being imposed by the low-rank connections between graph nodes (pair-wise models offer limited interaction between variables) that was mostly the outcome of lacking efficient optimization methods to address higher order interactions.
- absence of leveraging large scale manually annotated data towards defining the optimal inference problem. Obviously recovering the optimal solution of an arbitrary objective function

that does not correspond to the one that should be used towards getting the optimal visual perception answer is not that useful.

- absence of introducing high-level knowledge on the structure of the considered graphs towards facilitating the process of describing the space of solutions and making inference feasible at reasonable computational time with strong optimality guarantees.

An interesting future perspective will be to introduce a novel discrete, data-driven, higher order, structured computational framework to address visual perception and its applications in large scale modeling and biomedical image analysis. Medicine is a field where human understandable predictions could have a much stronger adoption rate from the clinical experts. Therefore coupling the advances of machine learning able to provide quite powerful individual predictions (like for example deep learning), with human understandable structured representations (graphs learned from examples, tree structures being the derivation of structured models like for example grammars) and efficient distributed optimization methods could be a major breakthrough towards reproducing or even surpassing human intelligence.

Acknowledgments

This research was partially supported from the European Research Council Grant 259112. The authors are grateful to former and current collaborators who have contributed to the works presented in this papers.

References

- [1] D. Koller, N. Friedman, Probabilistic Graphical Models - Principles and Techniques, MIT Press, 2009.
- [2] S. Geman, D. Geman, Stochastic relaxation, gibbs distributions, and the bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* 6 (6) (1984) 721–741.
- [3] J. H. Kappes, B. Andres, F. A. Hamprecht, C. Schnörr, S. Nowozin, D. Batra, S. Kim, B. X. Kausler, T. Kröger, J. Lellmann, N. Komodakis, B. Savchynskyy, C. Rother, A comparative study of modern inference techniques for structured discrete energy minimization problems, *International Journal of Computer Vision* 115 (2) (2015) 155–184.
- [4] C. Wang, N. Komodakis, N. Paragios, Markov random field modeling, inference & learning in computer vision & image understanding: A survey, *Computer Vision and Image Understanding* 117 (11) (2013) 1610–1627.

- [5] D. Greig, B. Porteous, A. Seheult, Exact maximum a posteriori estimation for binary images, *Journal of the Royal Statistical Society. Series B (Methodological)* (1989) 271–279.
- [6] Y. Boykov, O. Veksler, R. Zabih, A variable window approach to early vision, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (12) (1998) 1283–1294.
- [7] J. Pearl, Reverend bayes on inference engines: A distributed hierarchical approach, in: *Proceedings of the National Conference on Artificial Intelligence*. Pittsburgh, PA, August 18-20, 1982., 1982, pp. 133–136.
- [8] J. S. Yedidia, W. T. Freeman, Y. Weiss, Generalized belief propagation, in: *Advances in Neural Information Processing Systems 13, Papers from Neural Information Processing Systems (NIPS) 2000*, Denver, CO, USA, 2000, pp. 689–695.
- [9] V. Kolmogorov, Convergent tree-reweighted message passing for energy minimization, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (10) (2006) 1568–1583.
- [10] N. Komodakis, G. Tziritas, N. Paragios, Performance vs computational efficiency for optimizing single and dynamic mrfs: Setting the state of the art with primal-dual strategies, *Computer Vision and Image Understanding* 112 (1) (2008) 14–29.
- [11] C. Rother, V. Kolmogorov, V. S. Lempitsky, M. Szummer, Optimizing binary mrfs via extended roof duality, in: *2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, 18-23 June 2007, Minneapolis, Minnesota, USA, 2007.
- [12] P. Kohli, M. P. Kumar, P. H. S. Torr, P^3 & beyond: Move making algorithms for solving higher order functions, *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (9) (2009) 1645–1656.
- [13] H. Ishikawa, Transformation of general binary MRF minimization to the first-order case, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (6) (2011) 1234–1249.
- [14] B. Potetz, T. S. Lee, Efficient belief propagation for higher-order cliques using linear constraint nodes, *Computer Vision and Image Understanding* 112 (1) (2008) 39–54.
- [15] N. Komodakis, N. Paragios, G. Tziritas, MRF energy minimization and beyond via dual decomposition, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (3) (2011) 531–552.
- [16] A. Fix, A. Gruber, E. Boros, R. Zabih, A hypergraph-based reduction for higher-order binary markov random fields, *IEEE Trans. Pattern Anal. Mach. Intell.* 37 (7) (2015) 1387–1395.
- [17] C. Arora, S. Banerjee, P. K. Kalra, S. N. Maheshwari, Generalized flows for optimal inference in higher order MRF-MAP, *IEEE Trans. Pattern Anal. Mach. Intell.* 37 (7) (2015) 1323–1335.

- [18] A. Osokin, D. P. Vetrov, Submodular relaxation for inference in markov random fields, *IEEE Trans. Pattern Anal. Mach. Intell.* 37 (7) (2015) 1347–1359.
- [19] A. Shekhovtsov, Higher order maximum persistency and comparison theorems, *Computer Vision and Image Understanding* 143 (2016) 54–79.
- [20] D. Khandelwal, K. Bhatia, C. Arora, P. Singla, Lazy generic cuts, *Computer Vision and Image Understanding* 143 (2016) 80–91.
- [21] N. Komodakis, P. M. Kumar, N. Paragios, (hyper)-graphs inference through convex relaxations and move making algorithms: Contributions and applications in artificial vision, *Foundations and Trends in Computer Graphics and Vision* 10 (1) (2016) 1–102.
- [22] Y. Boykov, M. Jolly, Interactive organ segmentation using graph cuts, in: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2000, Third International Conference, Pittsburgh, Pennsylvania, USA, October 11-14, 2000, Proceedings, 2000*, pp. 276–286.
- [23] Y. Boykov, G. Funka-Lea, Graph cuts and efficient N-D image segmentation, *International Journal of Computer Vision* 70 (2) (2006) 109–131.
- [24] L. Grady, Random walks for image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (11) (2006) 1768–1783.
- [25] M. G. Uzunbas, C. Chen, D. N. Metaxas, An efficient conditional random field approach for automatic and interactive neuron segmentation, *Medical Image Analysis* 27 (2016) 31–44.
- [26] A. Besbes, N. Komodakis, G. Langs, N. Paragios, Shape priors and discrete mrfs for knowledge-based segmentation, in: *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA, 2009*, pp. 1295–1302.
- [27] D. R. Chittajallu, N. Paragios, I. A. Kakadiaris, An explicit shape-constrained mrf-based contour evolution method for 2-d medical image segmentation, *IEEE J. Biomedical and Health Informatics* 18 (1) (2014) 120–129.
- [28] P. Baudin, N. Azzabou, P. G. Carlier, N. Paragios, Prior knowledge, random walks and human skeletal muscle segmentation, in: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2012 - 15th International Conference, Nice, France, October 1-5, 2012, Proceedings, Part I, 2012*, pp. 569–576.
- [29] P. Baudin, D. Goodman, P. Kumar, N. Azzabou, P. G. Carlier, N. Paragios, M. P. Kumar, Discriminative parameter estimation for random walks segmentation, in: *Medical Image Com-*

- puting and Computer-Assisted Intervention - MICCAI 2013 - 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part III, 2013, pp. 219–226.
- [30] S. Parisot, A. Darlix, C. Baumann, S. Zouaoui, Y. Yordanova, M. Blonski, V. Rigau, S. Chemouny, L. Taillandier, L. Bauchet, H. Duffau, N. Paragios, A probabilistic atlas of diffuse who grade ii glioma locations in the brain, *PLoS ONE* 11 (2016) 1.
- [31] D. Zikic, B. Glocker, O. Kutter, M. Groher, N. Komodakis, A. Kamen, N. Paragios, N. Navab, Linear intensity-based image registration by markov random fields and discrete optimization, *Medical Image Analysis* 14 (4) (2010) 550–562.
- [32] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, N. Paragios, Dense image registration through mrfs and efficient linear programming, *Medical Image Analysis* 12 (6) (2008) 731–741.
- [33] B. Glocker, N. Paragios, N. Komodakis, G. Tziritas, N. Navab, Optical flow estimation with uncertainties through dynamic mrfs, in: 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA, 2008.
- [34] Y. Ou, A. Sotiras, N. Paragios, C. Davatzikos, DRAMMS: deformable registration via attribute matching and mutual-saliency weighting, *Medical Image Analysis* 15 (4) (2011) 622–639.
- [35] B. Glocker, N. Komodakis, N. Navab, G. Tziritas, N. Paragios, Dense registration with deformation priors, in: Information Processing in Medical Imaging, 21st International Conference, IPMI 2009, Williamsburg, VA, USA, July 5-10, 2009. Proceedings, 2009, pp. 540–551.
- [36] B. Glocker, A. Sotiras, N. Komodakis, N. Paragios, Deformable medical image registration: Setting the state of the art with discrete methods, *Annual Review of Biomedical Engineering* 13 (2011) 219–244.
- [37] E. Ferrante, N. Paragios, Non-rigid 2d-3d medical image registration using markov random fields, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2013 - 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part III, 2013, pp. 163–170.
- [38] C. Wang, O. Teboul, F. Michel, S. Essafi, N. Paragios, 3d knowledge-based segmentation using pose-invariant higher-order graphs, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2010, 13th International Conference, Beijing, China, September 20-24, 2010, Proceedings, Part III, 2010, pp. 189–196.
- [39] B. Xiang, J. Deux, A. Rahmouni, N. Paragios, Joint model-pixel segmentation with pose-invariant deformable graph-priors, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2013 - 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part III, 2013, pp. 267–274.

- [40] N. Komodakis, B. Xiang, N. Paragios, A framework for efficient structured max-margin learning of high-order MRF models, *IEEE Trans. Pattern Anal. Mach. Intell.* 37 (7) (2015) 1425–1441.
- [41] S. Kadoury, H. Labelle, N. Paragios, Automatic inference of articulated spine models in CT images using high-order markov random fields, *Medical Image Analysis* 15 (4) (2011) 426–437.
- [42] S. Kadoury, H. Labelle, N. Paragios, Spine segmentation in medical images using manifold embeddings and higher-order MRFs, *IEEE Trans. Med. Imag.* 32 (7) (2013) 1227–1238.
- [43] V. Fecamp, A. Sotiras, N. Paragios, Modular linear iconic matching using higher order graphs, in: *12th IEEE International Symposium on Biomedical Imaging, ISBI 2015, Brooklyn, NY, USA, April 16-19, 2015, 2015*, pp. 1097–1101.
- [44] E. Ferrante, V. Fecamp, N. Paragios, Slice-to-volume deformable registration: efficient one-shot consensus between plane selection and in-plane deformation, *Int. J. Computer Assisted Radiology and Surgery* 10 (6) (2015) 791–800.
- [45] S. Parisot, W. M. W. III, S. Chemouny, H. Duffau, N. Paragios, Concurrent tumor segmentation and registration with uncertainty-based sparse non-uniform graphs, *Medical Image Analysis* 18 (4) (2014) 647–659.
- [46] A. Sotiras, N. Komodakis, B. Glocker, J. Deux, N. Paragios, Graphical models and deformable diffeomorphic population registration using global and local metrics, in: *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2009, 12th International Conference, London, UK, September 20-24, 2009, Proceedings, Part I, 2009*, pp. 672–679.
- [47] S. Alchatzidis, A. Sotiras, N. Paragios, Discrete multi atlas segmentation using agreement constraints, in: *British Machine Vision Conference, BMVC 2014, Nottingham, UK, September 1-5, 2014, 2014*.