Energetic lattice based optimization

Bangalore Ravi Kiran

March 25, 2015

Hierarchies are a powerful tool for dimensionality reduction and has been in a area of intense research for the past decade. In the domain of image segmentation we have seen it applied to hierarchical segmentation with its prominence in the Berkeley dataset in the form of the Ultrametric Contour Map [2]. Hierarchies or trees are also an efficient model in the domain of statistical learning theory, development due the seminal work of Breiman [4] on classification and regression trees (CART). CART was later developed into the domain of source coding [11] for quantizers and also in the domain of segmentation by pruning partition trees [10]. In a related work on describing an hierarchy-energy dependent image descriptor called λ-cuts, Guigues [6] provides a way to calculate an optimal hierarchy of partitions. The thesis begins with this problem of selection an optimal cut by dynamic programming and develops the various conditions on the energy and the space of partitions to obtain a tractable solution.

We can quickly summarize the optimization problem on the space of hierarchies as follows:

$$\underset{\pi \in \Pi(E,H)}{\text{minimize}} \sum_{S \in \pi} \omega_{\varphi}(S) \;\text{ subject to} \sum_{S \in \pi} \omega_{\partial}(S) \leq C \tag{1}$$

where π is any cut of H (partition composed of subsets from the hierarchy H, $\Pi(E,H)$ is the set of all possible cuts/partitions extractable from the hierarchy H, $\omega_{\varphi}(S)$ is the objective function or energy, while $\omega_{\partial}(S)$ is the constraint function, while S is any element of a partition in H. C is a constraint value imposed on a cut $\omega_{\partial}(\pi) < C$. We are searching for the least energetic cut from a hierarchy if we ignored the constraint. For the constrained problem we resolve it by considering the Lagrangian like in [10].

1 Braids and their minimization

1.1 Dynamic Programming

Dynamic program consists in solving a structured complex problem by decomposing it into smaller simpler subproblems. These subproblems require to be overlapping in nature, such that a solution to a subproblem is calculated only once and serves to solve a larger subproblem. Furthermore this partial solution when aggregated with others, produces a global solution. The aggregation is quite often has a successive approxi- mation interpretation. The algorithm will implicitly have to solve subproblems appearing at higher scales before some lower scales. This means, some nodes appearing early lower down in the hierarchy might be part of a global solution. The optimal cut in [4], [10], [6], is calculated by aggregating local optima. The local optimum at class S either choses the parent $\{S\}$, or the disjoint union of the optimums over the its children as shown in equation 2.

$$\pi^*(S) = \begin{cases} \{S\}, & \text{if } \omega(S) \leq \text{comp}(\omega(\pi^*(a))), a \in \pi(S) \\ \bigsqcup_{a \in \pi(S)} \pi^*(a), & \text{otherwise} \end{cases} \tag{2}$$

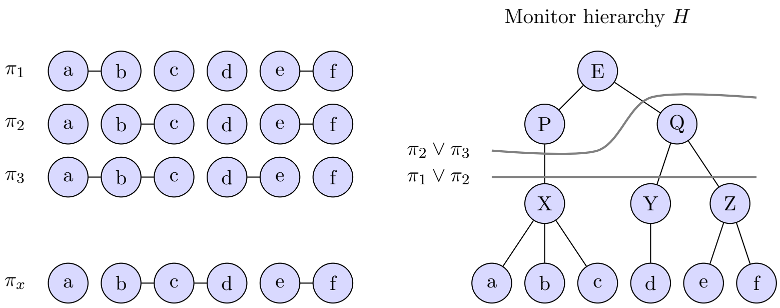

Figure 1: Space E is partitioned into leaves $\{a, b, c, d, e, f\}$. The family $B_1 = \{\pi_1, \pi_2, \pi_3\}$ forms a braid, whose pairwise supremum is indicated on the dendrogram. Note that $\pi_1(X), \pi_2(X)$ have a common parent X, but $\pi_2(Q), \pi_3(Q)$ a common grand parent Q. However the family $\pi_x \cup B_1$ is not a braid since $\pi_3 \vee \pi_x$ gives the whole space E.

1.2 Extension to a new family: Braids

We prove in the thesis that the dynamic program in [4], [10], [6], is in fact valid across a larger family of partitions, the braids. A braid is a family of partitions B, where the pairwise refinement supremum of any two elements is a cut of in some hierarchy $\Pi(E, H)$. This leads to the more formal definition:

Definition 1 *Let $\Pi(E)$ be the complete lattice of all partitions of set E; let H be a hierarchy in $\Pi(E)$. A braid B of monitor H is a family in $\Pi(E)$ where the refinement supremum of any pair of distinct partitions $\pi_1, \pi_2 \in B$ is a cut of H, other than $\{E\}$, that is in, $\Pi(E, H) \setminus \{E\}$:*

$$\forall \pi_1, \pi_2 \in B \quad \Rightarrow \pi_1 \vee \pi_2 \in \Pi(E, H) \setminus \{E\} \tag{3}$$

In figure 2 we show a simple example how a hierarchy is augmented with more partial partitions to produce a braid of partitions. In figure 1 we demonstrate a simple example of a braid family with the dendrogram corresponding to its monitor hierarchy. As we can see the classes of partitions π_1, π_2 are neither nested nor disjoint, and basically correspond to different segmentation hypotheses that exist in the stack of segmentations. The set of all cuts of a braid B is denoted by $\Pi(E, B)$. A braid may also contain its monitor H, though this is not necessary. On the other hand, any hierarchy is a braid with itself as monitor. A braid cannot be represented by a single saliency function, except when it reduces to a hierarchy whose classes are connected sets.

1.3 Compositions for DP: h-increasingness

h-increasingness is a property of energies, which preserves the optimal substructure in extracting the minimal cut so that one can use a dynamic program to solve it. It states that the ordering of energies is preserved under concatenation of partial partitions (Figure 4).

Definition 2 *(h-increasingness) Let $\pi_1(S)$, $\pi_2(S)$ be two different p.p. of the same support $S \in E$, be a family of disjoint supports over E. Let π_0 be any partial partition in $\mathcal{D}$ other than $\pi_1(S), \pi_2(S)$. A finite singular energy ω on the partial partitions $\mathcal{D}(E)$ is h-increasing when for every triplet $\{\pi_1(S), \pi_2(S), \pi_0\}$ one has:*

$$\omega(\pi_1(S)) \leq \omega(\pi_2(S)) \Rightarrow \omega(\pi_1(S) \sqcup \pi_0) \leq \omega(\pi_2(S) \sqcup \pi_0) \tag{4}$$

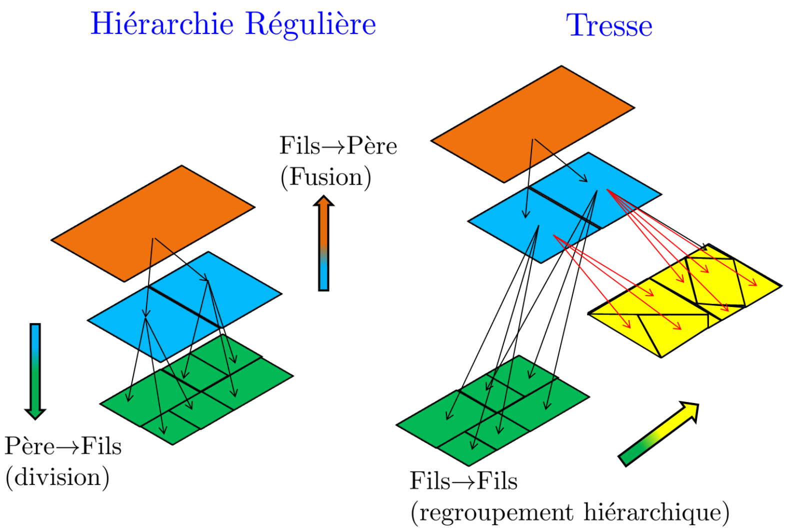

Figure 2: Comparing hierarchies and Braids: demonstration of partition structure. Hierarchies are characterized by two types of refinements: Parent→Children (division/fragmentation) and children→parent (fusion/merging). The braid introduces a third type of refinement operation which conserves the parent, namely the hierarchical re-merging or reorganization. This operation helps keep the substructure required to solve the optimal cut problem while providing a richer space to partitions to search for an optimum.

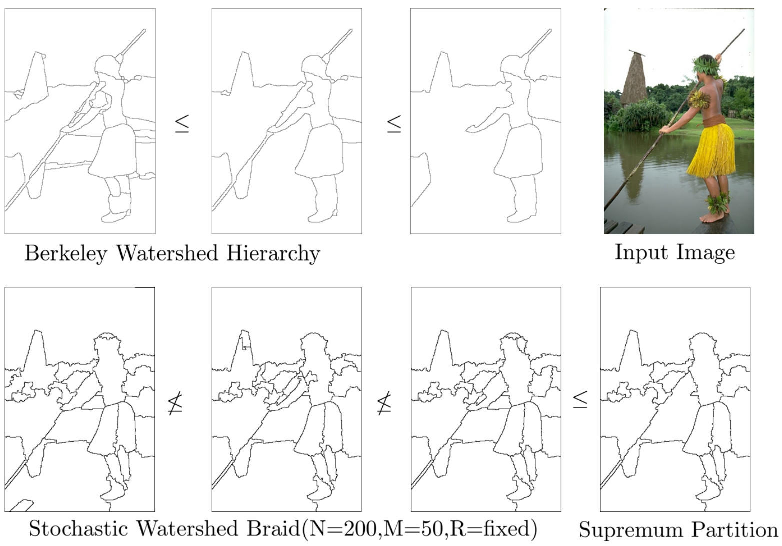

Figure 3: Top: Ultrametric contour map [2], Bottom: braid of partitions. Braids of partitions were produced from multiple instances of random marker based stochastic watershed, with same number of regions. The supremum or monitoring partition, corresponding to these unordered family of partitions is shown. Braids help reorganize partial refinement between partitions.

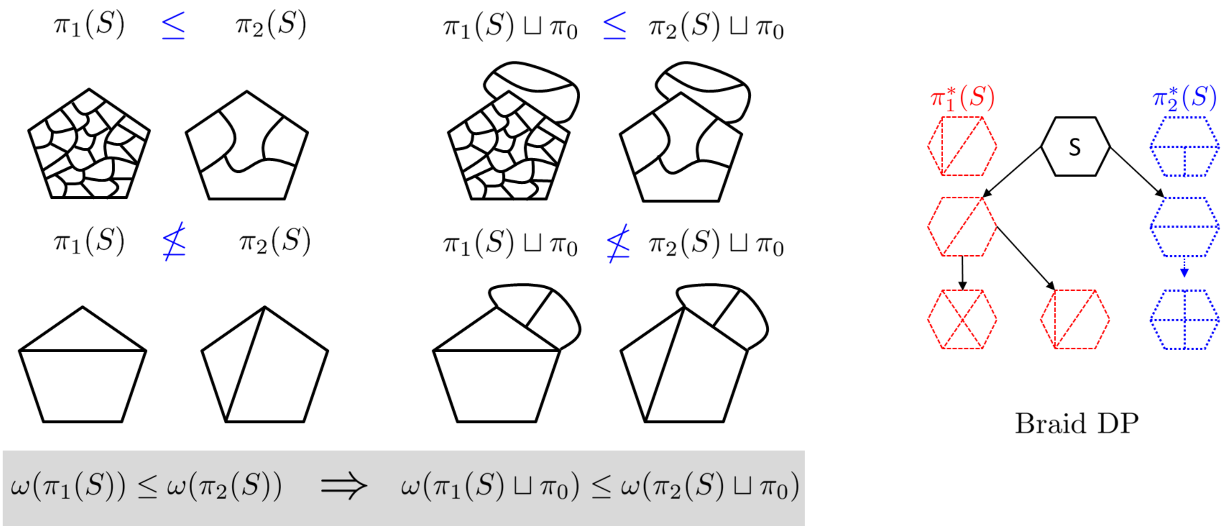

Figure 4: Left: h-increasingness for HOP (top) versus BOP (bottom). Right: An elementary step of the dynamic program on a partial braid over a support S. The optimal partial partition is the minimum across the partial optima and S, i.e. $\omega(S), \omega(\pi_1^*(S)), \omega(\pi_2^*(S))$.

In implication (4) when the inequality is made strict, we have what we call strict h-increasingness. h-increasingness was first introduced in [7], which generalized the condition of *separable energies* of Guigues [6]. Separability in equation (2), is obtained by replacing comp$(\cdot)$ by a sum of the energies of the constituent classes of a partial partition, to calculate the energy of the partial partition. We can also perform a composition by supremum [12], [13].

Both laws are indeed particular cases of the classical Minkowski expression

$$\omega(\pi(S)) = \left[\sum_{u=1}^{q} \omega(T_u)^{\alpha} \right]^{\frac{1}{\alpha}} \tag{5}$$

which is a norm in $\mathbb{R}^n$ for $\alpha \geq 1$. Even though over partial partitions $\mathcal{D}(E)$, it is no longer a norm, it yields strictly h-increasing energies for all $\alpha \in]-\infty, +\infty[$:

1.4 Modeling the minimum of DP: Energetic Lattice

Given the problem of finding an optimal cut, we review separately the requirement of obtaining a unique solution. On the HOP, this has been enforced by many authors [4], [10], [6], [13], [1] as a partition which is either the largest or the smallest, amongst optimal cuts with the same energy. A hierarchy can have multiple cuts with the same minimal energy, and to ensure a unique solution we use the general condition of singularity, which requires that any energy ω defined on the partial partitions of a braid B be such that the energies $\omega(\pi(S))$ of all p.p. $\pi(S)$ of H are either strictly smaller, or strictly greater, than the energies of their supports S: $\forall\, \pi(S) \in \Pi(S), \omega(\{S\}) < \omega(\pi(S))\}$ or $\omega(\{S\}) > \omega(\pi(S))\}$, [9], [8].

Consider now two partial partitions $\pi(S), \pi'(S)$ over support S, which is also their refinement supremum $S = \pi(S) \vee \pi'(S)$ (see Figure 5). One may assess that partition π_1 to be less energetic than π_2 for an energy ω i.e. $\pi_1 \preceq_\omega \pi_2$ when in each class of supremum $\pi_1 \vee \pi_2$ the energy of the partial partition of π_1 is smaller or equal to that of π_2.

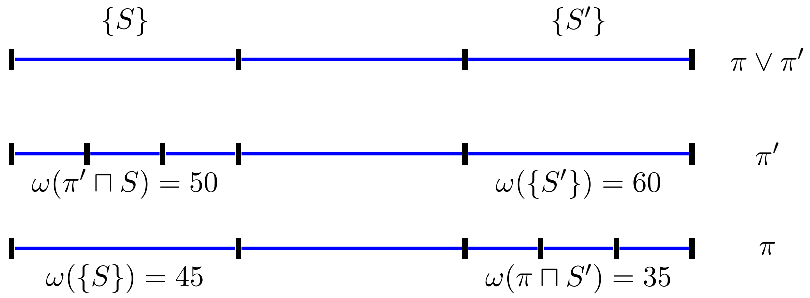

Figure 5: An example of energetic ordering: We have $\pi \preceq_\omega \pi'$ since in each class of $\pi \vee \pi'$, the energy ω of π is less than or equal to that of π'.

$$\pi_1 \preceq_\omega \pi_2 \Leftrightarrow \{S \in \pi_1 \vee \pi_2 \Rightarrow \omega(\pi_1 \sqcap \{S\}) \leq \omega(\pi_2 \sqcap \{S\})\} \tag{6}$$

The relation $\preceq_\omega$ called energetic ordering, is an ordering relation for all singular energies ω, if and only if the family Π is the set $\Pi(\omega, E, B)$ of all cuts of a braid B [8].

2 Conclusion

The thesis generalizes the dynamic program used first in pruning decision trees by Breiman et al. [4], Salembier [10], Guigues et al. [5, 3] for non-linear energies with the *h-increasingness* property by firstly introducing the braids of partitions to model the solution space, secondly the h-increasing energies to generate valid energies that preserve the DP substructure, thirdly the energetic lattice structure, that models the minimum obtained by the DP and explicating its local-global structure.

In the second part of the thesis we use the Everett's theorem to show that λ-cuts in case of Guigues [6], and optimal prunings of Salembier [10] provide only an upper-bound on minimal objective energy. This was explicated further by the dependence of the constraint function values on the Lagrangian multiplier, and also the possibility of the non-existence of multipliers for certain constraint values and vice versa. The constraint function values were shown to be lattice structured and not varying continuously. This motivated the use of an energetic lattice framework. This gave us two ways to enforce of a constraint providing us new models of constrained optimization on hierarchies and braids.

Lagrangian Minimization model enforces a numerical constraint on energy, without referring to the partition structure. The Energetic lattice generalized the Lagrangian model while working in an extended space of braids. Secondly a lattice structured Lagrange Minimization and a Class dependent constraint models were introduced, which enforce the constraint in the form a partition, by using the energetic ordering. This does not involve any numerical constraint function, but one that is based on the energetic lattice. [8]

References

[1] Huseyin Gokhan Akcay and Selim Aksoy. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE T. Geoscience & Remote Sensing*, 46(7):2097–2111, 2008.

[2] Pablo Arbelaez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. PAMI*, 33(5):898–916, may 2011.

[3] Coloma Ballester, Vicent Caselles, Laura Igual, and Luís Garrido. Level lines selection with variational models for segmentation and encoding. *Journal of Mathematical Imaging and Vision*, 27(1):5–27, 2007.

[4] Leo Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees.* Wadsworth, 1984.

[5] L. Guigues. *Modèles multi-échelles pour la segmentation d'images.* PhD thesis, Universitè de Cergy-Pontoise, Dec 2003.

[6] Laurent Guigues, Jean Pierre Cocquerez, and Hervé Le Men. Scale-sets image analysis. *International Journal of Computer Vision*, 68(3):289–317, 2006.

[7] B. Ravi Kiran and Jean Serra. Global-local optimizations by hierarchical cuts and climbing energies. *Pattern Recognition*, 47(1):12 – 24, 2014.

[8] Bangalore Ravi Kiran. *Energetic Lattice based optimization.* PhD thesis, Université Paris-Est, LIGM-A3SI, ESIEE, 2014.

[9] Bangalore Ravi Kiran and Jean Serra. Braid of partitions. *Accepted to ISMM 2015*, 2015.

[10] P Salembier and L Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. on Image Processing*, 9(4):561–576, 2000.

[11] Y. Shoham and A Gersho. Efficient bit allocation for an arbitrary set of quantizers. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(9):1445–1453, Sep 1988.

[12] Pierre Soille. Constrained connectivity for hierarchical image partitioning and simplification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(7):1132–1145, 2008.

[13] S. Valero. *Hyperspectral image representation and Processing with Binary Partition Trees.* Ph.d., Universitat Politècnica de Catalunya (UPC), 2011.