Pattern Recognition 47 (2014) 12–24

Contents lists available at ScienceDirect

Pattern Recognition

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

Global–local optimizations by hierarchical cuts and climbing energies

B. Ravi Kiran *, Jean Serra

Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, A3SI, ESIEE

ARTICLE INFO

Available online 23 May 2013

Keywords:
Hierarchical segmentation
Climbing optimization
Mathematical morphology
Energy minimization
Dynamic programming

ABSTRACT

Hierarchical segmentation is a multi-scale analysis of an image and provides a series of simplifying nested partitions. Such a hierarchy is rarely an end by itself and requires external criteria or heuristics to solve problems of image segmentation, texture extraction and semantic image labelling. In this theoretical paper we introduce a novel framework: hierarchical cuts, to formulate optimization problems on hierarchies of segmentations. Second we provide the three important notions of *h-increasing*, *singular*, and *scale increasing* energies, necessary to solve the global combinatorial optimization problem of partition selection and which results in linear time dynamic programs. Common families of such energies are summarized, and also a method to generate new ones is described. Finally we demonstrate the application of this framework on problems of image segmentation and texture enhancement.

1. Introduction

To segment an image by a global constraint classically means to associate a numerical energy with every possible partition of the space where this image is defined. The best partition is then that which minimizes the energy. Is this meaningful? Let us suppose that the energies range from 0 to 10^3. Using the formula for the classical Bell's number, a digital square of 5×5 pixels has 4.6×10^{18} different partitions possible [9]. Each value of energy thus maps onto millions of billions (4.6×10^{15}) of partitions. What do we minimize here? Which implicit assumptions underlie the methods which give a unique minimal cut?

There are only two ways for obtaining (or hoping) uniqueness: by limiting the number of partitions and by imposing constraints to the energy. To limit the number of partitions, we can think of cuts on hierarchies, which provide strong restrictions. To constrain the energy, we can try and replace the lattice of the integers by another one, more comprehensive, e.g. a lattice of partitions, and make hold the minimizations on it. But how to create a lattice of partitions from a given energy? Which conditions must we introduce? And if uniqueness is finally ensured (the lattice structure is precisely made for that), how to reach the minimal cut in the maze of all partitions? By means of which vital thread?

There have been several approaches to global constraints for optimization. There are two methods we contrast here: First, the graph cuts based optimization, popularized by Boykov [7], second, partition selection from hierarchies of partitions. The former emphasize the use of seeds, in addition, they view the space as a one scale structure. This perspective is illustrated by the search for a maximum flow in a directed graph, whose segmentation applications include the optimization of conditional random field (CRF) [22]. The latter approaches emphasize the scaling of the space by means of hierarchies, and attach less importance to labelling questions, in a first step at least.

This paper focusses on the second type of global constraints, which are approached from the viewpoint of hierarchical cuts (*h*-cuts) theory.[1] A hierarchy, or pyramid, of image segmentations is understood as a series of progressive simplified versions of an initial image, which result in increasing partitions of the space. How can these partitions cooperate and summarize the hierarchy into a unique cut, optimal in some sense. Three questions arise here, namely:

1. Given a hierarchy H of partitions and an energy ω on the partial partitions, how to combine the classes of this hierarchy for obtaining a new partition that minimizes ω, and which can be determined easily?
2. When one energy ω depends on an integer j, i.e. $\omega=\omega^j$, how to generate a sequence of optimal partitions that increase with j, which therefore should form a optimal hierarchy?
3. Most of the segmentations involve several features (colour, shape, size, etc.) that we can handle with different energies. How to combine them?

These questions have been taken up by several authors, over many years, and by various methods. The most popular energies ω for hierarchical partitions derive from that of Mumford and Shah

* Corresponding author. Tel.: +33 145926775; fax: +33 145926699.
E-mail addresses: kiranr@esiee.fr (B. Ravi Kiran), j.serra@esiee.fr (J. Serra).

[1] This work received funding from the Agence Nationale de la Recherche through contract ANR-2010-BLAN-0205-03 KIDIKO.

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.05.012

[25], in which a data fidelity term is summed up with a boundary regularization term. The optimization turns out to be a trade-off between these two constraints. In [20], for example, Koepfler et al. build a pyramid of segmentations, from fine to coarse, by progressively giving more and more weight to the second term of Mumford and Shah functional. They stop the region growing when a certain number of regions is reached. The method initiated by Salembier and Garrido for generating thumbnails rests on the same type of energy [33] and also in [34]. They interpret the optimal cut as the most accurate image simplification for a given compression rate. The approach has been extended to additive energies by Guigues et al. [17]. It is always assumed, in all these studies, that the energy of any partial partition equals the sum of the energies of its classes, which considerably simplifies the combinatorial complexity, and answers the above two questions 1 and 2.

However, one can wonder whether additivity is the very underlying cause of the simplifications, since Soille's constrained connectivity [38], where the addition is replaced by the supremum, satisfies similar properties. Finally, one finds in literature a third type of energy, which holds on nodes only, and no longer on partial partitions. It appears in the method for labeling of Arbeláez [3], or in the studies of Akçay and Aksoy, in [1]. And again, it yields optimal cuts.

Is there a common denominator to all these approaches, more comprehensive than just additivity, and which explains why they always lead to unique optima? The following paper is a theoretical attempt to delimit this central concept, and to give answers to the above questions from (1) to (3). The theory is established in Sections 3–5, Section 6 presents the algorithms, which are then applied to the two main families of climbing energies in Section 7. Before the conclusion, the approach is extended to partial optimization in Section 9, and some bridges between graph cuts and hierarchical optimizations are given in Section 10.

2. Basic notions: hierarchies and partitions

This section provides the background required to understand this paper. The usual distinction between continuous and digital spaces is not appropriate for the general theory developed in Sections 3–6. What is actually needed reduces to the two following hypotheses, which are assumed over the whole paper:

1. the space E to partition is topological and
2. the smallest partition π_0 of E has a finite number of classes.

The first assumption allows us to speak of frontiers between classes or edges (This may not be necessary always). The second one aims to avoid fractalities and to permit various inductions, in Proposition 3.2 and in Algorithm 1, among others. Some additional hypotheses are introduced when the energies are particularized in Section 7, e.g. "the classes are connected sets", or "the edges are simple arcs of $\mathbb{R}^2$". None of these assumptions are specific to image analysis. Space E may be the concern of parameters, semantic entities, grammars, NASDAQ quotations, or chamber music as well.

2.1. Partitions, partial partitions

Intuitively, a partition of E of the space under study (Euclidean, digital, graph, or else) is a division of E into regions that do not overlap, and whose union restores E in its entirety. These regions are called classes. More formally, one obtains the classes of a partition by means of an extensive mapping $S : E \to \mathcal{P}(E)$ such that

$$x, y \in E \Rightarrow S(x) = S(y) \text{ or } S(x) \cap S(y) = \varnothing.$$

Below, the symbols S, T stand for classes, and π for partitions. Partition π_1 is smaller than partition π_2 when each class of π_1 is included in a class of π_2. This condition provides an ordering on the partitions, called refinement, which in turn induces a complete lattice. This is equivalent to the ultrametric [29].

Following Ronse [30], a partition $\pi(S)$ associated with a set $S \in \mathcal{P}(E)$ is called *partial partition* (*in short p.p.*) of E of support S. In particular, the partial partition of S into the single class S is denoted by $\{S\}$. The family of all partial partitions of set E is denoted by $\mathcal{D}(E)$, or simply by $\mathcal{D}$.

2.2. Hierarchies of partitions

A hierarchy H is a chain of partitions π_i, i.e.

$$H = \{\pi_i, 0 \le i \le n \mid i \le k \le n \Rightarrow \pi_i \le \pi_k\}, \tag{1}$$

where π_0 is the finest partition and π_n is the partition $\{E\}$ of E in a single class. The classes of π_0 are called the *leaves* and E is the *root*. Since the number of leaves of π_0 is finite (as we have assumed above), the number n of different partitions of H is also finite.[2] The intermediary classes are called *nodes*. If the q classes of the partition $\pi(S)$ are $\{T_u, 1 \le u \le q\}$, one writes

$$\pi(S) = T_1 \sqcup .. T_u .. \sqcup T_q,$$

where the symbol $\sqcup$ indicates that the classes are concatenated. Given two p.p. $\pi(S_1)$ and $\pi(S_2)$ having disjoint supports, $\pi(S_1) \sqcup \pi(S_2)$ is the p.p. whose classes are either those of $\pi(S_1)$ or those of $\pi(S_2)$.

Let $S_i(x)$ be the class of partition π_i of H at point $x \in E$. Expression (1) means that at each point $x \in E$ the family $\{S_i(x), x \in E, 0 \le i \le n\}$ of those classes $S_i(x)$ that contain x forms a finite chain of nested sets from the leaf $S_0(x)$ to E

$$\mathcal{S}(x) = \{S_i(x), 0 \le i \le n\}. \tag{2}$$

Conversely, according to a classical result [8], a family $\{S_i(x), x \in E, 0 \le i \le n\}$ of indexed sets generates the classes of a hierarchy iff, for $i \le j$ and $x, y \in E$

$$S_i(x) \subseteq S_j(y) \text{ or } S_i(x) \cap S_j(y) = \varnothing, \tag{3}$$

conditions which mean that the classes form an ultra-metric space [4,23]. The partitions of a hierarchy may be represented by their classes, via a dendrogram, i.e. a tree where each node of bifurcation is a class S, or by their frontiers, via the saliency map of the edges, which indicates the level in the hierarchy when an edge disappears [27,14]. The first representation is depicted in Fig. 1 and the second one in Fig. 2. The classes of π_{i-1} at level $i-1$ which are included in class S_i of level i are said to be *the sons* of S_i. Clearly, the descendants of each node S form in turn a hierarchy $H(S)$ of root S, which is included in the complete hierarchy $H = H(E)$. One denotes by $\mathcal{S}(E)$, or just $\mathcal{S}$, the set of all classes S of all partitions involved in H.

The hierarchy can be loosely seen as a set of partitions containing superpixels of increasing sizes. Here we do not use the superpixel terminology, and prefer to distinguish between the class and partial partition.

2.3. Generating hierarchies of segmentations

In the paper, the focus is not on the methods for obtaining hierarchies of segmentations, they are considered as inputs. The main techniques for hierarchical segmentation include the various Matheron semi-groups of connected filters (openings, alternating

[2] One could argue that all components being finite, the underlying space E does not need to be infinite. However, some problems require a finite number of leaves embedded in a continuous space, e.g. ground truth by distance function [21]. Similar applications on distance function found in [15].