Pattern Recognition 84 (2018) 237–250

Contents lists available at ScienceDirect

Pattern Recognition

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

Binary Partition Tree construction from multiple features for image segmentation☆

Jimmy Francky Randrianasoa[a,b,*], Camille Kurtz[c], Éric Desjardin[b], Nicolas Passat[b]

[a] CNRS, ICube, Université de Strasbourg, France
[b] CReSTIC, Université de Reims Champagne-Ardenne, France
[c] LIPADE, Université Paris Descartes, France

ARTICLE INFO

Article history:
Received 4 September 2017
Revised 24 May 2018
Accepted 1 July 2018

Keywords:
Binary Partition Tree
Morphological hierarchies
Multiple features
Graph-based image processing
Image segmentation

ABSTRACT

In the context of image analysis, the Binary Partition Tree (BPT) is a classical data structure for the hierarchical modelling of images at different scales. BPTs belong both to the families of graph-based models and morphological hierarchies. They constitute an efficient way to define sets of nested partitions of image support, that further provide knowledge-guided reduced research spaces for optimization-based segmentation procedures. Basically, a BPT is built in a mono-feature way, i.e. for one given image, and one given metric, by merging pairs of connected image regions that are similar in the induced feature space. Our goal is to design a new family of BPTs, dealing with the need to directly manage multiple features within its building process. Then, we propose a generalization of the BPT construction framework, allowing one to embed multiple features. The cornerstone of our approach relies on a collaborative strategy used to establish a consensus between different metrics, thus enabling to obtain a unified hierarchical segmentation space. In particular, this provides alternatives to the complex issue of metric construction from several —possibly non-comparable— features. To reach that goal, we first revisit the BPT construction algorithm to describe it in a graph-based formalism. Then, we present the structural and algorithmic evolutions and impacts when embedding multiple features in BPT construction. Final experiments illustrate how this multi-feature framework can be used to build BPTs from multiple metrics computed through the (potentially multiple) image content(s).

1. Introduction

1.1. Context

In image processing and analysis, segmentation is a crucial task. The concept of segmentation is also quite generic from various points of views: in terms of semantics (from low-level definition of homogeneous areas to high-level extraction of specific objects), in terms of definition (object vs. background or total partition of the image support), and in terms of algorithmics (region-based or contour-based approaches).

In this context, morphological hierarchies propose a wide range of data structures for modelling images at various scales, allowing for the definition of connected operators [1]. Mainly based on the theoretical frameworks of graphs and mathematical morphology [2] [3, Chapters 3, 7, 9], these approaches have already proved their efficiency in many imaging applications. (The algorithms to build and handle them are generally of linear or quasi-linear time and space complexity.) Their very principle is to embed images in a dual spatial / spectral representation space, composed of shapes (i.e. spectrally homogeneous and spatially coherent regions) together with their spatial (neighbouring) and hierarchical (inclusion) relations. These representations offer a structured space to find the best regions / scales according to the applicative objective using, e.g., high-level features to describe the image regions and their content.

Among these representations, the Binary Partition Tree (BPT) [4] is a hierarchical representation of an image modelled as a tree structure, where each node is a connected region. Each of these nodes is either a leaf —an elementary region— or models the union of the regions of its two children nodes. The root is the node corresponding to the entire support of the image. Practically, a BPT is built from its leaves, provided by an initial partition of the image support, to its root, in a bottom-up fashion. This is done by iteratively choosing and merging two adjacent regions which minimize

☆ This work was supported by the French *Agence Nationale de la Recherche* under Grants ANR-12-MONU-0001, ANR-15-CE23-0009 and ANR-17-CE23-0015.

* Corresponding author at: ICube UMR 7357 - Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie, 300 bd Sébastien Brant - CS 10413 - F-67412 Illkirch Cedex (France).

E-mail address: randrianasoa@unistra.fr (J.F. Randrianasoa).

https://doi.org/10.1016/j.patcog.2018.07.003
0031-3203/

a merging criterion, based on a given metric, computed between them. The BPT structure allows users to explore the image at different scales. It can be used for various tasks such as segmentation, image retrieval, object recognition and visual browsing.

As other hierarchical structures, the BPT was mainly designed to process one image at a time. Furthermore, in contrast with most of them (e.g., component-trees, trees of shapes) that are intrinsically defined from the image content, the BPT is also designed to embed an extrinsic metric that is used, together with the image, to build a mixed image / knowledge model. In other words, a BPT is generally built for one image and one metric.

1.2. Motivations and contributions

The BPT is already known as a relevant data structure for the design of image processing and analysis tools, e.g. for video analysis, remote sensing or medical imaging. Until now, the way to build this data structure has remained mostly limited to a *one image, one metric* paradigm.

Indeed, on the one hand, the metric —i.e., the merging criterion for successive node merging— is a scalar function. When several kinds of information (e.g., colorimetric and geometric heterogeneity) are relevant, it is then required to find a way of fusing them as a unique metric. This complex task has a strong influence on the data structure construction. Thus, it has to be carefully carried out by the expert-user before the very construction process.

On the other hand, the BPT construction deals with one input image. This means that the handling of several images generally has to be dealt with by artificially creating a "super-image", or by fusing beforehand multiple information from various spectral bands into a single metric.

Our purpose is to provide an algorithmic process for BPT construction that goes beyond these current —image and metric— limitations, leading to a notion of multi-feature BPT (MBPT, for brief).

The main idea of our approach is to consider that k images, coupled with l metrics, lead to $n = k.l$ features. The cornerstone of the construction of a multi-feature BPT then consists of providing a way to involve all these features into the construction process, i.e. to make them interact for the definition of a relevant node merging order, and to obtain as output a unified hierarchical segmentation space.

In other words, our main contribution is the proposal of a generalized algorithmic framework for (M)BPT construction, for handling *many metrics and / or many images*. This algorithmic contribution extends the standard BPT construction (retrieved by setting $k = 1$ and $l = 1$). It also encompasses a wide range of cases, e.g., various metrics on one image ($k = 1$ and $l > 1$), one metric on many images ($k > 1$ and $l = 1$), or any combination of metrics on many images. Although being algorithmic, it is important to notice that our contribution is not directly related to image processing. We aim at providing an extended way to build a data structure that describes a hierarchy of partitions of an image. In other words, this provides a reduced, knowledge-guided research space, that can be further involved in various kinds of image analysis procedures: segmentation, classification, simplification, browsing... Such procedures fall, however, out of the scope of this article, and indeed constitute application cases of our proposal.

To reach our goal, some secondary contributions are proposed:

- First, we revisit the standard algorithmic process of BPT construction. In particular, we split its fundamental graph-based expression (namely a graph-collapsing procedure) from its knowledge-based layers (image topology, metric definition, merging policies, etc.) This preliminary analysis is developed in Section 3.
- Second, we describe how the basic BPT construction framework can be generalized to handle multiple features. To this end, we identify the data structure requirements (Section 4.1), and their algorithmic side effects (Section 4.2). This leads to a theoretical algorithmic framework.
- Third, we propose a practical description of this algorithmic framework. Some technical implementation details are provided in Section 5. For the sake of reproducibility, we also provide an open-source library[1] for the creation of MBPTs, which constitutes a technological contribution of this work. A complexity analysis (Section 5.2) describes the induced space and time cost increases of MBPT vs. BPT construction.

This work is concluded by application examples in the domain of satellite image analysis, in Section 6. In particular, we compare the behaviour of standard BPTs to MBPTs handling several images or several metrics.

The remainder of this article —which is an extended and improved version of the conference paper [5]— is composed of a synthetic state of the art of graph-based, hierarchical and multi-image segmentation, in Section 2; and a conclusion that emphasizes the perspectives of this work, in Section 7.

2. Related works

2.1. Graph-based and hierarchical image segmentation

Image analysis problems, and in particular segmentation, are often considered in a discrete way via some concepts of graph theory. Practically, image points (i.e. pixels, voxels) are considered as the vertices of a graph, whereas the spatial / neighbouring relations between them are modelled by graph edges. This paradigm, democratized since the early 1970's [6], led to the development of a wide range of segmentation approaches, based on basic graph manipulations.

In this context, image segmentation could be viewed as a partial (e.g., region growing [7]), or total partitioning problem with monotonic (e.g., watersheds [8]) or non-monotonic transformations (e.g., split and merge [9]). Some of these approaches led to the development of optimization schemes (e.g., graph-cuts [10], random walks [11], power watersheds [12]). In the framework of mathematical morphology, these graph-based approaches gave rise to the notion of connected operators [13].

Graph-based segmentation allows us to obtain one segmentation result from a given image. In order to tackle the ill-posed problem of segmentation, hierarchical approaches were developed to compute families of nested partitions, providing adapted solutions at different scales. These notions of hierarchies take their origin in image models initially dedicated to optimize the access and space cost of the carried information (e.g., octrees [14]). These regular models were progressively shifted toward image / content-guided, irregular hierarchies [15].

From this point on, several hierarchical image models were developed, mainly in the framework of mathematical morphology. The most popular are component-trees [16], trees of shapes [17,18], hierarchical watersheds [19], hyperconnected component-trees [20], and Binary Partition Trees [4] (see Section 2.2). Since they provide a space of potential segmentations, instead of a single result, these hierarchical models were progressively involved in attribute-based [21] or optimization schemes [22,23] for segmentation purpose.

Based on these image models, generally designed as trees, further developments were proposed to allow for a better flexibility in image and parameter handling. The case of multiband images was

[1] https://bitbucket.org/agat-team/agat-v0.3.