Knowledge-Based Systems 44 (2013) 101–111

Contents lists available at SciVerse ScienceDirect

Knowledge-Based Systems

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

Multiscale edge detection based on Gaussian smoothing and edge tracking

C. Lopez-Molina [a,b,*], B. De Baets [b], H. Bustince [a], J. Sanz [a], E. Barrenechea [a]

[a] Dpto. Automatica y Computacion, Universidad Publica de Navarra, 31006 Pamplona, Spain
[b] Dept. of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure Links 653, 9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 5 July 2012
Received in revised form 22 January 2013
Accepted 25 January 2013
Available online 4 February 2013

Keywords:
Edge detection
Gaussian scale-space
Multiscale image processing
Sobel operators
Edge tracking

ABSTRACT

The human vision is usually considered a multiscale, hierarchical knowledge extraction system. Inspired by this fact, multiscale techniques for computer vision perform a sequential analysis, driven by different interpretations of the concept of *scale*. In the case of edge detection, the *scale* usually relates to the size of the region where the intensity changes are measured or to the size of the regularization filter applied before edge extraction. Multiscale edge detection methods constitute an effort to combine the spatial accuracy of fine-scale methods with the ability to deal with spurious responses inherent to coarse-scale methods. In this work we introduce a multiscale method for edge detection based on increasing Gaussian smoothing, the Sobel operators and coarse-to-fine edge tracking. We include visual examples and quantitative evaluations illustrating the benefits of our proposal.

1. Introduction

Edge detection is a multistage process, in the sense that it cannot be performed in a single step [1,2]. Different breakdown structures can be found in the literature [2,3], but most of them include a stage where the local properties of a pixel and its neighborhood are evaluated. When performing a neighborhood-based evaluation, one first has to define its size, i.e., the number of pixels included in the evaluation of the neighborhood. Whereas some edge detection methods are based on fixed-size neighborhoods (such as FIRE [4] or the Sobel methods [5]), others modify the size depending upon the values of their parameters (such as the Canny [6] or the Marr-Hildreth methods [7]). The size of the neighborhood determines the scope of the intensity changes one is able to characterize. In this sense, it can be understood as the *scale* at which the edge detection is performed. In general, fine scales are expected to provide spatially accurate results, but also to be particularly sensitive to noise [8–11]. In fact, the relationship between spatial accuracy and sensitivity has been studied for some edge detection methods. The most relevant case is the work by Canny [6], who grounds his development in the modeling and optimization of three criteria: *low error rate*, *good localization* and *uniqueness of the response*. He concludes that there is a trade-off between the last two criteria: accurate location of the edges and robustness against spurious responses. This work has been revisited by different authors, such as Demigny [12] and McIlhagga [13], criticizing the way in which the criteria are modeled. Nonetheless, it is commonly accepted that edge detection performed at coarse scales is more robust against noise, textures and spurious edges, but tends to suffer from displacements of the edges from their actual position [14,10]. Some authors have investigated how to determine automatically the optimal scale for a given edge detection method [9,15,16], but no consensus has been reached.

Instead of working with a single scale, some authors use information obtained at different scales. These methods are generally called *multiscale* methods. In this work we present a multiscale edge detection method based on the Sobel operators for edge extraction and the concept of Gaussian scale-space. More specifically, we apply the Sobel edge detection method on increasingly smoothed versions of the image. Then, we propose to combine the edges appearing at different scales by performing coarse-to-fine edge tracking. We include experimental results illustrating how the proposed multiscale method has benefits single-scale methods fail to provide.

In Section 2 we analyze the concept of multiscale image processing, as well as its application to the edge detection problem. Then, in Section 3 we review some basic concepts of the Gaussian scale-space. Section 4 covers the details of the proposal, including the Sobel method for edge detection, the proposed coarse-to-fine edge tracking algorithm and the implementation details. Finally, Section 5 includes practical experiments, while some conclusions are drawn in Section 6.

* Corresponding author at: Dpto. Automatica y Computacion, Universidad Publica de Navarra, 31006 Pamplona, Spain.

E-mail addresses: carlos.lopez@unavarra.es (C. Lopez-Molina), bernard.debaets@ugent.be (B. De Baets), bustince@unavarra.es (H. Bustince), joseantonio.sanz@unavarra.es (J. Sanz), edurne.barrenechea@unavarra.es (E. Barrenechea).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.01.026

2. Multiscale techniques for image processing

This section covers the existing multiscale techniques in the image processing literature.

2.1. Multiscale image processing

The ultimate goal of a computer vision system is to perform exactly the way a human does. The success of such a system should be measured in terms of its output, and how similar that is to the one a human produces [17]. However, some computer systems do not only intend to obtain the same output, but also to simulate the underlying perceptual process a human performs. An example of simulation of human thinking is multiscale computer vision, which we consider in this paper. The human visual system is based on selective attention and focus, performing interpretations of different depth at each part of a scene [18,19]. The image in Fig. 1 can be used to illustrate such behavior. First, a human is likely to analyze the picture as a whole, understanding that it is a scene at a beach. Then, he or she focuses attention on the mid-size objects, such as the boat stranded in the forefront or the rocks in the background. At last, a finer-detail observation provides information that was overlooked, such as the name of the boat on its back or the rows it contains. Hence, a progressive, hierarchical process of picture understanding is carried out, varying the scope and level of attention at each stage.

Note that this process cannot be characterized in terms of zooming, since no new information is demanded by (or provided to) the human. However, the level of attention (and the amount of information to be handled at each step) forces the human to be aware or unaware of certain details. Avoiding details is nevertheless induced by the human, since the same information is available at each step of the image understanding.

Multiscale image processing methods are inspired by the human interpretation of a scene by considering different amounts of information at different parts of the scene. In order to imitate this behavior, they process the information at different stages, simulating a hierarchical image understanding. The key concept in this process is the scale, whose variations modify the scope and characteristics of the knowledge acquisition. In the image processing literature the scale is interpreted in many different ways. A large number of proposals are based on wavelets or signal decomposition methods [20]. These approaches are close to some scale-space methods, such as time-evolving snakes [21,22] or object tracking with increasing smoothing [23–25]. Some authors, such as Tabb and Ahuja [26], criticize the use of a fixed-scale method (employed by all of the previous approaches) and propose a recursive segmentation algorithm based on a structural, tree-like decomposition of an image. Zhang et al. [27] and Zhu et al. [28] use a hierarchical approximation as well, based on layering the knowledge to generate a complex interpretation from basic cues.

Fig. 1. Example image taken from the BSDS500 image dataset.

There are many different interpretations of the multiscale concepts, since the very purpose of the *scale* is not understood in a homogeneous way. In the literature we find examples where the scale relates to domains such as time [14,21,29], information complexity [27], signal frequency [6,20] or smoothing degree [25], among others. The multiscale concept always incorporates a sense of progressive comprehension of the image, varying the coarseness of the interpretation, but results in different computational paradigms.

2.2. Multiscale edge detection

In edge detection, we assume that the scale of an edge detector is related to the scope of the search for intensity changes [30]. Despite its importance, there is no evident solution (and maybe no solution at all) to the scale-determination problem. In fact, Torre and Poggio stated that, in order to correctly detect all meaningful edges appearing in an image, *derivatives of different types, and possible different scales* would be needed [1].

Starting from the assumption that using a single scale is not sufficient, multiscale methods intend to use varying-scale versions of the same image. Ideally, the use of multiple scales leads to a reduction of the ambiguity inherent to single-scale methods [31]. In the multiscale edge detection literature we find three families of methods, characterized by the way they manage the information obtained at different scales:

1. *Detecting the edges explicitly at each scale.* These methods perform edge detection at different scales and establish some kind of correspondence between them. That is, they perform a complete edge extraction at each scale and combine the results to produce a single edge image [8].
2. *Fusing information from each scale.* These methods do not detect the edges at different scales, but collect edge cues and information instead. An aggregated expression of the multiscale information is associated with each pixel, and later used for discriminating the edge pixels. This is the case for the multiscale morphological operators by Demin [32] or the classifier based on polytopes by Laligant et al. [33].
3. *Using a single, non-homogeneous scale at each pixel or subregion.* These methods determine the scale to be used at each pixel or subregion based upon local characteristics. Examples are the works by Jeong and Kim [34], minimizing an energy function, or Elder and Zucker [35], defining the *minimum reliable scale* for edges affected by focal and penumbral blur.

In this work we focus on the first family, which is the most explored one in the literature. Witkin was the first author using multiscale edge identification and tracking, featuring a coarse-to-fine procedure [31]. In fact, the motivation of this work was to create a framework where the edges at different scales *can be related to each other in an organized, natural and compact way*. Hence, the author avoided algorithmic details and focused on the behavior of the edges when the original signal is increasingly smoothed with a Gaussian filter. Canny also acknowledged the problems derived from single-scale methods and suggested to either aggregate the result at different scales (*feature synthesis*) or to filter the image selectively depending upon local features, such as the noise-to-signal ratio [6,36]. In a rather different way, Mallat and Zhong [20] used image decomposition grounded in the fact that (a) the Canny method is analogous to maxima-finding in a wavelet transform and that (b) multiscale edges are meant to characterize an image uniquely. See [37] for further developments of this idea. Bergholm introduced a proposal for combining the edges perceived at different scales, using coarse-to-fine edge tracking (*focusing*) [8]. Bergholm's work can be seen as a pragmatic extension of that of