Computer Vision and Image Understanding 114 (2010) 813–826

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Unsupervised colour image segmentation using dual-tree complex wavelet transform

Turgay Celik, Tardi Tjahjadi *

School of Engineering, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, United Kingdom

ARTICLE INFO

Article history:
Received 9 February 2009
Accepted 7 March 2010
Available online 15 March 2010

Keywords:
Unsupervised colour image segmentation
Multiscale edge detection
Dual-tree complex wavelet transform
Discrete wavelet transform
Multiscale analysis
Region growing
Colour difference

ABSTRACT

In this paper we present an effective unsupervised colour image segmentation algorithm which uses multiscale edge information and spatial colour content. The multiscale edge information is extracted using the dual-tree complex wavelet transform. Binary morphological operators are applied to the edge information to detect seed regions which are large enough to exclude boundary-only regions. The segmentation of homogeneous regions is obtained using region growing followed by region merging in the CIE $L^*a^*b^*$ colour space. We also present an edge preserving smoothing filter as a pre-process for the algorithm. We compare our algorithm with state-of-the-art algorithms and show its superior performance.

1. Introduction

Image segmentation is the process of subdividing an image into homogeneous regions. With colour images, the objects are segmented with respect to their colour and spatial features. Image segmentation techniques can be categorised into two groups [1]: soft segmentation and hard segmentations. In soft segmentation, each point in the feature space is associated with a label whose confidence value is computed using some function related to the distance of each converged cluster. For example, Tai et al. uses a Markov network to optimise a global objective function that combines the advantages of global colour statistics and local contextual image statistics [1]. The method is computationally complex. The method we proposed in this paper falls into the latter category. There are four main approaches to hard image segmentation. The thresholding approach is based on the assumption that clusters in the histogram correspond to either background or objects of interest that can be extracted by separating these clusters [2–5]. Boundary-based methods assume that there is significant change in pixel properties, such as intensity, colour and texture between different regions [6,7]. Region-based methods assume that neighbouring pixels within the same region have similar pixel properties [8,9]. Hybrid methods tend to combine boundary detection and region growing to achieve better segmentation [10–18].

One of the milestones in the hybrid methods is the unsupervised algorithm called JSEG [10]. JSEG uses colour quantization and local windows to compute *J*-images (corresponding to texture segmentation) at different scales. The *J*-images are combined via multiscale region growing to obtain image segmentation. Since the computation of *J*-images requires non-linear numerical computations, the JSEG algorithm has computational complexity which increases with the number of scales used. One drawback of the algorithm is its sensitivity to noise. Since JSEG is based on the results of the colour quantization, another drawback is that its performance is essentially limited by the colour quantization. The values of *J*-images cannot differentiate the regions with similar distribution of textural patterns but different colour contrast results in another disadvantage of JSEG. A modified version of JSEG proposed by Wang et al. [11] uses an adaptive mean-shift clustering (AMS) for non-parametric clustering of the image data instead of the colour quantization algorithm used in JSEG. They use Gaussian mixture modelling (GMM) of the image data constructed with classifications obtained by AMS in the calculation of *J*-images. The segmentation results are better in terms of better boundary representation and object segmentation. Since the method is a mixture of JSEG, AMS and parameter estimation using GMM, it is computationally expensive.

The multiscale nature and robustness to noise of the wavelet transform makes it an attractive tool for colour image segmentation. Multiscale edge information in non-decimated wavelet transform domain has been used with a watershed transformation at each scale and a hierarchical region merging procedure to connect

* Corresponding author. Fax: +44 24 76418922.
E-mail addresses: t.celik@warwick.ac.uk, celikturgay@gmail.com (T. Celik), t.tjahjadi@warwick.ac.uk (T. Tjahjadi).

1077-3142/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2010.03.002

segmented regions at different scales [12]. Although the method achieves good performance against noise, the algorithm produces oversegmented results. In [13], multiscale image representations and watershed transforms are used. The scale-space is based on a vector-valued diffusion scheme, and colour gradients in the YUV colour space are computed at each scale. The dynamics of the contours in scale-space are used to validate detected contours. The segmentation results are visually pleasant but the performance of the algorithm against noise is not determined.

Recently, decimated wavelet and watershed transforms are used to increase the robustness of an unsupervised colour image segmentation algorithm against noise [14]. Using the horizontal and vertical subbands, a colour gradient magnitude image is computed at the lowest resolution to detect edges. An initial segmentation map is obtained by applying a watershed transform to the gradient magnitude image. The inverse wavelet transform is then used to project this initial segmentation to finer scales, until the full resolution image is achieved. Finally, a region merging based on the CIE $L^*a^*b^*$ colour distances is applied to obtain the final segmentation. Since the performance of the algorithm mainly depends on the correctness of the detected edges in the scale in which watershed algorithm is applied, further improvement can be achieved with a better edge detector. Also, its performance depends on the watershed algorithm which requires a threshold selection process.

Multiscale edge detection could also be performed in the image domain using wavelet transform to obtain the multiscale edge features which have good time-spatial characteristics. Mallat and Zhong's approaches [19] classify singularity points as the local maxima of gradient modulus or the zero-crossings of wavelet coefficients as edges. Sun et al. [20] used Hidden Markov Model with wavelet transform to detect multiscale edges and fuse them to generate an edge map at the pixel resolution level for object recognition. Zhang and Bao [21] defined a scale product as the multiplication of two adjacent scales of wavelet coefficients to amplify edge structures while reducing noise, and using a single threshold classified the local maxima of the product as edges. DWT-based methods generally suffers from the shift-variance and lack of directionality required in multiscale analysis to detect multiscale edges. Wang and Jiao [22] utilised the better directionality and shift invariance of the complex wavelets, and the scale product similar to that in [21] to detect edges. But as with most of DWT-based methods, they used wavelet coefficients multiplication and thresholding to detect multiscale edges. There are two drawbacks with this approach. The first and the most important is that coefficients multiplication using intra-scales may obscure low coefficients which contains information associated with true edges. The second drawback is the dependency to thresholding. To overcome these drawbacks we propose a Bayesian approach to detect multiscale edges using complex wavelet coefficients.

Region-based merging is generally the final process for most colour image segmentation algorithms. Region merging usually involves statistical parameters for inter and intra region correlations [23,24] which are expensive to compute. Region adjacency graph (RAG) [25] based merging represents each region as a graph node and an edge exists between two adjacent nodes. A cost is assigned to each graph edge expressing the dissimilarity between two adjacent nodes. The most similar pairs of adjacent nodes have the edges with the minimum cost and are merged. The initialization of RAG, the calculation of edge costs and the iteration involved in merging are computational expensive. To reduce computational complexity Shih and Cheng [26] used two criteria for region merging: colour similarity and size of region. In the first pass, the criteria are applied sequentially as follows. Firstly, if the colour distance between two adjacent regions is less than a threshold, i.e., 0.1, the two regions are merged and the mean colour of the merged region is computed. This comparison is repeated for all regions including the newly merged regions until no more merging is possible. Secondly, if the number of pixels in a region is smaller than a threshold, i.e., 1/150 of the image size, the region is merged to its neighbouring region with the smallest colour distance. This procedure is repeated until there are no regions with size smaller than the threshold. In the second pass, any two adjacent regions with size smaller than 1/10 of the image size and the colour difference between them is greater than 0.2 are merged. This process is repeated until no more merging is possible. The algorithm is simple and provides good results.

We proposed an unsupervised colour image segmentation algorithm to address the aforementioned drawbacks, which uses the multiscale structure of the dual-tree complex wavelet transform (DT-CWT) [27] to detect representative edge information. The motivation for using DT-CWT is due to its ability to detect directional edges better than DWT, increased redundancy and the dual-tree implementation for real-time performance. The edge information is used to generate the representative seed region for each homogeneous region in the finest resolution of the image. We employ morphological binary operators rather than a watershed algorithm to detect seed regions in order to significantly reduce the computational complexity of the overall algorithm. Since edge information is used to find seed regions, only a few pixels are not assigned to any seed region. Thus, the pixel assignment performed by the region growing is computational inexpensive. A region growing and region merging are then applied to obtain the overall segmentation. The region merging is performed in the uniform CIE $L^*a^*b^*$ colour space [28] to address over-segmentation. We also introduced a new edge preserving smoothing filter which is used as a pre-process for the algorithm.

The paper is organised as follows. Section 2 presents the proposed algorithm. Section 3 presents the experimental results and discussions. Finally, Section 4 concludes the paper.

2. Proposed algorithm

Fig. 1 shows the overview of the proposed algorithm. The first stage of the algorithm transforms the image from the RGB to the CIE $L^*a^*b^*$ colour space. The second stage smoothes the image using a non-linear pre-filtering operation (see Section 2.1). The aims of the pre-filtering are two folds: to remove small variations in the input image while retaining a contrast between different regions; and to reduce the noise level.

A multiscale edge map is found using DT-CWT (see Sections 2.2 and 2.3) and combined with the result of a Canny edge detection to detect finer boundary edges. The resulting edge map is used to find seed regions that correspond to regions to be segmented (see

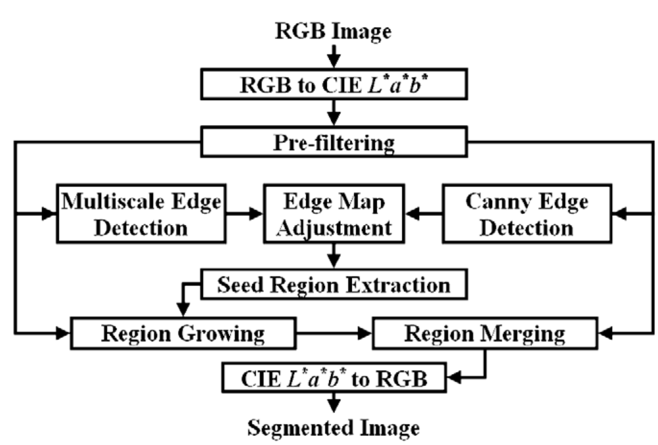

Fig. 1. Proposed colour image segmentation algorithm.