Neurocomputing 179 (2016) 186–201

Contents lists available at ScienceDirect

Neurocomputing

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

Region of interest extraction in remote sensing images by saliency analysis with the normal directional lifting wavelet transform

Libao Zhang*, Jie Chen, Bingchang Qiu

The College of Information Science and Technology, Beijing Normal University, Beijing, China

ARTICLE INFO

Article history:
Received 10 June 2015
Received in revised form
29 November 2015
Accepted 30 November 2015
Communicated by Y Gu.
Available online 15 December 2015

Keywords:
Image processing
Region of interest
Saliency analysis
Normal directional lifting wavelet transform

ABSTRACT

Region of interest (ROI) extraction techniques based on saliency comprise an important branch of remote sensing image analysis. In this study, we propose a novel ROI extraction method for high spatial resolution remote sensing images. High spatial resolution remote sensing images contain complex spatial information, clear details, and well-defined geographical objects, where the structure, edge, and texture information has important roles. To fully exploit these features, we construct a novel normal directional lifting wavelet transform to preserve local detail features in the wavelet domain, which is beneficial for the generation of edge and texture saliency maps. We also improve the extraction results by calculating the amount of self-information contained in the spectra to obtain a spectral saliency map. The final saliency map is a weighted fusion of the two maps. Our experimental results demonstrate that the proposed extraction algorithm can eliminate background information effectively as well as highlighting the ROIs with well-defined boundaries and shapes, thereby facilitating more accurate ROI extraction.

1. Introduction

Increases in the ability to acquire high spatial resolution remote sensing images by various satellites and sensors have led to great challenges in the detection of valuable targets from high spatial resolution remote sensing images [1–3]. Region of interest (ROI) extraction techniques based on saliency have been introduced into the remote sensing image analysis field and they have become a research hotspot in recent years [4–9]. In addition, these techniques are employed as an efficient information processing method to handle the rapidly growing volume of remote sensing images. After providing a potential ROI, the viewer can search for specific objects in the region and computing resources can be allocated in a reasonable manner to enhance the operating efficiency of an image processing system [10].

In remote sensing images, typical ROIs include residential areas, airports, airplanes, wharfs, and ships. Compared with the background, they have salient features that immediately grab human attention; hence, it is suitable to extract ROIs via saliency models. In particular, the salient characteristics are as follows.

a) Abundant and complex structure, edge, and texture information, which is typical of the interior of a residential area.
b) Unique shapes, particularly for airplanes, which are not as texture rich as residential areas, but their unique shape makes them stand out.
c) Orientation information, e.g., the ships usually head in the same direction because of the similar ocean currents and weather patterns in nearby waters.
d) Their distinct spectra compared with the surrounding environment.

The ROIs possess these characteristics, whereas the background does not, so high contrast stimuli are generated in receptive fields of the human visual system and human cortical cells may be hardwired to respond preferentially to these stimuli [11]. Visual saliency refers to distinctive parts of a scene that immediately attract significant attention without any prior information, thus it is flexible in adapting to different ROI extraction tasks, in which retraining is unnecessary.

Saliency-based methods were originally designed for natural scene images [12–15] by utilizing the intensity, color, orientation, texture, and other low-level features to determine contrast for saliency computation. One of the earliest computational models, which was built on a biologically plausible architecture [16], was proposed by Itti et al. (IT) [17]. This model obtains saliency maps based on the intensity, color, and orientation channels, and computes the final master saliency map by combining these three conspicuity maps based on center-surround differences.

Various computational models have been inspired by the biological concept of center-surround contrast in the IT model. These

---

* Corresponding author. Tel.: +86 13366069027.
E-mail address: libaozhang@bnu.edu.cn (L. Zhang).

http://dx.doi.org/10.1016/j.neucom.2015.11.093
0925-2312/

estimation models can be broadly classified as biologically-based, purely computational, and a combination. Harel's graph-based visual saliency method (GB) [18] is a combination model that employs an idea from graph theory to concentrate mass in activation maps and to obtain activation maps from raw features.

Among the purely computational models, Achanta et al. [14,15] attempted to build a saliency model (FT) using color contrast information. The feature vector was acquired in the CIE Lab color space and the absolute difference between the Gaussian-blurred image and the arithmetic mean vector was then calculated to obtain the saliency map. Goferman et al. [19] proposed a novel algorithm called context-aware saliency detection (CA). Supported by psychological evidence, CA uses a detection algorithm that relies on four basic principles reported in the psychological literature.

Cheng et al. [20] proposed a histogram-based contrast (HC) method to measure saliency for image pixels using color statistics determined for an input image. They also presented a regional contrast-based saliency extraction algorithm (RC), which simultaneously evaluates the global contrast differences and spatial coherence. In RC, the input image is first segmented into regions, before estimating saliency for each region as the weighted sum of the region's contrasts compared with all of the other regions in the image. The weights are set according to the spatial distance, where more distant regions are assigned smaller weights. RC obtains high precision and recall rates with natural images. In addition to these models, saliency models have been proposed in the spatial domain, such as an information theory-based computational model [21] and contrast-based filtering for salient region detection [22].

Recently, researchers have also tried to obtain solutions in the transform domain. The Fourier transform can be expressed in polar form using two different components: phase and amplitude spectra. By analyzing the log amplitude spectrum, Hou et al. [23] defined the spectral residuals (SR) algorithm, where the saliency map is derived by applying the inverse Fourier transform to an exponential function that combines spectral residual and phase spectrum information. In addition, Guo et al. [24] proposed a computational model based on the quaternion Fourier transform. Compared with the Fourier transform, the wavelet transform can perform multi-scale spatial and frequency analyses simultaneously, and thus it has begun to attract more attention from researchers. Murray et al. [25] computed weight maps from the high-pass wavelet coefficients of each level and the saliency map was obtained by the inverse wavelet transform of the weight maps. To improve this model, Imamoglu et al. [26] proposed a wavelet transform-based computational model (WT) that uses low-level features, which considers both local center-surround differences and the global contrast, thereby obtaining better results than the method of Murray et al. [25].

It should be noted that these standard saliency detection methods were not designed specifically for remote sensing images and differences exist between ROI extraction from remote sensing image and natural scenes. The ROIs in natural scenes have less complex textures and their distinct colors make them instantly recognizable from the surroundings. In addition, when shooting a picture, photographers manually set the lens to blur the background and focus on the ROI, which helps to highlight the ROIs. Furthermore, there is a strong center bias because human photographers tend to place one or two objects of interest in the center of photographs [27], which significantly narrows the search when locating ROIs. By contrast, ROIs such as airplanes and ships are scattered in the background of remote sensing images, where their positions and number are unpredictable. Moreover, the structure, shape, and texture information is abundant and complex in a high spatial resolution remote sensing image. To ensure that the input is accurate for subsequent applications, such as object recognition, image compression, and image retrieval, the principles followed to achieve good ROI extraction are stricter in remote sensing images. For example, the ROIs should be uniformly highlighted with well-defined boundaries to ensure the integrity of ROIs. In addition, the final object maps should retain full resolution without any loss of detail to preserve the fineness of remote sensing images. Thus, there is a difference between our method and standard saliency methods for objective evaluation, as described in Section 3.2.

In general, standard saliency methods may ignore the fact that high spatial resolution remote sensing images contain complex spatial information, clear details, and well-defined geography objects. Thus, they are likely to extract these complex structure, edge, and texture features in a coarse manner. For example, the IT model [17] obtains the saliency map based on the intensity, color, and orientation channels. In addition, the orientation information is obtained using oriented Gabor pyramids $O(\sigma, \theta)$ with $\theta \in \{0°, 45°, 90°, 135°\}$. GB [18] is derived from IT, so it obtains orientation information in the same manner. However, only four directions are used, which makes this method less accurate than ours. Moreover, the saliency map is 1/256 of the original image size, which inevitably leads to the loss of texture details. The same problem affects SR [23], which down-samples the input image to 64×64 pixels. Other methods such as FT [14], CA [19], HC [20], and RC [20] neglect texture features and they focus preferentially on the color and luminance features during saliency calculations. WT [26] represents different features that range from edges to textures by wavelets, but it employs Daubechies wavelets (Daub.5), which are not suitable for approximating image features with an arbitrary orientation that is not vertical or horizontal.

In recent years, saliency computational models have also been introduced into the remote sensing image processing field and they have become a research hotspot. A saliency computation approach (RS) was introduced [4] to select perceptually salient and highly informative regions that represent the main contents of high-resolution remote sensing images. Zhang et al. [7] used high-frequency filters for frequency domain analysis (FDA) to detect ROIs in high spatial resolution remote sensing images, but the model produces an attenuated interior in ROIs, thereby yielding incomplete ROI extraction. Zhang et al. [8] then proposed another model based on multi-scale feature fusion (MFF) of the intensity saliency result and the orientation saliency result to obtain one saliency map for ROI extraction, where the orientation saliency is based on the conventional lifting wavelet transform (LWT). Furthermore, Wang et al. [28] successfully applied the saliency technique to airport detection. Ding et al. [29] also attempted to implement ship detection using a saliency technique. Another efficient ROI extraction method based on spectral analysis was introduced for saliency testing in remote sensing images [6].

Based on the four characteristics of ROIs mentioned above, we propose a novel directional wavelet called normal directional LWT (ND-LWT) to fully exploit the information contained in texture, shape, and orientation. This method is designed specifically for high spatial resolution remote sensing images and it also makes use of spectral information to facilitate accurate extraction. The proposed method can preserve the local details of features in the wavelet domain, which is beneficial for generating the edge and texture saliency map. Our experimental results demonstrate that the proposed extraction algorithm can eliminate the background information in an effective manner as well as highlighting the ROIs with well-defined boundaries and shapes, thereby allowing more accurate ROI extraction.

The remainder of this paper is organized as follows. The proposed ROI extraction algorithm is introduced in Section 2. Section 3 presents the experimental results and discussion. In Section 4, we give our conclusions.