ISSN No. 0976-5697

Volume 4, No. 7, June 2013 (Special Issue)

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Analyzing the Effects of Image Segmentation in Compound Structures

Prakriti Trivedi
Assistant Professor CS and IT
Govt. Engineering College
Ajmer, Rajasthan India
niyuvidu@rediffmail.com

Himanshu Swarnkar
M.Tech IT
Govt. Engineering College
Ajmer, Rajasthan India
himashu.swarnkar23@gmail.com

Prafull Naruka
Research Scholar: MJRP University, Jaipur
Jaipur, India
narookaprafull@gmail.com

Abstract: Image segmentation is a vital part in today's technological changes. We proposed an algorithm for deriving the compound structure which consists of primitive image objects. Our algorithm works in two phases. Firstly it produces image regions with homogeneous spectral content using the segmentation process. Secondly the segmentation is translated into relational graph structure .In relational graph structure each node represents the specified region and the edges between them acts as a relation between the regions. The region objects that appear together frequently can be considered as strongly related. Experiments are carried out using an Ikonos image which shows that subgraphs found within the graph representing the whole image correspond to parts of different high-level compound structures.

Key words: Compound structure, Image segmentation, object detection, graph-based analysis.

I. INTRODUCTION

The objective of object-based image analysis is to partition the images into homogeneous regions and classify these regions. But the obtained image provides the less details .To overcome this problem we have to find the regions that are intrinsically heterogeneous. These image regions consist of primitive objects of different types which can also be referred to as compound objects Bhagavathy and Manjunath [1] build a texture motif model for harbors and golf courses from training examples for detecting compound objects of predefined types. Dogrusoz and Aksoy [2] detect organized and unorganized urban areas by clustering a scene graph whose nodes correspond to individual buildings for the same reason. Stasolla and Gamba [3] detect built-up areas in high resolution SAR images using local autocorrelation. In this paper, we propose an algorithm for discovering interesting and significant compound objects regardless of their types. The method translates image segmentation into a relational graph, and applies a graph-based knowledge discovery algorithm to find the interesting and repeating substructures that may correspond to compound objects. The first step is image segmentation where the resulting regions correspond to primitive objects that have relatively uniform spectral content. The next step is the translation of this segmentation into a relational graph structure where the nodes represent the regions and the edges represent the relationships between these regions. We assume that the region objects that appear together frequently can be considered as strongly related. This relation is modeled using the transition frequencies between neighboring regions. Each transition is represented by a point in a multi-dimensional space. This space is modeled by a non-parametric probability distribution, and the local maxima found from the density function are assumed to correspond to the most frequently occurring and hence the most significant and important transitions. Finally, a graph whose edges encode this frequent spatial co-occurrence information is constructed, and a sub graph analysis algorithm is used to discover substructures that often correspond to groups of region objects that occur together in high-level compound structures. Proof-of-concept experiments illustrate the proposed algorithm on an Ikonos image.

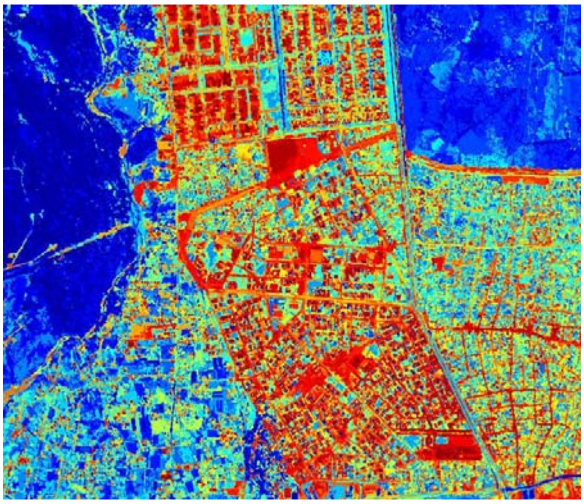
II. SEGMENTATION AND FEATURE EXTRACTION OF AN IMAGE

Image segmentation is the initial step where relatively uniform spectral content regions are found which contains primitive objects. For this segmentation step we use the Recursive Hierarchical Segmentation (RHSEG) algorithm [4]. (a) Ikonos image (b) Segmentation Fig. 1. An Ikonos image of Antalya, Turkey and its segmentation. RHSEG is a promising choice because of three key factors:

(i) It produces high spatial fidelity of image segmentations. (ii)It automatically group the spatially connected region objects into region classes, and (iii) Its automatic production of a hierarchical set of segmentations. It is possible to examine how the regions change at each level and choose the level of detail at which the particular regions of interest are delineated. Figure 1 shows a multi-spectral Ikonos image of Antalya, Turkey with 4 m spatial resolution and 700 × 600 pixel size, along with its segmentation in false color.

The segmented regions obtained are represented using their spectral and size information. The spectral features for each region are computed using the average red, green and blue values of the pixels in that region. The size information corresponds to the number of pixels in each region. In our work, we use size as a feature to be able to distinguish regions with similar spectral content but significantly different sizes. All features are normalized to the [0, 1] range using linear scaling. Finally, each region Ri is represented using the feature vector yi = (ri, gi, bi, si) with 4 components.

(a) Ikonos image



(b) Segmentation

Figure 1. An Ikonos image of Antalya, Turkey and its segmen-tation.

III. MODELING OF REGION CO-OCCURRENCE

After the segmentation step the translation of this segmentation into a relational graph structure is performed where the nodes correspond to the individual regions, and the edges model their spatial relationships. In this paper, we model the region relationships using the transition frequencies between neighboring regions in the image by assuming that the region objects that appear together frequently in the image can be considered as strongly related. We can calculate the inter-region transition frequency is by determining the types of the regions and by counting the transitions involving the same types of region pairs. However, the determination of region types is a challenging and important classification problem, and errors at this step will result in misleading transition types. In this paper we propose to use a spatial co-occurrence model that enables transition frequency calculation without a preceding transition or region type assignment. This model involves a multi-dimensional space where each point corresponds to an inter-region transition, and enables the incorporation ration of region transition frequencies together with region features. The space is modeled by a non-parametric probability distribution so that the probability value for each transition

point corresponds to the frequency of its occurrence in the image. The details of this model are described below.

A. Spatial co-occurrence space

Each inter-region transition is defined by the features of the corresponding regions so that their contents can be incorporated in the model. In an image with NR regions Ri, i =1, . . . ,NR, the transition Tij involving the regions Ri and Rj is represented by the concatenation of the feature vectors of the two regions as yij = (yi, yj). Given the region feature vectors with 4 components, the feature vector for a transition corresponds to a point in the 8-dimensional spatial co occurrence space. For simplicity, we refer to these points as xk 2 Rd, k = 1, . . . ,NT where d = 8 and NT is the number of transitions.

We assume that the transitions that involve two similar region pairs fall close to each other in the spatial co-occurrence space because regions with similar spectral content and sizes are expected to be similar in terms of their features. Consequently, the transitions that occur frequently cause the accumulation of points in the space. While similar transitions are pooled together to form dense clusters, seldom transitions are located sparsely. This model provides tolerance to small variations and noise in the region features. Furthermore, it can easily be extended with additional region features.

The significance of a particular transition can be determined according to its location relative to the dense areas in the spatial co-occurrence space. We model this space with a Parzen window-based probability density estimate

$$P(x) = \frac{1}{N_T}\sum_{K=1}^{N_T}\frac{1}{(2\Pi)^{d/2}|H|^{1/2}}e^{-1/2(X-Xk)^T H^{-1}(X-Xk)} \quad (1)$$

B. Deriving important relations

We assume that the dense regions in this space correspond to the most frequently occurring and hence the most significant and important transitions. These dense regions can be found by locating the modes (local maxima) of the estimated density. We obtain these modes using the mean-shift algorithm [6]. Starting from a randomly selected set of points, the algorithm computes the mean-shift vector at each point x as

$$\mathrm{m(x)} = \frac{\sum_{k=1}^{N_T} X_k e^{-1/2(X-X_k)^T H^{-1}(X-X_k)}}{\sum_{k=1}^{N_T} e^{-1/2(X-X_k)^T H^{-1}(X-X_k)}} \quad (2)$$

using the Parzen density gradient estimate at that point, and moves along this vector by iterating until the difference between two successive means is less than a convergence threshold or the number of iterations reaches a maximum value. The points at which the algorithm converges are considered as the candidate modes.

The convergence of the mean-shift algorithm is affected by the convergence threshold and the number of maximum iterations allowed. Due to local details in the spatial cooccurrence space, starting at points that actually belong to the same mode may result in convergence at slightly different locations. To eliminate such noisy convergence, we merge the candidate modes that are closer to each other than the bandwidth. Further elimination can be done due to the symmetric nature of the co-occurrence space. Since the transition Tij is equivalent to transition Tji, we compare the corresponding parts of the feature vectors of the candidate modes, and eliminate one of such mode pairs corresponding to symmetric transitions. The resulting set of modes provide an implicit clustering of the spatial co-occurrence space as any point in this space can be assigned to its closest mode.

VI. DERIVING COMPOUND STRUCTURES

After deriving the important relations this information is employed in the translation of the image segmentation to the relational graph structure. The details of graph construction and subgraph analysis for finding compound structures are described below.

A. Graph construction

A relational graph is constructed from the segmentation of the whole scene so that the nodes represent the regions and there is an edge between the nodes that correspond to the adjacent regions. In particular, for each region Ri there is a corresponding vertex Ri, and for each transition Tij there is an edge connecting the nodes Ri and Rj . It is common to use an unweighted graph and let the edges represent only the spatial adjacency [7]. However, by using this approach we may lose the detailed contextual information and the results may also suffer from the errors in segmentation. As described in Section 3.2, we assume that the modes of the density estimate of the spatial co-occurrence space correspond to the most significant and important transitions. This information is reflected in the constructed graph edges. First, the candidate modes with a probability smaller than a threshold are eliminated as such modes are likely to correspond to noisy, rare or insignificant transitions in sparse regions of the co-occurrence space. Then, the graph edges corresponding to the transitions that belong to the eliminated modes are removed. Furthermore, the graph can also be extended so that it reflects the transition type information. The transitions that are assigned to the same mode are accepted as a relation of the same type, and each transition (and the corresponding edge) is assigned an integer label between 1 and NM (the number of selected modes). As a result, the relationship information is fully encoded in the graph edges and their labels.

B. Subgraph analysis

The goal is to find compound structures that are comprised of the subgraphs of the complete scene graph. In this paper, we use a method that was introduced in [8] and was implemented in the Subdue system for graph-based knowledge discovery. In our case, the input to the system is an undirected graph with labeled edges (the nodes are not labeled as we do not perform any classification of the regions after segmentation). Subdue searches for substructures (subgraphs) of the input graph that best compress this graph. The compression of the graph by a subgraph is defined as the replacement of this subgraph by a single node in the graph. The compression ability of a subgraph during the search is computed b the minimum description length heuristic [8]

$$\text{Compression} = \frac{DL(S) + DL(G|S)}{DL(G)} \quad (3)$$

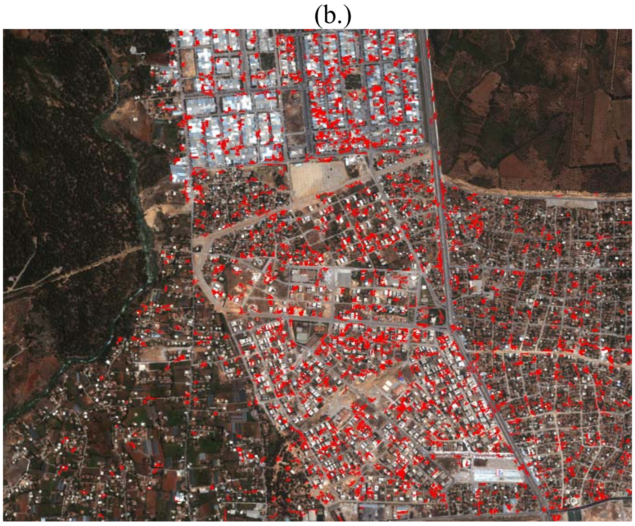
Where S is the subgraph being evaluated, DL(S) is the description length of S, DL(G|S) is the description length of the input graph G after it has been compressed using S, and DL(G) is the description length of G. The description length is computed in terms of the number of bits required to encode a graph. The best subgraph is the one that minimizes (3). The search is performed iteratively by compressing the graph with the best subgraph found in each iteration. The output is a list of subgraphs (in terms of the nodes and the edges they contain) that represent the discovered patterns together with all occurrences of each subgraph in the input graph. These subgraph instances are expected to constitute parts of compound structures in the complex urban scene.

V. EXPERIMENTS

To illustrate the effectiveness of the proposed method, we performed proof-of-concept experiments on the multi-spectral Ikonos image shown in Figure 1(a). The third segmentation scale (Figure 1(b)) was chosen among the 11 scales produced by RHSEG. The 51,558 regions present in this scale resulted in 263,246 transitions forming the points in the spatial cooccurrence space. By using these data, the bandwidth parameter was estimated as _ = 0.0188. The convergence threshold for the mean-shift algorithm was empirically set to 10−6 and the maximum number of iterations allowed was 4,000. We ran the algorithm 1,400 times starting at different sets of randomly selected points. This resulted in 1,197 unique candidate modes.

(a.)

(b.)



(c.)

Figure 2. Substructures obtained by graph analysis

After mode merging and the elimination of the symmetric modes, the number of modes was reduced to 271. 95 modes were chosen as significant (NM = 95) by applying a threshold to the corresponding probability values. The Subdue algorithm was applied to the constructed graph, and the resulting substructures (subgraphs) were examined.

some example substructures and the corresponding region groups are shown in Figure 2. Even though a single substructure does not exclusively correspond to a particular compound structure, we can observe that different substructures constitute parts of different compound structures. For example, the substructure instances in Figure 2(a) mostly constitute the parts of residential areas with low height buildings. Similarly, the instances in 2(b) mainly correspond to parts of an industrial area and a residential area with high buildings, and the instances in 2(c) are contained within a forest. We observed that the quality of the initial segmentation strongly influences the effectiveness of the following graph analysis. Future work includes improving the segmentation results and evaluating other graph clustering techniques for finding the interesting sub graphs.

VI. CONCLUSIONS

Different from the conventional object-based image analysis approach of finding homogeneous regions, an unsupervised method is presented in this very work in a way of discovering compound image structures that were content with complex groups of simpler primitive objects. We have worked by considering the primitive region objects that appeared together frequently could be considered as strongly related. Such potentially important relations were discovered using the modes of a probability distribution estimated using the features of the transitions between the neighboring regions in the image. The resulting modes were used to construct the edges of a graph in which the primitive regions form the nodes. In order to obtain the substructures of interest a sub graph analysis was used. Experiments on an Ikonos image showed initially that the algorithm has shown immense potential for discovering different high-level compound structures for very high resolution images and in high spatial mode as well.

VII. REFERENCES

[1] S. Bhagavathy and B. S. Manjunath, "Modeling and detection of geospatial objects using texture motifs," IEEE Trans. on GRS, vol. 44, no. 12, pp. 3706–3715, 2006.

[2] E. Dogrusoz and S. Aksoy, "Modeling urban structures using graph-based spatial patterns," in IGARSS, 2007.

[3] M. Stasolla and P. Gamba, "Spatial indexes for the extraction of formal and informal human settlements from high-resolution SAR images," IEEE JSTARS, vol. 1, no. 2, pp. 98–106, 2008.

[4] J. C. Tilton, "Parallel implementation of the recursive approximation of an unsupervised hierarchical segmentation algorithm," in High-performance Computing in Remote Sensing, A. Plaza and C.-I. Chang, Eds. 2007.

[5] R. P. W. Duin, "On the choice of smoothing parameters for parzen estimators of probability density functions," IEEE Trans. on Computers, vol. C-25, no. 11, pp. 1175–1179, 1976.

[6] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," IEEE Trans. On PAMI, vol. 24, no. 5, pp. 603–619, 2002.

[7] J. C. Tilton, D. J. Cook, and N. Ketkar, "The integration of graph based knowledge discovery with image segmentation hierarchies for data analysis, data mining and knowledge discovery," in IGARSS, 2008.

CONFERENCE PAPER
"National Conference on Spatial Data Mining" On 20th March 2013
Organized by
Dept. of Computer Science, MDS University, Ajmer (Raj) India