VillageFinder: Segmentation of Nucleated Villages in Satellite Imagery

Kashif Murtaza[1]
kashifm@lums.edu.pk
Sohaib Khan[1]
sohaib@lums.edu.pk
Nasir Rajpoot[2]
nasir@dcs.warwick.ac.uk

[1] Computer Vision Lab
LUMS School of Science and Engineering, Lahore, Pakistan
http://cvlab.lums.edu.pk

[2] Department of Computer Science
University of Warwick, UK

The importance of geo-spatial information for development planning cannot be over-emphasized. Yet, in most developing countries, local administrators have limited, if any, access to geo-spatial information. While direct access to recent satellite imagery may be expensive for local authorities in developing countries, a wealth of image data is contained in online tools such as Google Earth™ and Microsoft's Live Search Maps™. However, such images, while suitable for online viewing, do not yield the required geo-spatial information because of limited viewport and due to the fact that the relevant geographical features have to be extracted and processed before they are of use to the decision makers.

In this paper, we look at the problem of segmenting villages over a large area taken from Google Earth™. We have developed a crawler to save high resolution image blocks of a large area from Google Earth™. Once these images are saved, they are automatically stitched into a single image, which is then used for further processing. This provides a simple way for resource strapped users in developing countries to generate and access large scale geo-spatial datasets.

Robust segmentation of villages is a difficult problem because of the huge amount of variability in the shape, features and layout of a village. A mud-house village in Nigeria is very different in appearance from, say, a farmhouse community in the Netherlands. Segmentation of villages is made even more complicated by the fact that the images in Google Earth™ are taken from different sensors, at different times and under varying atmospheric conditions. Moreover, the exact demarcation of where the village ends and the peripheral areas start is not clear. For the purposes of this paper, we consider the village extents to be delimited by the contiguity of the houses in it; adjoining farmland or open spaces are not considered part of the village class. Even under this definition, the variability in what constitutes a village is very high between human observers.

The primary features used in our work are phase gradient features. Phase information, somewhat overlooked by researchers in texture analysis, can be used an effective cue for texture analysis since its gradient can both represent textural properties and point to the edge of textural regions. Let us model the visual texture in an image $v(\mathbf{x})$ as a linear combination of 2D sinusoidal wave functions and some additive white noise $w(\mathbf{x})$, as given below:

$$v(\mathbf{x}) = \sum_i a_i(\mathbf{x}) \cos\phi_i(\mathbf{x}) + w(\mathbf{x}) \tag{1}$$

where $\mathbf{x} = [x, y]^T$ and $\phi_i = [\phi_{ix}, \phi_{iy}]$ denotes the local phase.

In 1D, computing the gradient of local phase yields instantaneous or local frequency. We utilize log-Gabor filters to decompose a given image into its scale+orientation components, before computing the phase gradients for each of the components.

Let $v_i(\mathbf{x})$ denote the ith log-Gabor component of a given image $v(\mathbf{x})$. It can be represented as follows,

$$v_i(\mathbf{x}) = |v_i(\mathbf{x})| e^{j\phi_i(\mathbf{x})} \tag{2}$$

where $|v_i(\mathbf{x})|$ denotes the magnitude of v_i at a particular value of $\mathbf{x}$. Differentiating both sides of the above equation and re-arranging terms, we get the following expression for local phase gradient,

$$\phi_i'(\mathbf{x}) = j\left[\frac{|v_i(\mathbf{x})|'}{|v_i(\mathbf{x})|} - \frac{v_i'(\mathbf{x})}{v_i(\mathbf{x})}\right]. \tag{3}$$

The magnitude of $\phi_i'(\mathbf{x})$ as given below,

$$|\phi_i'(\mathbf{x})| = \sqrt{\frac{d\phi_i^2}{dx} + \frac{d\phi_i^2}{dy}} \tag{4}$$

gives local frequency in a particular direction, namely the direction perpendicular to the radial direction in the log-Gabor domain. To the best of our knowledge, ours is the first method to estimate the local frequency in this way.

Figure 1: Village segmentation of a 50 km^2 image of a rural area. All 19 villages marked in the ground truth are accurately extracted, with a false positive rate of 2.3% and 0.01% false negatives.

In addition to the phase gradient features, we also added "cornerness" measure as a feature. We compute the scatter matrix of the gradient vector over a 15×15 neighborhood and take the smaller eigenvalue as a feature. This captures the notion that we expect higher number of corner-type features in a village than in surrounding areas. This feature is computed at eight levels of the pyramid. We also added a color feature which is simply the green component of a pixel divided by its red plus blue components, to distinguish trees and foilage from the village structures. For classification, we generate weak classifiers based on our features and combined them through Adaboost.

We collected a reasonably representative dataset of nucleated villages for training and testing our segmentation algorithm. We have created a dataset of 60 images, having more than 345 million pixels and covering more than 100 km^2 of area, at a resolution of 0.54 meters per pixel. The data consists of nucleated villages in four continents and 15 different countries. Manual annotations by six different persons were used as ground truth, and were fused together using a Kalman Filtering framework. Five fold cross validation, with 25% test data, was used during training. A total of 69 features were used. The features were given as input to Adaboost algorithm to yield the final classifier. Our results show an Equal Error Rate (EER) of around 3.4%. Once the classifier was trained, we ran it on a large image extracted from Google Earth™ spanning approximately 50 km^2, no portion of which was used in training the classifier. Our classifier extracted all villages reasonably accurately, with approximately 2.3% false positives at less than 0.01% false negatives when compared to the ground truth labeling.

We have demonstrated a system to accurately segment nucleated villages from freely available satellite imagery available in online geographic information systems such as Google Earth™. The system employs phase gradient features along with color and cornerness measures to characterize the nucleated villages. To deal with the within-class variability of nucleated villages, we trained the system on images taken from varied locations and by different satellite sensors, and used six manual annotations to generate the ground truth, merged using a Kalman filtering framework. Our approach shows excellent results when tested on a large image captured using our crawler from Google Earth™.

The datasets used in this paper, along with the ground truths and the code are available at `http://cvlab.lums.edu.pk/villagefinder`