

## Segmentation and Contour Extraction in Biological Transmission Electron Microscope Images with ‘Bag-of-Features’ Method in Machine Learning

Gen Maeda<sup>1</sup>, Shoki Tezuka<sup>1</sup>, Shohei Sakamoto<sup>1</sup>, Misuzu Baba<sup>2</sup> and Norio Baba<sup>1</sup>

<sup>1</sup> Major of Informatics, Graduate School, Kogakuin University, Shinjuku, Japan

<sup>2</sup> Research Institute for Science and Technology, Kogakuin University, Hachioji, Japan

When observing minute objects and structures with a transmission electron microscope (TEM), it is necessary to search a very large specimen area for them and keep them inside the field of view. Because of the relatively small field of view of a TEM this manual work is very time consuming. Several research works to overcome the problem have been reported. For example, an automatic method picking up of particles in cryo-electron microscopy [1], and unique one that finds out localized fine crystalline samples such as graphitized carbon [2]. However, at present, researches of the general-purpose technique of automatically and mechanically searching for any biological object appearing in thin section images or in tomographic images with electron CT is seemed to be few.

For the above purpose, we have attempted some different types of machine learnings. The electron microscopic images have no colour information, and then, each biological object must be recognized with only morphological image features. The biological thin section images are almost formed with distribution of staining granules. If the difference of the distribution pattern for every object is captured by a computer, the automatic identification may be achieved. It is thinkable that the second order statistical feature of the distribution of granules is rather effective than the first order feature (granular image pattern itself). For the second order statistics, the texture analysis and the local image frequency spectrum with a wavelet were adopted. Based on these things, we propose the following method. The method consists of two techniques.

The first is an image classification technique which is an application of the method called ‘Bag-of-Features’ [3]. Image features (2nd order statistics) of very large number of image patches are previously analyzed by using the texture analysis and the Gabor wavelet spectral analysis, and their patches are classified into important elementary categories as a supervised learning. In a real TEM observation by a user, an acquisition image is classified according to the database obtained by the above. Then, if it is judged that the target biological objects exist in the field of view, the image is taken for saving.

The second is a contour extraction technique. This is realized by applying the Gabor wavelet transform, which has the property of responding strongly to a specific frequency and angle sensitive to contours. By using this technique, it becomes possible to automatically collect the object region in images. In this technique, “rough” segmentation is previously performed over the entire image. At first, the entire image is divided into many block images, and then the divided images are fed into a discriminator, and they are classified with the first technique mentioned above. By estimating the approximate position and type of the contour of the biological object based on the result of rough segmentation, the automatic trace of the contour is effectively carried out. The contour extraction clearly distinguishes each object.

The proposed method applied to thin section images of yeast cells taken with a TEM. Various organelles appeared in the images. In the experiments, it was examined that how accurately images of each cell organelle were classified and contours were traced using the rough segmentation.

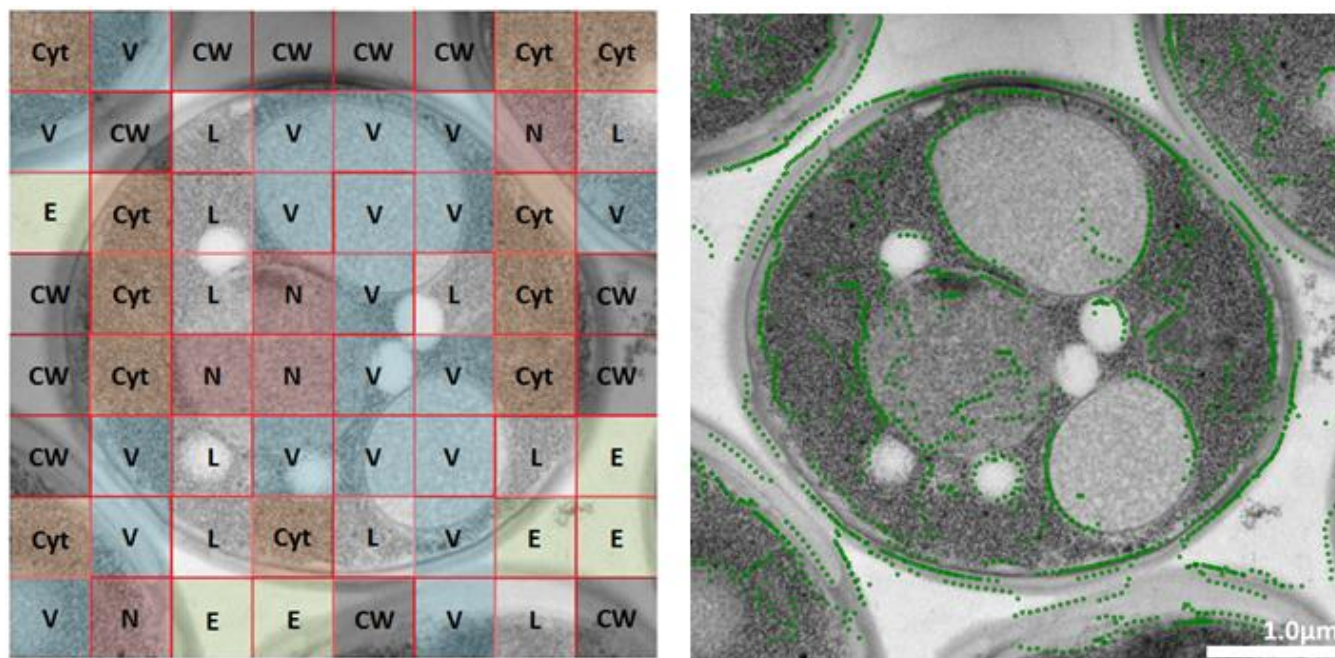
As a result, with regard to the image classification of each object, by using a combination of the Gabor wavelet feature and the statistical texture feature, seven cell organelles were classified at the average accuracy rate 65.2%. In addition, every cell organelle was classified at the accuracy rate of 60 to 80%. Also, with regard to the contour tracing, not only clear cell walls and vacuole membrane contours, but also relatively ambiguous nucleus membrane contours could be automatically tracked with high accuracy.

References:

- [1] Toshihiko Ogura and Chikara Sato, *Journal of Structural Biology* **146** (2004), pp.233.
- [2] P.Toth *et al*, *Ultramicroscopy* **129** (2013), pp.53.
- [3] Taichi Joutou and Kenji Yanai, *IPSI SIG Technical Reports* **42** (2007), pp.201.

feature name \ class name	cell wall	nucleus	vacuole	glycogen granule	lipid	cytosol	mitochondria
only Gabor	90.54	78.43	31.34	90.54	61.64	76.00	31.58
Gabor + statistical texture	77.03	62.75	70.15	82.43	71.23	56.00	36.84

**Table 1.** Result of image classification with only Gabor feature or combination of Gabor and statistical texture feature (%).



**Figure 1.** Left: Result of “rough” segmentation. Right: Result of automatic contour search using “rough” segmentation.