Feature selection methods applied to multiscale texture

DANIELA M. U. SABINO[1], LUCIANO DA F. COSTA[1]

[1]IFSC - Instituto de Física de São Carlos, Av. Trabalhador Sancarlense, 400, 13560-970 São Carlos, SP, Brasil
{dani,luciano}@if.sc.usp.br

Abstract. Texture is often considered as a repeating pattern of local variations in image intensity which are too fine to be distinguished as separate objects at the observed resolution. A multiscale approach is often necessary to determine particular spatial scales more suitable for pattern recognition.

1 Introduction

This paper intends to characterize texture of blood images based on gray level cooccurrence matrices (GLCM) and to compare that characterization to a human subjective evaluation of texture in similar images. This study does not attempt to segment region of an image based on texture, however it attempt to classify texture regions according to features, proposed by Haralick [1], in order to identify different textures, enhancing the characterization of blood smear images. Feature selection helps to determine of relevant features for normal leukocyte discrimination.

2 Materials and Methods

The program *Leuko* processes leukocyte segmentation [2] and the color cytoplasm region is preprocessed using: (i) RGB to grayscale transformation; (ii) median filter ($w = 2$); (iii) histogram equalization; (iv) quantization in 16 gray levels, required to reduce computational effort and reduce noisy gray level flutuations. In the current multiscale approach, a window (w) varies as we calculate GLCM [1], relating two pixels in a given direction and distance using w = 2x2 to 20x20. Measurements are extracted from the isotropic GLCM addressing contrast (amount of local variations) and orderliness (regularity of the pixel values within the w) of the block. Contrast, entropy, energy, correlation and homogeneity parameters are extracted from GLCM for each w and interpreted in terms of its mean value.

We suggest the use of feature selection in order to identify spatial scales more appropriate for texture description of different blood cell types. Cytoplasm of all leukocytes can present fine grains, however only lymphocyte and monocyte have high occurrence of such structures against eosinophils, basophils and neutrophils, which have large granules (granulocitic lineage).

3 Results

Textural criteria can discriminate among different leukocytes only at particular scales (Figure1). Inertia (i), energy (ene) and homogeneity (h) are most discriminative

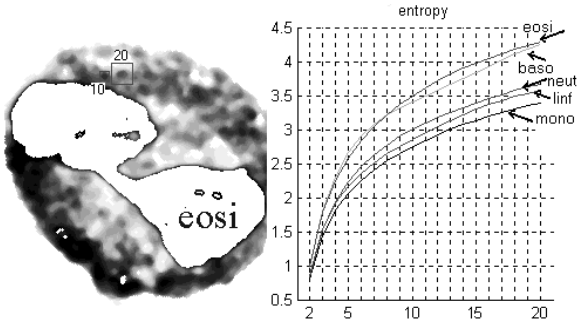

Figure 1: Eosinophil and multiscale entropy: the greater the w, the more discriminative the entropies.

at smaller scales while entropy (ent) and correlation (c), the opposite. These parameters regard the local and global aspects of the texture: c, ene and h curves for different cells intercepts themselves for greater w and they are most discriminative for smaller w. On the other side,a small w privileges a relatively uniform distribution of the texture, therefore both c and ent curves exhibit similar behavior, improving discrimination as w increases.

Preliminary mean correct classification rates regarding the 6 best features follow: i: 88%, ent: 86%, ene: 90%, c: 90% and h: 92%. These are promising rates using Bayesian classification, however a larger database must be considered in advance.

Acknowledgements

We thank FMRP (USP) Hematology Laboratory for biological information and Giemsa-stained blood smear images.

References

[1] R. M. Haralick , Statistical and structural approaches to texture, *Proc. Ieee*, 67(5), pp. 786-804, (1979).

[2] D. M. U. Sabino, L da F Costa, R T Calado, M A Zago, Automatic leukemia diagnosis, Acta Microscopica (2003) - *under publication.*