AN MCF BASED ON GABOR FUNCTIONS TO ANALYZE REMOTE SENSED IMAGES

J. García-Consuegra, G. Cisneros** and A. Fernandez Caballero**

*Departamento de Informatica
Universidad de Castilla-La Mancha
Campus Universitario, 02071- Albacete, Spain
e-mail: {jdgarcia,caballer}@info-ab.uclm.es

**Grupo de Tratamiento de Imagenes
Universidad Politecnica de Madrid
ETSI Telecomunicacion, Ciudad Universitaria, 28040 - Madrid, Spain

ABSTRACT

In this paper, a comparative study of two widely used MCFs (Multi-Channel Filters) based on Gabor Functions, with the interest in segmenting remote sensed imagery, is carried out. Specially, with high spatial resolution images such as aerial photography and those images scanned by the Daedalus-1268 multispectral scanner on board an airborne platform, in which textons (trees) are shown as isolated elements. Furthermore, a new MCF, based on the performance yielded by other MCFs and the spectral characteristics of the images, is proposed.

1. INTRODUCTION

Remote sensed image segmentation involves locating and identifying the different cover types on the earth's surface. The segmentation process is accomplished in virtue of the spectral, textural and contextual features that characterize these cover types. When the interest is focused on the woody crop (almond and olive tree) location, two main problems must be faced. The crop does not cover the entire plot, that is to say, in the same plot, trees are commonly isolated from the rest. The second problem is inherited from the spatial and spectral resolution of the remote sensed images used in their discrimination and location, since they determine the techniques to be applied. So, in low-resolution images, pixels represent two or more cover types, in this case, soil and trees; i. e., there are mixed pixels. There is a great number of elements which affect the digital scanned levels of these crops: soil type, percentage of covered area, phenological state of the crop, sensor IFOV, etc., making the classification of this kind of crops a very complicated, and some times, an impossible task. There are many techniques used with different performances varying from the application of mixed pixel treatment [9,12,13] to the use of the phenological evolution [11], as displayed by the temporal change in their spectral signature, in multitemporal data studies. Nevertheless, in high-resolution images, the plots appear as textured regions with the soil as the background and trees as the textons (figures 1 and 2).

In the scientific community, a great number of approaches to carry out the textural region segmentation can be found [3,13,14]. While these approaches may work in specific circumstances, there is no doubt that the most efficient and robust texture analyzer is the human visual system.

General visual system performance could be accounted for by a number of channels tuned to different spatial frequencies in the early stages of the human vision system [4]. There are a number of representations that adopt a compromise between resolution in spatial and frequency domains. These include radially symmetric and directionally tuned offset Gaussian differences [5], as well as the use of cartesian and polar separable filters, prolate spheroidal functions, and wavelets. In particular, this paper is focused on the study of high spatial resolution-remote sensed image segmentation through different MCF based on multi-scale Gabor image representation [6, 8,10].

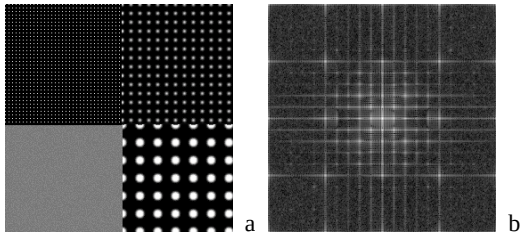


Figure 1. (a) A synthetic image showing different textural and non-textural patterns; and (b), its power spectrum.

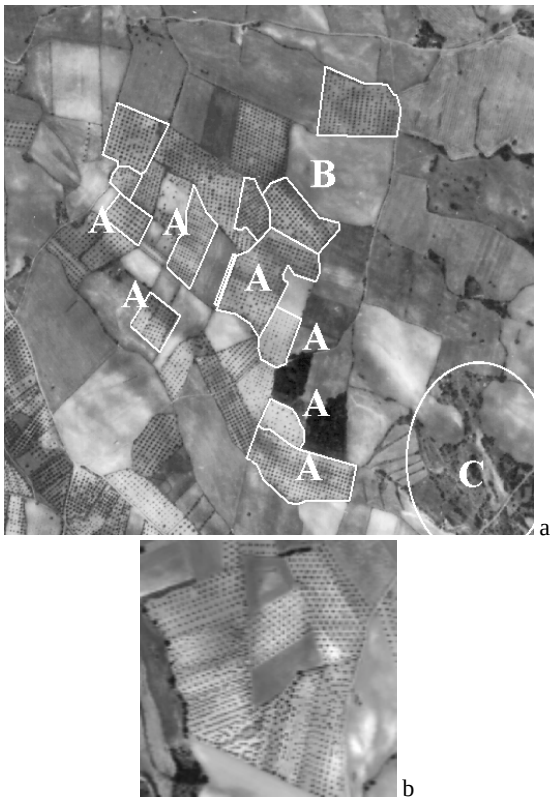
2. MCF CHARACTERIZATION

The main issues involved in the multi-channel filtering approach to texture analysis are: (1) functional characterization of the channels and the number of channels, (2) extraction of appropriate texture features from the filtered images, (3) relationship between channels (dependent vs. independent), and (4) segmentation of texture features. Different multi-channel filtering techniques proposed in the literature differ in their approach in one or more of the above issues.

Daugman, in [2], showed that 2-D Gabor filters are conjointly optimal in providing the maximum possible resolution for information in both the spatial and frequency domains. The filters are also a good approximation to the receptive field profiles of simple cells in the striate cortex, Marcelja [7]. The filters are Gaussian in nature, having a center frequency, an orientation and two spatial extent parameters. By varying these parameters a filter may be made to pass any elliptical region of the complex frequency domain [1]. By choosing a set of such filters, the

frequency domain may be uniformly sampled. In the spatial domain, any filter takes the form of a sinusoidally modulated Gaussian profile, rotated at a particular orientation.

A Gabor filter samples the frequency space of an image providing information about oriented, band-pass, localized textures. Each Gabor filter has a Gaussian profile in the frequency domain of an image. The Gaussian has a center frequency, (u, v), two spatial extent parameters (σ_x, σ_y) and the filter is rotated through an angle of θ degrees about the origin. In the spatial domain the filter is a Gaussian modulated sine wave grating. The spatial extent parameters of the Gaussian are inversely proportional to σ_x and σ_y and the sine wave grating has a frequency $\sqrt{u^2+v^2}$ and orientation θ. These five parameters uniquely determine the filter. The radial frequencies are usually determined 1 octave apart [4,8,15], due to several experiments about the simple cells in the human visual cortex. The rest of parameters depend on each particular case.



a

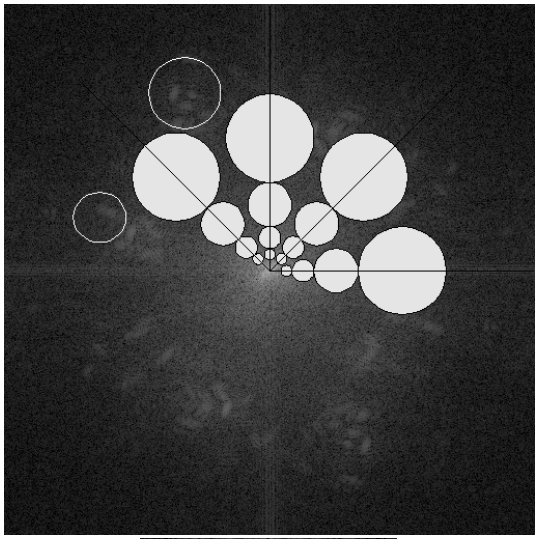
b

Figure 2. (a) An aerial photograph with some test regions with the boundaries highlighted and (b) a high spatial-resolution band of a multispectral image acquired by a Daedalus-1268 sensor.

Usually, the channel set configuration is accomplished by two approaches: to cover some specific areas in, or searching an uniformly sample of, the frequency domain. The former approach involves a previous knowledge of the spectral characteristics of the textured region to be segmented. This knowledge can be obtained in a supervised or non-supervised way. The non-supervised way is usually carried out finding the local maximum in the power spectrum [16], and filters centered around them are chosen. In the latter approach, the amount of overlap between the filters in the set can be designed, either to be as small as possible so that the decomposition is nearly orthogonal, allowing the reconstruction of the input image from the filtered images [8], or to be high enough, to obtain a flat overall response [15].

When the number and characteristics of the channels are established by finding the spectral peaks in the power spectrum of the image [16], the number of channels can be increased as the harmonics from the texture pattern are more representative (figure 1b). This involves a great number of redundant channels, which makes the feature selection/extraction as well as the classification process more difficult.

Once the automatic selection of Gabor filters has been discarded, the only option is a set of filters selected in an ad-hoc manner with pre-determined parameters. The best choices are those configurations in which the frequency domain is uniformly sampled, in case an automatic segmentation algorithm is desired, since no previous knowledge is needed. There have been a number of attempts to optimize the number of used channels, although evidence suggests that segmentation error increases as the number of filters decreases, particularly near texture boundaries. A widely-used set of channels is the 4x4 [4,8,15] with 4 frequencies at $f_N/2$, $f_N/4$, $f_N/8$ and $f_N/16$ cycles/pixel (f_N being the Nyquist frequency=1/2 cycles/pixel); and 4 orientations at 0º, 45º, 90º and 135º; with the radial frequency bandwidth equal to 1 octave, and with an aspect ratio ($\gamma=B_a/B_r=1$) equal to 1, since the computational simplicity afforded by circular Gabor filters outweighs the slight benefits achieved by using elliptical filters.


a

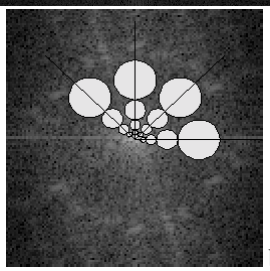

b

Figure 3. Power spectrum of figures 2a and 2b, respectively.

On the other hand, there are other sets where the radial frequencies are determined 1/2 octave apart [15], in order to reach a flatness of overall response. The set proposed by Swapp, in [15], has got 8 frequencies $1/2\sqrt{2}$, 1/4, $1/4\sqrt{2}$, 1/8, $1/8\sqrt{2}$, 1/16, $1/16\sqrt{2}$, 1/32 cycles/pixel; and 6 orientations 0º, 30º, 60º, 90º, 120º and 150º; with an aspect ratio ($\gamma=B_a/B_r=1$) equal to 1. An MCF like this one could represent a more appropriate solution to our necessities for two reasons: a flat cover of the frequency domain involves a better discrimination of the textured regions versus the background; and, the classification process is improved since the number of features is increased. However, a new problem appears with this MCF: the mask size increases, and therefore, the computational cost as well as the bandwidth decreases.

In figure 3a, the highest frequency channels of the MCF 4x4 almost cover all the frequency range of interest. Only those areas highlighted by a circle in figure 3a are not covered. Nevertheless, in figure 3b, these channels do not cover the interest areas. At this point, it is important to remember that the Gabor function is represented, in the frequency domain, by a summation of two Gaussians symmetrically located at $-\omega/2\pi$ and $\omega/2\pi$, where ω is the radial frequency. So, the effective area occupied by a Gabor filter in the frequency domain is not equally weighted.

In this paper, an MCF based on the four channels of higher frequency of the MCF 4x4, and the four channels of higher frequency with orientations 30º, 60º, 120º and 150º of the MCF 8x6 is proposed. This channel selection is based on the energy concentration in the medium-high frequency range in these kinds of images, making an analysis of the low-medium frequency range useless. The channels with 0º and 90º orientation in MCF 8x6 were discarded due to the square form of the domain in frequency covered by the image. So, the frequency range is the minor for the orientations of 0º and 90º whilst the range is superior for 45º and 135º.

This filter bank proposes an ad-hoc solution to the problem of establishing the number and characterization of the channels. So, only the medium-high frequency range is sampled. In that way, there is a computational cost saving as the useless channels are erased. Unfortunately, the advantages of the hierarchical approach [10] are lost since MCF works with the full spatial resolution. On the other hand, the problems arisen from the pyramidal segmentation process are avoided [8, 10].

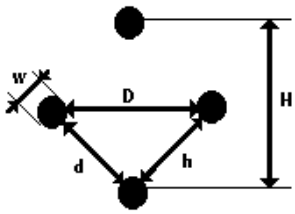


Figure 4. Almond and olive trees are cultivated in diamond-like shape, where w is the tree-canopy diameter, and d, h, D and H represent the distance between the nearest trees in four directions.

3. DATA AND RESULTS

Our interest in this paper is focused on the discrimination of crops with a regular textural pattern, such as woody crops (almond and olive trees, vineyard, and so on). These crops are cultivated in a diamond-like shape (figure 4), which in a high spatial resolution remotely sensed image can be seen as the test plots shown in figures 1 and 2. Each tree can be considered as the individual texture element (texton) that is repeated in a structured way in each region. The distances d, D, h and H are directly related with the plot representation in the frequency domain, whilst w is related with the amount of energy of the plot.

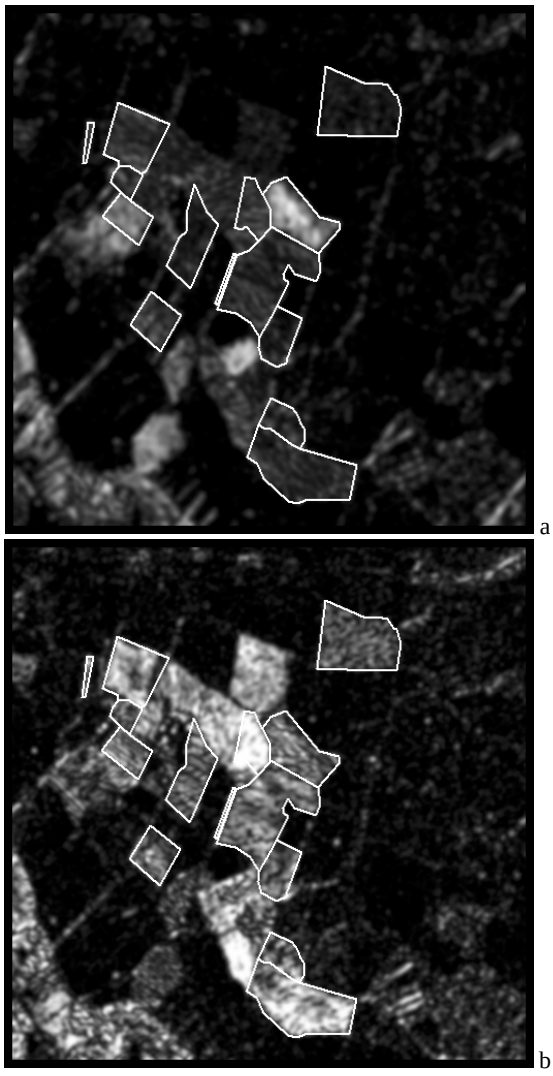


Figure 5. (a) MCF 4x4 output for 1/2 cycles/pixel and 135º; and, (b) MCF 8x6 output for $1/2\sqrt{2}$ cycles/pixel and 150º from figure 2a.

The chosen images, with the interest zones (figure 2), are an aerial photography taken at a scale of 1:30,000 and an image scanned with a multispectral scanner Daedalus 1268, also called Airborne Thematic Mapper (ATM), with a spatial resolution of

3,5 m. approximately. The scanner separates incoming radiation into eleven spectral bands ranging from visible blue to thermal infrared depending on the spectral configuration selected. The configuration of the Daedalus-1268 multispectral scanner is extremely flexible (variable IFOV, four scanning speed, spectral configuration Thematic Mapper compatible, CZCS simulator, 3-5µm and 8.5-13µm high resolution TIR channels, etc.). This choice was been taken because the zone edges are easily distinguished by a photo-interpreter, besides the low cost and high availability of aerial photographs and the possibilities offered by multispectral images.

The covered areas are located in the Guadalajara and Albacete provinces (Spain), respectively. The above mentioned olive and almond crops are very extended in central and south Spain, as in the Mediterranean countries in general, and are of great social importance. Plots were determined by fieldwork.

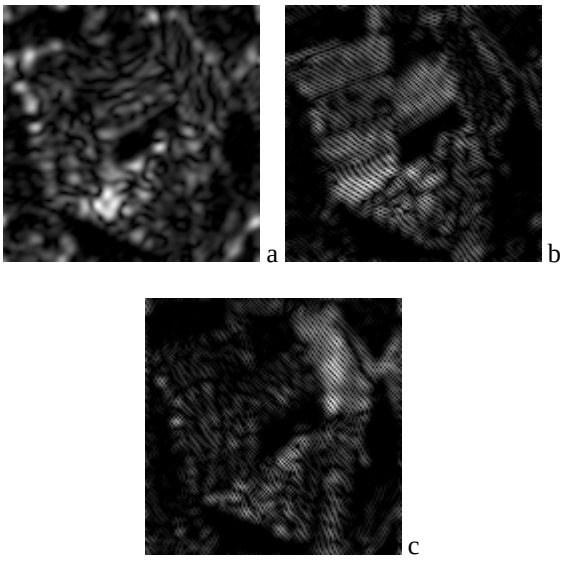


Figure 6. (a) MCF 4x4 outputs for 1/2 cycles/pixel with 135º; and, (b - c) MCF 8x6 output for $1/2\sqrt{2}$ cycles/pixel with 120º and 150º from figure 2b.

4. CONCLUSIONS

The main frequencies of the image regions are in the medium-high frequency range (see figure 3). Thus, the low frequencies rarely contribute to a significant information. Even, in some cases, these channel outputs can be elements of confusion (in function of the similarity criterion and of the characteristic weight).

Anyone could suppose that these conclusions arise of the analysis of a reduced number of images. However, there are two events that reinforce them. The cultivation frame of these kinds of crops is as small as possible in order to increase the number of trees in a plot as well as its productivity. The ratio economic cost/spatial resolution/surface covered by the image makes preferable to reduce the spatial resolution as much as possible. Both events involve in an energy increase in the high-frequency range, making useless any effort to analyze the low-medium frequency range. At the same time, this analysis suggests the appearance of new channels in the high frequencies beyond that one proposed in the MCF 4x4.

As shown in figures 5 and 6, the selection of the higher-frequency filters of the MCF 8x6 gives a more compact output in some regions (test plots labeled with A, in figure 2), facilitating the following steps. On the other hand, the four channels of higher frequency of the MCF 4x4 cover the medium-high frequency range. This is the area where there is the main concentration of energy in plots cultivated with woody crops. Thus, the proposed MCF takes the advantages of both MCFs.

5. FUTURE WORK

There is an increase in the number of little regions (noise) in the result obtained by the four channels of higher frequency of the MCF 8x6. In some cases, these little regions are isolated trees, tree rows or open forest. The aim of our future work is to expand this methodology to those cases in which trees are isolated or the texture does not form a regular structure.

ACKNOWLEDGEMENTS

The authors wish to thank the Departamento de Teledetección y Aeronomía of the Instituto Nacional de Técnica Aeroespacial for the transferred images.

REFERENCES

[1] A.C. Bovik, M. Clark and W.S. Gesiler," Multichannel texture analysis using localized spatial filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, 1990, pp 55-73.

[2] J. Daugman, "Uncertainty Relation for Resolution in the Space, Spatial-Frequency adn Oriented Optimized by two Dimensional Visual Cortical Filters", *Journal of the Optical Society of America, A,* 2, 1985, pp. 1160-1169.

[3] R. M. Haralick, K. Shammugam and I. Distein, "Textural Features for Image Classification", *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, 1973, pp. 610-621.

[4] A.K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters", *Pattern Recognition*, vol. 24, 1991, pp. 1167-1186.

[5] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms", *Journal of the Optical Society of America, A*, vol. 7, no. 5, 1990, pp. 923-932.

[6] S.G. Mallat, "A theory for multirresolution signal decomposition: the wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7, vol. 11, 1989, pp. 674-693.

[7] S. Marcelja, "Mathematical description of the response of simple cortical cells", *Journal of the Optical Society of America, A.*, 70, 1980, pp. 1297-1300.

[8] O. Nestares et al., "Automatic computation of the area irradiated by ultrashort laser pulses in Sb materials through texture segmentation of TEM images", *Ultramicroscopy*, 66(1-2), 1996, pp. 101-115.

[9] N.A. Quarmby, J.R.G. Townshend, J.J. Settle, K.H. White, M. Milnes, T.L. Hindle and N. Silleos, "Linear mixture modelling applied to AVHRR data for crop area estimation," *International Journal of Remote Sensing,* vol 11, 1992, pp. 415-425.

[10] E. Salari and Z. Ling, "Texture segmentation using hierarchical wavelet decomposition", *Pattern Recognition*, vol. 28, 1995, pp. 1819-1824.

[11] D. Segarra et al., "El viñedo en Castilla-La Mancha y su seguimiento mediante imágenes Landsat-5 TM", *Proceedings of the VII Congreso Nacional de Teledetección,* Santiago de Compostela, pp. 70-73, 1997.

[12] J.J. Settle and N.A. Drake, "Linear mixing and the estimation of ground cover proportions," *International Journal of Remote Sensing,* vol. 14, 1993, pp. 1159-1177.

[13] Y.E. Shimabukuro and J.A. Smith, "The least-squares mixing models to generate fraction images derived from remote sensing multispectral data," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 29, 1991, pp. 16-20.

[14] J. Strand and T. Taxt, "Local frequency features for texture classification", *Pattern Recognition*, vol. 27, no. 10, pp. 1397-1406, 1994.

[15] D. Swapp, "*Estimation of Visual Textural gradient using Gabor functions*", PhD, Aberdeen University, 1992.

[16] T.N. Tan, "Texture edge detection by modelling visual cortical channels", *Pattern Recognition*, vol. 28, 1995, pp. 1283-1298.