

NEURAL BASED ROTATION AND SCALE INDEPENDENT DETECTION OF TARGETS IN A HYPERSPECTRAL WATERWAY MONITORING SYSTEM

Blanca Priego, Richard J. Duro, Francisco Bellas and Daniel Souto

Integrated Group for Engineering Research, Universidade da Coruña, A. Coruña, Galicia, Spain

Keywords: Hyperspectral images, Rotation and scale detection, Neural classifiers, Autonomous surveillance.

Abstract: This paper is devoted to the presentation of the orientation and scale invariant detection subsystem within the current development of Hywacoss (Hyperspectral waterway control and security system). A neural network ensemble based identification and rotation detection module is considered in order to be able to detect and classify objects in waterways from hyperspectral image cubes in a fast and efficient manner. The neural approach followed is inspired by the orientation detection structures in the visual processing cortex. The system is tested over two different hyperspectral image cubes extracted from simulated waterways to verify its adequate operation.

1 INTRODUCTION

The objective of the *Hywacoss* project is to produce a real time small, light and easy to transport visible and near infrared hyperspectral detection and recognition system that autonomously monitors waterways, especially port and bay areas, and detects and classifies all the traffic, producing alerts when previously unknown objects or behavior patterns arise. *Hywacoss* is a part of a multisensory intelligent monitoring and protection system for ports and waterways called *Watchman* that is aimed at capturing real time information on what is going on in the area being monitored from multiple sensorial sources, both static and moving, and fusing this information to provide a coherent view of all the activities that are taking place within the designated area, identifying, tracking and profiling all targets found. It comprises dedicated hardware and software modules, some of them neural based. Here we will provide a global description of the whole system and a detailed analysis of the neural based modules related to rotation and scale independent overhead target detection from hyperspectral images.

A hyperspectrometer obtains images in which the spectral information of every pixel is collected in hundreds of contiguous discrete spectral bands. Thus, each hyperspectral image contains a large amount of information that can be perceived as a

cube with two spatial and one spectral dimension. The availability of such detailed spectral information for each pixel allows the classification of different materials or targets with an accuracy and discriminative power that are much better than in the case of lower dimensional color representations, such as RGB.

Hyperspectrometers were originally designed as remote sensing instruments operated from highflying planes (Glackin, 1999) and, therefore, presented the handicap of a low spatial resolution. Consequently, analysis methods were developed to provide the segmentation of the images in terms of the ratio of endmembers present in every pixel so as to improve the spatial discrimination of these systems when analyzing different types of covers. Currently, due to improvements in the spatial resolution of the systems and to the new requirements that have come about due to the expansion of the applications for which these systems are used, an increasing demand for spatial-spectral processing techniques has been observed. This is especially patent in ground-based applications (Pan, 2003) (de Juan, 2004), where images are taken close to the subject producing a relatively detailed view. Thus, the emphasis in hyperspectral image processing is no longer placed only on extracting subpixel information, but also on detection and classification of elements within these images based on multiple pixel combinations taking

into account their geometric layout (Li, 2010), that is, their spatial-spectral signatures.

This is the case in this work. The objective of the *Hywacoss* system is to detect and classify maritime traffic within its assigned monitoring area in real time in order to discriminate between normal traffic and possible intruders. Consequently it must be able to obtain hyperspectral images and from them quickly detect and classify whatever objects (ships or others) are present independently from scale or orientation. In addition it must be able to provide a value for the orientation of the object if it is classified as a ship. As hyperspectral images involve hundreds of values per pixel within the image and as we want to detect and classify objects independently from scale and orientation in a timely fashion, in this paper we have considered a system that uses a neural network trained to extract the abundance of a target and a second neural architecture, loosely based on the visual orientation processing structures of certain animals, to perform the detection and orientation determination.

Neurons in the visual cortex display orientation selectivity, which basically leads to the detection of local bars and edges in the visual images and the subsequent encoding of their orientations (Hubel, 1962) (Hubel, 1974). In fact, neighboring neurons in the visual cortex have similar orientation tunings, producing orientation columns or iso-orientation domains (Hubel, 1974). The mechanism through which precise orientation detection is achieved is rather complex, but it can be summarized by saying that each neuron is tuned to a given orientation for which it provides the maximum response. It is through the integration of the responses of the different sets of neurons with different orientation tunings that a decision can be made on the precise orientation of a feature of the image. Some authors have based their approximations for orientation detection on these types of models of the visual cortex over 1 channel or three channel (RGB) images. A clear example of this is (Han 2010), where the authors implement a detailed model using spiking neurons in order to detect the orientation of cars. However, in the field of high dimensional or multichannel imaging orientation detection has mostly been carried out using traditional deterministic or statistical approaches and not brain inspired or neural techniques (Plaza, 2009).

The paper is organized as follows. In the following section we provide a general description of the hyperspectral system that was developed for the *Hywacoss* system. Section 3 will be devoted to the presentation of the neural processing system. In

section 4 we discuss the results of the application of the system in the simulated environments used for its validation and, finally, section 5 provides the main conclusions of this work.

2 HYPERSPPECTROMETER

In terms of the sensing element and associated hardware, we have designed and constructed a small, light and easy to transport push-broom type spectrometer. It is shown in Figure 1. It consists of a moving mirror that captures light that is focalized on a 10 mm long, 10 μ m wide slit which selects a single line from the image each instant of time. The light corresponding to this pixel line is passed through a diffraction grid and its image is focalized on the sensing element of a front-illuminated interline CCD camera. This arrangement produces the images we are using that have a size of 1392x1392 pixels, where each pixel is represented using 1040 spectral bands in the 400-1000 nm wavelength interval (visible to near-infrared). The information obtained is directly sent to a processing computer for image processing and the other tasks required by the *Hywacoss* system.



Figure 1: Hyperspectrometer designed and built for this project.

3 NEURAL PROCESSING SYSTEM

The basic elements of the neural processing system that is going to be used for processing the images obtained from the hyperspectrometer are represented in the block diagram of Figure 2. The system is

made up of two main elements: a target abundance extraction artificial neural network (ANN) and a rotation and target identification ANN.

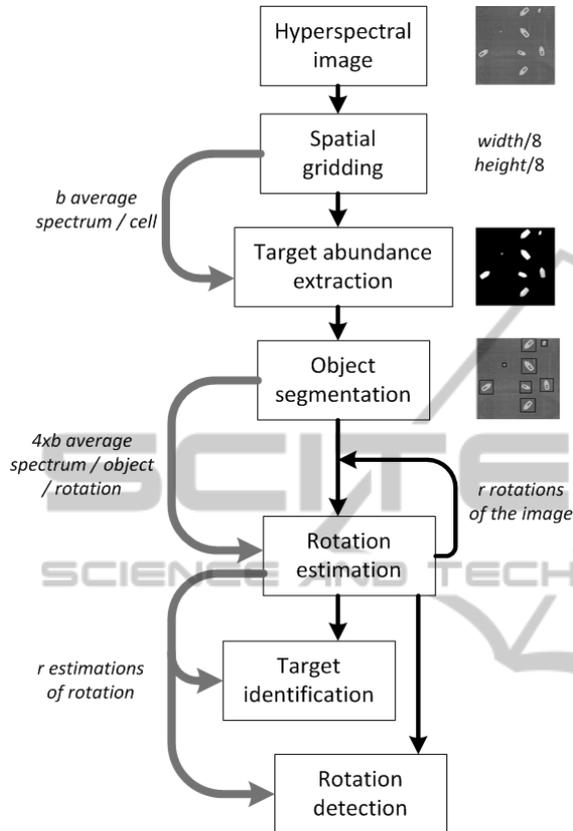


Figure 2: Block diagram representing the basic elements that make up the neuronal processing system and their relations.

Initially, a spatial gridding is performed over the whole image and the proportion of a target present within each grid cell is estimated by the target abundance ANN. The segmented objects are checked by a rotation detection and target identification ANN ensemble. They are able to decide if a particular target is present in a hyperspectral image and to detect its rotation. Obviously, this ensemble is particular to each target.

The main two elements of this system will be described with more detail in the following subsections.

3.1 Object Segmentation

Starting from a hyperspectral image of dimensions $w \times h \times b$ (width w , height h and b number of spectral bands), the first stage of the system is a target abundance extraction step. The hyperspectral image

is spatially downsampled, by means of a grid of $w/8 \times h/8$ cells. Then, an ANN is in charge of deriving the percentage of target present in every cell. Basically this ANN has b inputs corresponding to the average spectrum of the points for the b spectral bands considered. The output of this ANN is the target percentage, a value between 0 and 1. As the target may be present within a very small part of the cell, this is basically an endmember extraction ANN tuned to the particular average spectral features of the target.

The image areas with a percentage higher than 0 are segmented. For each area, the center of mass and the surface are calculated (weighted by the percentage values provided by the ANN). Figure 3 displays an example of this segmentation process. The left image shows an input image with 8 targets (grey areas) and the right image shows the ANN output corresponding to the areas that will be segmented.

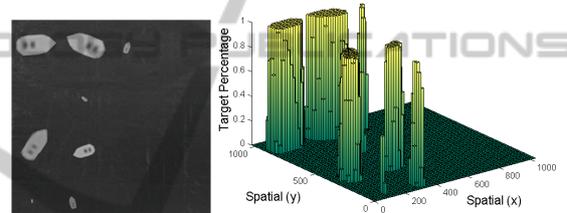


Figure 3: Input image (left) and a 3D representation of the target percentage provided by the ANN (right).

3.2 Rotation Estimation and Identification

Whenever an object has been segmented in the previous step, it is necessary to check if the target can be identified independently of rotation. In order to do this, first of all, a rectangle area that surrounds the segmented object is estimated. The region determined by this rectangle is subdivided into four areas (right image of Figure 4) and the average spectrum is calculated for each area. These $4 \times b$ values are run through the ANN based identification and rotation detection ensemble. Figure 5 depicts the structure of this ensemble. It draws inspiration from the orientation columns within the visual cortex and it is implemented as a set of orientation detectors for the target. Each of these detectors has been trained using the curve displayed in Figure 6 with the peak centered on the angle for which the detector is trained.

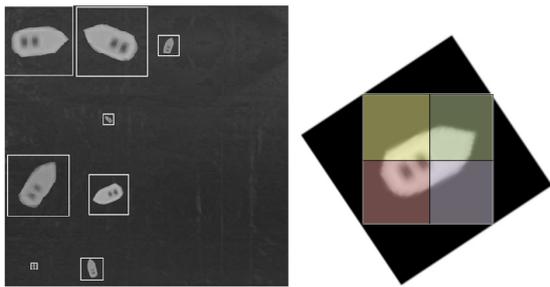


Figure 4: Rectangle areas surrounding the segmented objects are estimated (left) and divided into four parts where the average spectrum is calculated (right).

The outputs of the ANNs in the first layer of the ensemble are used as inputs to a second layer that contains two ANNs, one that decides if the target is correctly identified and another one that provides the rotation angle of the target. Basically, if the target is present at the scale under analysis, the ANNs in the first layer should provide outputs that look very much like those in Figure 7, where the left and central graphs correspond to two positive identifications at different angles and the graph on the right to a negative identification. The second layer ANNs, which have as many inputs as ANNs there are in the first layer, are in charge of both, deciding on the angle depending on the values of the first layer and deciding on whether the identification is positive.

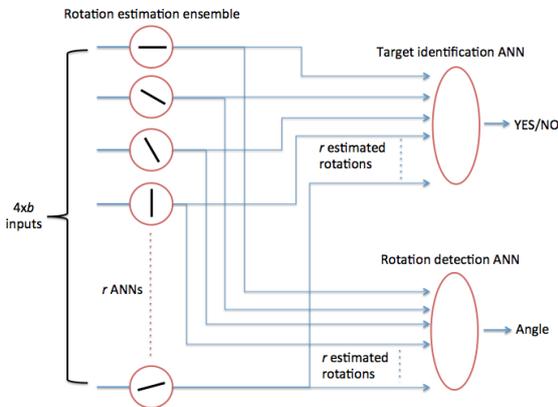


Figure 5: Representation of the ANN based identification and rotation detection ensemble.

4 EXPERIMENTAL RESULTS

To clarify the operation of the system and to show the capabilities of *Hywacoss*, two different scenarios have been selected. Both of them consist in the top-view of a simulated waterway, which contains

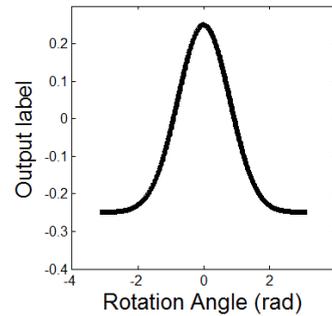


Figure 6: Curve used to train the rotation detection ANNs.

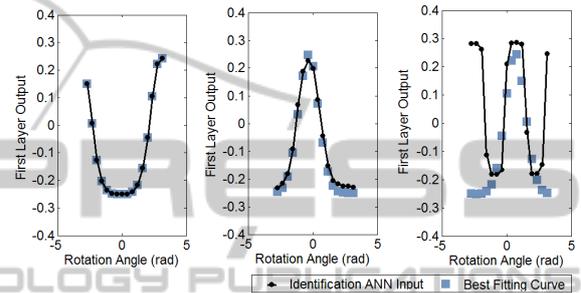


Figure 7: Points obtained by the different ANNs in the first layer of the ensemble (black) and theoretical points for that angle (grey). Two positive identifications (left and center) and a negative one (right) are shown.

different possible targets that must be found: different ships, a person in the sea and a buoy.

In the first scenario, we intend to identify a person in the water that can be confused with other objects present in water, like a buoy. In this case, the estimation of the rotation angle of the target is not relevant, but low error identification is crucial. The second scenario shows two very similar ships and the objective is to identify one of them and to estimate its rotation angle. An image containing only the 588 nm band of the hyperspectral cube of this scenario is shown in Figure 8 (left) while the right image displays the details of the hyperspectral cube (from 493nm band to 1000nm band) of the two ships.

In the following subsections we discuss the specific processing parameters that have been used, the ANN architectures and we detail the training of each phase of the neural system. We also show the results after applying the whole algorithm to the hyperspectral captures.

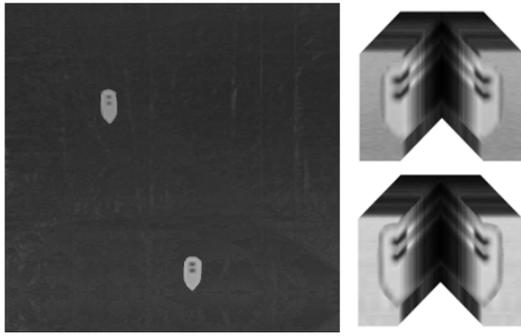


Figure 8: Left: 588nm band of the hyperspectral capture. Right: Interior detail of the hyperspectral cube (bands from 493nm to 1000nm).

4.1 Identification of a Person in the Water

In the first detection problem we try to identify a shipwrecked person with the presence of a buoy and a small ship within the image capture (Figure 9). The hyperspectral image has been spectrally reduced to 64 bands through binning of the 1040 initial bands since we have checked that 64 bands are sufficient to achieve good ANN training results, and increasing the number of inputs to the neural system complicates and slows down the detection. In this case, we have used a feed-forward ANN for the Target abundance extraction, Rotation estimation, Target identification and Rotation detection stages.

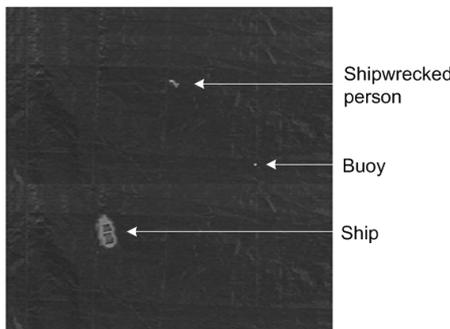


Figure 9: 588nm band of the hyperspectral capture of the first scenario.

The Target abundance extraction ANN has to detect, for any region of the image, what percentage of the shipwrecked person is present. This ANN has 64 inputs that correspond to the average spectrum of every band of all the pixels of the region that is being analyzed and an output that oscillates between 0 and 1 (percentage). The training of the detection ANN was carried out by selecting rectangles of different sizes that could contain or not the target

(shipwrecked person) from different hyperspectral captures. The number of samples we used for the training was around 500.

The Rotation detection ensemble is devoted to returning a specific value depending on the rotation angle of the target as explained in section 3. As commented above, in order to obtain the inputs of this ANN, the region under analysis is divided into four sub-regions (see Figure 4 right) and for each of them the average spectrum is calculated. Consequently, the system will have as inputs the 256 (64x4) values of these four average spectra. The training was performed as described in the previous section.

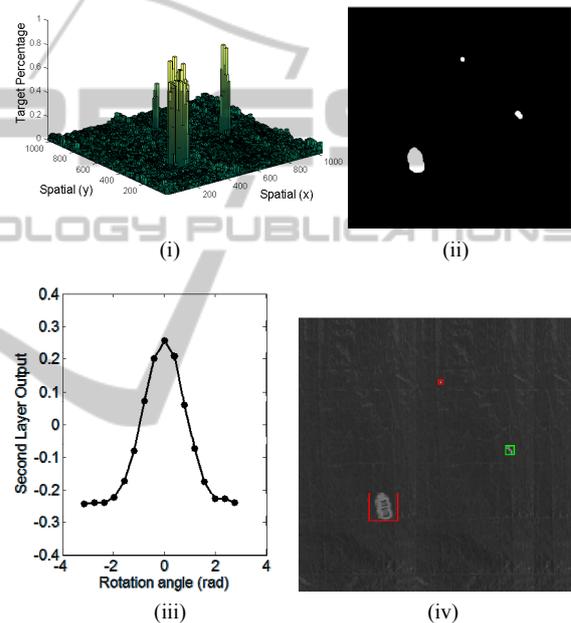


Figure 10: Application steps of the whole system to the hyperspectral image of scenario 1.

In the case of the ANNs in the second layer of the ensemble, that is, the Target identification and Rotation detection stages, a feed-forward ANN was selected. The structure of all the networks of the neural system and the training and validation mean squared errors (MSE) are shown in Table 1. As we can see from the last two columns, the detection errors are very low and the ANNs perform their task in a very successful way.

The whole neural system was applied to the image displayed in Figure 9. Image (i) of Figure 10 shows the output provided by the target abundance extraction ANN and image (ii) the corresponding segmented objects. Next, image (iii) displays the output provided by the rotation estimation ANN that corresponds to the person in the water. Finally,

Table 1: Network structure and training results of the neural system in scenario 1.

		NETWORK STRUCTURE (SHIPWRECKED PERSON)				TRAINING RESULTS		
		Neural Network type	Inputs	Outputs	Hidden layers		MSE	Validation MSE
					Neurons/layer			
	ABUNDANCE EXTRACTION	Feed Forward	64	1	2		$2.18 \cdot 10^{-4}$	$4.85 \cdot 10^{-4}$
					16	1		
	ROTATION CURVE FITTING	Feed Forward	256 (4-64)	1	2		$4.65 \cdot 10^{-5}$	$4.32 \cdot 10^{-4}$
					16	1		
	IDENTIFICATION	Feed Forward	16	1	2		$3.70 \cdot 10^{-4}$	$1.41 \cdot 10^{-4}$
					16	1		
	ROTATION ESTIMATION	Feed Forward	16	1	2		$2.65 \cdot 10^{-3}$	$1.01 \cdot 10^{-3}$
					16	1		

image (iv) represents the positive detection (green square) and the two negative detections (red squares) obtained in this case by the target identification and rotation detection ANNs.

4.2 Ship Discrimination and Rotation Detection

For the second scenario, we have arranged copies of two very similar ships, that, even though they present some different materials, they display the same apparent colour and shape. We have named the target ship as ship₁, and the other ship as ship₂. Two different captures of this scenario are displayed in Figure 11.



Figure 11: 588nm band of two hyperspectral capture of the second scenario.

The ANNs employed in this case are the same used in the previous example, also they use the same number of hidden layers and neurons. Table 2 contains the network details and the mean squared error levels achieved in training and validation. Figure 12 left shows in detail the division into four sub-regions of the original region carried out for the training and execution of the Rotation detection ANN and the average spectra of each of the sub-regions that are used as inputs to this ANN in this

case (right image).

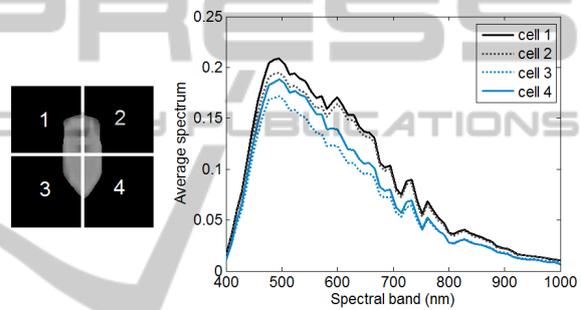


Figure 12: Four sub-regions of the original image used for the ANN training (left) and detail of the inputs for the rotation curve fitting ANN (right).

Again, we have applied the neural system to a set of three hyperspectral images corresponding to scenario 2, and the results obtained are shown in Figure 13. Left images correspond to the 588 band image in the three different cases. Middle images show the segmented objects and right images display the final result with the identification results. The mean squared error in the angle estimation was 1.22° for these test images. As shown, in all the images the target ship has been properly detected (green window).

5 CONCLUSIONS

In this paper we have presented an orientation and scale independent ANN ensemble based target identification and orientation determination system for targets within hyperspectral images. The system is inspired by the way orientation is processed in the visual cortex and provides a fast and efficient way to address the problem of finding objectives

Table 2: Network structure and training results of the neural system in scenario 2.

	NETWORK STRUCTURE (SHIPS)				TRAINING RESULTS		
	Neural Network type	Inputs	Outputs	Layers		MSE	Validation MSE
				Neurons/layer			
ABUNDANCE EXTRACTION	Feed Forward	64	1	2 16 1		$3.91 \cdot 10^{-6}$	$8.66 \cdot 10^{-5}$
ROTATION CURVE FITTING	Feed Forward	256 (4·64)	1	2 16 1		$4.32 \cdot 10^{-6}$	$5.32 \cdot 10^{-5}$
IDENTIFICATION	Feed Forward	16	1	2 16 1		$2.45 \cdot 10^{-4}$	$1.57 \cdot 10^{-4}$
ROTATION ESTIMATION	Feed Forward	16	1	2 16 1		$1.35 \cdot 10^{-4}$	$1.54 \cdot 10^{-3}$

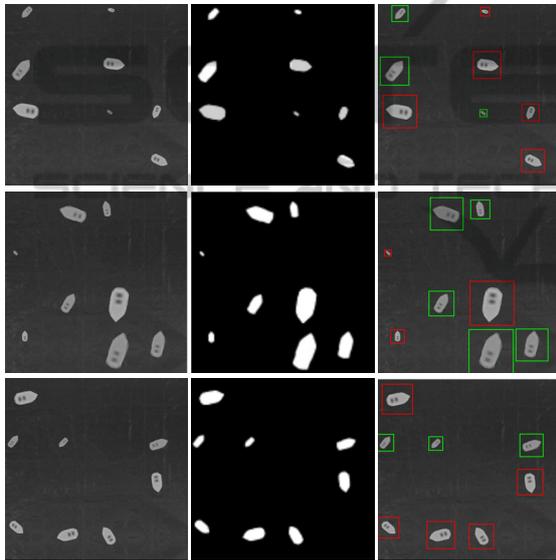


Figure 13: Application of the whole algorithm to hyperspectral images of scenario 2.

independently from scale and rotations and, at the same time provide an accurate estimation of the object's orientation. It has been tested with different hyperspectral images and has been shown to appropriately detect targets as well as differentiate between targets and non-targets that were very similar (similar ships). We are now in the process of implementing these algorithms and their extensions over GPUs in order to be able to run them in real time.

ACKNOWLEDGEMENTS

This work was partially funded by the Xunta de Galicia and European Regional Development Funds through projects 09DPI012166PR and 10DPI005CT

as well as the MCYT of Spain under project TIN2011-28753-C02-01.

REFERENCES

- Glackin D. L. and Peltzer G. R., 1999, Civil, Commercial, and International Remote Sensing Systems and Geoprocessing, *Aerospace Press, American Institute of Aeronautics and Astronautics*.
- Pan, Z., Healey, G. E., Prasad, M., Tromberg, B. J., 2003, Hyperspectral face recognition for homeland security. *Infrared Technology and Applications XXIX. Edited by Andresen, B. F.; Fulop, G. F. Proceedings SPIE, V 5074: 767-776*.
- de Juan, A., Tauler, R., Dyson, R., Marcolli, C., Rault, M., Maeder, M., 2004, Spectroscopic imaging and chemometrics: a powerful combination for global and local sample analysis, *Trends in Analytical chemistry*, 23: 70-79.
- Li, J., Bioucas-Dias, J. M., Plaza, A., 2010, Semisupervised Hyperspectral Image Segmentation Using Multinomial Logistic Regression with Active Learning, *IEEE Transactions on Geoscience and Remote Sensing*, V. 48, N.11, pp. 4085-4097.
- Hubel, D. H., T. N. Wiesel, 1962, Receptive Fields, Binocular Interaction And Functional Architecture In The Cat's Visual Cortex, *Journal of Physiology*, v. 160, pp. 106-154.
- Hubel, D. H., Wiesel, T. N., 1974, Sequence regularity and geometry of orientation columns in the monkey striate cortex, *J Comp Neurol*, Dec 1;158 (3): 267-93.
- Han, W. J., Kim, S. D., Han, I. S., 2010, Bio-inspired visual information processing – the neuromorphic approach *Wseas Transactions on Circuits and Systems*, Issue 7, Volume 9, pp. 441-452.
- Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J. C., Trianni, G., 2009, Recent advances in techniques for hyperspectral image processing, *Remote Sensing of Environment* 113, S110–S122.