African Journal of Agricultural Research Vol. 7(3), pp. 456-466, 19 January, 2012
Available online at http://www.academicjournals.org/AJAR
DOI: 10.5897/AJAR11.1229
ISSN 1991-637X ©2012 Academic Journals

Full Length Research Paper

Hierarchical temporal memory for mapping vineyards using digital aerial photographs

Perea A. J.[1]*, Meroño J. E.[2] and Aguilera M. J.[1]

[1]Agronomic engineer, Dpto. Física aplicada, Edificio Einstein, University of Cordoba, Campus Rabanales, 14071 Spain.
[2]Agronomic engineer, Dpto, Ingeniería Gráfica y Geomática, Edificio Gregor Méndel, University of Cordoba, Campus Rabanales, 14071 Spain.

Accepted 21 December, 2011

The responsibility to manage vineyards in the European Union belongs to a large range of organizations, which need detailed information about geographic data. For this purpose, most member states have developed vineyard registers. This paper has explored an inferential system for vineyard detection using digital aerial photographs. The system has been inspired by a recent memory prediction theory and models the high-level architecture of the human neocortex. In this study, the hierarchical architecture and recognition performance of this Bayesian model were described and applied. Using a photogram received by a photogrammetric UltraCamD® sensor of Vexcel, 96% of the parcels has been detected. The automatic process developed can be easily integrated into the final user's geographical information system and produces useful information for vineyard management.

Key words: Memory-prediction theory, nupic, ultracamd sensor, hierarchical temporal memory, vineyards.

INTRODUCTION

The responsibility to manage vineyards in the European Union belongs to a large range of organizations which need detailed information about geographic data and information systems to advise the decision-making process. The creation, maintenance and update of a vineyard register are assumed by the member states, where the responsibility is shared by public administrations and professional associations. These organizations have to maintain a register of activities about the vine-growing, make decisions related to common agricultural policy, analyze the development of the impact produced by politic decisions and develop the production of high-quality wine in a sustainable environment. Although Europe is the greatest wine producer worldwide, there is not a validated common methodology to update the inventories of vineyard distribution in the region or technical means for supporting the decision-making processes.

Until recently, the European vineyard inventories were produced from field visits and interviews with the farmers, using in some cases the photointerpretation of the aerial photograph. These processes need large time periods for their development and the results are not always satisfactory due to technique restrictions and incomplete information. The cartographic base to elaborate the vineyard register is the cadastre, often obsolete and not in accordance with the plot boundary. In some cases, as France, this cadastre is not available.

The considerable increase in digital technologies makes it possible to automatically analyze images, but also to understand them by providing high-level information on their content.

On the other hand, a considerable increase of very high spatial resolution (VHSR) remote-sensing data is observed and it offers a new potential application in the agricultural domain.

Several studies use advanced digital classification techniques combining with very high resolution remote-sensing data for detecting vine rows (Bobillet et al., 2003), or foliar density of vineyard (Hall et al., 2003).

Other research is focused on the use of active sensors such as radar to classify vineyards (Company et al., 1994; Budgen, 1999; Soria et al., 2010). The results were satisfactory, reaching an accuracy of 80% for

*Corresponding author. E-mail: g12pemoa@uco.es.

vineyard classification. However, these methods are very sensitive to the vine training system (goblet pruning, cordon, trellis, etc.). In contrast, the best results have been obtained by applying Fourier transform based techniques to high-resolution aerial colour photographs, with overall accuracies over 0.82 and Kappa statistic of 0.64 (Ranchin et al., 2001; Wassenaar et al., 2001). These techniques use the shape, texture and orientation, rather than by their spectral response and allow the automation of the process.

In La Peyne valley (Herault, France), Wassenaar et al. (2002) modeled and predicted the hidrological processes associated with French vine, cultivated in Mediterranean region. A method was developed to provide such information by special frequency analysis on very high spatial resolution data. A simple crop geometry model, based on general knowledge and field observations was applied to the Fourier power spectrum of aerial colour imagery.

Gong et al. (2003) compared a number of feature combination techniques in image classification using airborne multispectral digital camera in order to distinguish vineyard from non-vineyard land cover types in northern California. They used image processing techniques applied to raw images to generate feature images including grey level co-occurrence based textural measures, low pass and Laplacian filtering results, Gram-Schmidt orthogonalization, principal components, and normalized difference vegetation index (NDVI).

The maximum likelihood classifier was applied and the most successful result as determined by t-tests of the kappa coefficients was achieved based on the use of texture image of homogeneity obtained from the near infrared image band, NDVI and brightness generated through orthogonalization analysis, obtaining an overall accuracy of 81% for six frames of image tested.

Lately, in France, Delenne et al. (2009) developed a comprehensive and automatic tool for vineyard detection, delineation and characterization using aerial images. The proposed method computes a fast Fourier transform on an aerial image, providing the delineation of vineyards and the accurate evaluation of row orientation and interrow width. They used the red channel of an aerial image and they reach to detect 90% of the parcels; 92% were classified according to their rate of missing vine plants and 81% according to their cultural practice.

Rabatel et al. (2008) proposed an automatic methodology for vineyard detection in aerial images (pixel size: 0.5 m) using fast Fourier transform, resulting vine-plot segmentation, with boundaries in polygonal form and characterization with accurate estimation of interrow width and row orientation. About 84% of vineyard surface was detected.

Da Costa et al. (2007) applied a textural approach to meet this need. Even if the results obtained on several plots (less than 10) are good, it seems difficult to generalize this method as it is applied on a 0.15 cm resolution and needs the user to select a window inside the field he wants to process. Moreover, Delenne et al. (2008) compared two different approaches for vineyard detection and characterization. The first one used directional variations of the contrast feature computed from Haralick´s co-occurrence matrices and the second one was based con a local Fourier transform. 70.8 and 86% of the 271 plot of the study area were correctly classified using the co-occurrence and the frequency method, respectively.

Rodríguez et al. (2008) review some projects related to vineyard identification. The Vinident study use aerial photographs to identify vineyards in areas of France. A more recent work is the Bacchus Project, this project is trying to perform a methodology for vineyard location, parcel identification and vine description, using a high resolution remote sensing data and GIS. However, as Rodríguez et al. (2008) point out, this project could obtain optimum results although the procedure is unfeasible for extensive areas.

On the other hand, new progresses in neuroscience have increased the knowledge about the organization and operation of the cerebral cortex. Therefore it is possible to apply its operation algorithms to the software, which was simplistic and had limited results using neuronal networks up to now.

For decades most artificial intelligence researchers tried to build intelligent machines that did not closely model the actual architecture and processes of the human brain. One of the reasons was that neuroscience provided many details about the brain, but an overall theory of brain function that could be used for designing such models was conspicuously lacking.

A new theory called memory-prediction theory offers a large-scale framework of the processes in the human brain and invites computer scientists to use it in their quest of machine intelligence (Hawkins and Blakeslee, 2004).

The memory-prediction theory is based on the functioning of the human neocortex. It has a hierarchical network structure where each region performs the same basic operation (Hawkins and Blakeslee, 2004).

Hawkins and Blaskeslee (2004) focus his theory on a unified model of how the human neocortex works, but in truth you do not need to have deep interest in neurobiology to see the power of the model. The basic idea is as follow: the brain uses large amounts of memory to create a hierarchical model of the world and uses it to create, by analogy, continuous predictions about future events.

A hierarchical network structure guides the functioning of each region in the cortex. All regions in the hierarchy perform the same basic operation. The inputs to the regions at the lowest levels of the cortical hierarchy come from our senses and are represented by spatial and temporal patterns. The neocortex learns sequences of patterns by storing them in an invariant form in a

Figure 1. Study area.

hierarchical neural network. It recalls the patterns auto-associatively when given only partial or distorted inputs. The structure of stored invariant representations captures the important relationships in the world, independent of the details. The primary function of the neocortex is to make predictions by comparing the knowledge of the invariant structure with the most recent observed details.

Parts of this theory, known as the memory-prediction theory (MPT), are modeled in the hierarchical temporal memory or HTM technology developed by a company called Numenta®.

The new technology of hierarchical temporal memory is able to develop processes of recognition and pattern classification in images with good results for the requirements discussed. Perea et al. (2009) carried out a land use classification of digital aerial photographs using a network based on the hierarchical temporal memory. Good results were reached but this network was limited because the classification used an only pattern in an image.

The general goal of this paper is to propound a methodology based on the hierarchical temporal memory model, proposed by Numenta®, to improve the methodologies used nowadays in the vineyard registers using digital aerial photograph. For this propose a supervised classification and HTM classification were made and compared.

MATERIALS AND METHODS

The area of study was located in Huelva Province, Spain, and includes the municipality of Villalba del Alcor (37° 23 'N, 6° 25' O) (Figure 1).

This is a rectangular area of 6 × 10 km and covers 6 000 ha which is representative of Andalusian dryland crops and has a typical continental Mediterranean climate, characterized by long dry summers and mild winters.

The vineyards of this region are *Denominación de Origen Condado de Huelva* (designation of origin), which covers 4 000 ha of planted vineyards and 2 800 ha of vineyards producing. There are 36 winery producing. The growing of the vineyards documented in the region “El Condado” (Huelva, Spain) is dated in XIV. However there are references about the exchange between Tartessos and Greeks, the grape goods sent to Rome and the tolerance of the Muslims with the growing and producing of vineyard. The wine region, which nowadays is known as

Figure 2. Classified categories.

Denominación de Origen Condado de Huelva covers a large area located in the Southeast of Huelva, bounded by El Andévalo to the north, by the Atlantic Ocean to the South, by the regions of Seville and Cadiz to the East, and by the county town of Huelva to the west. It extends in the lowlands of the Guadalquivir River, from the watershed of its affluent, the Guadiamar river, to the Tinto River.

The dataset used in this research was a photogram received by a photogrammetric Ultracamd® sensor of Vexcel on 23 May 2007, with dimensions of 7 500 × 11 500 pixels. Its bands combination was formed by red, green and blue. The digital aerial photographs had a spacial resolution of 30 cm and were composed of four bands: blue (B), green (G), red (R) and near infrared (IR).

Digital vector maps, color orthophotos and digital terrain models were used to orthorectify the image, select training areas and validate the classifications. Data map was projected using the UTM system (ED-1950, UTM-Zone30N). Also the study area was visited to determinate land uses.

The system was developed to distinguish the following land covers: vineyards (*Vitis vinifera* L.); other uses: bare soil, irrigated land, olive groves (*Olea europaea* L.) and urban soil (Figure 2).

The ERDAS Imagine 9.0 software (Leica Geosystems Geospatial Imaging, Norcross, Georgia, USA) was used to carry out the supervised classification. In the case of HTM classification, Nupic® (Numenta platform for intelligent computing), software for implementing HTMs developed by Numenta was used.

The methodology begins with the calculation the normalized difference vegetation index NDVI index. We obtained an image with the desired combination of bands and proceeded to make classifications. Finally, we validate the results of such classifications.

Obtaining the normalized difference vegetation index (NDVI) index

Vegetation has very characteristic spectral behavior. It shows high absorption of red wavelengths, yet it exhibits high reflectivity with respect to the near infrared ones.

The NDVI (normalized difference vegetation index) was obtained so as to highlight the different spectral behaviors of each type of ground cover. The reflectivity image was obtained by calculating this index following a study of the influence of the calculation of apparent reflectance as a reference in obtaining the green vegetation index (NDVI) and its cartographic expression, which showed a positive effect (Marini, 2006).

This index is based on the difference between the maximum absorption in the red (690 nm), owing to chlorophyll pigments, and the maximum reflection in the near infrared (800 nm), owing to the cellular structure of leaves (Haboudane et al., 2004). Using narrow hyperspectral bands, this index is quantified according to the following equation:

$$NDVI = \frac{(R_{NIR} - R_{RED})}{R_{NIR} + R_{RED}}$$

Where R_{NIR} and R_{RED}, are reflectance in the near infrared band (R_{800} nm) and the red band (R_{690} nm), respectively.

Supervised classification

The Bayesian Classifier of maximum probability was used to classify the image. This algorithm is the most exact of the classifiers in the ERDAS Imagine 9.0® system because it takes into consideration the largest number of parameters for its analysis and because of the variability of the classes using a covariance matrix. For this type classification, an image composed of IRGB bands and NDVI index was used.

Hierarchical temporal memory (HTM)

Hierarchical temporal memory is a technology that replicates the

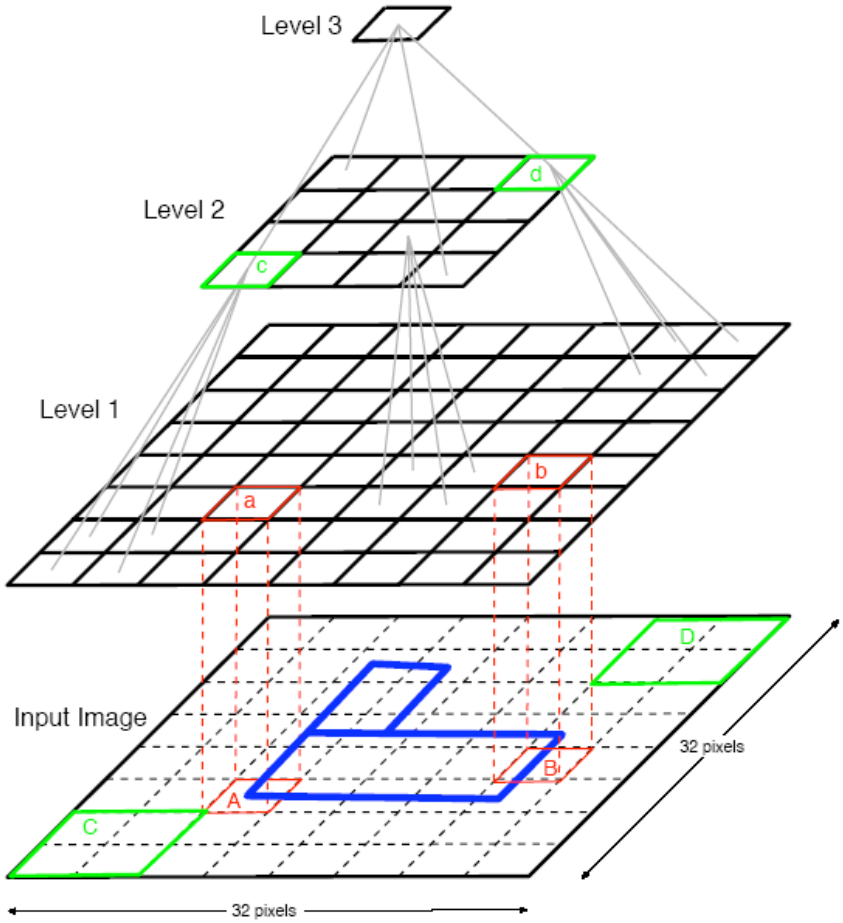

Figure 3. The HTM (hierarchical temporal memory) model with three layers of nodes.

structural and algorithmic properties of the neocortex (Hawkins and George, 2007). HTM is organized as a tree shaped hierarchy of nodes (Figure 3). The HTM (hierarchical temporal memory) model with three layers of nodes. Each sub region in level 1 receives image fragment of size 4×4 pixels. Each subregion in level 2 receives input from 4 children in level 1. A single sub region in level 3 receives input from all level 2 subregions (George and Jaros, 2007, Figure 4). All objects in the world have a structure. This structure is hierarchical in both space and time. HTM is also hierarchical in both space and time, and therefore it can efficiently represent the structure of the world.

Hierarchical temporal memory (HTM) networks consist of several layers or levels of nodes, with one node at the top level. HTM networks operate in two stages: the learning stage and the inference stage. During the learning stage, the network is exposed to training patterns, and it then builds a model of this data. During the inference stage, the network recognizes the new, usually unseen, test patterns. More concretely, during a (supervised) learning stage, the network learns what pattern belongs to what category, while during the inference stage the network will generate a belief distribution over these categories for every new pattern it sees. Belief distributions (represented by belief vectors) are a measure of belief that the input pattern belongs to one of the categories.

All of the nodes (except the top node used in supervised learning) process information in the same way, so we will now explain the operation of such a node.

Operation of nodes during learning

During the learning mode, the node is receiving inputs and measuring their statistics. The spatial pooler learns a mapping from a potentially infinite number of input patterns to a finite number of quantization centers. The output of the spatial pooler, which is considered as an input to the temporal pooler, is expressed in terms of its quantization centers. This stage can be seen as a preprocessing step for the temporal pooler, simplifying its input. The temporal pooler learns temporal groups, which are groups of quantization centers that frequently occur close together in time. The output of the temporal pooler is in terms of the temporal groups that it has learned (George,and Jaros, 2007).

Operation of spatial pooler during learning

The spatial pooler has two stages of operation:

1. During the learning stage it quantizes the input patterns and memorizes the quantization centers.
2. Once these quantization centers are learned, it produces outputs in terms of these quantization centers during the inference stage (George and Jaros, 2007).

The spatial poolers from nodes at the first level receive raw data from the sensor, while the spatial poolers from nodes higher in the

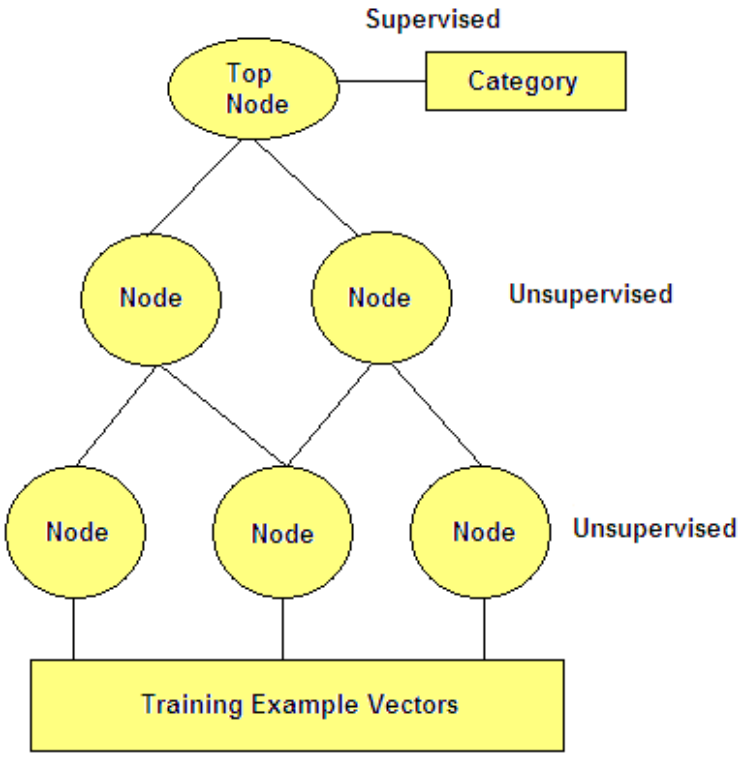

Figure 4. HTM network.

hierarchy receive the outputs from child nodes. The inputs to the spatial poolers of nodes higher in the hierarchy are the concatenations of the output of their child nodes. The input to the spatial pooler is represented by a row vector, and the role of the spatial pooler is to quantize this vector and build a matrix from these quantization centers. This matrix is empty before training. The vectors in this matrix (the quantization centers) are called *coincidences*, and hence the matrix is called a coincidence matrix. There are three spatial pooler algorithms: Gaussian, dot and product. during learning, the dot and product algorithms work the same. The Gaussian spatial pooler algorithm is used for nodes at the first level, whereas the dot/product learning algorithm is applied at level >1. The input of the spatial pooler at level n+1 is a probability distribution over the temporal groups of the nodes at level n. A spatial pooler algorithm parameter specifies which algorithm to use, although it is common to use the same algorithm for every node up the hierarchy.

Operation of temporal pooler during learning

The objective of the temporal pooler is to create temporal coherent groups from a sequence of spatial patterns. This mechanism pools patterns using their temporal proximity. If pattern A is frequently followed by pattern B, the temporal pooler can assign them to the same group.

To this end, it builds a first order time adjacency matrix; after learning, this can be used to derive how likely a certain transition between each of the coincidences is.

When a new input vector is presented during training, the spatial pooler represents it as one of its learned coincidences *i*. The temporal pooler then looks back in history a certain number of steps, which is represented by the parameter *transition memory*.

After the learning stage and before inference, when the time-adjacency matrix is formed, the temporal pooler uses this matrix to create temporal groups.

Training the network

To be able to make classifications a supervised mapper is used that replaces the temporal pooler at the highest level of a HTM network. For every training input pattern, the supervised mapper receives two inputs during learning: the coincidence from the spatial pooler and the category of the input vector from the category sensor. It has a mapping matrix, which stores how many times a coincidence *i* belongs to a category *c* by incrementing element (*c*, *i*) every time it receives these inputs together.

Operation of nodes during inference

After training a node, it can be switched to inference mode. During inference, the level already has a model of the world (stored in the spatial and temporal pooler nodes). When the level receives an input from its children, it uses its internal model of the world to create an output to send to its parent(s).

Spatial pooler during inference

The three spatial pooler algorithms: Gaussian, Dot and Product work differently during inference stage, but they all convert an input vector into a belief vector over coincidences. As stated before, the Gaussian spatial pooler algorithm is used in first level nodes and the Dot or Product algorithms are used in the nodes higher in the hierarchy.

Operation of temporal pooler during inference

During inference, the temporal pooler receives a belief vector over coincidences from the spatial pooler. It will then calculate a belief distribution over groups. In this mode, two different algorithms exist for temporal pooler: *maxProp* and *sumProp*, governed by the parameter *temporal pooler algorithm*.

In maxProp inference mode the maximum value per temporal group is set as output.

When set to sumProp, computes a smoother score for the group based on the current input only.

Operation of top node during inference

During inference of the top node, the spatial pooler works as described earlier. The supervised mapper receives a belief vector over coincidences from the spatial pooler and a category from the category sensor. It calculates a belief distribution over these categories. In this stage, it's necessary to choose between two different temporal pooler algorithms during inference: *maxProp* and *sumProp*, controlled by the parameter *mapper algorithm*.

Hierarchical temporal memory (HTM) design and implementation

A platform for implementing HTMs called NUPIC and developed by Numenta® was used to implement our HTM network.

HTM networks are built and configured by writing Python scripts. While the majority of the scripts follow a standard pattern, each network requires customization. One must leverage in-depth knowledge of data to design and configure the hierarchy of nodes. Each node algorithm need to be customized based on the input values it is encountering. Because of the large number of node parameters, node configuration values will most likely be 'tweaked' after each iteration in order to improve accuracy. The network structure usually remains the same, reducing the amount of code that must be changed.

Our HTM consists of 7 levels, three levels each with two sub-levels (the level which analyzes the spatial component and other level which analyze the temporal component) and a final classifier. It is the final element of the hierarchy and classifying the image into common categories. Through the parameter output element count, the number of categories can be defined, five in this case.
The parameter configuration was as follows:

MaxDistance on the first level defines the minimum value that the squares of the Euclidean distances between an input (x) and all the previously memorized inputs (yi) have to take in order for x to be considered novel. maxGroupSize sets an upper limit for the number of quantized inputs that can form a group in the temporal pooler. The pooler algorithm used by the spatial pooler of higher levels is 'product', which means that the belief that an input during inference is similar to a given vector (previously memorized by the spatial pooler) is calculated as follows:

$$belief_i = \prod_{j=1}^{nchildren} y_i[child_j] * x\,[child_j] \quad (1)$$

Where nchildren is the number of children the node has, x is the input vector, yi are the vectors previously stored by the spatial pooler, and a[childn] is the part of vector a that is received from the nth child.

Finally, the temporal pooler at each level uses the sumProp algorithm, which takes the highest belief from each group to generate a distribution of beliefs over temporal groups during inference.

Other parameters related with the scale of the images are:

1. ScaleRF- An integer specifying the number of scales (resolutions) in the multi-resolution topology from which each node should receive input. For example, a value of 2 means that each node should receive input from 2 scales. Note that unless scaleRF is 1, the number of resolutions seen by the parent level will be lower than the number seen at the current level.
2. Scale overlap- An integer specifying how many scales neighbouring nodes should share in common. For example, if scaleRF is 2, scaleOverlap is 1, and there are 3 resolutions in the level underneath, some nodes will see the smaller and middle resolutions, and some nodes will see the middle and larger resolutions.

Training phase

Once the network is built, defining the architecture through which information flows, we set up the training process and the information processing. Thus, the key parameter is the number of iterations performed using the training images. In this case we have performed 2000 iterations in three levels. It has been shown experimentally that, if the iterations are increasing to the double value (4000), it is not observed a significant increase of accuracy in the analysis.

NuPIC has a user interface that allows interacting with the network while the analysis process is carried out. The Nupic platform has a module called GaborNode which analyzes the shape and texture of the input patterns.
We used images (128×128 pixels) composed of IRGB bands and NDVI index.

In Figure 5 the training of temporal pooler of level 1, sub-level 2 is showed. Next to the training image, a representation of information received by the spatial node of the first spatial pooler: GaborNode is also presented.

Inference phase

Once the network has been trained with the database provided, stating the categories, the inference stage is starting, where unknown images are analyzed by the network, according to the learned and memorized in the previous stage. Table 1 shows the number of training and testing images for the architecture 'demo'.

For each one of the classifications the overall accuracy, the kappa statistic and the producer´s and user's accuracy were calculated. The overall accuracy was calculated through the plot ratio correctly classified divided by the total number included in the evaluation process. The kappa statistic is an alternative measure of classification accuracy that subtracts the effect from random accuracy. Kappa quantifies how much better a particular classification is in comparison to a random classification. Some authors suggested the use of a subjective scale where kappa values < 40% are poor, 40 to 55% fair, 55 to 70% good, 70 to 85% very good and > 85% excellent (Monserud and Leemans, 1992).

RESULTS AND DISCUSSION

We investigated the effect of the parameters Maxdistance, ScaleRF, ScaleOVERLAP on overall accuracy, kappa coefficient, and the average number of coincidences and temporal groups learned in the bottom-level nodes. The other parameters (transition Memory and

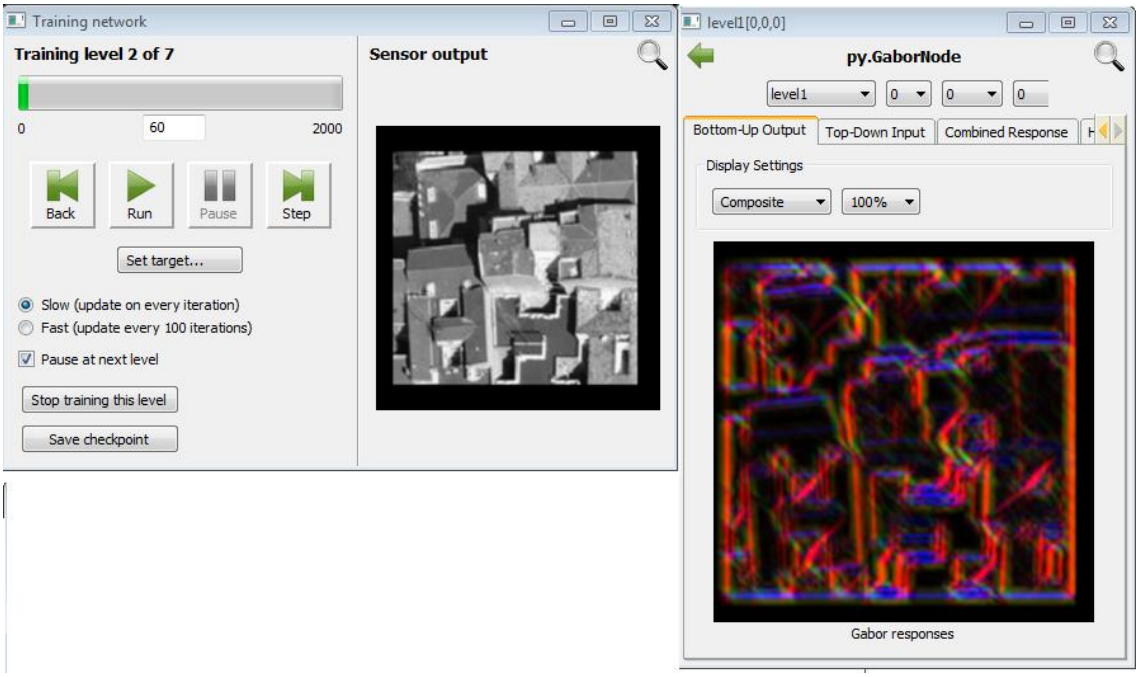

Figure 5. Training stage of level 1, sub-level 2.

Table 1. The number of training and test images.

Category	Training images	Testing images
Vitis Vinifera L.	300	150
Other land covers	300	150

Table 2. Overall accuracy, average number of coincidences and temporal groups learned in the 16 bottom nodes for different values of maxDistance and ScaleRF and ScaleOverlap.

MaxDistance	scaleRF	scaleOverlap	Overall accuracy (%)	#Coincs	#Groups
1	1.00	1.00	87.00	55.00	25.00
3	1.43	1.00	96.00	44.79	20.00
6	2.65	2.00	83.13	17.94	11.88
9	4.00	3.00	76.35	12.20	7.45

topNeighbors) were set to 5 and 1, respectively. These are default values, and different values had a negative effect on the performance of the system. We varied across different values for Maxdistance and set Sigma to the square root of Maxdistance. This is a reasonable starting value for Sigma, because distances between coincidences are calculated as the squared Euclidean distance instead of the standard Euclidean distance. The results are shown in Table 2.

The higher overall accuracy was obtained with an intermediate value for Maxdistance: 3 and values of 1.43 and 1.00 for ScaleRF and ScaleOVERLAP respectively. This might indicate that with a lower value for Maxdistance, the HTM would see variations in input patterns due to noise as different coincidences. On the other hand, when Maxdistance is higher than the optimal value, the spatial pooler will pool together patterns that have different causes. In the confusion matrix can be seen successes experienced by the system in each of the categories for the higher overall accuracy (Figure 6).

The improvement induced by the introduction of textural and contextual features was significant for all classes with respect to the pixel-based analysis. The highest producer's accuracies were for 'Irrigated land', 'bare soil', 'urban soil' and 'olive groves' categories, all with the value of 100%. In contrast, the lowest value was

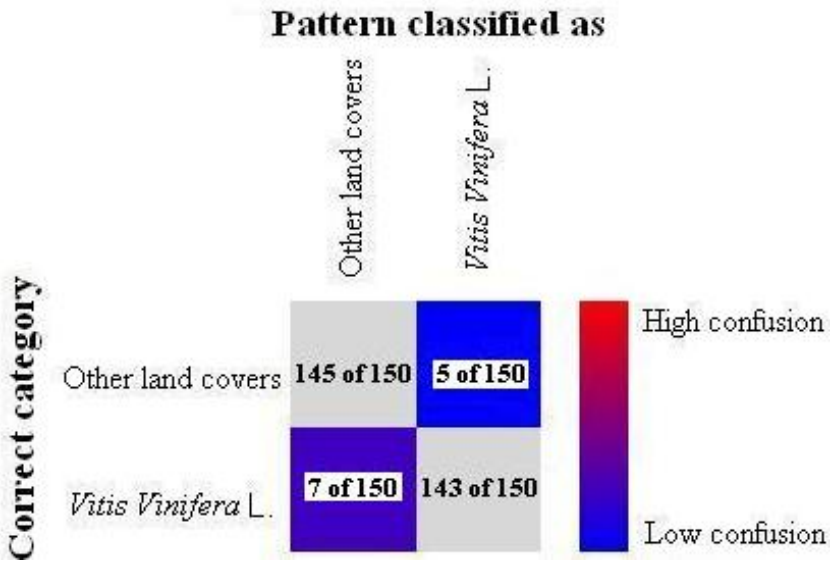

Figure 6. Confusion matrix of the best performing system. Pa: Producer's accuracy. Ua: user's accuracy; NDVI: normalized difference vegetation index; IRGB: near Infrared, green and blue bands.

Table 3. Producer's and user's accuracy, overall and Kappa statistic for supervised classifications and HTM classification.

Category	Supervised classification image IRGB and NDVI		HTM classification	
	Pa (%)	Ua (%)	Pa (%)	Ua (%)
Bare soil	84.5	92.3	100.0	92.9
Vineyards	91.6	89.2	87.4	95.3
Irrigated land	99.1	63.6	100.0	82.34
Urban soil	91.0	86.5	100.0	100.0
Olea europaea L.	93.4	96.8	100.0	100.0
Overall accuracy (%)	87.9		96.0	
Kappa statistic (%)	75.9		93.8	

for 'Vineyards' (87.4%) owing to the spectral similarity to '*Olea europaea* L.' but this value is higher than that one obtained in the supervised classification. Referring to the user's accuracy, the best results were achieved again for the categories 'urban soil' (100%) and '*Olea europaea* L.' (100%) and the lowest value was for the category 'irrigated land' (82.34%) also higher than the value obtained in the supervised classification (Table 3).

As for the overall accuracy and kappa statistic, they have been very successful, reaching the value of 96 and 93.8% respectively. In addition, the HTM classification significantly narrowed down the variation of class-based accuracies compared with the result of the pixel-based classification method. The problems associated to the use of high spatial resolution images have been resolved to a large extent, as in the case of the salt and pepper effect. This effect makes difficult to obtain a clean classified image, and different land uses in a plot have been observed where would be just one. In Figure 7 a map obtained from the HTM classification is presented.

The accuracy values obtained with the algorithm based on the hierarchical temporal memory were similar to and/or higher than the values obtained by other authors, which shows that the methodology is adequate for vineyard mapping.

Granger et al. (2005) carried out a classification based on spatial patterns in panchromatic Ikonos images on the following categories: vineyards and orchards, obtaining an overall accuracy of 95.4%, which is lower than that obtained in this work. Aitkenhead and Wright (2004) classified urban areas, crops and bare soil using neuronal networks in Landsat Thematic Mapper (TM) images and obtained 60% of accuracy for urban areas, 100% for water and forests, 90% for bare soil and 95% for agricultural crops. Vaudour et al. (2010) proposed to map viticultural soils using bootstrapped regression trees

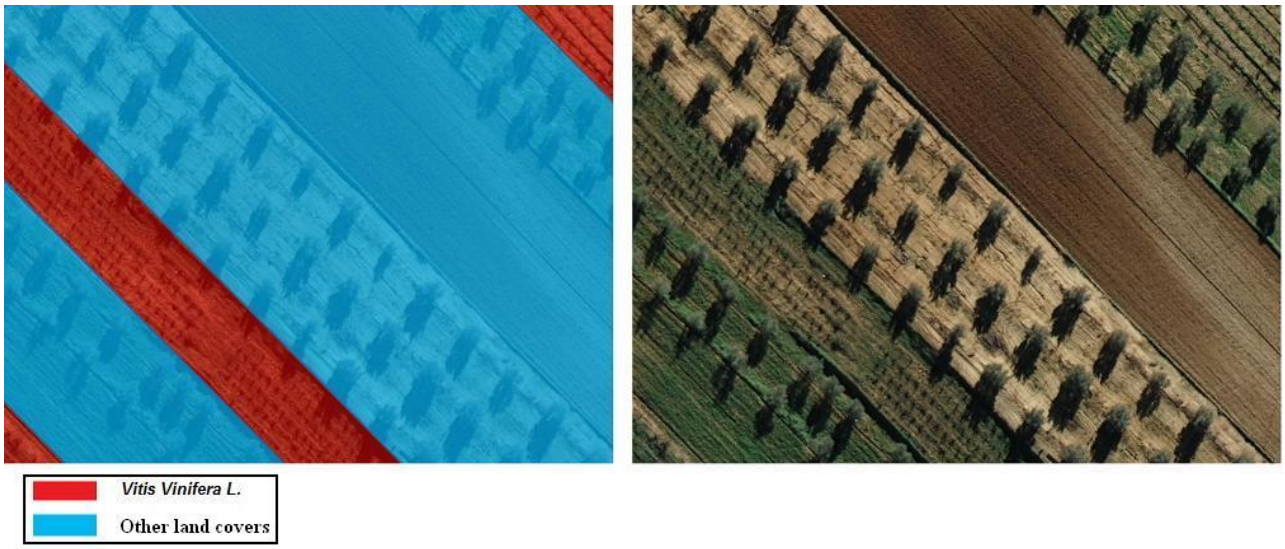

Figure 7. Classified image obtained from a system based on hierarchical temporal memory.

on distinct combinations of morphometric data and SPOT satellite images over the Stellenbosch viticultural area (South Africa), obtaining a median accuracy of 52 to 78%. This percentage is lower than that obtained with the HTM network developed in this work.

Conclusion

In this paper a complete process has been proposed for detecting vineyards, their demarcations and their characteristics in the plot. The main advantages of this model are: an easier implementation, a faster processing and a limited quantity of parameters. The model used is the hierarchical temporal memory (HTM), which is a Bayesian network that assumes a node hierarchy where each node learns spatial and temporal coincidences of patterns which give information about the world. This model has a similar hierarchy to the cortical region and the nodes of this model correspond with little regions of the cerebral cortex. The HTM network has been developed using the platform Nupic of Numenta®. Good results were achieved, obtaining an overall accuracy of 96% and problems associated to the use of high spatial resolution images have been resolved, vineyard mapping. These results show that HTM approach provide new promises for vineyard registration.

REFERENCES

Aitkenhead MJ, Wright GG (2004). Mapping land use in NE Scotland with neural networks from remote sensing imagery. In Proceedings of Remote Sensing and Photogrammetry Society Annual Conference, Aberdeen, UK.

Bobillet W, Da Costa JP, Germain C, Lavialle O, Grenier G (2003). Row detection in high resolution remote sensing images of vine fields. In: J. Stafford and A. Werner (eds), European Conference on Precision Agriculture, Wageningen academic publishers, Berlin, pp. 81-87.

Budgen JL, Salinas G, Howarth PJ (1999). Vineyard Identification in the Tulum Valley, Argentina, using RADARSAT imagery. Paper presented at: 4th Int. Airborne Remote Sensing Conference and Exhibition and 21st Canadian Symposium on Remote Sensing. Ottawa: Canada, 2: 375-382.

Company A, Delpont G, Guillobez S, Arnaud M (1994). Potentiel des données radar ERS-1 pour la détection des surfaces contributives au ruissellement dans les vignobles méditerranéens du Roussillon (France). In Proc. 6ème Symposium International Mesures Physiques et Signatures en Télédétection ; 1994 Jan 17-21; Val d'Isère, France. Orléans: BRGM, pp. 375-382.

Da Costa JP, Michelet F, Germain C, Lavialle O, Grenier G (2007). Delineation of vine parcels by segmentation of high resolution remote sensed images. Precision Agric., 8(1): 95-110.

Delenne C, Durrieu S, Rabatel G, Deshayes M (2009). From pixel to vine parcel: A complete methodology for vineyard delineation and characterization using remote sensing data. Comput. Electron. Agric., 70(1): 78-83.

Delenne C, Durrieu S, Rabatel G, Deshayes M, Bailly JS, Lelong C, Couteron P (2008) Textural approaches for vineyard detection and characterization using very high spatial resolution remote-sensing data. Int. J. Remote Sens., 29(4): 1153-1167.

George D, Jaros B (2007). The HTM learning algorithms. Available from http://www.numenta.com/htm-overview/education/Numenta_HTM_Learning_Algos.pdf.

Gong P, Mahler SA, Biging GS, Newburn DA (2003). Vineyard identification in an oak woodland landscape with airborne digital camera imagery. Int. J. Remote Sens., 24:1303-1315.

Haboudane D, Miller JR, Pattey E, Zarco-Tejada PJ, Strachan I (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. Remote Sens. Environ., 90: 337-352.

Hall A, Louis J, Lamb D (2003). Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images. Comput. Geosci., 29: 813-822.

Hawkins J, Blakeslee S (2004). On Intelligence. New York: Henry Holt.

Hawkins J, George D (2007). Hierarchical Temporal Memory, Concepts, Theory, and Terminology. Available from: http://www.numenta.com/Numenta_HTM_Concepts.pdf

Marini MF (2006). Influencia del cálculo de la reflectancia aparente en la obtención del índice verde (NDVI) y en su expresión cartográfica. Paper presented at: XXIII Reunión Científica de la Asociación Argentina de Geofísicos y Geodestas. 2006 August 14-18; Bahía Blanca, Argentina.

Monserud RA, Leemans R (1992). Comparing global vegetation maps

with the Kappa statistic. Ecol. Model., 62: 275-293.
Numenta Inc (2008). Advanced Nupic Pogramming. Available from: http://www.numenta.com/archives/education/nupic_prog_guide.pdf.
Perea AJ, Meroño JE, Aguilera MJ (2009). Application of Numenta® Hierarchical Temporal Memory for land-use classification. S. Afr. J. Sci., 105(9-10): 370-375.
Rabatel G, Delenne C, Deshayes M (2008). A non-supervised approach using Gabor filters for vine plot detection in aerial images. Comput. Electron. Agric., 62(2): 159-168.
Ranchin T, Naert B, Albuisson M, Boyer G, Astrand P (2001). An automatic method for vine detection in airborne imagery using wavelet transform and multiresolution analysis. Photogramm. Eng. Rem. S., 67: 91-98.
Rodríguez-Pérez JR, Álvarez-López CJ, Miranda D, Álvarez MF (2008) Vineyard area estimation using medium spatial resolution satellite imagery. Span. J. Agric. Res., 6(63): 441-452.
Soria-Ruiz J, Fernandez-Ordoñez Y, Woodhouse IH (2010). Land-cover classification using radar and optical images: a case study in Central Mexico. Int. J. Remote Sens., 31(12): 3291-3305.
Vaudour E, Carey VA, Gilliot JM (2010). Digital zoning of South African viticultural terroirs using bootstrapped decision trees on morphometric data and multitemporal SPOT images. Remote Sens. Environ., 114(12): 2940-2950.
Warner T, Steinmaus K (2005). Spatial Classification of Orchards and Vineyards with High Spatial Resolution Panchromatic Imagery. Photogramm. Eng. Rem. S., 71(2): 179-187.
Wassenaar T, Baret F, Robbez-Masson JM, Andrieux P (2001). Sunlit soil surface extraction from remotely sensed imagery of perennial, discontinuous crop areas; the case of Mediterranean vineyards. Agron. Sustain. Dev., 21: 235-245.
Wassenaar T, Robbez Masson JM, Andireux p, Baret F (2002). Vineyard identification and description of spatial crop structure by per-field frequency analysis. Int. J. Remote Sens., 23(17): 3311-3325.