UNIVERSIDAD DE CÓRDOBA

ESCUELA TÉCNICA SUPERIOR INGENIEROS AGRÓNOMOS Y MONTES

DEPARTAMENTO DE INGENIERÍA GRÁFICA Y GEOMÁTICA

TESIS DOCTORAL

Algoritmos avanzados de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo

Presentada por:

D. Alberto Jesús Perea Moreno

Ingeniero Agrónomo

Dirigida por:

D. José Emilio Meroño de Larriva

Dr. Ingeniero Agrónomo

Córdoba, Marzo 2012

TÍTULO: *Algoritmos avanzados de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo*

AUTOR: *Alberto Jesús Perea Moreno*

© Edita: Servicio de Publicaciones de la Universidad de Córdoba. 2012
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

TÍTULO DE LA TESIS: Algoritmos avanzados de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo.

DOCTORANDO/A: Alberto Jesús Perea Moreno

INFORME RAZONADO DEL DIRECTOR DE LA TESIS

José Emilio Meroño de Larriva, como director de la tesis titulada «**Algoritmos avanzados de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo**» realizada por **Alberto Jesús Perea Moreno**,

INFORMA:

Que dicha Tesis Doctoral ha sido realizada bajo mi dirección.

Que ha tenido como principal objetivo la evaluación de las nuevas técnicas de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo. En el desarrollo de esta Tesis Doctoral se abarca el estudio y aplicación de nuevos algoritmos de clasificación digital en imágenes de alta resolución espacial, obtenidas tanto de satélites de alta resolución como de sensores digitales aerotransportados, con el objetivo de obtener clasificaciones de patrones geométricos en elementos del territorio. Se proponen tres metodologías avanzadas de clasificación digital: i) una empleando imágenes de satélite de alta resolución espacial combinadas con algoritmos de clasificación experta que se auxilian de resultados obtenidos por clasificadores orientados a objetos, índices de vegetación y transformaciones de bandas. ii) el empleo de información espectral de fotogramas captados por sensores digitales aereotransportados combinado con algoritmos de clasificación experta, índices de vegetación y transformaciones de bandas. iii) empleo de imágenes de sensores digitales aerotransportados y algoritmos basados en el funcionamiento del neocortex humano.

Que tanto la metodología como el trabajo de investigación, las conclusiones y los resultados obtenidos son satisfactorios.

Que derivado de esta Tesis Doctoral se han publicado los siguientes trabajos:

REVISTAS INCLUIDAS EN EL SCI

Perea, A.J., Meroño, J.E., Aguilera, M.J. 2009. Algorithms of expert classification applied in Quickbird satellite images for land use mapping. Chilean Journal of Agricultural Research; 69, (3), 400-405.

Perea, A.J., Meroño, J.E., Aguilera, M.J. 2009. Application of Numenta® Hierarchical Temporal Memory for land-use classification. South African Journal of Science; 105, (9-10), 370-375.

Perea, A.J., Meroño, J.E., Aguilera, M.J., de la Cruz, J.L. 2010. Land-cover classification with an expert classification algorithm using digital aerial photographs. South African Journal of Science; 106, (5-6) 82-87.

Perea, A.J., Meroño, J.E., Aguilera, M.J. 2012. Hierarchical Temporal Memory for mapping vineyards using digital aerial photographs. African Journal of Agricultural Research; 7(3) 456-466.

REVISTAS NO INCLUIDAS EN EL SCI (con revisión por pares)

Perea, A.J., Meroño, J.E., Aguilera, M.J. 2011. Application of the Memory-Prediction Theory in olive grove inventory. Agro-ciencia; 27(1) 29-39.

Que considero que el trabajo realizado cumple los requisitos necesarios para su presentación y lectura.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 23 de Marzo de 2012

Firma del director

Fdo.: José Emilio Meroño de Larriva

UNIVERSIDAD DE CÓRDOBA

Departamento de Ingeniería Gráfica y Geomática

DEPARTAMENTO DE INGENIERÍA GRÁFICA Y GEOMÄTICA

UNIVERSIDAD DE CÓRDOBA

Título de la tesis: Algoritmos avanzados de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo.

Autor: Alberto Jesús Perea Moreno

Presentada en satisfacción de los requisitos necesarios para la obtención del grado de

DOCTOR INGENIERO AGRÓNOMO

Dirigida por el Doctor: El Doctorando:

José Emilio Meroño de Larriva Alberto Jesús Perea Moreno

Córdoba, Marzo de 2012

A mis padres

Agradecimientos

Agradezco a mis compañeros y amigos del Grupo de Investigación «Nuevas tecnologías aplicadas a la agricultura y el medioambiente», en especial a María Jesús Aguilera y José Luis de la Cruz, por su colaboración en el desarrollo de esta tesis, ya que con sus numerosos consejos, sugerencias y muestras de ánimo han contribuido a crear este documento.

Mención especial en estos agradecimientos para José Emilio Meroño que, en su calidad de tutor, ha dirigido en estos años mi carrera investigadora, ha aportado numerosas ideas en este trabajo y hace un esfuerzo continuado para su difusión y aplicación práctica. A mis compañeros de la empresa TRAGSA y del Departamento de Física Aplicada, en especial Federico Manzano, Juan Chaves y Ana Laguna, por su apoyo durante estos años.

A mi familia y amigos, en especial a mis padres y hermanos, que bien de forma directa o indirecta, con su cariño y compañía, me han ayudado tanto a realizar este trabajo como a evadirme de él cuando era necesario.

Y por último, quiero agradecer y dedicar esta Tesis a Carmen, ya que nada de esto habría sido posible sin tu apoyo y ayuda incondicionales.

Indice

Indice de figuras

Indice de tablas

Abreviaturas

B: blue

CP: Componentes Principales

DEM: Digital Elevation Model

DN: Número Digital

G: green

GIS: Geographic Information System

HM: Hierarchical Temporal Memory

IRGB: near Infrared, green and blue bands

JRC: Join Research Center

MTJ: Memoria Temporal Jerárquica

NDVI: Normalizad Difference Vegetation Index

NDWI: Normalizad Difference Water Index

NI: Near Infrared

NuPIC: Numenta Platform for Intelligent Computing

Pa: producer’s accuracy

PCA: Principle component analysis

R: red

SIG: Sistema de información Geográfica

Ua: user’s accuracy

Resumen

En las últimas décadas, se ha producido una tendencia a la clasificación automatizada de usos del suelo en imágenes de alta resolución espacial para la verificación y control de las ayudas económicas en la Unión Europea.

Sin embargo, la precisión global de los mapas producidos de esta forma es normalmente inferior a las necesidades del usuario. Por lo tanto, la mayoría de trabajos de clasificación de usos del suelo todavía dependen en cierta medida de la fotointerpretación, que es menos rentable y más subjetiva que el método anterior. A medida que la atención se centra cada vez más en el seguimiento, en lugar de un simple mapa de cultivos, hay una necesidad de mejorar la cuantificación y modelización. Pero esta necesidad no puede ser cumplida con los tradicionales métodos automatizados. Uno de los objetivos de esta tesis es contribuir en la resolución de este dilema mediante la integración de los nuevos algoritmos de clasificación digital. El enfoque tradicional para el análisis de imágenes aéreas digitales para la cartografía de cubiertas vegetales, el cual se basa principalmente en la clasificación de las firmas espectrales (muestras multivariantes procedentes de píxeles), conduce a resultados poco satisfactorios, principalmente porque apenas aprovecha la estructura espacial de las imágenes.

El objetivo general de esta Tesis Doctoral es la evaluación de las nuevas técnicas de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo. Para ello se emplean tres algoritmos avanzados de clasificación digital sobre imágenes aéreas digitales.

Dos motivos principales impulsan el desarrollo del análisis de imágenes con los nuevos algoritmos antes citados: 1) Imágenes de alta resolución espacial ya están disponibles y las herramientas de

computación están constantemente mejorando; 2) El análisis de la imagen basado en píxeles es limitado.

Se presentan tres metodologías para la clasificación digital de imágenes de alta resolución espacial que proporcionen información de detalle sobre los usos del suelo presentes en la zona de estudio. La primera está basada en el uso de imágenes multiespectrales QuickBird con una resolución espacial de 30 cm, el empleo de la clasificación orientada a objetos y un algoritmo de clasificación experta creado con información adicional procedente del análisis orientado a objetos y resultados obtenidos en clasificaciones supervisadas empleando la imagen formada por los componentes principales y el índice de vegetación NDVI. Los resultados son satisfactorios, mejorando en gran medida los obtenidos por la clasificación tradicional basada en píxel.

La segunda metodología combina el uso de imágenes captadas por sensores digitales aéreotransportados, el empleo del análisis orientado a objetos y un algoritmo de clasificación experta creado con las mismas premisas que en la metodología anterior. Los resultados obtenidos demuestran que las imágenes procedentes de sensores aereotransportados contienen una información muy útil en el campo de la clasificación digital y que mejoran sustancialmente los resultados con respecto a los obtenidos empleando imágenes de satélite.

La tercera metodología presenta un nuevo algoritmo basado en la corteza cerebral, Memoria Temporal Jerárquica, con fotografías aéreas digitales. Este algoritmo se basa en el funcionamiento de la mente que almacena patrones y hace predicciones sobre los patrones que encuentra o espera encontrar. En este modelo, las relaciones temporal y espacial entre patrones de señales sensoriales forman una arquitectura de memoria jerárquica durante el proceso de aprendizaje. El aprendizaje puede ser supervisado y no supervisado. Cuando aparece un nuevo patrón el proceso de reconocimiento se activa y elige entre los patrones almacenados el que mejor lo representa.

Abstract

During the last decades, there is a trend toward automated classification of land use in high spatial resolution images to check and monitor financial aids of the European Union.

However, the overall accuracy of the maps produced in this way is usually less than the user's requirements. Therefore, most studies of land use classification still depend to a certain extent on photo interpretation, that is less profitable and more subjective that the previous method.

As attention is increasingly focused on monitoring rather than on a simple map of crops, there is a need to improve the quantification and modeling. But this requirement should not be fulfilled using traditional automated methods. The objective of this Doctoral Dissertation is to contribute in solving this dilemma by integrating new digital classification algorithms.

Traditional approach to digital aerial imagery analysis to mapping vegetation cover, which is primarily based on the classification of the spectral signatures (multivariate samples from pixels), leads to slightly unsatisfactory results, mainly because it barely makes use of the spatial structure of images.

The overarching aim of this Doctoral Dissertation is to evaluate the new digital classification techniques of high spatial resolution imagery to land use mapping. For this purpose, three advanced algorithms of digital classification are used on digital aerial imagery.

Two main reasons drive the development of the imagery analysis with the new aforementioned algorithms: 1) High spatial resolution images are already available and computing tools are constantly improving; 2) The pixel-based analysis is limited.

Three methodologies are proposed to digital classification of high spatial resolution images which provide detail information on land uses of the study area. The first methodologie is based on Quickbird multispectral images with a spatial resolution of 30 cm, the employment of object-based classification and an algorithm of expert classification developed with additional information from object-based analysis and the results obtained in supervised classification using the image formed by the principle components and the vegetation index NDVI. The results are satisfactory, greatly improving those obtained by traditional pixel-based classification.

The second methodology combine the use of images received from airbone digital sensors, the object-based analysis and an algorithm of expert classification developed with the same previous premises. The results obtained show up images from airbone sensors contain useful information in the field of digital classification and greatly improve the results with regard to those obtained using satellite images.

The third methodology launches a new algorithm based on human brain cortex, Hierarchical Temporal Memory, with digital aerieal photographs. This algorithm is based on the brain functioning which stores patterns and makes predictions about the patterns that it finds o waits to find. In this model, temporal and spatial relations between sensory signal patterns form hierarchical memory architecture during the learning process. The learning process can be supervised or unsupervised. When a new pattern appears the recognition process stars working and chooses between the stored patterns that best represents it.

Capítulo I. Introducción General

«Toda la información que entra en nuestra mente lo hace como patrones espaciales y temporales en los axones»

Jeff Hawkins (2004)

Actualmente los países miembros de la Unión Europea emplean la teledetección para la verificación y control de ayudas económicas. Esta actividad denominada MARS-PAC (Monitoring Agriculture with Remote Sensing – Common Agricultural Policy) consiste en la interpretación digital y clasificación automática para comprobar las ayudas percibidas por los agricultores y ha evolucionado de forma considerable en los últimos años.

En las estimaciones de rendimiento, los componentes climáticos y agronómicos pueden integrarse en un modelo agrometeorológico, que servirá para evaluar la respuesta de los cultivos a los cambios en las condiciones meteorológicas y/o en las técnicas de cultivo: El proyecto MARS desarrolló los modelos agrometeorológicos CGMS y OLIWIN. Dichos estimadores de rendimiento necesitan como dato de partida las áreas ocupadas por los cultivos agrícolas.

La clasificación digital de imágenes es el proceso por el que los píxeles que tienen características espectrales similares y que, por lo tanto, se supone que pertenecen a una misma clase, se identifican y se asignan a un único color (Gibson y Power, 2000). Entre los procedimientos de clasificación dirigidos a datos de alta resolución espacial, el clasificador máxima probabilidad es el más extendido y se utiliza debido a su efectividad y a la robustez estadística (Strahler, 1980). En 1991, Giovacchini y Brunetti, destacan la utilización de los datos de teledetección para la elaboración de estadísticas agrícolas en Italia a través de la clasificación de imágenes por patrones de máxima probabilidad.

La extracción de información útil de las imágenes de satélite (clasificación) es el problema técnico principal de la teledetección. Los datos obtenidos son de difícil utilización debido a (Wilkinson et al. 1991): (1) La información espectral contenida en los píxeles no es suficiente en la mayoría de los casos, como para identificar especies de vegetación o tipos de cubiertas de la superficie; (2) Normalmente los píxeles incluyen una mezcla radiométrica de sus vecinos y por lo tanto pocas zonas tienen homogeneidad total.

A pesar de ello, el enfoque comúnmente usado para analizar las imágenes de satélite con fines cartográficos da lugar a resultados no óptimos debido principalmente al uso exclusivo de patrones espectrales de los píxeles, de forma que no se considera la estructura espacial de la imagen. Este enfoque se basa en la discriminación de firmas espectrales. Estas firmas están normalmente constituidas por los valores que adopta cada píxel en cada una de las bandas que constituyen la imagen multiespectral, la cual se obtiene siguiendo el siguiente proceso: el sensor adquiere datos que son agrupados espacialmente en una matriz o ráster en la que cada celdilla corresponde a una medición de una señal eléctrica que depende de la cantidad de radiación recibida en esa posición y momento. La medición es discretizada por un convertidor analógico-digital a una escala finita o rango dinámico (de 0 a 255 para la mayoría de los sensores ópticos), y el valor resultante es introducido en esa celdilla en forma de Número Digital (DN). La radiación incidente es separada antes de alcanzar los detectores del sensor por medio de un sistema de prismas y filtros, de forma que cada banda de una imagen multiespectral corresponde a la radiación capturada en un intervalo particular del espectro electromagnético. Los valores de cada celdilla, representados a lo largo del espectro, se pueden interpolar dando lugar a una curva o firma espectral, que aunque más grosera tiene una cierta similitud con la que se obtiene de los espectrómetros de sobremesa, de ahí que cada píxel se considere como una muestra individual. Estas firmas se pueden también representar como puntos de un espacio multivariante en el que cada eje ortogonal se refiere a una banda y está constituido por el conjunto ordenado de valores del rango dinámico. La clasificación espectrométrica de las imágenes consiste, por tanto, en delimitar las regiones de ese espacio asociadas a cada clase y que deben cumplir una serie de requisitos:

- Exhaustividad: debe haber una clase que asignar a cada píxel de la imagen.

- Separabilidad: las clases deben ser separables en el espacio multivariante con el clasificador empleado.

- Utilidad: las clases deben cubrir las necesidades de información del usuario.

El requisito de separabilidad implica que las firmas de clases diferentes estén relativamente distantes las unas de las otras, de forma que su grado de solape sea despreciable. Sin embargo para que esto se cumpla, cada cluster (nube de puntos) del espacio multivariante debe contener una clase mayoritaria. Por otro lado, la cuadrícula de terreno sobre la cual el sensor realiza la medida de cada píxel debe ser suficientemente grande como para incluir los elementos que producen la firma espectral típica de cada clase. El tamaño mínimo necesario para realizar una clasificación correcta sobre el terreno recibe el nombre de resolución espacial de la clasificación. Por tanto, una premisa básica del enfoque espectrométrico es que la extensión de la cuadrícula sobre la que se extrae la muestra (es decir, el tamaño del píxel) supere esa resolución. Ahora bien, cuanto mayor sea el tamaño de la cuadrícula, mayor será el porcentaje de píxeles «mezclados», es decir, píxeles incluyen bordes entre teselas. Como la firma espectral de éstos es una mezcla de las firmas típicas de las clases de esas teselas, su posición en el espacio multivariante corresponderá a las zonas que separan clusters de clases diferentes. Sin embargo, el propio concepto de cluster requiere que éste esté separado de otros por una discontinuidad, eso es, por una región del espacio multivariante casi vacía. Por tanto la premisa del tamaño suficiente y del solape despreciable no pueden ser satisfechas simultáneamente cuando las unidades analizadas son píxeles individuales, a no ser que se estudie un territorio relativamente simple (como un paisaje agrícola con grandes campos de cultivo) con unas clases de cubierta muy generales. A pesar de esto, el enfoque basado en píxeles sigue siendo el paradigma comúnmente aceptado en esta disciplina (Castilla, 2003).

La teledetección espacial ha ido evolucionando desde los años setenta, cuando comenzó. El tamaño de píxel de las imágenes captadas por los satélites con que ésta comenzó (80 m) era compatible con la resolución espacial de la mayor parte de las clasificaciones. A esa escala, era más natural considerar las clases de cubierta como materiales homogéneos distribuidos sobre el territorio en parcelas mayores que un píxel, por tanto era razonable asimilar cada píxel individual a una muestra introducida en un espectrómetro que puede ser analizada por separado. Con el paso de los años la evolución técnica permitió mayores resoluciones y la variabilidad radiométrica de las clases aumentó. Para mejorar el resultado de las clasificaciones se hizo necesario, por un lado, incorporar al análisis alguna característica espacial como la textura, que pudiera paliar esta heterogeneidad, y por otro, realizar un pre y/o un post-procesamiento basado en filtros, que redujese la inconsistencia espacial de las imágenes clasificadas.

La aparición a principios de este siglo de satélites civiles de muy alta resolución (< 5m.) ha puesto de manifiesto las deficiencias del enfoque espectrométrico cuando no se cumple la premisa del tamaño.

Por todo ello, el JRC (Join Research Center) responsable de la investigación en Europa de todo lo relacionado con esta materia, actualmente está interesado en aplicar las siguientes tecnologías: redes neuronales artificiales, sistemas expertos o inteligencia artificial, integración de la información de los SIG en el análisis de la imagen y desarrollo de segmentación de la imagen (Wilkinson et al. 1991).

1. Clasificación basada en objetos

El análisis de imágenes basado en objetos es una nueva sub-disciplina que está recibiendo una gran atención por parte de numerosos investigadores como Galli y Malinverni (2005), Benz et al. (2004), Blaschke et al. (2004) y firmas comerciales como eCognition, ITT VIS o Erdas que han desarrollado aplicaciones específicas para este tipo de análisis.

Las técnicas de clasificación tradicionales, basadas en rasgos de la imagen a nivel de píxel, presentan ciertos problemas, como son la aparición del efecto «sal y pimienta» o la dificultad para extraer objetos de interés. Al aplicarse en imágenes de alta resolución espacial este problema se agrava. Una alternativa a dichos sistemas de clasificación pasa por un proceso previo de segmentación de la imagen.

Hay y Castilla (2006) definen el análisis de imágenes basado en objetos como una sub-disciplina de la ciencia de los Sistemas de Información Geográfica dedicada a dividir las imágenes en objetos con significado propio y al mismo tiempo, obtener sus características desde un punto de vista espacial, espectral y temporal.

Según Lizarazo y Elsner (2008), el análisis de imágenes basado en objetos está basado en los datos captados por sensores y produce resultados aptos para los SIG. Por tanto, puede considerarse como el puente entre el dominio ráster de las imágenes y el dominio predominantemente vectorial de los SIG.

Los clasificadores por objetos están diseñados para abordar la clasificación de paisajes heterogéneos y han mostrado su efectividad incrementando la precisión de las clasificaciones (Aplin et al., 1999).

El objetivo principal de esta sub-disciplina (Hay y Castilla, 2006) es desarrollar métodos, teorías y herramientas para emular la interpretación humana de imágenes aéreas o espaciales, de una forma automática o semiautomática que permita aumentar la producción de cartografía temática, reduciendo los costes y la subjetividad propia de la interpretación humana.

La clasificación basada en objetos tiene su punto inicial y fundamental en la creación de objetos; definidos como agrupaciones de píxeles contiguos con niveles digitales similares y que tienen tamaño, forma y relación geométrica con el componente del mundo real que modela. Lizarazo y Elsner (2008) introducen el término grixel para referirse a los objetos formados como agrupaciones de píxeles.

El primer paso es la segmentación de la imagen que debe hacerse teniendo en cuenta la resolución de la imagen y el tamaño de los objetos a identificar. El objetivo de la segmentación es simplificar la representación de una imagen a una forma con más significado y más sencilla de analizar (Shapiro y Stockman, 2001). El resultado de la segmentación de una imagen es un conjunto de regiones que cubren completamente la imagen. Todos los píxeles de una región son similares con respecto a alguna característica, al mismo tiempo que son diferentes de los píxeles situados en regiones adyacentes. Dentro del tratamiento digital de imágenes existen numerosos algoritmos de segmentación basados en crecimiento de regiones, análisis del histograma, detección de bordes, agrupamientos basados en distancias, etc. Al no existir una solución general al problema de la segmentación de imágenes, estas técnicas deben combinarse con el conocimiento del problema a resolver. Una imagen puede segmentarse en objetos de mayor o menor tamaño, influyendo considerablemente el tamaño de la segmentación que se utilice, en las características derivadas de los objetos de la imagen. En función del objetivo de la clasificación, el tamaño de los objetos a delimitar y la resolución de la imagen se determinarán la o las escalas en las que se segmentará la imagen. El hecho de segmentar una imagen varias veces con distintas escalas, da lugar a que surja una estructura jerárquica entre los objetos de los distintos niveles, ya que un objeto puede incluir objetos de niveles inferiores y estos a su vez abarcar la superficie ocupada por polígonos de otro nivel inferior. Esta estructura jerárquica entre los objetos es especialmente útil para caracterizar el paisaje que también tiene una estructura jerárquica. Según la escala de observación que utilicemos para analizar el paisaje podremos distinguir las regiones que lo forman con diferentes tamaños de detalle.

La característica más valiosa de la clasificación de imágenes orientada a objetos es la posibilidad de obtener un gran número de características descriptivas de los objetos y de las relaciones existentes entre los mismos que permitirán describirlos mejor y por lo tanto, diferenciarlos y obtener resultados más precisos y específicos. El conjunto de características utilizadas para describir los objetos se pueden clasificar en las categorías siguientes:

- Características espectrales: Mediante parámetros estadísticos como la media aritmética, la desviación típica, valores máximos o mínimos, etc., se describe la distribución de los niveles digitales de los píxeles, que forman el objeto, en las bandas espectrales o índices utilizados. La descripción espectral del propio objeto se puede completar con su comparación con los objetos vecinos, con los objetos de menor tamaño incluidos en el objeto de estudio o con objetos de mayor tamaño que lo engloben.

- Características de forma: La descripción de la morfología y dimensiones de un conjunto de píxeles se realiza tanto con propiedades como el área, el perímetro, el largo o el ancho como con índices de forma que describen la elongación, compacidad, asimetría, etc. Estas características por sí solas, facilitan la discriminación entre objetos con formas aproximadamente lineales, como vías de comunicación o cursos de agua, de objetos más compactos.

- Características relativas a la posición: La posición de un objeto en el espacio puede ser relevante en el caso de paisajes cuya distribución siga un determinado patrón geométrico, como es el caso de los árboles ordenados según un marco de plantación, o bien, en un paisaje fragmentado en regiones con usos y coberturas distintos entre estas regiones.

La distancia a un determinado objeto puede ser un claro indicador de la clase a la que corresponde el objeto. Por ejemplo, la distancia a la línea de costa puede ser una variable interesante para clasificar objetos correspondientes a playas, marjales, albuferas, etc.

Propiedades como las coordenadas del centro del objeto, sus coordenadas extremas o la distancia existente a un objeto determinado aportan información significativa para la descripción del objeto.

- Características de textura: Existen numerosos métodos para describir las propiedades texturales en una imagen. Si bien, las variables más utilizadas para describir la textura de un objeto son las extraídas de la matriz de coocurrencias de niveles de gris (Haralick et al., 1973) que cuantifican numéricamente propiedades como la homogeneidad, el contraste, la rugosidad, etc.

- Características relativas a objetos vecinos: En el caso de trabajar con clasificadores iterativos, en los cuales en la asignación de la clase a un objeto se consideran las clases asignadas a sus vecinos, es necesario definir propiedades que reflejen las relaciones de vecindad entre los objetos. Ejemplos de este tipo de propiedades serían el porcentaje de frontera común con una clase determinada, distancia a un objeto de una clase, existencia de una clase en el vecindario de un objeto, etc. Con estas características se puede, por ejemplo, clasificar las sombras según su posición respecto a los elementos que las pueden producir como edificios o árboles.

- Características relativas a objetos en un nivel de segmentación distinto al del objeto de estudio: Al realizar varias segmentaciones, con distintas escalas, en una imagen se crea una jerarquía entre los objetos de distintas segmentaciones ya que unos están englobados en otros de mayor tamaño, al mismo tiempo que engloban a objetos de menor tamaño. La consideración de esta jerarquía puede aportar información eficaz para describir el objeto y su contexto. Así por ejemplo, una zona de vegetación incluida en un objeto mayor clasificado como zona urbana deberá clasificarse como parque o jardín. De igual modo, la clase correspondiente a los polígonos contenidos en una parcela agrícola, resulta básica para asignar el cultivo a la parcela.

Ejemplos de estos descriptores pueden ser la proporción de área de una clase en un objeto, la existencia de una clase en él, la clase del objeto que lo engloba, etc.

Esta metodología facilita la interpretación del paisaje mediante la implantación de sistemas de producción basados en reglas que aplican el conocimiento experto disponible, o bien mediante sistemas de extracción de conocimiento propios de la inteligencia artificial (Lang et al., 2006). No obstante, aunque los métodos basados en reglas son los más frecuentes, este enfoque de clasificación puede aplicarse por medio de cualquiera de los clasificadores existentes.

Los motivos que han motivado la aparición de esta nueva metodología son:

- El aumento en la disponibilidad de imágenes de alta resolución.

- El incremento constante en las necesidades de los usuarios de SIG.

- Las limitaciones que tiene el análisis de imágenes por píxel.

- El desarrollo de las herramientas de programación orientadas a objetos.

- La opinión generalizada de que este enfoque puede aprovechar mejor la información espacial contenida en las imágenes.

- La adecuación de esta metodología al enfoque multiescala del análisis del medio ambiente.

2. Sistemas Expertos

Los sistemas expertos son una serie de condiciones organizadas en forma jerárquica a modo de árbol (Hernández Orallo et al., 2004). Son muy útiles para encontrar estructuras en espacios de alta dimensionalidad y en problemas que mezclen datos categóricos y numéricos. La estructura de los árboles de decisión, donde las reglas son evaluadas para comprobar hipótesis, es la más adecuada para expresar un sistema experto (Jensen, 2005).

Hoy día tenemos al alcance imágenes de alta resolución espacial que proporcionan gran detalle de los objetos contenidos en una escena. Este gran detalle puede dar a un exceso de variabilidad dentro los límites de una zona que pertenece a una misma cobertura. Este exceso de variabilidad llevan consigo una disminución de la separabilidad entre las distintas coberturas presente en la zona de estudio. Para paliar estos problemas está en auge el desarrollo de nuevos algoritmos de clasificación (Ayala y Menenti, 2002). En las clasificaciones basadas únicamente en la reflectancia de los píxeles es la insuficiente información para aislar completamente los objetos debido a la complejidad de los mismos y sus interacciones radiativas con otros objetos adyacentes en las imágenes de alta resolución espacial. Diversos autores han demostrado que los datos espectrales no son suficientes para la extracción de información de detalle en imágenes de alta resolución espacial (Matizan et al., 2007; Antunes et al., 2003). Para solucionar estos problemas, Abkar (2000) propone la combinación de los datos espectrales con otras fuentes de datos auxiliares con objeto de mejorar las clasificaciones. Sin embargo, Gong y Howarth (1990), afirman que es importante reconocer que los clasificadores convencionales no clasifican los modelos espaciales del mismo modo que lo hace el intérprete humano y proponen la incorporación de sistemas auxiliares para mejorar los resultados de dichos clasificadores convencionales.

Johnsson (1994) consiguió mejorar los resultados de clasificaciones espectrales mediante una segmentación previa de las imágenes en función de sus características espaciales, como por ejemplo el DEM (Digital Elevation Model).

Los sistemas expertos permiten la integración de datos de teledetección con otras fuentes de información georeferenciada, como los datos de uso del suelo, la textura espacial y los modelos digitales de elevación (DEM) para obtener mayor precisión en la clasificación (Lidov et al. 2000). El uso de información auxiliar para aumentar la precisión de la clasificación digital puede verse como un método que involucra el uso de una base de conocimiento anterior, con la información extraída de las imágenes (Trotter, 1991). Según Herman (2003), para mejorar los procedimientos automáticos de clasificación tenemos que introducir un conjunto de parámetros para la valoración de la clasificación más allá de los valores digitales de los píxeles. Podemos además, con los datos auxiliares, corregir los resultados iniciales de los procedimientos a los que estamos acostumbrados en las clasificaciones a través de reglas basadas en el conocimiento (Wicks et al. 2002). Stefanov et al. (2001), empleó un sistema experto para realizar la ordenación según reglas de decisión de una clasificación supervisada inicial de cobertura del suelo. Para ello, empleó como información adicional datos espaciales tales como la textura, el uso del suelo, reservas de agua, límites urbanos, etc. La precisión global de esta técnica fue del 85%.

Otra técnica empleada consiste en emplear la información contenida en el realce de imágenes mediantes los índices de vegetación. Esta información puede ser aportada a los clasificadores expertos, como es el caso del índice de vegetación NDVI (Normalizad Difference Vegetation Index), con el objetivo de separar zonas con y sin vegetación (Giannetti et al., 2001). En la Provincia de Jiangsu, China, Xiao et al. (2002), realizaron una clasificación no supervisada empleando como información auxiliar el índice NDVI y el índice NDWI

(Normalizad Difference Water Index) calculado en imágenes obtenidas por el sensor VGT (Vegetation – 1km de resolución espacial) a bordo del satélite SPOT-4. Los resultados fueron comparados con imágenes Landsat clasificadas y datos del censo agrícola.

El sistema experto se puede entender como una aplicación informática, o conjunto de ellas, capaz de solucionar un conjunto de problemas que exigen un gran conocimiento sobre un determinado tema. Sobre una base de conocimientos, posee información de uno o más expertos en un área específica.

Se puede entender como una rama de la inteligencia artificial, donde el poder de resolución de un problema en un programa de un ordenador viene del conocimiento de un dominio específico previamente facilitado a la aplicación. Estos sistemas imitan las actividades de un humano para resolver problemas de distinta índole (no necesariamente tiene que ser de inteligencia artificial).

Los sistemas expertos son herramientas desarrolladas con un objetivo muy claro y es dar fluidez y rapidez al análisis de información y casos en base a un conocimiento experto en temas concretos. Por esto debe ser una aplicación que fácilmente interactue con el usuario cumpliendo esto con dos características fundamentales:

- Explicar sus razonamientos o base del conocimiento: los sistemas expertos se deben realizar siguiendo ciertas reglas o pasos comprensibles de manera que se pueda generar la explicación para cada una de estas reglas, que a la vez se basan en hechos, es decir, no solo deben dar una solución al problema sino ser capaces de demostrar o justificar la elección lógica de la respuesta

- Adquisición de nuevos conocimientos o integrador del sistema: Son mecanismos de razonamiento que sirven para modificar los conocimientos anteriores. Sobre la base de lo anterior se puede decir que los sistemas expertos son el producto de investigaciones en el campo de la inteligencia artificial ya que ésta no intenta sustituir a los expertos humanos, sino que se desea ayudarlos a realizar con más rapidez y eficacia todas las tareas que realiza.

Básicamente un sistema experto esta compuesto por una base de conocimiento, que es el conocimiento aportado por un experto y procesado para su correcto uso, y otra base de hechos que es una memoria de trabajo donde almacena la información obtenida de los casos analizados durante el trabajo.

Después posee un motor de inferencia donde se programa y modela el proceso de razonamiento que queremos que siga y unos módulos de justificación que nos dan la explicación del porque se ha elegido esa solución en base a la lógica programada y la información experta disponible.

En la clasificación de usos del suelo existen múltiples variaciones en la forma de aplicar estos sistemas, siendo las dos maneras más habituales mediante la clasificación de una imagen incorporando en el proceso datos auxiliares y la reclasificación de una clasificación existente considerando la información auxiliar.

Huang y Jensen (1997) comparan sobre un mismo conjunto de datos, formado por imágenes y datos temáticos, los resultados de tres clasificaciones realizadas mediante un sistema experto basado en conocimiento, ISODATA y Máxima Probabilidad. Los mejores resultados corresponden al sistema experto con incrementos de fiabilidad global del 13% y del 9% respecto a los otros métodos.

Cohen y Shoshany (2000) emplearon el conocimiento experto para mejorar una clasificación no supervisada de cultivos en Israel. Para ello, añaden información auxiliar como son los mapas de tipos de suelo y de pluviometría. Del análisis de las características de una serie de muestras, deducen reglas basadas en la información auxiliar con las que reclasifican las imágenes aumentando las precisiones obtenidas.

Li et al. (2000) mejoraron los resultados de una clasificación de Máxima Probabilidad mediante el empleo de reglas en la que atributos como la elevación, la posición y los valores de probabilidad obtenidos en la clasificación son de utilidad. Con esto consiguieron aumentar la fiabilidad de las clases más similares espectralmente.

La construcción de un sistema experto basado en conocimiento está siempre condicionada a la adquisición del conocimiento necesario. Esto es, a la definición de una serie de reglas que permitan clasificar el conjunto de objetos en sus correspondientes clases informacionales. Un experto puede disponer de conocimiento útil para el sistema que deberá ser expresado en forma de reglas o condiciones aplicables por el sistema. No obstante, este conocimiento del experto será limitado y deberá ampliarse para conseguir el objetivo propuesto.

El sistema experto se compone por medio de reglas, que pueden constar de una o varias condiciones unidas por medio de operadores lógicos: SI se cumple una condición ENTONCES ocurre una acción.

Según Jensen (2005), la mejor forma de conceptualizar un sistema basado en conocimiento es mediante un árbol de decisión o sistema experto, donde los datos son evaluados mediante reglas para determinar las hipótesis. Los pasos a seguir para construir de forma automática una base de conocimiento para un sistema experto dedicado a la clasificación de imágenes de satélite, por medio de aprendizaje artificial inductivo serían:

- obtención de datos de entrenamiento

- creación del árbol de decisión (uno o varios)

- extracción de reglas

- clasificación por medio de reglas.

- evaluación del rendimiento del clasificador

Las ventajas principales de los algoritmos de los sistemas expertos son las siguientes: (Hernández Orallo et al., 2004)

- Son aplicables a distintas tareas: clasificación, regresión, agrupamiento, etc.

- Tratan con atributos continuos y discretos.

- Son flexibles. No hacen ninguna suposición sobre la distribución de los datos, al contrario de lo que hacen algunos métodos estadísticos. Esta característica permite incorporar datos discretos a la clasificación de imágenes, independientemente de la distribución y la correlación que exista entre ellos.

- Son fáciles de usar.

- Son tolerantes al ruido, a atributos no significativos y a valores faltantes.

- Las condiciones extraídas son inteligibles por el usuario

- Existe software para su aplicación y en algunos casos es gratuito.

- Permiten tratar relaciones no lineales entre características y clases.

Algunos de los softwares comerciales de tratamiento de imágenes más utilizados en teledetección como son ERDAS IMAGINE, ENVI o IDRISI disponen de herramientas componer algoritmos de clasificación experta.

Una vez establecida la base de conocimiento expresada en forma de árbol, se procede a la clasificación de la imagen de entrada obteniendo una imagen temática.

El clasificador experto de Erdas Imagine tiene permite que el usuario experto cada una de las reglas y su confianza asignada. De modo que si un píxel cumple varias reglas, será asignado a la hipótesis cuyas reglas tengan una confianza mayor.

3. Algoritmo basado en la corteza cerebral

Con los recientes avances que se han realizado en el campo de la neurociencia, existen más información sobre la organización y funcionamiento de la corteza cerebral, por lo que podemos aplicar los algoritmos de su funcionamiento al software, cosa que hasta ahora sólo se había conseguido de forma muy simplista y con resultados muy limitados usando redes neuronales.

El algoritmo basado en el córtex cerebral que se presenta en este Tesis Doctoral está basado en un nuevo concepto de tecnología de computación desarrollado por el equipo de trabajo encabezado por el ingeniero informático Jeff Hawkins, inventor del Palm Pilot y del teléfono inteligente Treo, además de fundador de las empresas Palm y Handspring. Este ingeniero ha trabajado en el campo de la neurociencia y es presidente del Instituto de Neurociencia de Redwood, fundado por él en 2002. Junto con Donna Dubinsky y Dileep George ha fundado la empresa Numenta, con el objetivo de desarrollar un nuevo tipo de memoria basada en el funcionamiento del cerebro humano.

Según Hawkins expresa en su libro Sobre la inteligencia, el cerebro funciona sobre la base de la memorización y el reconocimiento de patrones, de forma que la tarea que realiza el cerebro (o al menos la parte del cerebro denominada córtex) es la predicción, es lo que el autor llama Marco de Memoria-Predicción (memory-prediction framework). Según el autor, «El papel de cualquier región del córtex es averiguar qué relación hay entre sus entradas (inputs), memorizarla y usar esa memoria para predecir cómo se comportarán las entradas (inputs) en el futuro».

Tras varios años de estudio nace una nueva teoría basada en el córtex cerebral, lugar donde reside realmente la memoria y por tanto la inteligencia. El córtex es básicamente igual en todas partes, y aunque se sabe que diversas zonas tienen relación con ciertos sentidos y atributos (ej. la visión, el tacto, el lenguaje, etc.) en realidad todas sus partes, capas celulares, densidad, etc., son básicamente iguales.

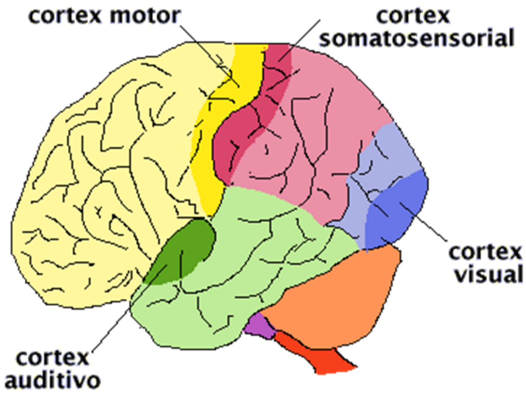

Figura 1. 1 Esquema de la corteza cerebral (George, 2008)

Esta nueva teoría, llamada Memoria-Predicción, argumenta que todas las capas del córtex pueden realizar las mismas funciones (Hawkins and Blakeslee, 2004), lo cual explicaría por qué funciones en zonas que se dañan por accidentes o enfermedad (por ejemplo, el lenguaje), se pueden volver a desarrollar en otra zona distinta.

El número y disposición de las conexiones al córtex, así como la diferencia de tamaño relativo (cuerpo/cerebro) y el tipo y el número de neuronas en el córtex humano y en el de los animales es lo que nos hace básicamente diferentes: los humanos somos inteligentes porque somos en cierto modo más conscientes del mundo que nos rodea y porque tenemos mayor capacidad de predicción y simulación mental (Hawkins and Blakeslee, 2004).

El algoritmo que utiliza la mente es un proceso que almacena patrones y hace predicciones sobre los patrones que encuentra o espera encontrar (Hawkins and Blakeslee, 2004). La exposición a los diversos estímulos (inputs) se guarda en el córtex. Pero el córtex es muy diferente a lo que habitualmente conocemos como «la memoria» de un ordenador. Las diferencias fundamentales de cómo funciona la memoria en el cerebro respecto a otros tipos de memorias (como un archivo o un ordenador) son las siguientes (Hawkins and Blakeslee, 2004):

La memoria almacena secuencias de patrones en vez de los datos en sí. Esto es como almacenar las diferencias de una nota a otra en una canción en vez de almacenar las notas en sí.

La memoria es accesible de forma auto-asociativa. Todos los recuerdos están asociados unos con otros de algún modo: ver una parte de una cara está asociado con que esa cara corresponde a una cabeza completa.

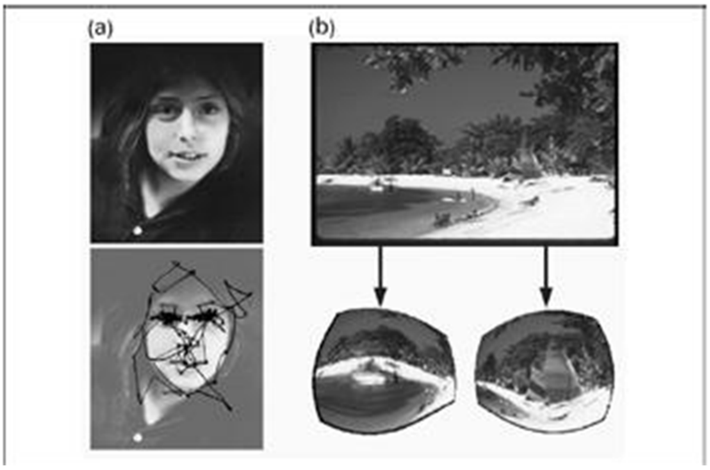

Figura 1. 2 (a) Cómo efectúa sacudidas oculares el ojo por una cara humana. (Hawkins and Blaskee, 2004).

Figura 1. 3 (b) Distorsión causada por la distribución irregular de receptores en la retina (Hawkins and Blaskee, 2004).

La memoria almacena los patrones en formato invariante. Una vez que has aprendido a leer, puedes reconocer las palabras y leerlas en cualquier ángulo, perspectiva, condición luminosa o aunque cambie el tipo de letra.

Los patrones se almacenan en una jerarquía. El concepto de jerarquía tiene que ver con las diversas capas del córtex, y básicamente consiste en una estructura que podría denominarse fractal, bellamente simétrica porque además cuenta con retroalimentación, en el que las regiones de menor nivel transmiten los patrones y los «nombres» que asignan a los patrones a las zonas de nivel superior. Pero todas son iguales y equivalentes en realidad. Esta jerarquía se corresponde, curiosamente, con el hecho de que también el Mundo en sí sea jerárquico.

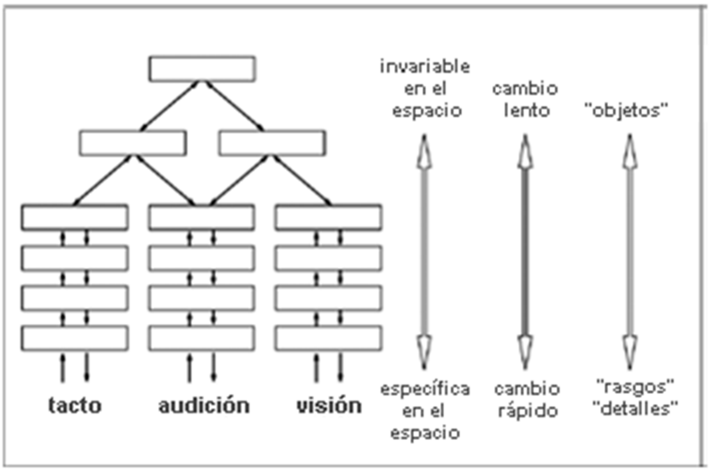

Figura 1. 4 Formación de representaciones invariables en el oído, la visión y el tacto (Hawkins and Blaskee, 2004).

En parte de esta teoría se basa el modelo de Memoria Temporal Jerárquica (MTJ) desarrollado por la empresa Numenta®. En este modelo, las relaciones temporal y espacial entre patrones de señales sensoriales forman una arquitectura de memoria jerárquica durante el proceso de aprendizaje. El aprendizaje puede ser supervisado y no supervisado. Cuando aparece un nuevo patrón el proceso de reconocimiento se activa y elige entre los patrones almacenados el que mejor lo representa.

Este algoritmo denominado Memoria Temporal Jerárquica se ha utilizado dentro del proyecto NUMENTA para desarrollar una plataforma de análisis de patrones geométricos, reconocimiento y clasificación llamada NuPIC (Numenta Platform for Intelligent Computing).

Este tipo de sistema de reconocimiento de patrones y clasificación de los mismos se puede entender, conceptualmente, como una red de neuronas artificial (o red neuronal) con una serie de diferencias conceptuales que convierten a este algoritmo en herramienta con un potencial muy fuerte en su campo de trabajo.

La diferencia fundamental entre las memorias jerárquicas temporales y las redes neuronales es que la primeras no están programadas para ejecutar una tarea que nos devuelve un resultado lógico emulando la conducta humana o la inteligencia humana a la hora de realizar un análisis; las memorias jerárquicas temporales están desarrolladas y diseñadas bajo una lógica de diseño obtenida a partir del estudio de la inteligencia humana en profundidad, es decir, no imitan la conducta inteligente si no que funcionan bajo una arquitectura similar a la que lo hace el cerebro humano.

Los sistemas clásicos similares a las redes neuronales están diseñados para resolver una (o varias tareas) para los que son programados de una forma específica buscando analizar los patrones de sus flujos de datos de una forma concreta. Por el contrario la tecnología NUMENTA no se programa para ningún fin concreto si no que con una arquitectura que emula el sistema de análisis de flujos de datos sensoriales del cerebro humano, se le entrena con un flujo de información suficiente del cual es capaz el sistema de sacar conclusiones en el análisis, predicción y clasificación de patrones.

Los algoritmos de análisis no están diseñados con un fin concreto o para resolver una tarea específica con lo que otra ventaja de esta novedosa tecnología es que se puede aplicar a un amplio conjunto de problemas que involucran modelos sensoriales de datos complejos.

4. La Memoria Temporal Jerárquica

Es un sistema de memoria que analiza la información de entrada y la procesa en busca de encontrar patrones conocidos y poder clasificarla. Tiene, a nivel de concepto y de arquitectura, mucho en común con las redes bayesianas ya que es tiene una estructura en forma de árbol con una serie de nodos interconectados que procesan la información bajo una serie de algoritmos de clasificación. En cuanto a su arquitectura, tiene una estructura en forma de árbol pero en este caso es una estructura jerárquica especial, con forma piramidal donde cada nodo implementa parte de un aprendizaje común. El objetivo de esto es imitar la naturaleza de la corteza cerebral a la hora de procesar y transmitir la información procedente de los sentidos, ya que la naturaleza de la realidad es jerárquica. Cualquier objeto se puede descomponer en otros objetos menores y estos a su vez en formas sencillas, es decir, tenemos una jerarquía de pequeños patrones que agrupados dan una forma, que a su vez agrupadas dan un objeto, y así sucesivamente. Tenemos así una red jerárquica en la que los nodos ejecutan un mismo algoritmo que permite el descubrimiento patrones en la información. En la figura 1.5 vemos una simplificación de una red con tres capas de nodos.

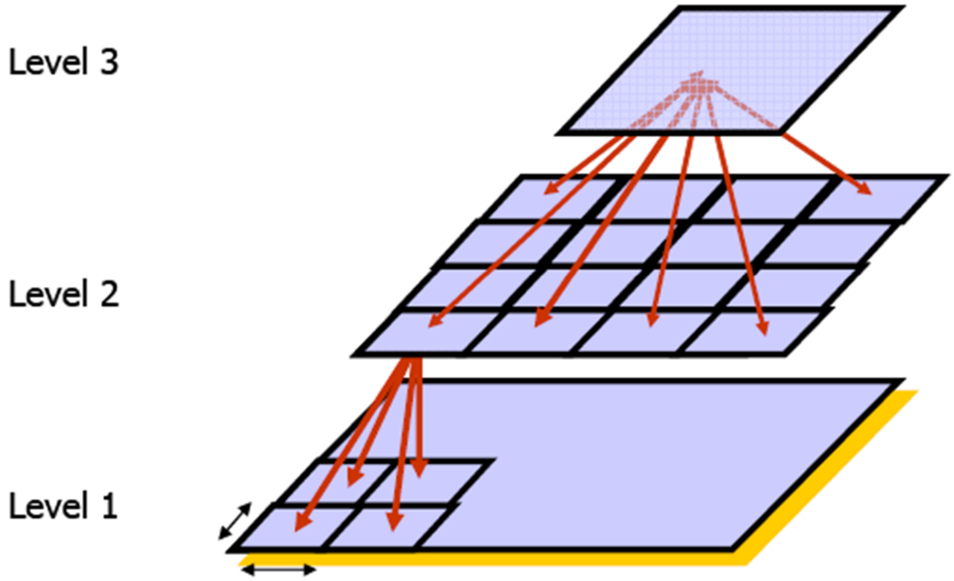

Figura 1. 5 Simplificación de una red de MTJ de tres niveles de nodos.

Cuando se propaga la información a través de una red de este tipo estamos imitando la naturaleza jerárquica de la información que nos llega a través de los sentidos ya que los nodos de los niveles inferiores solo recibirían una pequeña cantidad de información reconociendo pequeños patrones que combina con los nodos circundantes para entregarlos en el nivel superior, el cual puede reconocer patrones mayores y así de forma ascendente en la jerarquía.

Este es el punto clave de la MTJ (y su mayor diferencia con las demás tecnologías de inteligencia artificial existentes): imitan en su arquitectura y lógica de análisis el comportamiento de la inteligencia humana, es decir, el comportamiento de la corteza cerebral (que hemos visto en apartados anteriores con detalle). Con esta jerarquía y a través del análisis de la información con un único algoritmo, se busca en esta tecnología obtener una herramienta de inteligencia artificial versátil, capaz de analizar cualquier información.

Este es otro punto que diferencia las MTJ de las demás tecnologías existentes en temas de inteligencia artificial. No tiene algorítmicos específicos, ni funciones direccionales, es decir, no programamos una aplicación para la resolución de un problema concreto. Este sistema de memorias jerárquico puede procesar cualquier tipo de información y no tiene porque ser necesariamente correspondiente la de los sentidos humanos. Puede analizar flujos de datos de muchas fuentes (temperatura, humedad, datos bursátiles, etc.). La esencia de esto reside en que con la algorítmica diseñada por NUMENTA Inc. para esta aplicación tiene como objetivo encontrar patrones básicos que conforman la realidad de la información que está recibiendo (si hablamos de la información sensorial clásica de los sentidos humanos es más fácil comprender el concepto de modelo de la realidad), construyendo un modelo de la realidad. Es evidente que tenemos una fase inicial de entrenamiento donde aportamos la información necesaria para la red construya el modelo de la realidad que esta «observando» mediante el análisis de patrones, tras la cual podremos realizar clasificaciones e predicciones sobre información nueva.

Es en realidad una imitación de la forma de trabajar que tiene la corteza cerebral en el análisis de la información procedente de los sentidos, es decir, se ha creado una arquitectura artificial que funcione bajo los mismo preceptos de la zona del cerebro que maneja los procesos de memoria e inteligentes donde son puntos clave en su funcionamiento la estructura jerárquica y el papel del tiempo en el análisis de flujos de datos.

4.1. Funciones de la Memoria Temporal Jerárquica

Esta tecnología puede analizar un flujo de datos de diferente naturaleza ya que trabaja bajo un algoritmo único con lo que podemos analizar datos de cualquier sentido equivalente humano e incluso datos que carecen de correlación con una percepción sensorial humana. Sobre estos flujos de datos las MTJ realizan básicamente dos tareas que además son las que marcan los dos estados en los que se puede encontrar el sistema (George D., 2007).

Por un lado tenemos una etapa inicial, que podríamos llamar de entrenamiento, donde la red está recibiendo y analizando información de una cierta naturaleza de forma que busca conocer los patrones comunes de esta y los que se repiten en secuencias similares. De esta manera estamos creando en nuestro sistema de memoria una representación de la realidad a la que corresponden esos datos. Si hablamos de información equivalente a un sentido humano (como por ejemplo la visión) es más sencillo comprender el concepto de representación de la realidad ya que lo que estaría haciendo el sistema es almacenar todos los diferentes patrones que se le están aportando, si hablamos de la visión, pues estaría analizando y memorizando los diferentes objetos que se le presentan en secuencia temporal.

En esta etapa inicial de funcionamiento, la función principal del sistema es descubrir causas en la información que recibe para crear una representación del «mundo» que se le presenta. Analizando los paralelismos con el funcionamiento de la inteligencia humana, lo que el sistema está haciendo aquí es memorizar las representaciones invariantes como hacía la corteza cerebral.

Existe una segunda fase o estado del sistema de memoria en el que este ya conoce la realidad que se le presenta a través de un flujo de información, es decir, ha creado una representación de esta mediante representaciones invariantes y está preparado para recibir información nueva, nuevos inputs. Sobre esta información el sistema realizará un proceso de inferencia basándose en el modelo del «mundo» que ha creado, analizando los nuevos patrones y comparándolos con las representaciones invariantes que posee de forma que realiza una clasificación de los nuevos patrones acorde a los que ya conoce. En este caso, también estamos imitando, mediante la algorítmica y la arquitectura con la que se ha construido el sistema, la forma en que la corteza cerebral analiza constantemente la información que le llega a través de los sentidos. En este caso es algo diferente ya que en nuestro cerebro no tenemos una fase de entrenamiento que podemos dar por finalizada para que comience la fase de reconocimiento pero si se imita la forma de trabajar con información (de analizarla, compararla, memorizarla. Aquí la función principal de las memorias jerárquicas temporales es la inferencia de patrones a partir de las representaciones invariantes creadas en la etapa anterior (George D., 2007).

Estas son las dos principales funciones de una red MTJ: descubrir causas e inferir causas en información «nueva».

4.2. Importancia de la jerarquía y el tiempo en las redes MTJ

Ya se ha visto, en parte, que hay dos características claves para el correcto funcionamiento de las redes de memoria jerárquica temporal y estas son la estructura jerárquica de la red y la componente temporal de los datos que recibe la red a través de los sensores.

- Importancia de una arquitectura jerárquica

Existen varios motivos por los que es necesario trabajar bajo un arquitectura jerárquica en la redes MTJ pero el más básico es que con una estructura de este tipo aprovechamos la naturaleza jerárquica de la realidad. Todos los objetos que podemos percibir de la realidad (o cualquier información de otros sentidos que no sean la vista, no tenemos por qué limitarnos al análisis de datos visuales) tienen una estructura jerárquica ya que se pueden descomponer en sub-patrones (por ejemplo, un rosto se compone de ojos, nariz, boca, etc. y unos ojos se componen de iris, pupilas, párpados… y así sucesivamente). Además, los patrones que estén cerca espacialmente hablando estarán fuertemente correlacionados, propiedad la cual se explota con fuerza por este tipo de red debido a una arquitectura jerárquica que busca relaciones espaciales entre patrones que se presentan cercanos. El concepto de cercanía aumenta conforme ascendemos en la jerarquía pues en principio patrones muy distantes no tienen ninguna correlación ni forman parte de una misma entidad. Pero esto es si lo analizamos desde una perspectiva cercana, de los nodos de niveles inferiores ya que cuando ascendemos en la red aumenta la parte de la realidad analizada y patrones muy distantes ahora pueden cobrar sentido como parte de una misma entidad (por ejemplo, y siguiendo el caso del rostro, los patrones del contorno de un labio no tienen mucha relación con los de una ceja para nodos de nivel bajo pero si para el nodo último en la jerarquía que activa el patrón de rostro).

Pero no solo aprovechamos la estructura jerárquica espacial de la realidad sino que también tenemos que explotar la jerarquía temporal de los patrones que se presentan desde los sensores, es decir, gracias a una arquitectura jerárquica podemos crear secuencias de patrones estableciendo relaciones entre estos por su cercanía en el tiempo. Toda la información de los sentidos como el tacto o el oído tiene una evidente componente temporal pero la visión también tiene una importante fuerza en este sentido ya que nuestros sensores no paran de moverse y los objetos también lo hacen dentro de la realidad. Gracias a esto podemos crear secuencias de patrones ya que supondremos que los que suceden continuados en el tiempo tienen una fuerte correlación. Al igual que ocurre con la jerarquía espacial, en los nodos inferiores encontraremos correlaciones temporales muy cercanas en esta dimensión mientras que conforme ascendemos en la jerarquía de la red y comenzamos a analizar secuencias de secuencias de patrones, el campo de análisis temporal aumenta, encontrado secuencias mayores con una relación temporal de mayor tamaño.

Pero tenemos también un motivo técnico clave para usar este tipo de arquitectura y es que usar este tipo de jerarquía nos permite optimizar la memoria y reducir los tiempos de análisis de la información.

Muchos sistemas clásicos de inteligencia artificial tienen grandes problemas ya que conforme comienzan a aumentar el volumen de datos analizado así como la dimensión de la información, se producen requerimientos muy altos de memoria para procesar esta. Las redes MTJ requieren también de una gran cantidad de memoria y de un tiempo de aprendizaje alto pero no sufren problemas de crecimiento exponencial de las necesidades de memoria y del tiempo de procesado, es decir, requieren un sistema relativamente potente para ejecutarse pero no sufren problemas de escala.

Esto lo conseguimos con las representaciones compartidas o secuencias compartidas. Existen causas sencillas en los niveles bajos que formarán parte de muchas causas complejas en niveles superiores con lo que no se requiere que sean aprendidas cada vez que se reconoce un objeto o secuencia en los niveles superiores sino que se comparten estas secuencias ya conocidas por lo niveles bajos, reduciendo significativamente el tiempo de análisis de la información así como de la memoria necesaria para ello.

5. Presentación del trabajo

En el desarrollo de esta Tesis Doctoral se abarca el estudio y aplicación de nuevos algoritmos de clasificación digital en imágenes de alta resolución espacial, obtenidas tanto de satélites de alta resolución como de sensores digitales aerotransportados, con el objetivo de obtener clasificaciones de patrones geométricos en elementos del territorio. Se proponen tres metodologías avanzadas de clasificación digital:

i) una empleando imágenes de satélite de alta resolución espacial combinadas con algoritmos de clasificación experta que se auxilian de resultados obtenidos por clasificadores orientados a objetos, índices de vegetación y transformaciones de bandas.

ii) el empleo de información espectral de fotogramas captados por sensores digitales aereotransportados combinado con algoritmos de clasificación experta, índices de vegetación y transformaciones de bandas.

iii) empleo de imágenes de sensores digitales aerotransportados y algoritmos basados en el funcionamiento del neocortex humano.

Por tanto, el trabajo se compone de cuatro capítulos que se resumen a continuación:

5.1. Tercer Capítulo.

El objetivo del presente capítulo es poner a punto una metodología de clasificación de imágenes de satélite, que auxiliada por la clasificación orientada a objetos y el NDVI (Normalized Difference Vegetation Index), permita cuantificar las áreas agrícolas de la región utilizando algoritmos de clasificación experta, con vistas a mejorar los resultados finales de las clasificaciones temáticas. Se utilizaron imágenes satelitales Quickbird y datos de 2532 parcelas en Hinojosa del Duque, España, para validar las clasificaciones, consiguiendo una precisión total del 91,98 % y un excelente estadístico Kappa (87,65%) para el algoritmo de clasificación experta.

5.2. Cuarto Capítulo.

El propósito de este capítulo fue evaluar la utilidad de la información espectral de sensores digitales aéreos en la determinación de la clasificación de la cubierta vegetal mediante nuevas técnicas de clasificación digital. Los usos del suelo que han sido evaluados son los siguientes: (1) suelo desnudo, (2) los cereales, incluido el maíz (Zea mays L.), avena (Avena sativa L.), centeno (Secalecereale L.), trigo (Triticum aestivum L.) y cebada (Hordeun vulgareL.), (3) los cultivos de alto valor proteico, tales como los guisantes (Pisum sativum L.) y habas (Vicia faba L.), (4) Alfalfa(Medicago sativa L.), (5) bosques y matorrales, incluidas las encinas (Quercus ilex L.) y retama común (Retama sphaerocarpaL.), (6) Suelo urbano, (7) olivos (Olea europaea L.) y (8) retirada cubierta. El mejor resultado se obtuvo mediante un algoritmo de clasificación de experta, logrando un índice de fiabilidad del 95%. Este resultado mostró que las imágenes de sensores digitales aerotransportados mantienen una promesa considerable para el futuro en el ámbito de las clasificaciones digitales debido a que estas imágenes contienen información valiosa que solo se estaba aprovechando desde el punto de vista geométrico. Por otra parte, las nuevas técnicas de clasificación reducen los problemas relacionados con imágenes de alta resolución alcanzando mejores resultados que los obtenidos con los métodos tradicionales.

5.3. Quinto Capítulo

El objetivo de este capítulo es presentar la teoría Memoria-Predicción, aplicada en forma de Memoria Temporal Jerárquica (MTJ), para la clasificación de usos del suelo. Numenta® MTJ es una nueva tecnología informática que imita la estructura y función del neocórtex humano. En este estudio, un fotograma, captado por el sensor fotogramétrico UltraCamD de Vexcel®, y datos sobre 1513 parcelas en Manzanilla (Huelva, España) fueron utilizados para validar la clasificación, logrando una precisión en la clasificación del 90,4%.

Los resultados obtenidos abren un nuevo campo de algoritmos avanzados de clasificación digital con unos resultados prometedores para la discriminación de usos del suelo.

5.4. Sexto capítulo

Este artículo presenta un sistema de inferencia para la detección de viñedo con fotografías aéreas digitales. El sistema se inspira en la reciente teoría memoria-predicción y en los modelos de arquitectura de alto nivel de la neocorteza humana. El documento describe la arquitectura jerárquica y el reconocimiento de la actuación de este modelo Bayesiano. Los resultados indica que utilizando un fotograma procedente del sensor fotogramétrico Ultracamd® de Vexcell, el 96% de las parcelas se han detectado. Como conclusión se indica que el proceso automático desarrollado puede ser integrado fácilmente en el Sistema de Información Geográfica del usuario final y produce información útil para la gestión de los viñedos.

Cada uno de estos capítulos ha dado lugar a un artículo científico que ha sido enviado a revistas internacionales de reconocido prestigio. Tras la consecuente revisión, han sido aceptados y publicados de forma definitiva, los siguientes artículos:

2009. Perea, A.J., Meroño, J.E., Aguilera, M.J., *Algorithms of expert classification applied in Quickbird satellite images for land use mapping*. Chilean Journal of Agricultural Research, 69, (3), 400 - 405.

2009. Perea, A.J., Meroño, J.E., Aguilera, M.J., *Application of Numenta® Hierarchical Temporal Memory for land-use classification*. South African Journal of Science; 105, (9-10), 370 - 375.

2010. Perea, A.J., Meroño, J.E., Aguilera, M.J., de la Cruz, J.L., *Land-cover classification with an expert classification algorithm using digital aerial photographs*. South African Journal of Science; 106, (5-6) 82-87.

2012. Perea, A.J., Meroño, J.E., Aguilera, M.J., *Hierarchical Temporal Memory for mapping vineyards using digital aerial photographs*. African Journal of Agricultural Research; 7(3) 456-466.

6. Referencias

Abkar, A.A, Sharifi, M.A. y N.J. Mulder. 2000. *Likelihood-based image segmentation and classification: a framework for the integration of expert knowledge in image classification procedures*. International Journal of Applied Earth Observation and Geoinformation; 2:104-119.

Ayala, R.M. y M. Menenti. 2002. *Alternativas a los problemas presentados en un proceso de clasificación basado en el reconocimiento espectral de patrones*. MAPPING; 75:72-76.

Carpenter, A.B. y B. Malcolm. 1994. *Neuroanatomía: fundamentos*. Panamericana.

Castilla, G. 2003. *Análisis orientado a objetos de imágenes de teledetección para cartografía forestal: bases conceptuales y un método de segmentación para obtener una partición inicial para la clasificación* (Tesis Doctoral). ETSI Montes (Universidad Politécnica de Madrid), España.

Cohen, Y. y Shoshany, M. 2000. *Integration of remote sensing, GIS and expert knowledge in national knowledge-based crop recognition in Mediterranean environment*. IAPRS; Vol. XXXIII, Parte B7; p. 280-286.

Deleep, G. y J. Hawkins. 2007. *Hierarchical Temporal Memory, Concepts, Theory, and Terminology*. Disponible en: http://www.numenta.com/htm-overview/education/Numenta_HTM_Concepts.pdf

Mantifar, H.R, Sarmadian, F, Alavi Panah, S.K. y R.J. Heck. 2007. *Comparisons of Object-Oriented and Pixel-Based Classification of Land Use/Land Cover Types Based on Landsat7, Etm+ Spectral Bands (Case Study: Arid Region of Iran)*. American-Eurasian Journal of Agricicultural and Environmental Science; 2(4): 448-456.

Antunes, A.F.B, Lingnau, C. y J. Centeno da Silva. 2003. *Object oriented analysis and semantic network for high resolution image classification*, Anais XI SBSR, Belo Horizonte, Brasil, 5-10 abril.

Aplin, P, Atkinson, P.M. y P.J. Curran. 1999. *Fine Spatial Resolution Satellite Sensor Imagery for Land Cover Mapping in the United Kingdom*. Remote Sensing of Environment; 68(3):206-216.

Giannetti, F, Montanarella, L. y R. Salandin. 2001. *Integrated use of satellites images, DEMs, soil and substrate data in studying mountainous lands*. International Journal of Applied Earth Observation and Geoinformation; 3:25-29.

Gibson, P.J. y C.H. Power. 2000. *Introductory Remote Sensing – Digital image processing and applications*. Taylor & Francis Ltd.

Giovacchini, A. y A. Brunetti. 1991. *Agricultural Statistics by Remote Sensing in Italy an Ultimate Cost Analysis. Proceedings*. Conference on The Application of Remote Sensing to Agricultural Statistics, Villa Carlotta, Belgirate, Lake Maggiore, Italy, 1991 26-27 Nov; p. 87-93

George, C. 2008. *General psychology*. Shippensburg University. Disponible en: http://webspace.ship.edu/cgboer/genpsy.html

Gong, H. y P.J. Howarth. 1990. *An Assessment of Some Factors Influencing Multispectral Land-Cover Classification*. Photogrammetric Engineering and Remote Sensing; 56:597-603.

Haralick, R.M., Shanmugan, K. y I. Dinstein. 1973. *Texture features for image classification*. IEEE Transactions on Systems, Man and Cybernetics; Vol. 3(1), p. 610-621.

Hay, G.J. y G. Castilla. 2006. *Object-based Image Analysis: Strengths, weakness, opportunities and threats (SWOT)*. En: S. Lang, T. Blaschke y E. Schöpfer (Eds.), OBIA. 4-5 Julio, Salzburg, Austria. Disponible en: http://www.commission4.isprs.org/obia06/papers.htm

Hawkins, J. y S. Blakeslee. 2004. *On intelligence*. Nueva York, (EE.UU.): Times books, Henry Holt and Co.

Hernández Orallo, J., Ramírez Quintana, M.J. y C. Ferri Ramírez. 2004. *Introducción a la minería de datos*. Madrid: Pearson Educación S.A.

Heyman, O. 2003. *Automatic Extraction of Natural Objects form 1-m Remote Sensing Images*. Disponible en: http://www.cobblestoneconcepts.com/ucgis2summer/heyman/heyman.htm

Huang, X. Y J.R.A. Jensen. 1997. *Machine-learning approach to automated knowledgebase building for remote-sensing image analysis with GIS data*. Photogrammetric Engineering and Remote Sensing; 63(10):1185-1194.

Johnsson, K. 1994. *Segmented-Based Land-Use Classification from SPOT Satellite Data*. Photogrammetric Engineering and Remote Sensing; 60:47-53.

Jensen, J.R. 2005. *Introductory digital image processing* (3ª ed.). Upper Saddle River (EE.UU.): Pearson Education, Inc.

Lang, S. y T. Blaschke. 2006. *Bridging remote sensing and GIS - which are the main supportive pillars?* En: S. Lang, T. Blaschke y E. Schöpfer (Eds.), OBIA. 4-5 Julio, Salzburg, Austria. Disponible en: http://www.commission4.isprs.org/obia06/papers.htm

Li, D, Sui, H. y P. Xiao. 2002. *Automatic change detection of geo-spatial data from imagery*. International Archives of Photogrammetric and Remote Sensing (IAPRS; Vol. XXXIV, Part 2, p. 245-251.

Lidov, L, Miller, R, Wormer, D.M. y K.A. Tilley. 2000. *Interview: Understanding the Future of Commercial Remote Sensing*. Photogrammetric Engineering and Remote Sensing; Vol. 66(1):5-14.

Lizarazo, I. y Elsner, P. 2008. *From pixels to grixels: A unified functional model for geographic object-based image analysis*. Proceedings of GEOBIA, 2008 - Pixels, Objects, Intelligence: Geographic Object Based Image Analysis for the 21st century. 5-8 Agosto, Calgary, Canada. Disponible en: http://homepages.ucalgary.ca/~gjhay/geobia/Proceedings/Sessions/Session10/6664_Lizarazo_Proc.pdf

Shapiro, L.G. y G.C. Stockman. 2001. *Computer Vision*. London: Prentice Hall.

Snell, R.S. 2007. *Neuroanatomía clínica*. Washigton (EE.UU): Ed. Médica Panamericana.

Stefanov, W.L, Ramsey, M.S. y P.R. Christensen. 2001. *Monitoring urban land cover change: An expert system approach to land cover* *classification of semiarid to arid urban centers*. Remote Sensing of Environment; Vol. 77:173-185.

Strahler, A.H. 1980. *The Use of Prior Probabilities in Maximum Likelihood Classification of Remotely Sensed Data*. Remote Sensing of Environment; 10: 135-165.

Trotter, C.M. 1991. *Remotely-sensed data as an information source for geographical information systems in natural resource management: a review*. International Journal Geographical Information Systems; 5:225-239.

Wicks, T.E, Smith, G.M y P.J. Curran. 2002. *Polygon-based aggregation of remotely sensed data for regional ecological analyses*. International Journal of Applied Earth Observation and Geoinformation; 4:161-173.

Wilkinson, G.G, Kanellopoulos, I, Kontoes, C. y R. Schoenmakers. 1991. *Advances in the Automatic Processing of Satellite Images. Proceedings*. Conference on The Application of Remote Sensing to Agricultural Statistics, Villa Carlotta, Belgirate, Lake Maggiore, Italy, 26-27 Nov; p.125-132.

Xiao, X, Boles, S, Frolking, S, Salas, W, Moore, B, Li, C, He, L. y R. Zhao. 2000. *Landscape-scale characterization of cropland in China using Vegetation and Landsat TM images*. International Journal of Remote Sensing; 23:3579-3594.

Capítulo II. Objetivos Generales

«Nuestro cerebro efectúa predicciones constantes sobre la misma estructura del mundo en que vivimos»

Jeff Hawkins (2004)

El objetivo general de esta Tesis es la evaluación de las nuevas técnicas de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo.

El objetivo general se puede dividir en los siguientes objetivos específicos:

1. Análisis y estudio de las nuevas técnicas de clasificación digital basadas en el análisis de objetos mediante segmentación de imágenes.

2. Análisis y estudio de los algoritmos de clasificación experta, los cuales utilizan información complementaria a la puramente espectral.

3. Análisis de la Teoría Memoria-Predicción, evaluación de la Memoria Temporal Jerárquica, algoritmos empleados y plataforma de programación. Aplicación a la clasificación de patrones geométricos para la discriminación de usos del suelo.

4. Estudio de la aplicabilidad de las nuevas técnicas de clasificación digital a las imágenes fotogramétricas.

5. Evaluación de los resultados obtenidos aplicando diferentes técnicas de clasificación en imágenes de sensores fotogramétricos.

Capítulo III. Algorithms of expert classification applied in Quickbird satellite images for land use mapping

Autores: Perea, A.J.[a], Meroño, J.E.[b], Aguilera, M.J.[a]

[a] Departamento de Física Aplicada, Universidad de Córdoba (UCO), Córdoba, Spain

[b] Departamento de Ingeniería Gráfica y Geomática, Universidad de Córdoba (UCO), Córdoba, Spain

Publicado en: Chilean Journal of Agricultural Research, 69, (3), 400 - 405.

«La corteza cerebral posee un algoritmo de aprendizaje inteligente que descubre y capta de forma natural cualqier estructura jerárquica que exista»

Jeff Hawkins (2004)

1. Resumen

El objetivo del presente trabajo es poner a punto una metodología de clasificación de imágenes de satélite, que auxiliada por la clasificación orientada a objetos y el índice de vegetación de diferencia normalizada (normalized difference vegetation index, NDVI), permita cuantificar las áreas agrícolas de la región utilizando algoritmos de clasificación experta, con vistas a mejorar los resultados finales de las clasificaciones temáticas. Se utilizaron imágenes satelitales Quickbird y datos de 2532 parcelas en Hinojosa del Duque, España, para validar las clasificaciones, consiguiendo una precisión total del 91,9% y un excelente estadístico Kappa (87,6%) para el algoritmo de clasificación experta.

Palabras clave: clasificación experta, índice de vegetación, cobertura de tierra, clasificación orientada a objetos.

2. Abstract

The objective of this paper was the development of a methodology for the classification of digital aerial images, which, with the aid of object-based classification and the Normalized Difference Vegetation Index (NDVI), can quantify agricultural areas, by using algorithms of expert classification, with the aim of improving the final results of thematic classifications. QuickBird satellite images and data of 2532 plots in Hinojosa del Duque, Spain, were used to validate the different classifications, obtaining an overall classification accuracy of 91.9% and an excellent Kappa statistic (87.6%) for the algorithm of expert classification.

Key words: expert classification, vegetation index, land cover, object-based classification

3. Introduction

The digital classification of images is the process in which pixels with similar spectral characteristics, and therefore assumed to belong to the same class, are identified and assigned a color (Gibson and Power, 2000).

In the high resolution images from satellites, such as Ikonos and QuickBird, each pixel no longer refers to a complete object, character or area, but rather to a portion of the components of these, which means that classic techniques of classification based on pixels present some limitations (Wilkinson et al., 1991): (1) the spectral information contained in pixels is not sufficient in the majority of cases, such as to identify vegetation species or the types of surface cover; and (2) normally pixels include a radiometric mixture from their neighbors and consequently few zones have total homogeneity. In the area of the digital treatment of images, there is currently great interest in the development of new classification algorithms (Ayala and Menenti, 2002). The combination of spectral data with other sources of auxiliary data allows for the use of more information that can improve classifications (Abkar et al., 2000).

Expert systems consider the use of data other than spectral characteristics in order to improve the results of classification (Lidov et al., 2000). The use of auxiliary information to increase the accuracy of digital classification involves combining an existing knowledge base with information extracted from images (Trotter, 1991). To improve automatic classification procedures, it is necessary to introduce a set of parameters to inform the classification beyond the digital values of the pixels (Heyman, 2003). As well, with the use of auxiliary data, we can correct the initial results of the procedures through knowledge-based rules (Wicks et al., 2002). Johnsson (1994) showed that the accuracy of spectral classifications can be improved by segmenting the images in function of their spatial characteristics. Stefanov et al. (2001) developed a system of expert classification with the main objective of reclassifying the initial classification of maximum probability of urban zones and thus reduce errors of omission and commission. As well, the results of highlighting images using vegetation indices or other advanced classifiers, such as object-oriented classifiers, can be incorporated into expert classifiers.

The main objective of this work was to design and train an expert classification algorithm to classify a QuickBird satellite image of Hinojosa del Duque, Cordoba Province, Spain, obtained in April 2005, putting special interest in discriminating agricultural crops. The confusions that appeared in the different land uses were evaluated and the results obtained in the object-oriented classification were incorporated into the algorithm to study their effect in improving the classification compared to traditional techniques.

4. Materials and Methods

The area of study was located in Cordoba Province, Spain, in Pedroches Valley and includes the municipality of Hinojosa del Duque (38º33’ and 38º23’ N; 5º16’ and 5º50’ W). This is a rectangular area of 16 x 20 km and covers 32 000 ha.

Six multispectral images were used (QuickBird, Ortho Ready Standard Imagery, Digital Globe, Longmont, Colorado, USA), in UTM coordinates (Universal Transverse Mercator) and georeferenced in the WGS84 system. These images were ortorrectified and referred to the European Datum 1950 of the International Ellipsoid. The images were codified in 16 bits with a resolution of 2.4 m and were composed of four bands (blue, green, red and near infrared). The images were taken on 21 April 2005, beginning at 11:22, with a solar elevation angle of 62.9º.

To develop this work, information was used from field visits by the Public Enterprise for Agricultural Development. The uses of the land were: bare soil; cereal: corn (Zea mays L.), oats (Avena sativa L.), rye

(Secale cereale L.), wheat (Triticum aestivum L.), barley (Hordeum vulgare L.); removed cover; high protein crops: peas (Pisum sativum L.), beans (Vicia faba L.); alfalfa (Medicago sativa L.), woodlands/scrublands: holly oak (Quercus ilex L.), common retama (Retama sphaerocarpa [L.] Boiss.); constructed surfaces; olive (Olea europaea L.).

The ERDAS Imagine 9.0 software (Leica Geosystems Geospatial Imaging, Norcross, Georgia, USA) was used to carry out the supervised classifications and the expert classification algorithm. In the case of object-oriented classification, eCognition Professional 5.0 software (Definiens, München, Germany) was used.

The process began with the radiometric correction of the images. Following this treatment, the principle components and normalized difference vegetation index (NDVI) were calculated. Subsequently, images were obtained with the combination of desired bands and the different types of classification were made. Finally, the results of these classifications were validated.

4.1. Radiometric corrections of QuickBird images

Radiometric corrections modify the original digital levels to assimilate them to values that will present the image in the case of ideal reception. QuickBird images already have a series of radiometric corrections that the distributing company applies to its commercial products. The main corrections in the images are: restoration of lost pixels from the image or the possible loss or addition to the image.

A transformation of the digital levels at radiance values in the atmospheric ceiling was made and a reflectivity image was obtained. The conversion to the spectral radiance of the atmospheric ceiling can be done simply in two steps: the value of the corrected pixels is multiplied by the appropriate absolute calibration factor and the result is divided by the effective bandwidth to obtain spectral radiance. The radiometric calibration factor is included in the metadata files of the image.

4.2. Obtaining the principle components

The objective of principle component analysis (PCA) is to summarize a wide group of variables in a new and smaller set without losing a significant part of the original information (Chuvieco, 2000). For the final user of distance imaging products, the objective of PCA is to construct images with an increased capacity to differentiate types of cover.

4.3. Obtaining the NDVI index

Vegetation has very characteristic spectral behavior, with high absorption of red wavelengths and high reflectivity in the near infrared.

The NDVI index was obtained with the objective of highlighting 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 (Marini, 2006), which showed a positive effect.

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 RNIR and RRED, are reflectance in the near infrared band (R800 nm) and the red band (R690 nm), respectively.

4.4. Supervised classification

A series of images was obtained based on the different combinations of bands (Table 2.1). A supervised classification was made on all of these images.

Type of sensor	Image	Number of bands
QuickBird	Main components	4
	Main components + NDVI	5

NDVI: normalized difference vegetation index

Table 2. 1 Images used in 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.

4.5. Classification oriented to objects

The particularity of this type of analysis is that the classification is based on objects rather than pixels. Once the image is formed by pixels, the first step in object-oriented analysis is grouping the adjacent pixels through the region growing method to subsequently classify the extracted objects. In this way the number of parameter that can be valued is increased significantly, allowing for consideration of criteria such as size, shape, mean color, maximums and minimum, proximity to other objects, texture, etc. At the same time, segmentation reduces the number of objects to be classified, which also reduces the processing time.

The stopping criterion in the process of merging the regions is produced thanks to the parameter termed “scale” and can be defined by the user. It determines the maximum permitted in the global heterogeneity of the segments. The larger the parameters of scale for a database, the larger are the objects of the image. Given that the parameter of scale can be modified, we can obtain different types of segmented images. Thus, the objects generated in a broader segmentation inherit information from smaller objects with finer parameters of scale.

Subsequently, the classifications are trained using different training parcels and are validated using the same validation parcels used in the previous classifications. Finally, for the best results obtained in the segmentation process, the spatial and spectral characteristics are considered, as well as groups of pixels that define relatively homogenous areas. The “Multiresolution Segmentation” option was used, which automatically extracts homogenous objects. The number of objects to be created was 200, which is a parameter related to image resolution, the working scale and the heterogeneity of the data.

The program takes into account three criteria for segmentation: color, homogeneity and compactability. Color is the most important and has the greatest effect in defining objects in the majority of cases. The color criteria take into account the percentage of spectral homogeneity. However, shape and homogeneity are also important in extracting objects. The segmentation criteria of the image bands were 0.9 for color and 0.1 for shape, and under shape 0.5 was considered for “homogeneity” and 0.5 for “compactability”.

The nearest neighbor algorithm was used for the classification: some samples were chosen (training area) for each of the classes. The rest of the scene is classified in accordance to this. This is a very rapid and simple method, adequate when the classification of an object requires many bands/criteria. As well, it takes into account different parameters related to the objects (area, longitude, mean color, brightness, and texture).

4.6. Expert classification algorithm

The expert classification algorithm used in this work consisted of assigning the classes that make up the legend based on the area of coincidence among different types of images that had been classified previously. To do this, the following information was necessary: an image created based on field visit and the map of land use and vegetal cover for Andalucía for 2004 were used as true terrain. The ERDAS Imagine 9.0 system and the supervised classifications obtained from based on the image formed by the principle components and the image formed by the principle components + NDVI, as well as the object-oriented classification.

This algorithm was designed with the following decision-making criteria or rules: (1) when the pixels of each class of the classified image of the principle components + NDVI coincide with the image classified from principle components, they will be assigned to this class, (2) the other pixels where there is no coincidence belong to the which they are assigned by the object-oriented classification. To evaluate the quality of classifications, a total of 75,000 verification points were taken (approximately 2% of the area) for those that provide both real cover (true terrain) and those obtained by classification.

5. Results and Discussions

5.1. Results of the object-oriented classification

The result obtained from the segmentation was a new image that divided the original image into 9481 regions, such that the pixels included in each region are very similar among themselves and different from those from neighboring regions.

Use legend	Producer accuracy	User accuracy
	—————— % ——————	
Bare soil	89.4	94.4
Cereal	86.6	94.5
Removed Cover	94.4	85.0
High protein crops	99.5	71.4
Alfalfa	88.3	64.2
Woodlands and scrublands	100.0	83.3
Constructed surface	80.0	100.0
Total accuracy, %	90.9	
Kappa statistic, %	87.6	

Table 2. 2 Producer and user accuracies, overall and Kappa statistic for the object based classification.

With regard to the goodness-of-fit analysis of the classification that is shown in Table 2.2, the confusions that were observed between high protein crops and alfalfa give lower user accuracy than the other classes (64.2%). On the other hand, the producer accuracy was very high, with 88.3%, which indicates that 8.83 of every 10 pixels belonging to this cover were assigned correctly to the classification. As possible causes for this confusion of alfalfa, above all with high protein crops, is the incorrect delimiting of the training areas for the heterogeneity of the species throughout the image. Nevertheless, the level of confusion is not very high, in the majority of cases user and producer accuracy being 80%. As is normal, a class that is so easily discernable as constructed surfaces presents an accuracy of 100%. Consequently, the total accuracy of the classification obtained is very high, situated at 90.9%. Finally, the Kappa statistic also indicates a good classification, presenting a value of 87.6%.

5.2. Results of the expert classification algorithm

The methodology presented for the quantification of the agricultural areas of the studied region offered significant improvements in the results of thematic classification compared to the purely spectral classifications in which the image formed by the principle components and the NDVI index was used.

Classes	**Sup. class. image principle components**		**Sup. class. image principle components and NDVI**		**Expert classification algorithm**	
	Pp	**Pu**	**Pp**	**Pu**	**Pp**	**Pu**
	%					
Bare soil	70.3	87.6	93.5	92.3	94.2	95.3
Cereal	62.3	81.3	78.4	95.8	86.9	95.2
Removed Cover	47.8	68.7	33.3	99.0	72.7	100.0
High protein crops	94.6	66.6	100.0	42.8	100.0	78.1
Alfalfa	98.5	25.0	97.6	47.3	89.2	67.7
Woodlands and scrublands	87.5	50.0	100.0	87.5	100.0	86.6
Total accuracy, %	76.6		85.2		91.9	
Kappa statistic, %	68.5		77.3		87.6	

Sup. class.: supervised classification; NDVI: normalized difference vegetation index; Pp: producer accuracy; Pu: user accuracy.

Table 2. 3 Producer and user accuracies, overall and Kappa statistic for supervised classifications and algorithm of expert classification.

The results presented in Table 2.3 show the feasibility of the methodology employed in the designation of the expert classification algorithm, having improved the user and producer accuracy with all the classes except woodlands and scrublands and with alfalfa in comparison to purely spectral classifications. The category woodlands and scrublands presents a user accuracy of 86.6% owing to the similarity of the spectral response to the category removed cover. The category alfalfa presents confusion with the category high protein crops for aforementioned reasons. The levels of error for the remaining classes were relatively low in the test phase, which indicates a good capacity of generalization of the expert system. The highest reached for user accuracy was for removed cover with 100%, while the lowest value was for alfalfa with 67.7%. In terms of the producer accuracy, the categories of woodlands and scrublands and high protein crops reached a value of 100%, while the category removed cover presented the lowest value with 72.7% owing to the spectral similarity to woodlands and scrublands. The Kappa coefficient had a value of 87.6% and total accuracy was 91.9%.

The accuracy values obtained with the object-oriented classification and with the expert classification algorithm were similar to and/or higher than the values obtained by other authors, which shows that the methodology is adequate for the classification of land uses.

In the province of Milan, Italy, Marchesi et al. (2006) carried out a classification oriented to objects in QuickBird images of the following categories: vegetation in urban areas, sports infrastructures, highways and water, obtaining a total accuracy of 83% and a Kappa statistic of 79%. These values are lower than those obtained in the object oriented classification of this work.

In northern California, Yu et al. (2006) obtained a total accuracy of 60% in an object-oriented classification of digital aerial images captured by the Digital Airborne Imaging System in which a total of 52 vegetal uses were classified. This article also showed some results lower than those obtained in the present work, although it should be noted that a larger number of uses was classified.

In southern Mexico, Mas (2005) made a classification using expert classification of the following categories: jungle, mangrove swamp, agricultural crops, water and urban areas. A single image was employed, which included five bands of the Landsat Enhanced Thematic Mapper (ETM), obtaining a total precision of 67%.

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.

Triñanes et al. (1994) in Pontevedra, Spain, classified dense urban zones, non-dense urban zones, scrublands-grassland, water, forest and non-forest vegetal zones, using neuronal networks in images from the TM sensor of Landsat-5, obtaining an accuracy of 91.32%. This percentage is lower than that obtained with the expert classification developed in this work.

6. Conclusions

The statistical results show the validity of the methodology employed to design the expert classification algorithm. A total accuracy of 91.9% and an excellent Kappa statistic (87.6%) were obtained with this algorithm. The producer accuracy and user accuracy have been improved in all the classes except in woodlands and scrublands and alfalfa. This algorithm will facilitate updating the databases of agricultural crops, thus reducing the need for field visits. In relation to object-oriented classification, it has been shown to be of great help in isolating classes that are confused with others that have similar spectral responses (high protein crops and alfalfa).

7. Aknowledgements

This work was possible thanks to the images and field information provided by the Fondo Andaluz de Garantía Agraria (FAGA) through the public enterprise for Agricultural and Fisheries Development (DAP), Spain. We also thank Carmen Andú for her important collaboration.

8. References

Abkar, A.A, Sharifi, M.A. and N.J. Mulder. 2000. Likelihood-based image segmentation and classification: a framework for the integration of expert knowledge in image classification procedures. International Journal of Applied Earth Observation and Geoinformation; 2:104-119.

Aitkenhead, M.J., and G.G. Wright. 2004. *Mapping land use in NE Scotland with neural networks from remote sensing imagery.* In Remote Sensing and Photogrammetry Society Annual Conference, Aberdeen. 6-7 September 2004. The Macaulay Institute, Craigiebuckler, UK Workshop Output. Available at http://www.macaulay.ac.uk/workshop/remotesensing2004/MA_Full_paper.pdf.

Ayala, R.M. and M. Menenti. 2002. *Alternativas a los problemas presentados en un proceso de clasificación basado en el reconocimiento espectral de patrones*. Mapping; 75:72-76.

Chuvieco, E. 2000. *Fundamentos de teledetección espacial*. 3rd ed. p. 558. Madrid (Spain): Ed. Rialp.

Gibson, P.J. and C.H. Power. 2000. *Introductory remote sensing – Digital image processing and applications*. p. 249. London (UK): Routledge.

Haboudane, D., J.R. Miller, E. Pattei, P.J. Zarco-Tejada, and I. Strachan. 2004. *Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the* *context of precision agriculture*. Remote Sensing of Environment; 90:337-352.

Heyman, O. 2003. *Automatic extraction of natural objects form 1-m remote sensing images*. Oregon State University, Department of Geosciences, Corvallis, Oregon, USA. Available at http://www.cobblestoneconcepts.com/ucgis2summer/heyman/heyman.htm

Johnsson, K. 1994. *Segmented-based land-use classification from SPOT satellite data*. Photogrammetric Engineering and Remote Sensing; 60(1):47-53.

Lidov, L., R. Miller, D.M. Wormer, and K.A. Tilley. 2000. *Interview: Understanding the future of commercial remote sensing.* Photogrammetric Engineering and Remote Sensing; 66(1):5-14.

Marchesi A., R. Colombo and P. Valentini. 2006. *Application of high spatial resolution satellite imagery for urban environment mapping*. In Proceeding from ISPRS 1st International Conference on Object-based Image Analysis. July 2006. International Society for Photogrammetry and Remote Sensing (ISPRS), Salzburg, Austria.

Marini, M.F. 2006. *Influencia del cálculo de la reflectancia aparente en la obtención del índice verde (NDVI) y en su expresión cartográfica*. XXIII Reunión Científica de la Asociación Argentina de Geofísicos y Geodestas, Buenos Aires, Argentina. 14-18 agosto 2006. Asociación Argentina de Geofísicos y Geodestas, Buenos Aires, Argentina.

Mas, J.F. 2005. *Un método para combinar datos espectrales e información auxiliar en una red artificial neuronal*. p. 3543-3549. In Anais XII Simposio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil. 16-21 abril 2005. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo, Brasil.

Stefanov, W.L., M.S. Ramsey, and P.R. Christensen. 2001. *Monitoring urban land cover change: An expert system approach to* *land cover classification of semiarid to arid urban centers*. Remote Sensing of Environment; 77:173-185.

Triñanes, J.A., J. Torres, A. Tobar, and C. Hernández. 1994. *Clasificación de imágenes multiespectrales mediante redes neuronales*. Revista de Teledetección; 3:1-5.

Trotter, C.M. 1991. *Remotely-sensed data as an information source for geographical information systems in natural resource management: a review*. International Journal of Geographic Information System; 5:225-239.

Wicks, T.E., G.M. Smith, and P.J. Curran. 2002. *Polygon-based aggregation of remotely sensed data for regional ecological analyses*. International Journal of Applied Earth Observation and Geoinformation; 4:161-173.

Wilkinson, G.G., I. Kanellopoulos, C. Kontoes, and R. Schoenmakers. 1991. *Advances in the automatic processing of satellite images*. p. 125-132. In Toselli, F., and J. Meyer (eds.) Conference on the Application of Remote Sensing to Agricultural Statistics, Belgirate, Italy. 26-27 November. Office for Publications of the European Commission, Luxembourg (Luxembourg).

Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly, and D. Schirokauer. 2006. *Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery*. Photogrammetric Engineering and Remote Sensing; 72:799-811.

Capítulo IV. Land-cover classification with an expert classification algorithm using digital aerial photographs

Autores: Perea, A.J.[a], Meroño, J.E.[b], Aguilera, M.J.[a], de la Cruz, J.L.[a]

[a] Departamento de Física Aplicada, Universidad de Córdoba (UCO), Córdoba, Spain

[b] Departamento de Ingeniería Gráfica y Geomática, Universidad de Córdoba (UCO), Córdoba, Spain

Publicado en: South African Journal of Science, 106, (5-6), 82-87.

«Una región de la corteza cerebral no solo aprende secuencias conocidas, sino también cómo modificar sus clasificaciones»

Jeff Hawkins (2004)

1. Abstract

The purpose of this study was to evaluate the usefulness of the spectral information of digital aerial sensors in determining land-cover classification using new digital techniques. The land covers that have been evaluated are the following, (1) bare soil, (2) cereals, including maize (Zea mays L.), oats (Avena sativa L.), rye (Secale cereale L.), wheat (Triticum aestivum L.) and barley (Hordeun vulgare L.), (3) high protein crops, such as peas (Pisum sativum L.) and beans (Vicia faba L.), (4) alfalfa (Medicago sativa L.), (5) woodlands and scrublands, including holly oak (Quercus ilex L.) and common retama (Retama sphaerocarpa L.), (6) urban soil, (7) olive groves (Olea europaea L.) and (8) burnt crop stubble. The best result was obtained using an expert classification algorithm, achieving a reliability rate of 95%. This result showed that the images of digital airborne sensors hold considerable promise for the future in the field of digital classifications because these images contain valuable information that takes advantage of the geometric viewpoint. Moreover, new classification techniques reduce problems encountered using high-resolution images; while reliabilities are achieved that are better than those achieved with traditional methods.

2. Introduction

In recent years, the development of remote sensing technologies has increased exponentially. Until recently, high-resolution satellites could only obtain images to a size of 5 meters spatial resolution. Nowadays, these technologies have been improved. Data obtained from this sort of sensors have generated a large amount of environmental information (Moreira, 2005). Extracting useful information from high-resolution satellite imagery is a major technical problem of remote sensing, however, as the data obtained are difficult to use because the spectral information contained in pixels is not sufficient, in the majority of cases, to identify vegetation species or the types of surface cover. Pixels normally include a radiometric mixture from their neighbours and consequently few zones have total homogeneity (Wilkinson et al., 1991).

Currently, process improvements have enabled digital photogrammetry based on aerial photography to generate geometrically corrected products compatible with conventional mapping detail. They are able to provide decisions or potential territorial element analysis of natural resources surpassing those available from satellites. The production of digital orthophotos is an ideal complement to environmental assessment processes and spatial planning that heretofore made use only of satellite imagery (Moreira, 2005). Digital orthophotos constitute a basic tool in the task of managing the environment and they are also a basis of reference in spatial plans (Ayala and Menenti, 2002).

The launch of photogrammetry using digital cameras has made available multispectral information concerning large areas of territory. This information is being used solely from the geometric point of view, because there are no algorithms and models to exploit infrared information captured simultaneously with colour information. There is currently great interest in the development of new classification algorithms in the area of the digital treatment of images (Abkar, Sharifi and Mulder, 2000). The combination of spectral data with other sources of auxiliary data allows the use of more information to improve classifications (Wicks, Smith and Curran, 2002).

In recent years, and probably due to the availability of more powerful software, some researchers have reported that the segmentation techniques used in classifications reduce the local variation caused by textures, shadows and shape (Yu et al., 2006; Hay, Marceau and Bouchard, 2003). Object-based classification may be a good alternative to the traditional pixel-based methods. To overcome the H-resolution problem and the salt-and-pepper effect, it is useful to analyse groups of contiguous pixels as objects instead of using the conventional pixel-based classification unit.[6]

Expert systems use data other than spectral characteristics to improve the results of classification. The use of auxiliary information to increase the accuracy of digital classification involves combining an existing knowledge base with information extracted from images (Trotter, 1991). To improve automatic classification procedures, it is necessary to introduce a set of parameters to inform the classification beyond the digital values of the pixels (Heyman, 2003). With the use of auxiliary data, the initial results of the procedures can be corrected through knowledge-based rules (Wicks, Smith and Curran, 2002).

2.1. New techniques for classification

In high-resolution images from satellites or aerial digital cameras (UltracamD, DMC, ADS-40, etc.), each pixel does not refer to an object, character or area as a whole, but to a portion of some components, which limits the classic techniques of pixel-based classification (Sánchez, 2003). Similarly, the great detail in digital images obtained from airborne sensors can lead to excessive variability within an area that has the same coverage, associated with decreased separability of different types of coverages.

Alternative approaches to classification techniques involve the object-oriented analysis of images, which takes into account, inter alia, the shapes, textures, background information and spectral information in the image. Recent studies have demonstrated the superiority of the new concept of traditional classifiers (Leuker, Darwish and Reinhardt, 2003; Tansey et al., 2008; Geneletti and Gorte, 2003; Perea, Meroño and Aguilera, 2009). Its basic principle is to make use of important information (shape, texture, background information) that is only present in significant image objects and their mutual relations. This type of classification is called ‘object-oriented classification’ and requires a prior segmentation, defined as the search for homogeneous regions in an image and then classifying these regions (Mather, 1999). Software called eCognition® is available that allows segmentation and classification according to this concept. The influence the described parameters have on the segmentation is flexible and can be specified by the user through the manipulation of different parameters based on colour and shape (compactness and smoothness) factors (Flanders, Hall-Beyer, Pereverzoff, 2003). The second step is the classification of these regions based on examples (by nearest neighbourhood algorithm) or membership functions, allowing users to develop an expert knowledge base (based on fuzzy logic) and to assign regions to certain classes (Flanders, Hall-Beyer, Pereverzoff, 2003).

Another current trend is to develop algorithms that improve the classifications based solely on the reflectance of the pixels. It should be noted, however, that the neighbouring pixel radiometric mixture prevents the extraction of homogeneous regions of interest (Flanders, Hall-Beyer, Pereverzoff, 2003).

Gong and Howarth (1990) argue that it is important to recognise that conventional classifiers (maximum likelihood classifier, minimum distance classifier) do not recognise the spatial patterns in the same way as the human performer. An expert system was therefore developed to incorporate data other than the spectral features to improve the outcome of the purely spectral classification.

This work aims to evaluate the utility of spectral information from these photogrammetric sensors in determining land covers.

3. Material and Methods

The area of study was located in the Pedroches Valley of Cordoba Province, Spain (Figure 3.1) and includes the municipality of Hinojosa del Duque (38°23′ N – 38°33′ N; 5°16′ W – 5°50′ W). This rectangular area of 16 km × 20 km, covering 32 000 ha, is representative of Andalusian dryland crops and has a typical continental Mediterranean climate, characterised by long dry summers and mild winters.

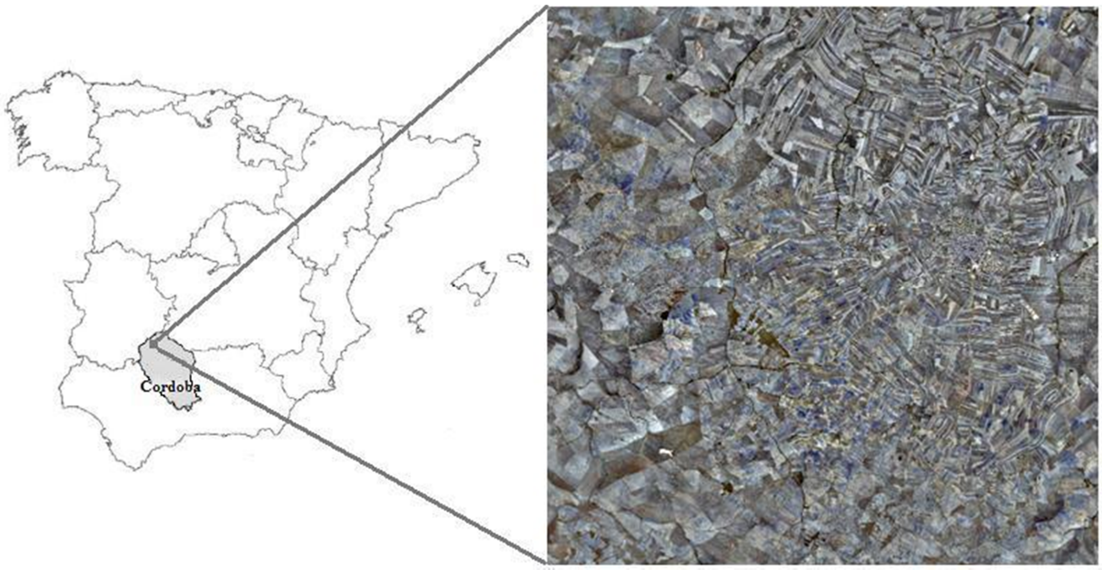

Figure 3. 1 A map showing the study area in Spain.

To carry out the study, 64 frames were captured by the sensor of Vexcel UltracamD photogrammetric on 23 May 2006, with dimensions of 7500 × 11 500 pixels and encoded in 8 bits. The frames had a spatial resolution of approximately 0.5 m and consisted of infrared, red, green and blue bands. These frames were orthorectified and referred to European Datum 1950 on the International Ellipsoid.

To develop this work, information was used from field visits by the Public Enterprise for Agricultural and Fisheries Development. Land covers evaluated included, (1) bare soil, (2) cereals, including maize (Zea mays L.), oats (Avena sativa L.), rye (Secale cereale L.), wheat (Triticum aestivum L.) and barley (Hordeun vulgare L .), (3) high protein crops, such as peas (Pisum sativum L.) and beans (Vicia faba L.), (4) alfalfa (Medicago sativa L.), (5) woodlands and scrublands, mainly holly oak (Quercus ilex L.) and common retama (Retama sphaerocarpa L.), (6) urban soil, (7) olive groves (Olea europaea L.) and (8) burnt crop stubble.

To perform the supervised classification and expert classification algorithm, the Erdas Imagine 9.0® system (Leica Geosystems Geospatial Imaging, Norcross, Georgia, USA) was used. In the case of object-oriented classification, the eCognition Professional 5.0® software (Definiens, München, Germany) was used.

The methodology began with the calculation of principle components and then calculated the normalised difference vegetation index (NDVI). Images were obtained with the desired combination of bands and classifications made. Finally, the results of the classifications were validated.

3.1. Obtaining the principle components

The objective of ‘principle component analysis’ (PCA) is to summarise a wide group of variables in a new and smaller set, without losing a significant part of the original information (Chuvieco, 2000). For the final user of distance imaging products, the goal of PCA is to construct images in order to increase their capacity to differentiate types of covers.

3.2. Obtaining the NDVI

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

The NDVI was obtained so as to highlight the different spectral behaviours 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:

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

[Eqn 1]

where R_{NIR} and R_{RED}, are reflectance in the near infrared band (R800 nm) and the red band (R690 nm), respectively.

3.3. Supervised classification

Starting from different combinations of bands (Table 3.1), a series of images was obtained. Next, a supervised classification was made from all these images.

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 analytical parameters and because of the variability of the classes using a covariance matrix.

Type of sensor	Image	Number of bands
UltracamD	Principle components	4
	Principle components + NDVI	5

Table 3. 1 Images used in supervised classification.

3.4. Object-oriented classification

As noted above, the particularity of this type of analysis is that the classification is based on objects rather than pixels. Being the image formed by pixels, the first step in object-oriented analysis is to group adjacent pixels through region-growing techniques, in order to classify objects subsequently extracted. In this way, the number of parameters that can be valued greatly increases, allowing criteria such as size, shape, colour stockings, highs and lows, proximity to other objects and texture. At the same time, segmentation reduces the number of objects to classify, so the processing time decreases.

The stopping criterion in the process of merging regions occurs through the so-called scale parameter, which can be defined by the user in relation to the maximum global heterogeneity of the segments. The larger the scale parameters for a database, the bigger the objects in the image and, since the scale parameter can be changed, different types of segmented images can be obtained. Thus, the generated objects in a coarser segmentation inherit information from smaller objects generated with finer scale parameters. Subsequently, the rankings are trained using the same plots of training, and validated using the same validation plots used in previous classifications.

The output of the segmentation process depends on specifications and weighting of input data and controlling parameters such as scale (control size parameter), color (spectral information) and shape (smoothness and compactness information) of the resulting image objects. The option ‘multiresolution segmentation’ was used, which performs automatic extraction of homogeneous objects. The scale parameter is an abstract term that determines the maximum allowed heterogeneity for the resulting image objects. Color parameter and shape parameter (smoothness and compactness) define the percentage that the spectral values and the shape of objects, respectively, will contribute to the homogeneity criterion. Finally the values of 211, 0.9, 0.1, 0.5 and 0.5 were defined for scale, color, shape, smoothness and compactness. For most cases, colour is the most important and has the greatest weight in the definition of objects.

The nearest-neighbour algorithm was used for the classification: some samples were chosen (training area) for each of the classes. The rest of the scene was then classified accordingly. This is a very rapid and simple method, adequate when the classification of an object requires many bands/criteria. It also takes into account different parameters related to the objects (area, longitude, mean colour, brightness, and texture).

3.5. Expert classification algorithm

The expert classification algorithm used in this work consisted of assigning the classes that made up the legend, based on the area of coincidence among different types of images that had been classified previously. To do this, the following information was necessary: an image created based on a field visit and the map of land cover and vegetal cover for Andalusia for 2005, used as the true terrain. The ERDAS Imagine 9.0® system and the supervised classifications were based on the image formed by the principle components, the image formed by the principle components and NDVI, as well as the object-oriented classification.

This algorithm was designed with the following decision-making criteria or rules, (1) when the pixels of each class of the classified image of the principle components and NDVI coincided with the image classified from principle components, they were assigned to this class and (2) in the case of the other pixels, where there was no coincidence, they were assigned by the object-oriented classification. To evaluate the quality of classifications, a total of 75 000 verification points were taken (approximately 2% of the area) for those that provided both real cover (true terrain) and for those obtained by classification.

The overall accuracy, kappa statistic and the producer's and user's accuracy were calculated for each one of the classifications. 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; it quantifies how much better a particular classification is in comparison with a random classification. Some authors have suggested the use of a subjective scale where kappa values < 40% are poor, 40% – 55% fair, 55% – 70% good, 70% – 85% very good and > 85% excellent (Monserud and Leemans, 1992).

For individual classes, two accuracies can be calculated, (1) the producer's accuracy is a measure of omission error and indicates the percentage of pixels of a given land-cover type that are correctly classified and (2) the user's accuracy is a measure of the commission error and indicates the probability that a pixel classified into a given class actually represents that class on the ground.

4. Results and Discussion

4.1. Results of the object-oriented classification

The result of segmentation is a new image that divides the original image into regions such that the pixels included in each of them are similar. After the process of segmentation, a new image was obtained and divided into 13 243 regions that were later classified (Figure 3.2).

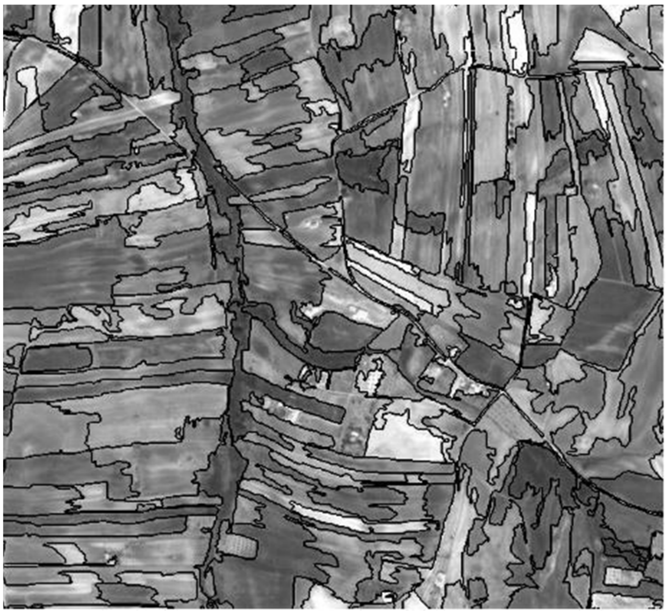

Figure 3. 2 Example of segmentation of the digital aerial photography at the scale of 211.

The accuracy assessment of this classification was measured using randomly selected points for which land cover was determined with an orthophoto mosaic that was geo-referenced to the image. Table 3.2 shows the accuracy of classification in the digital aerial classification according to its boundary analysis.

Category	Supervised classification				Object-oriented classification		Expert classification algorithm	
	Image principle components		Image principle components and NDVI					
	Pa (%)	Ua (%)	Pa (%)	Ua (%)	Pa (%)	Ua (%)	Pa (%)	Ua (%)
Bare soil	83.3	95.2	84.8	92.3	91.3	95.5	100.0	92.9
Cereal	90.7	86.7	91.9	89.1	90.0	96.4	87.4	95.3
Burnt cover stubble	75.0	100.0	77.1	100.0	100.0	83.3	95.0	100.0
High protein crops	98.7	57.1	99.1	63.6	98.3	81.6	100.0	68.7
Alfalfa	100.0	25.0	80.0	44.4	78.3	74.5	95.4	57.1
Woodlands and scrublands	85.7	85.7	97.7	80.0	95.4	83.3	100.0	96.6
Urban soil	89.5	86.5	91.0	86.8	100.0	100.0	100.0	100.0
Olive groves	85.4	82.3	93.8	96.8	100.0	100.0	100.0	100.0
Overall accuracy (%)	83.8		87.8		91.7		95.0	
Kappa statistic (%)	74.5		78.4		87.5		91.1	

NDVI, normalised difference vegetation index; Pa, producer's accuracy; Ua, user's accuracy.

Table 3. 2 Producer's and user's accuracy, overall and Kappa statistic for supervised classifications, object-oriented classification and expert classification algorithm.

The improvement achieved by the introduction of textural and contextual features was significant for all classes with respect to the pixel-based analysis. For some classes, the producer's and user's accuracy reached a value of 100% (e.g. for 'urban soil' and 'olive groves'). For others, the producer's and user's accuracy increased but remained low, for example, in the case of 'woodlands and scrublands'.

The highest producer's accuracies were for the 'burnt crop stubble', 'urban soil' and 'olive groves' categories, all with the value of 100%. In contrast, the lowest value was for 'alfalfa' (78.3%), because of its spectral similarity to 'high protein crops'. Referring to the user's accuracy, the best results were achieved for the categories 'urban soil' (100%) and 'olive groves' (100%) and, as in the case of the producer's accuracy, the lowest value was for the category 'alfalfa' (74.5%), due to misclassification of 'high protein crops' during image classification.

The overall accuracy and kappa statistics were excellent, reaching values of 91.7% and 87.5%, respectively. In addition, the object-oriented method significantly narrowed down the variation of class-based accuracies compared with the result of the pixel-based classification method.

A map obtained from the object-oriented classification is presented in Figure 3.3.

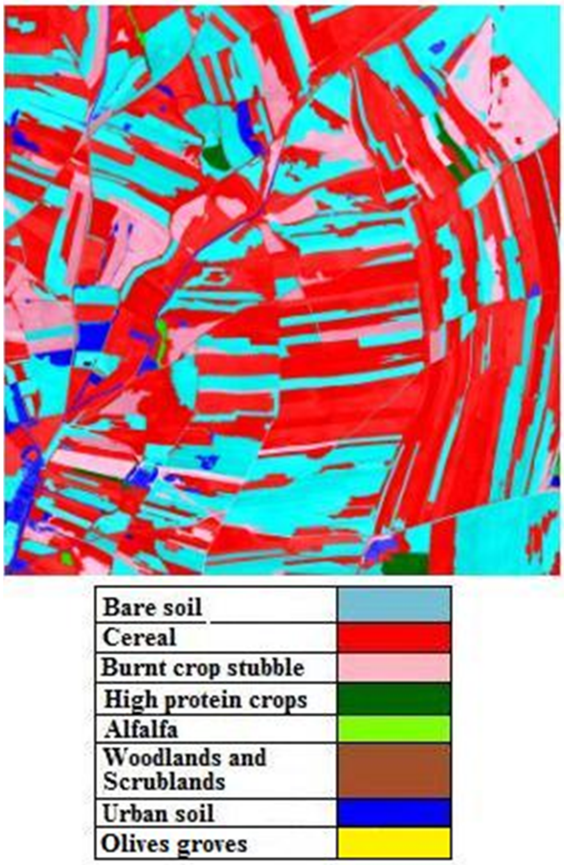

Figure 3. 3 Example of object-oriented classification.

4.2. Expert classification algorithm

The accuracy of the expert classification algorithm was higher than the pixel-based classification. Both the overall accuracy and kappa coefficient were significantly higher and the producer's and user's accuracy also gave better results in the expert classification algorithm.

The results of the expert classification algorithm (Table 3.2) showed a marked improvement in the reliability of both producer and user in most categories, when comparing them with purely spectral classifications. Besides, this algorithm achieved some accuracy rates and kappa statistics that were above 90%. The producer's accuracy increased in all cases, except in those of 'cereal' and 'alfalfa', but was nevertheless above 87%. The category 'alfalfa' was confused with the category 'high protein crops' for the reason already mentioned. The user's accuracy increased in all categories except that of 'bare soil'

(92.9%). The overall accuracy was 95% and the kappa statistic had a value of 91%, indicating strong agreement between the classification map and the ground reference information.

The accuracy values obtained with object-oriented classification and with the expert classification algorithm in digital aerial photography were similar to, and/or higher than, the values obtained by other authors using satellite images. The methodology is therefore adequate for the classification of land covers, which is presented in Figure 3.4 as a comparison between supervised classifications and the expert classification algorithm.

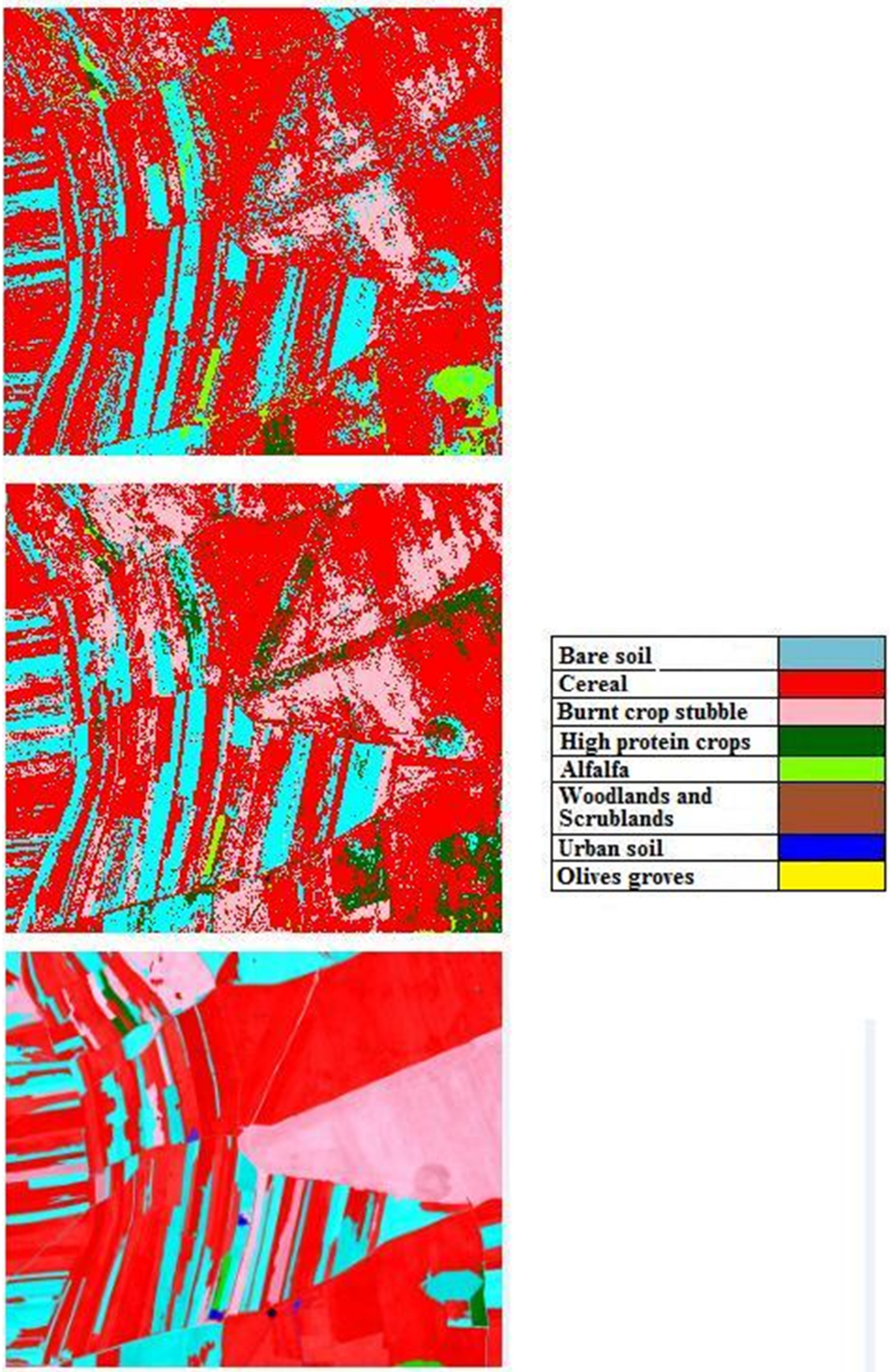

Figure 3. 4 Example of comparison between, (a) supervised classification of the image formed by the principle components, (b) supervised classification of the image formed by the principle component and the NDVI index and (c) the classification using the expert algorithm.

In the southern Baltic sea, Janas, Urbański and Mazur (2009) used object-oriented classification methods to classify seagrass landscape, composed of meadows, beds and patches/gaps, obtaining a total precision of 83%.

On the Gulf coast of Texas, Green and Lopez (2007) classified bivalve reef, sea grass, land, mangroves, emergent marsh, unconsolidated sediments and unknown benthic habitat using object-oriented classification in images from the ADS40 aerial sensor, obtaining an accuracy of 90%, lower than that obtained with the expert classification developed in this work.

In the Three Gorge area of Chongqin in China, Zhang et al (2008) made a classification using expert classification of 17 categories. SPOT5 XS and Pan data were acquired between 2004 and 2006 for cloud-free images, with two scenes of different seasons for each area being selected for vegetation detection, attaining a total precision of 86%, again lower than that obtained in the present work (although it should be noted that a larger number of uses were classified.)

5. Conclusions

The results obtained in the different classifications of digital aerial photographs show that the photographs from digital aerial sensors can be used in tasks that previously were only specific to satellite images, offering the ability to discriminate land cover with great precision. Moreover, the new classification techniques represent a breakthrough in agricultural field controls, as the quality of the results of digital aerial photography, together with the development of the new techniques described, allows the control and monitoring of various agricultural areas without making field visits. The combination of bands, which provides a better result in the supervised classification, is the image formed by the principle components and NDVI. Finally, it is noteworthy that the use of object-oriented classification and the expert classification algorithm yielded the best results, greatly reducing the problems associated with the use of high-resolution images, such as the salt-and-pepper effect. The best result was obtained with the expert classification algorithm, achieving a kappa index and a confidence rating of 91.07% and 95%, respectively.

6. References

Abkar, A.A, Sharifi, M.A. and N.J. Mulder. 2000. Likelihood-based image segmentation and classification: a framework for the integration of expert knowledge in image classification procedures. International Journal of Applied Earth Observation and Geoinformation; 2:104-119.

Ayala, R.M. and M. Menenti. 2002. *Alternativas a los problemas presentados en un proceso de clasificación basado en el reconocimiento espectral de patrones*. Mapping; 75:72-76.

Chuvieco, E. 2000. *Fundamentos de teledetección espacial*. 3rd ed. 558 p. Madrid (Spain): Ed. Rialp.

Flanders D, Hall-Beyer M. and J. Pereverzoff. 2003. *Preliminary evaluation of eCognition object based software for cut block delineation and feature extraction*. Canadian Journal of Remote Sensing; 29:441–452.

Geneletti D. and B.G.H Gorte. 2003. *A method for object-oriented land cover classification combining Landsat TM data and aerial photographs*. International Journal of Remote Sensing; 24:1273–1286.

Gong H. and P.J. Howarth. 1990. *An assessment of some factors influencing multispectral land-cover classification*. Photogrammetric Engineering and Remote Sensing; 56:597–603.

Green K. and C. Lopez. 2007. *Using object-oriented classification of ADS40 to map benthic habitats of the state of Texas*. Photogrammetric Engineering and Remote Sensing; 73:861–865.

Haboudane D., Miller J.R., Pattey E., Zarco-Tejada P.J. and I. Strachan. 2004. *Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture*. Remote Sensing of Environment; 90:337–352.

Hay G.J., Blaschke T., Marceau D.J. and A. Bouchard. 2003. *A comparison of three image-object methods for the multiscale analysis of landscape structure*. Photogrammetric Engineering and Remote Sensing; 57:327–345.

Heyman O. 2003. *Automatic extraction of natural objects from 1-m remote sensing images*. Available from: http://www.cobblestoneconcepts.com/ucgis2summer/heyman/heyman.htm

Janas U., Urbański J. and A. Mazur. 2009. *Object-oriented classification of QuickBird data for mapping seagrass spatial structure*. Oceanological and Hydrobiological Studies; 38:27–43.

Leukert K., Darwish A. and W. Reinhardt. 2003. *Urban land-cover classification: An object-based perspective*. Paper presented at: URBAN 2003. Proceedings of the 2nd Joint Workshop on Remote Sensing and Data Fusion over Urban Areas. 2003 May 22–23; Berlin, Germany. Berlin: IEEE Geoscience and Remote Sensing Society; p. 278–282.

Marini M.F. 2006. *Influence of calculating the apparent reflectance in obtaining the green index (NDVI) and its cartographic expression*. Paper presented at: XXIII Reunión Científica de la Asociación Argentina de Geofísicos y Geodestas; 2006 August 14–18; Bahía Blanca, Argentina.

Mather P. 1999. *Computer processing of remotely-sensed images: An introduction*. Chichester (UK): John Wiley and Sons.

Monserud R.A. and R. Leemans. 1992. *Comparing global vegetation maps with the Kappa statistic*. Ecological Modelling; 62:275–293.

Moreira J.M. 2005. *Digital orthophotos of Andalusia, an important environmental value. Journal of Environment*. Seville (Spain): Andalusian Departament of Environment; 49:35-37.

Perea A.J., Meroño J.E. and M.J. Aguilera. 2009. *Oriented-based classification in aerial digital photography for land-use discrimination*. Interciencia; 34:612–616.

Sánchez N. 2003. *Current overview of mixed techniques of image classification using spectral and texture segmentation. Application to high spatial resolution images*. Mapping; 88:32–37.

Tansey K., Chambers I., Anstee A., Denniss A. and A. Lamb. 2008. *Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas*. Applied Geography; 29:145–157.

Trotter C.M. 1991. *Remotely-sensed data as an information source for geographical information systems in natural resource management: a review*. International Journal of Geographical Information System; 5:225–239.

Wicks T.E., Smith G.M. and P.J. Curran. 2002. *Polygon-based aggregation of remotely sensed data for regional ecological analyses*. International Journal of Applied Earth Observation and Geoinformation; 4:161–173.

Wilkinson G.G., Kanellopoulos I., Kontoes C. and R. Schoenmakers. 1991. *Advances in the automatic processing of satellite images*. Paper presented at: Conference on the Application of Remote Sensing to Agricultural Statics; 1991 Nov 26–27; Villa Carlotta, Belgirate, Lake Maggiore, Italy. Luxembourg: Office for Publications of the European Commission; p. 125–132.

Yu Q., Gong P., Clinton N., Biging G., Kelly M and D. Schirokauer. 2006. *Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery*. Photogrammetric Engineering of Remote Sensing; 72(7):799–811.

Zhang L, Yueming Z and W. Bingfang. 2008. *Expert system based on object-oriented approach for land cover mapping*. Paper presented at: ISPRS Congress Beijing 2008, Proceedings of Commission VII; 2008 July 3–11; Beijing, China. Nschede: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences; Vol. 37, Part B7. p. 679–684.

Capítulo V. Application of Numenta® Hierarchical Temporal Memory for land-use classification

Autores: Perea, A.J.[a], Meroño, J.E.[b], Aguilera, M.J.[a]

[a] Departamento de Física Aplicada, Universidad de Córdoba (UCO), Córdoba, Spain

[b] Departamento de Ingeniería Gráfica y Geomática, Universidad de Córdoba (UCO), Córdoba, Spain

Publicado en: South African Journal of science, 105, (9-10), 370-375.

«No nacemos con el conocimiento del lenguaje, las casas o la música. La corteza cerebral posee un algoritmo de aprendizaje inteligente que descubre y capta de forma natural cualquier estrcutura jerárquica que exista»

Jeff Hawkins (2004)

1. Abstract

The aim of this paper is the application of Memory-Prediction Theory, implemented in the form of a Hierarchical Temporal Memory (HTM), for land use classification. Numenta's HTM is a new computing technology that replicates the structure and function of the human neocortex. In this project a photogram, received by a photogrammetric UltracamD® sensor of Vexcel, and data on 1 513 plots in Manzanilla (Huelva, Spain) were used to validate the classification, achieving an overall classification accuracy of 90.4%. The HTM approach appears to hold promise for land uses classification.

Key words: Memory-Prediction Theory, Nupic, Ultracamd sensor, Hierarchical Temporal Memory.

2. Introduction

Vision is the primary sensory modality for humans and most mammals to perceive the world. In humans, vision-related areas occupy about 30 per-cent of the neocortex (Douglas and Martin, 2004). Light rays are projected upon the retina, and the brain tries to make sense of the world by means of interpreting the visual input pattern. The sensitivity and specifity with which the brain solves this computationally complex problem cannot yet be replicated on a computer. The most imposing of these problems is that of invariant visual pattern recognition.

Recently it has been said that the prediction of future sensory input from salient features of current input is the keystone of intelligence (Hawkins and Blakeslee, 2004). The neocortex is the structure in the brain which is assumed to be responsible for the evolution of intelligence. Current sensory input patterns activate stored traces of previous inputs which then generate top-down expectations, which are verified against the bottom-up input signals. If the verification succeeds, the predicted pattern is recognised. This theory explains how humans, and mammals in general, can recognize images despite changes in location, size and lighting conditions, and in the presence of deformations and large amounts of noise. 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 (Hawkins and George, 2007); the model is an attempt to replicate the structural and algorithmic properties of the neocortex (Hawkins and George, 2007). Spatial and temporal relations between features of the sensory signals are formed in a hierarchical memory architecture during a learning process. When a new pattern arrives, the recognition process can be viewed as choosing the stored representation that best predicts the pattern. HTMs have been successfully applied to the recognition of relatively simple images (George and Jaros, 2007), showing invariance across several transformations and robustness with respect to noisy patterns.

We have applied the concept of HTM as implemented by Numenta to land use recognition, building and testing a system that learned to recognize five different types of land use.

2.1. Overview of the HTM learning algorithm

HTMs can be considered a form of Bayesian network where the network consists of a collection of nodes arranged in a tree-shaped hierarchy (George and Jaros, 2007). Each node in the hierarchy self-discovers a set of causes in its input through a process of finding common spatial patterns and then detecting common temporal patterns (George and Jaros, 2007). Unlike many Bayesian networks, HTMs are self-training, have a well-defined parent/child relationship between each node, inherently handle time-varying data, and afford mechanisms for covert attention. Sensory data are presented at the “bottom” of the hierarchy. To train an HTM, it is necessary to present continuous, time varying, sensory inputs while the causes underlying the same sensory data persist in the environment. In other words, you either move the senses of the HTM through the world, or the objects in the world move relative to the HTM's senses. Time is the fundamental component of an HTM, and can be thought of as a learning supervisor. HTM networks are made of nodes; each node receives as input a temporal sequence of patterns. The goal of each node is to group input patterns that are likely to have the same cause, thereby forming invariant representations of extrinsic causes.

An HTM node uses two grouping mechanisms to form invariants. The first one is called spatial pooling, which receive raw data from the sensor; spatial poolers of higher nodes receive the outputs from their child nodes. The input of the spatial pooler in higher layers is the fixed-order concatenation of the output of its children. This input is represented by row vectors, and the role of the spatial pooler is to build a matrix (the coincidence matrix) from input vectors that occur frequently. There are multiple spatial pooler algorithms, e.g. Gaussian and 'Product'. The Gaussian spatial pooler algorithm is used for nodes at the input layer, whereas the nodes higher up the hierarchy use the 'Product' spatial pooler. The Gaussian spatial pooler algorithm compares the raw input vectors to the existing coincidences in the coincidence matrix. If the Euclidean distance between an input vector and an existing coincidence is small enough, the input is considered to be the same coincidence, and the count for that coincidence is incremented and stored in memory.

The 'Product' spatial pooler is always part of a node higher up the hierarchy, and receives the concatenation of the outputs of its child nodes. This vector is divided up into N portions, which is the number of children of the node. The 'Product' spatial pooler sets the highest value in each of these N distributions to 1, while the other values are set to 0. These new vectors are stored in the coincidence matrix, and the counts of the coincidences that already exist are incremented.

The second mechanism is called temporal pooling, which groups together patterns that are temporally close. This way, patterns that are very different, but that have a common cause, can be in the same group.

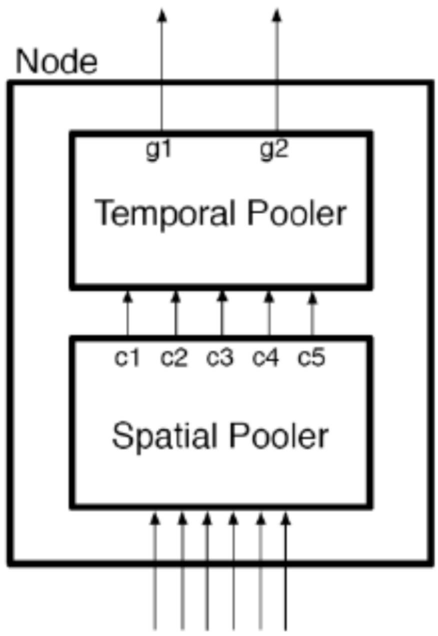

Figure 4. 1 HTM node structure (George and Jaros 2007, Figure 6).

Both the spatial and temporal poolers switch from learning to inference mode at some point. In the case of the spatial pooler, its output is a vector of length equal to the number of patterns pooled by the node, and the ith position in this vector corresponds to the ith pattern inside this spatial pooler. This output is a probability distribution of the similarity between the input pattern and the stored patterns, measured in terms of Euclidean distances. An assumption commonly made by the designers of HTM is that the probability that a pattern is closest to another pattern falls off as a Gaussian function of the Euclidean distance, therefore it can be calculated as proportional to $e^{-\frac{d^2_i}{\theta^2}}$ in a node, and the outputs of the spatial pooler are the inputs of the temporal pooler. As mentioned before, the temporal pooler forms groups of patterns that are likely to follow each other in time, since it would indicate that they are likely to have the same cause in the world.

The designers of HTM used a time-adjacency matrix partitioned with a ‘greedy’ algorithm. This algorithm creates groups by finding the most-connected pattern that is not part of a group, and picking the N most-connected patterns to this pattern recursively (George and Jaros, 2007). For every input from the spatial pooler, the temporal pooler outputs a probability distribution over its groups, propagating the uncertainties up in the hierarchy in a Bayesian Belief Propagation way. The ambiguous information propagated from the bottom of the hierarchy is resolved higher up in the hierarchy.

2.2. Materials and methods

The study area was located in the central plains of Huelva Province, Spain (Figure 4.2), in the sub-region known as “Manzanilla” (37º 23' N; 6º 25' O).

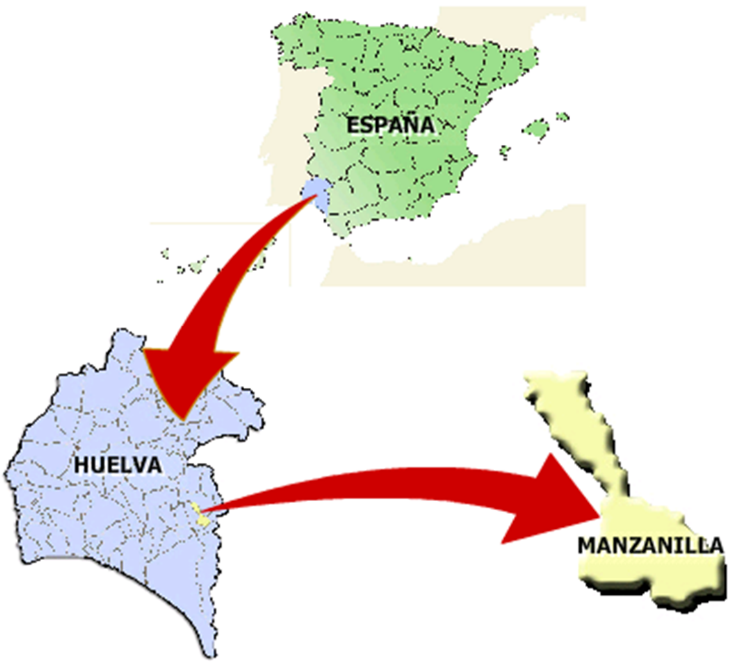

Figure 4. 2 Area of study

2.3. Digital aerial photograph

The dataset used in this research was a photogram received by a photogrammetric Ultracamd® sensor of Vexcel on 23 Octuber 2007, with dimensions of 7 500 x 11 500 pixels. Its bands combination was formed by red, green and blue. The digital aerial photographs had a special resolution of 30 cm. The photogram was segmented in small images of 128 x 128 pixels, as the HTM platform only classifies small images which contend only one pattern.

2.4. Map of crops and exploitation

A map of crops and exploitation of the region of Huelva (2007) was used to carry out the 'training' of the classification and its subsequent validation. The land uses of this area are: vineyards ('Vitis vinifera L'.), 'olive groves' (Olea Europaea L.), 'fallow land', 'irrigated land' and 'built-up surface'.

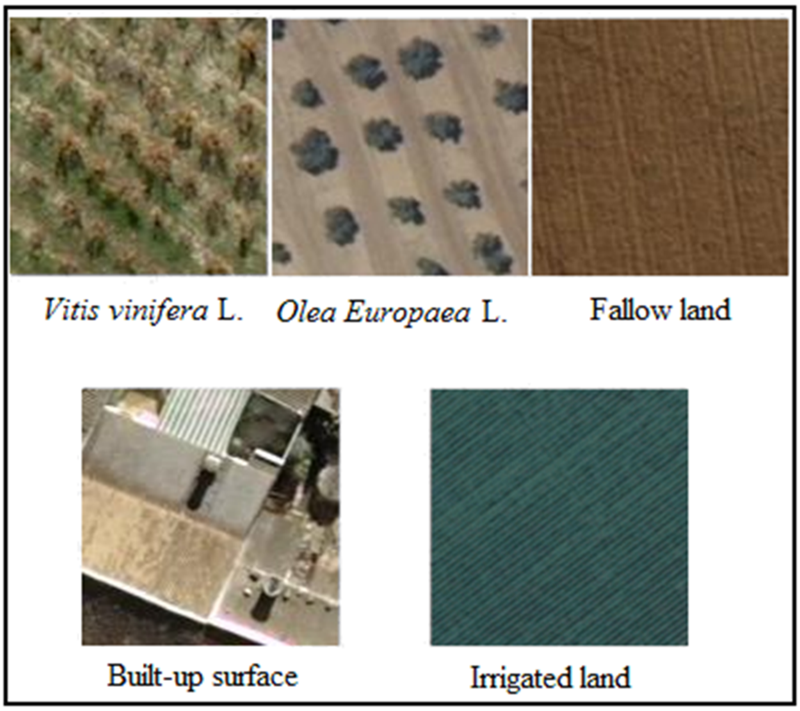

Figure 4. 3 Classified categories.

Table 4.1 shows the number of training and testing images for the architecture 'demo'.

Category	Training Images	Testing Images
***Vitis vinifera* L.**	300	150
Irrigated land	300	150
***Olea Europaea* L.**	300	150
Fallow land	300	150
Built-up surface	300	150

Table 4. 1 Number of training and testing images.

We used Nupic® (Numenta Platform for Intelligent Computing), software for implementing HTMs developed by Numenta to implement our HTM network. The company provides examples of how to create and use HTMs in various scenarios. One of these examples trains an HTM to recognize black and white pictures (one bit per pixel) with different levels of deformations. Another example uses an HTM to classify fruit images (grayscale, 8 bits per pixel). We adapted these examples to solve problems related to the classification of different land uses using small grayscale images (128 x 128 pixels), because HTMs only classify these kinds of images. To implement an HTM, two steps have to be taken: creating the architecture, and training it with a set of training patterns. After we created an architecture and trained the network on the digital aerial photographs train set, we tested the HTM with test set.

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 3 levels. The input level consists of 16 nodes, each receiving a feature and the corresponding delta. Level 2 consists of 4 nodes, each receiving the output of 4 input level child nodes. Level 3 consists of one top level node.

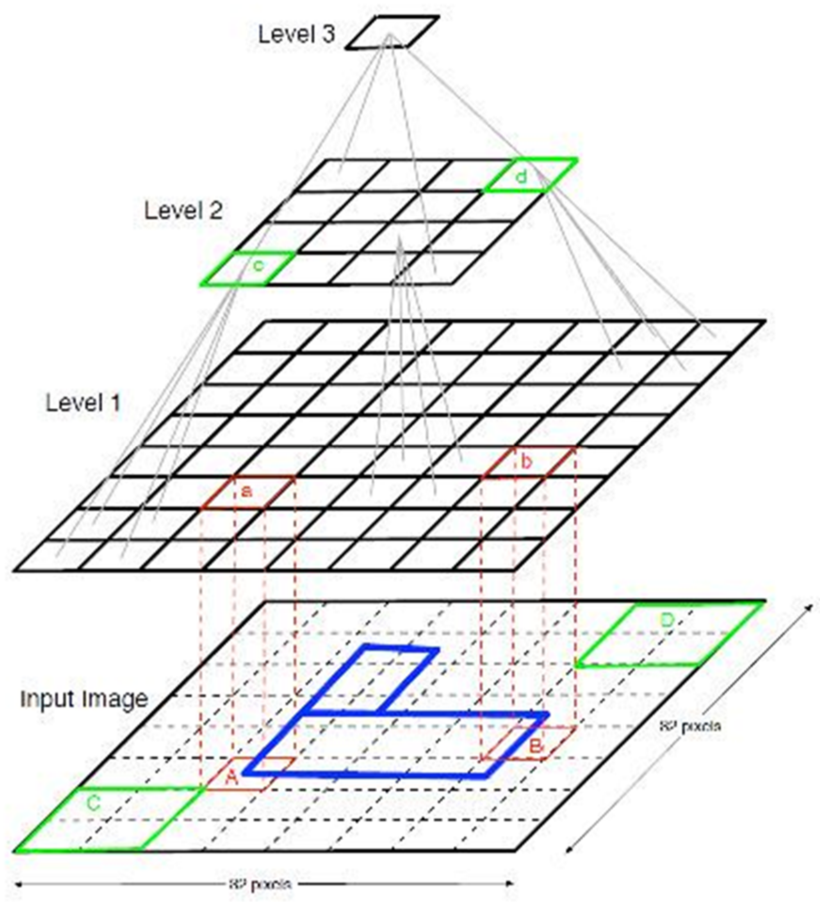

Figure 4. 4 HTM with three layers of nodes (George and Jaros 2007, Figure 4).

The parameters of the HTM network used were as follows:

<u>Level 1</u>:

- levelSize = 64
- pooler algorithm: gaussian; sigma = 0.4
- maxDistance = 5

- maxGroupSize = 1 435
- grouper algorithm: sumProp

Level 2:

- levelSize = 4
- pooler algorithm: product
- maxGroupSize = 1 435
- grouper algorithm: sumProp

Level 3:

- levelSize = 1
- pooler algorithm: product
- mapper algorithm: sumProp

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 (y_i) 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:

$$belief_i = \prod_{j=1}^{nchildren} y_i\left[child_j\right] * x\left[child_j\right]$$

where nchildren is the number of children the node has, x is the input vector, y_i are the vectors previously stored by the spatial pooler, and a[$child_n$] is the part of vector a that is received from the n^{th} 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 (Csapó and Baranyi, 2007).

This type of hierarchical network structure is analogous to the hierarchy of the visual regions in human neocortex, which is also organized as a hierarchy of cortical regions. The receptive field size in the cortical regions also gradually increases in the higher levels of the hierarchy. The neural structures in higher regions of the cortex represent increasingly complex structuresm and the structures in the top visual region represent visual objects just as they do in this model.

2.5. Accuracy Evaluation and Validation

The accuracy evaluation is a general term to compare the generated classification with known geographical information. Its main aim is therefore to determine the veracity of the classification process. A true-terrain image from the information containing in the crop maps and exploitations of the region of Huelva was prepared. The statistics used were: producer's accuracy, user's accuracy, overall accuracy, and Kappa statistic.

The Kappa statistic is a measure of the difference between the observed accuracy and the random possibility of chance agreement between the reference data and the classification (Lillesand and Kiefer, 1994). When the total number of correctly classified pixels in a class is divided by the total number of pixels that should have been classified in that class, it is known as producer's accuracy (Chuvieco, 2000). If the total number of correctly classified pixels in a class is divided by the total number of pixels that were actually classified in that class (both correctly and incorrectly), the result is a measure of user's accuracy (Chuvieco, 2000). The overall accuracy is the percentage of correctly classified pixels. We used Numenta Vision Test App® software to validate our HTM network.

3. Results and discussion

The methodology proposed was applied to the region of study obtaining a final classification of land use. Table 4.2 shows the accuracy of classification in the digital aerial photograph according to its boundary analysis. The highest producer's accuracy was achieved for the' built-up surface', having a value of 100%, while the lowest value has been the 'Vitis vinifera L.' (81.33%).

In the case of the user's accuracy, the highest value was also obtained for the 'built-up surface' class (100%) while the lowest corresponded to 'Olea Europaea L.' (73.03%). The HTM classification thus gave a high overall accuracy of 90.4% ,and the Kappa Statistic had a value of 0.80, showing that the classification was 80% better than a random one.

Category	Vitis vinifera L.	Irrigated land	Olea Europaea L.	Fallow Land	Built-up Surface	Total
Vitis vinifera L.	**122**		28	0	0	150
Irrigated land	1	**134**	12	3	0	150
Olea Europaea L.	20	0	**130**	0	0	150
Fallow Land	0	0	8	**142**	0	150
Built-up Surface	0	0	0	0	**150**	150

Producer´s Accuracy	81.33%	89.33%	86.66%	94.66%	100%
User´s Accuracy	85.31%	100%	73.03%	97.93%	100%
overall Accuracy	90.4%				
Kappa Statistic	0.80				

Table 4. 2 Producer´s Accuracy, User´s Accuracy, overall Accuracy and Kappa statistic of the HTM classification obtained from the digital aerial photograph.

We also verified the capability of the model to learn invariant representations from visual patterns and to store these patterns in the hierarchy and recall them auto-associatively. During the experimentation, we varied many internal constants affecting the learning process, and also made modifications to the algorithms and data structures themselves. The following figures illustrate the main recognition capabilities of the system trained to recognize 5 categories of images. One of the two original training images in the category ‘olives groves’ is shown in figure 4.5. The system easily recognized the shifted version of the original image shown in figure 4.6.

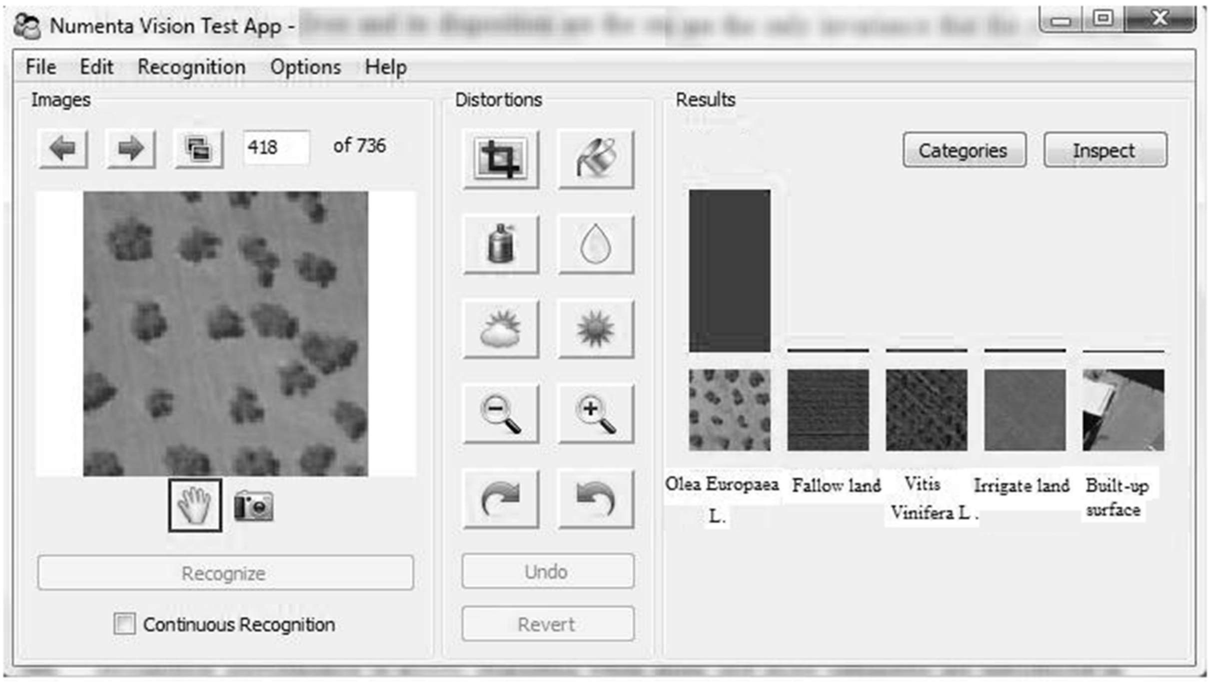

Figure 4. 5 Original Image.

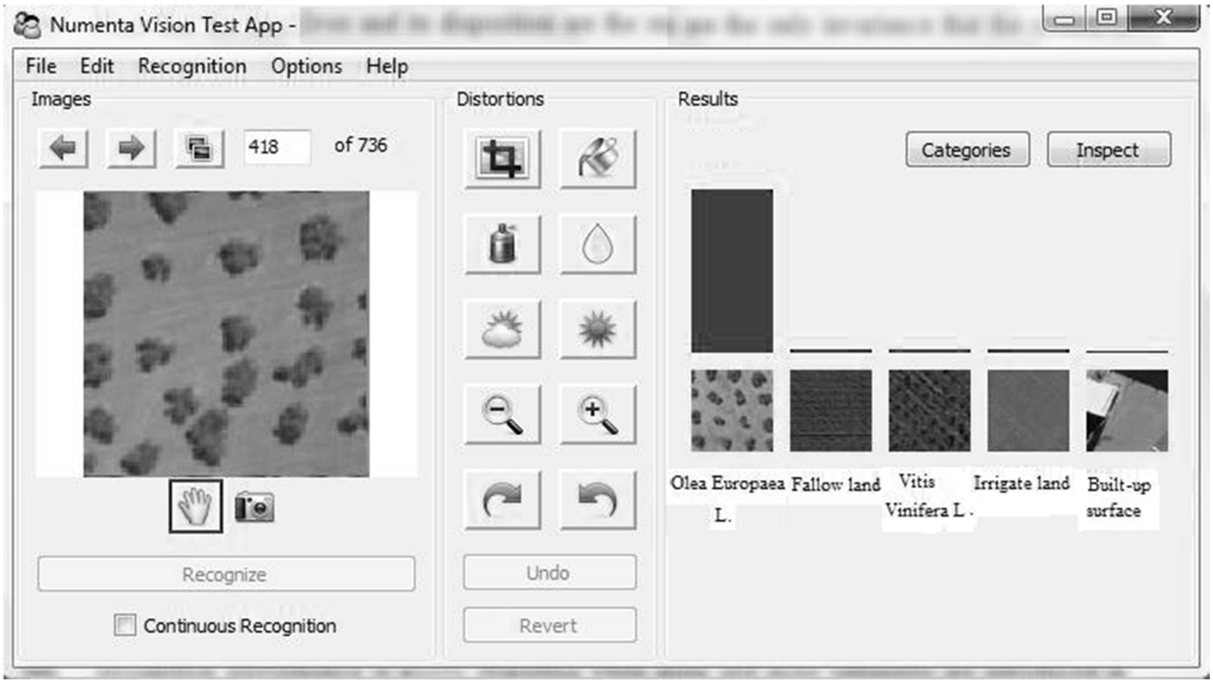

Figure 4. 6 Rotated Image.

Note that the number of the 'olive groves' and their disposition were the only invariance that the system was explicitly exposed to during the training; hence the other invariances described below were discovered automatically by the system.

The system can function as an auto-associative memory, as demonstrated by figure 4.7. Given a part of the original image, the missing information is reconstructed and the category is predicted correctly. This resembles a capability of the brain to recall missing information given only partial input.

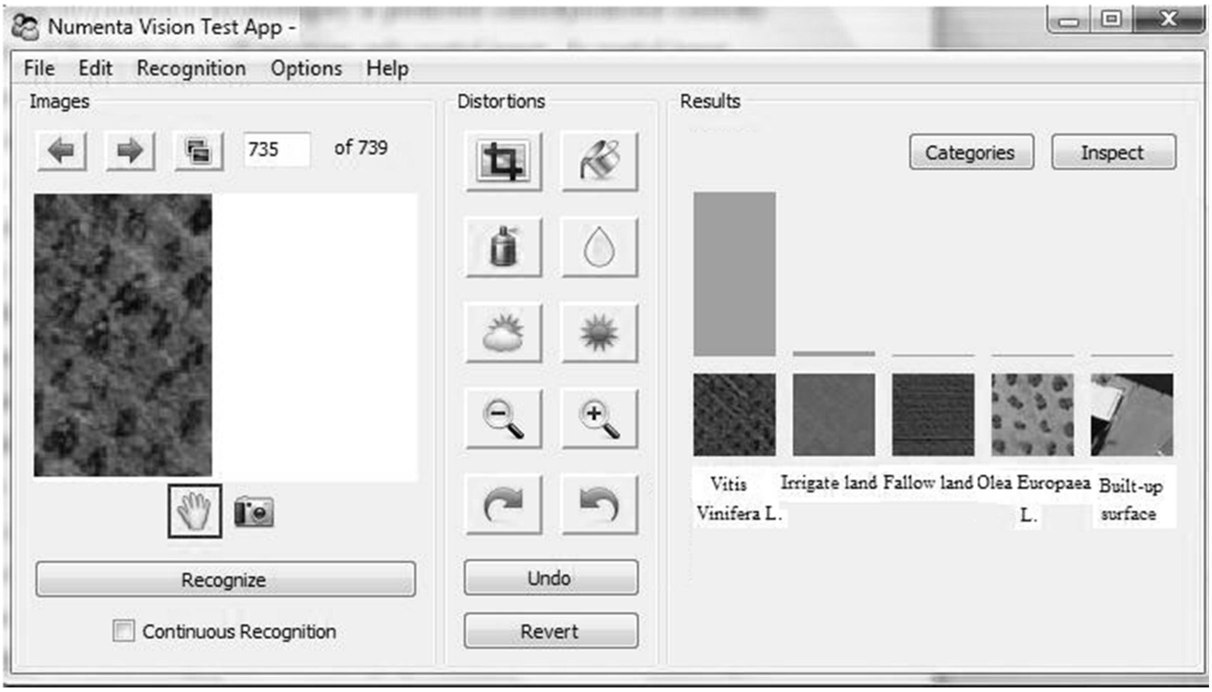

Figure 4. 7 Part of the original image recognition.

The system can also tolerate a substantial amount of noise of various types and still discern and correctly recognize the category as shown in figure 4.8.

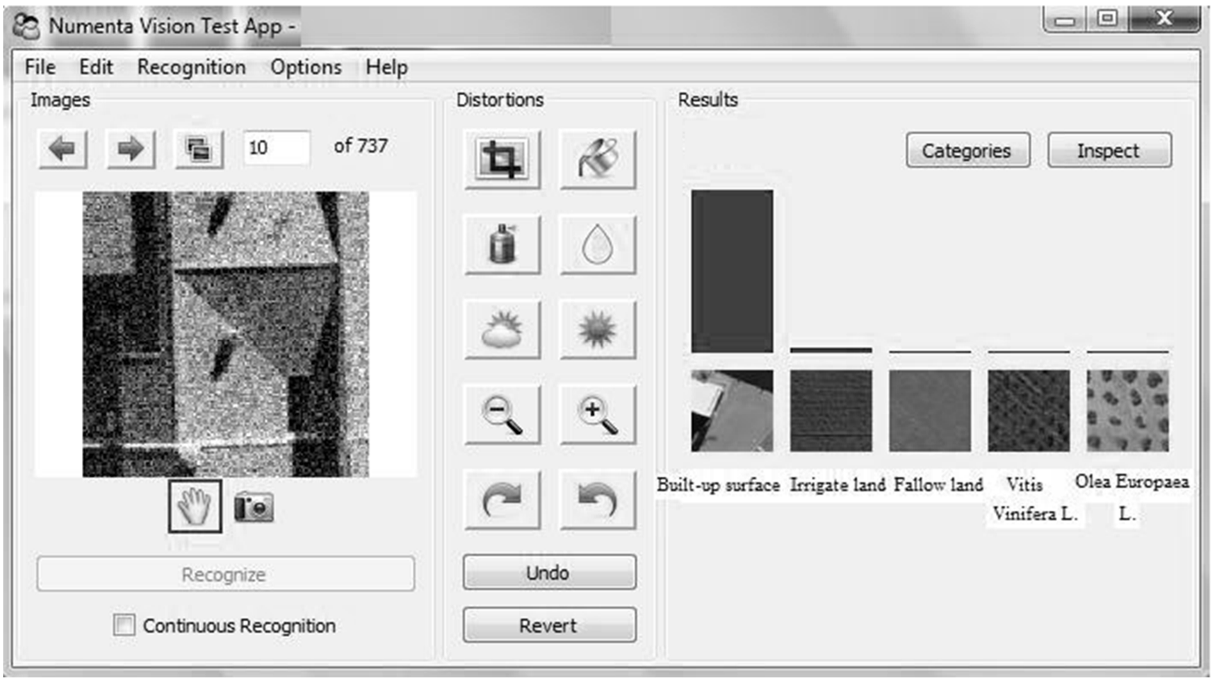

Figure 4. 8 Built-up surface with noise.

Overall, we observed that the system performed better while recognizing complex images having more discernible features such as corners and line intersections, so, for example, images of 'fallow land' and 'irrigated land' were not recognized as well as 'built-up surface; or 'Vitis vinifera L.'. The system also sometimes tended to confuse categories sharing many similar shapes, such as 'Vitis vinifera L.' and 'Olea Europaea L.'. We also observed that the recognition performance was slowly degraded when more and more categories were introduced in training, arising from the same confusion between similar images. For example, we tried to classify different irrigated crops (Phaseolus vulgaris L., Triticum aestivum L., Vicia faba L. and Pisum sativum L.), but the analysis and result showed a small overall accuracy of 69.65% and a Kappa Statistic of 0.72.

It is also useful to observe the relative strength of beliefs of the ten best-predicted categories that is displayed by the system as a bar graph. When input image is not heavily distorted, and resembles its true category much more than any other categories, we see the graph similar to the one shown in figure 4.9. We can judge from the graph that the winning prediction is very confident. When the input image is not readily recognizable or seems similar to several categories, the graph will look like the one in figure 4.10.

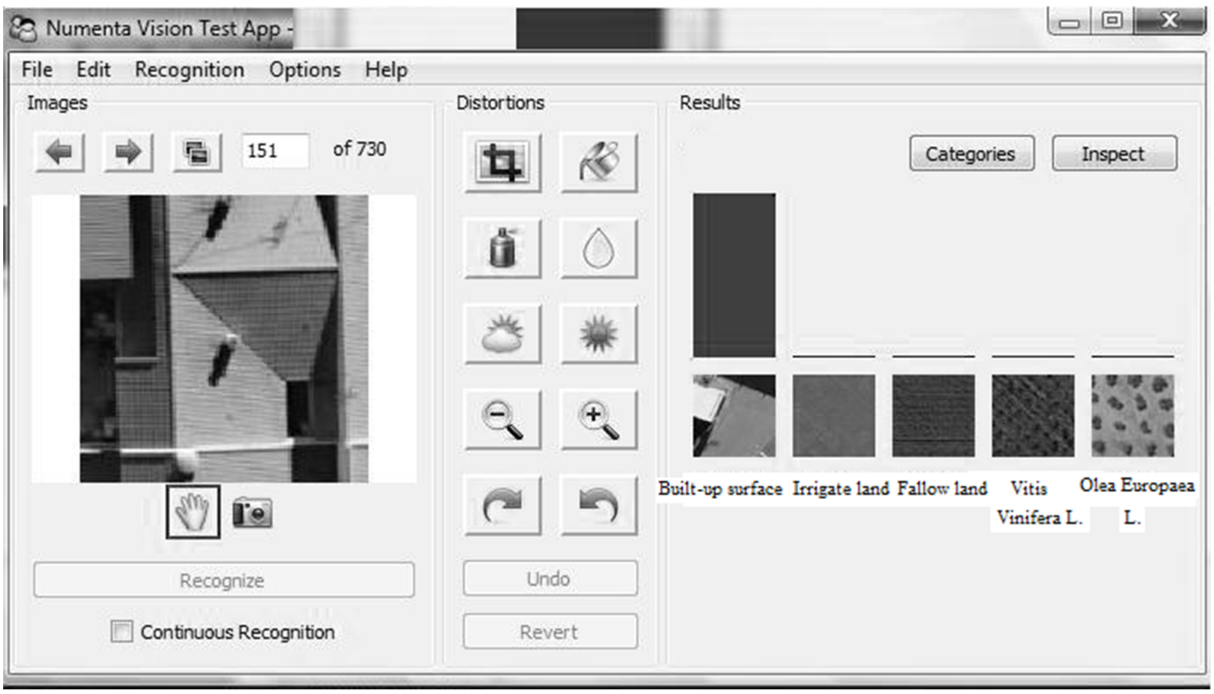

Figure 4. 9 100% true.

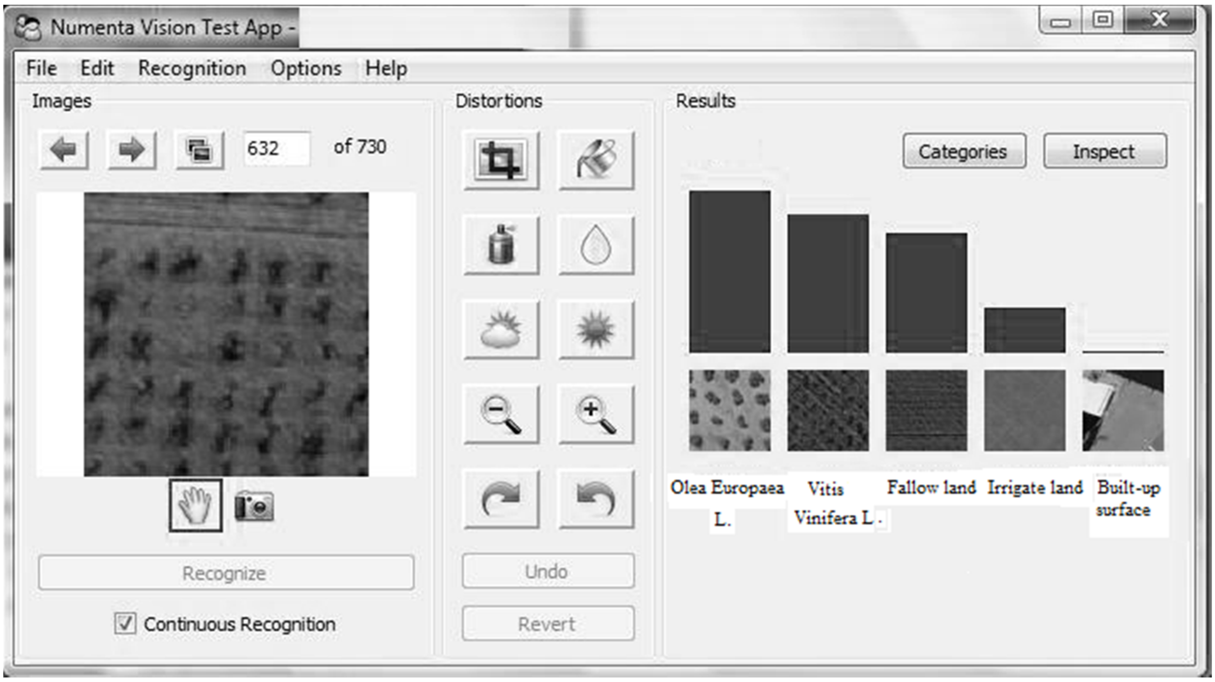

Figure 4. 10 Confuse classification.

4. Conclusions

The images from the digital aerial sensors in our model may be an extremely useful tool in the agriculture field, providing an accurate result about the uses of land in a fixed area under certain conditions. By contrast, traditional classification techniques, basically pixel-based approaches, are limited in that they typically produce a characteristic “salt and pepper” effect, and are unable to extract objects of interest. An HTM network considers spatial and temporal relations between features of the sensory signals which are formed in a hierarchical memory architecture during a learning process. The methods are not actually comparable, however, because HTM only classifies small grayscale images with only one pattern. In the future, we expect that the platform will be able to classify more than one pattern, and salt and pepper effect could be eliminated.

This model shares many common ideas with traditional neural networks. The hierarchy consists of many relatively simple units (subregions) that do the same basic operation and can be made to run in parallel. It solves problems by using cooperation between subregions without a centralized algorithm. The knowledge and beliefs in the system are distributed between the subregions in various hierarchy levels. It learns its skills by training and is able to generalize. The memory-prediction framework is an inferential system, however, that uses beliefs for learning and recognition. Nevertheless, due to the similarities, the model shares a number of advantages with neural networks; it clearly can function as an associative memory, can tolerate noise, and can generalize training images to similar ones.

Finally, it offers a greater promise of understanding what intelligence is by closely modeling the overall structure of the human neocortex.

5. References

Chuvieco E. 2000. *Fundamentos de teledetección espacial*. 3rd Ed. Madrid (Spain): Ed. Rialp; p. 453.

Csapó A. and P. Baranyi. 2007. *VFA-driven Hierarchical Temporal Memory Input for Object Categorization*. 8th International Symposium of Hungarian Researchers On Computational Intelligence and Informatics, Hungary. Available from: http://bmf.hu/conferences/cinti2007/3_CsapoAdam.pdf

Douglas R.J. and K.A.C. Martin. 2004. *Neuronal Circuits of the Neocortex*. Annual Review of Neuroscience; 27, 419-451.

George D. and B. Jaros. 2007. *The HTM Learning Algorithms, Numenta.* Numenta Incorporated, Redwood city, California. Available from: http://www.numenta.com/for-developers/education/Numenta_HTM_Learning_Algos.pdf

Hawkins J. and S. Blakeslee. 2004. *On Intelligence*; New York (USA): Henry Holt; chapter 4, p. 83-104.

Hawkins J. and D. George. 2007. *Hierarchical Temporal Memory, Concepts, Theory, and Terminology.* Numenta Incorporated, Redwood city, California. Available from: http://www.numenta.com/Numenta_HTM_Concepts.pdf

Lillesand T.M. and K.W. Kiefer. 1994. *Remote Sensing and Image Interpretation*, 3rd ed. USA: Ed. John Wiley & Sons, Inc.; p. 616-617.

Capítulo VI. Hierarchical Temporal Memory for mapping vineyards using digital aerial photographs

Autores: Perea, A.J.[a], Meroño, J.E.[b], Aguilera, M.J.[a]

[a] Departamento de Física Aplicada, Universidad de Córdoba (UCO), Córdoba, Spain

[b] Departamento de Ingeniería Gráfica y Geomática, Universidad de Córdoba (UCO), Córdoba, Spain

Publicado en: African Journal of Agricultural Research, 7, (3), 456-466.

«La inteligencia se mide por la capacidad de recordar y predecir patrones del mundo, incluidos lenguaje, matmática, propiedades físicas de los objetos y situaciones sociales»

Jeff Hawkins (2004)

1. Abstract

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.

2. 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 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 per cent 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. 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’s 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 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.

3. Materials and methods

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

This is a rectangular area of 6 x 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 hectares of planted vineyards and 2 800 hectares 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 *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 x 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).

Figure 5. 1 Area of study.

Digital Vector maps, color orthophotos and digital terrain models were used to orthorectify the image, select training areas and validate the classifications. Data map were 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.

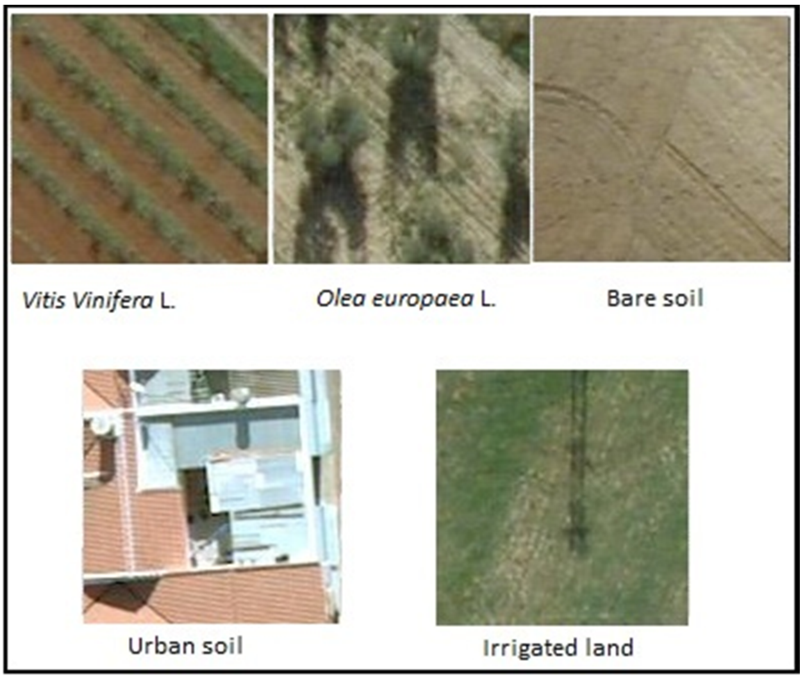

Figure 5. 2 Classified categories.

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 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.

3.1. Obtaining the 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.

3.2. 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.

3.3. Hierarchical Temporal Memory (HTM)

Hierarchical Temporal Memory is a technology that replicates the structural and algorithmic properties of the neocortex (Hawkins and George, 2007). HTM is organized as a tree shaped hierarchy of nodes (Fig. 3).

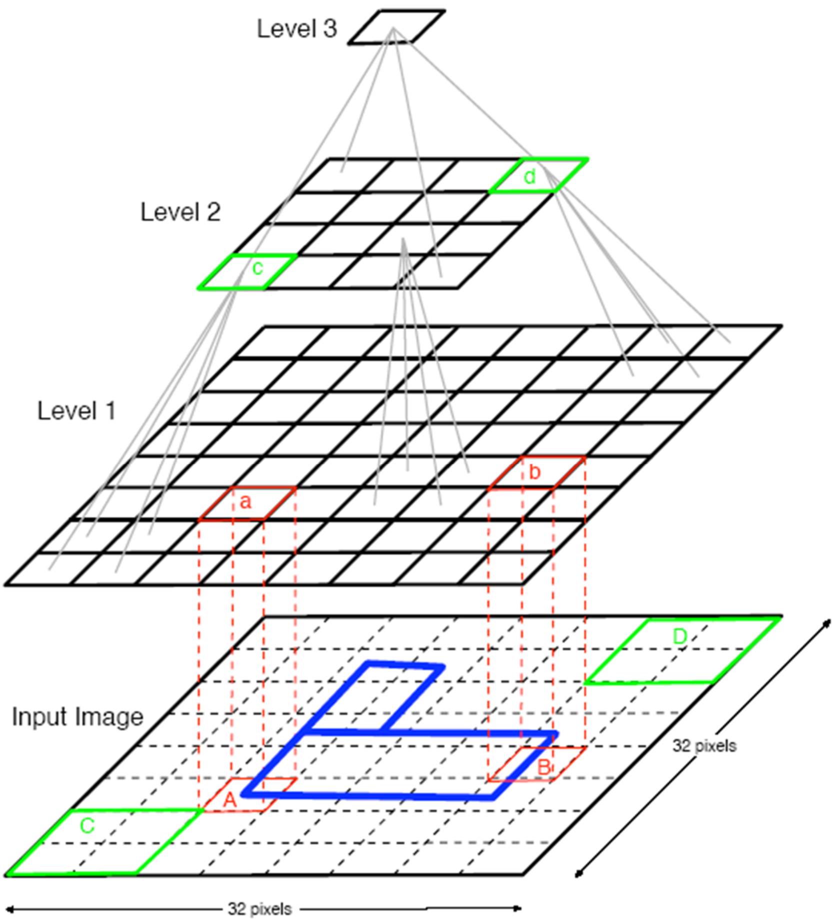

Figure 5. 3 The HTM (Hierarchical Temporal Memory) model with three layers of nodes. Each subregion in level 1 receives image fragment of size 4x4 pixels. Each subregion in level 2 receives input from 4 children in level 1. A single subregion in level 3 receives input from all level 2 subregions (George 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.

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.

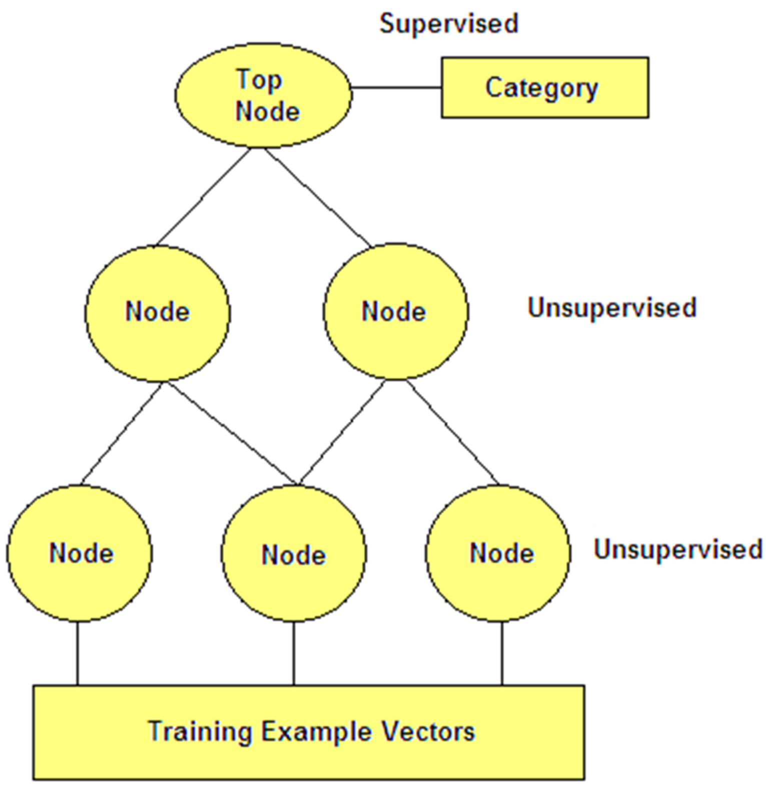

Figure 5. 4 HTM network

3.4. 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 D, Jaros B, 2007).

3.5. Operation of Spatial pooler during learning

The spatial pooler has two stages of operation:

• During the learning stage it quantizes the input patterns and memorizes the quantization centers.

• Once these quantization centers are learned, it produces outputs in terms of these quantization centers during the inference stage (George D and Jaros B, 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 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 imput 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.

3.6. 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 *transitionMemory*.

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

3.7. 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.

3.8. 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).

3.9. 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.

3.10. 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 *temporalPoolerAlgorithm*.

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.

3.11. Operation of Top node during inference

During inference of the top node, the spatial pooler works as described above. 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 choice between two different temporal pooler algorithms during inference: *maxProp* and *sumProp*, controlled by the parameter *mapperAlgorithm*.

3.12. 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 outputElementCount, 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:

$$belief_i = \prod_{j=1}^{nchildren} y_i\left[child_j\right] * x\left[child_j\right]$$

(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:

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.

scaleOverlap- An integer specifying how many scales neighboring nodes should share in common. For example, if scaleRF is 2, scaleOverlap is 1, and there are 3 resolutions in the level below, some nodes will see the smaller and middle resolutions, and some nodes will see the middle and larger resolutions.

3.13. 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 an 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 (128x128 pixels) composed of IRGB bands and NDVI index.

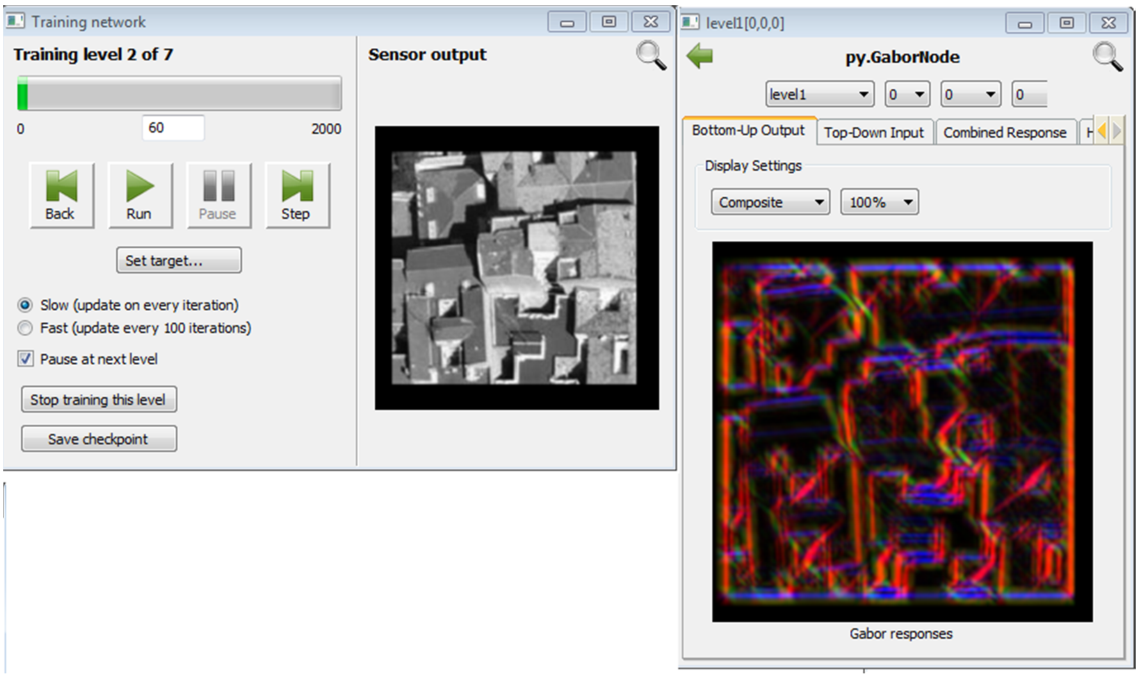

Figure 5. 5 Training stage of level 1, sub-level 2.

In figure 5.5 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.

3.14. 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 5.1 shows the number of training and testing images for the architecture 'demo'.

CATEGORY	TRAINING IMAGES	TESTING IMAGES
Vitis Vinifera L.	300	150
Other land covers	300	150

Table 5. 1 The number of training and test images.

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, 40-55% fair, 55-70% good, 70-85% very good and > 85% excellent (Monserud and Leemans, 1992).

4. 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 (transitionMemory and 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 in Table 5.2.

MAXDISTANCE	SCALERF	SCALEOVERLAP	OVERALL ACCURACCY (%)	#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

Table 5. 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.

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 (Fig. 6).

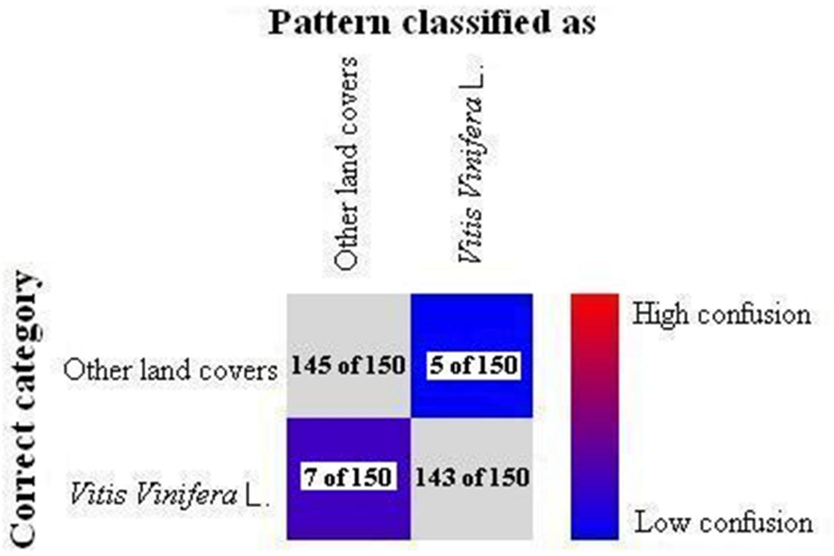

Figure 5. 6 Confusion matrix of the best performing system.

Pa: producer's accuracy. Ua: user's accuracy.

NDVI: normalized difference vegetation index.

IRGB: near Infrared, green and blue bands

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	

Table 5. 3 Producer's and user's accuracy, overall and Kappa statistic for supervised classifications and HTM classification.

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 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.

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 5.7 a map obtained from the HTM classification is presented.

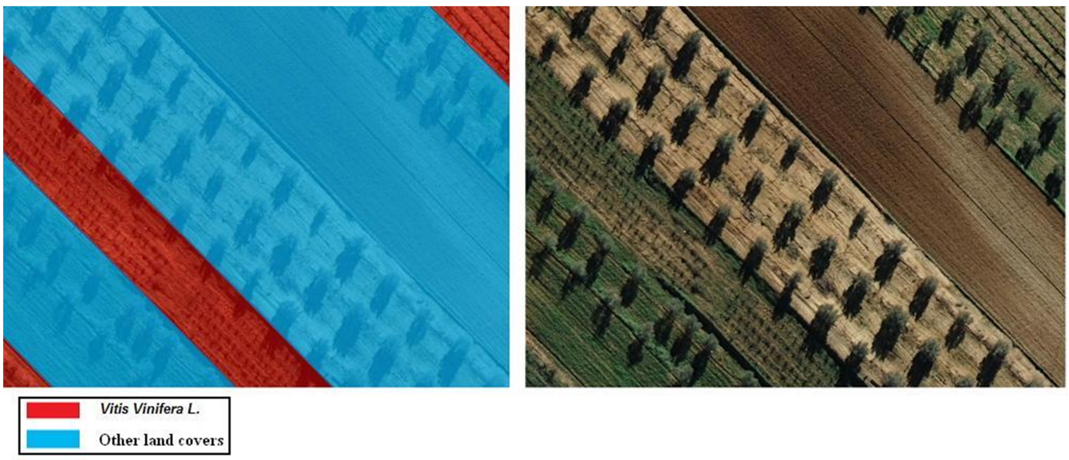

Figure 5. 7 Classified image obtained from a system based on Hierarchical Temporal Memory.

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.

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

5. Conclusions

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.

6. References

Aitkenhead M.J. and G.G. Wright. 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 J.P., Germain C., Lavialle O. And G. Grenier. 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; p. 81–87.

Budgen J.L., Salinas G. and P.J. Howarth. 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. and M. Arnaud. 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; p. 375-382.

Da Costa J.P., Michelet F., Germain C., Lavialle O. and G. Grenier. 2007. *Delineation of vine parcels by segmentation of high resolution remote sensed images*. Precision Agriculture; 8(1): 95–110.

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

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

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

Gong P., Mahler S.A., Biging G.S. and D.A. Newburn. 2003. *Vineyard identification in an oak woodland landscape with airborne digital camera imagery*. International Journal of Remote Sensing; 24:1303-1315.

Haboudane D., Miller J.R., Pattey E., Zarco-Tejada P.J. and I. Strachan. 2004. *Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture*. Remote Sensing of Environment; 90: 337-352.

Hall A., Louis J. and D. Lamb. 2003. *Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images*. Computers and Geosciences; 29: 813–822.

Hawkins J. and S. Blakeslee. 2004. *On Intelligence*. New York (USA): Henry Holt.

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

Marini M.F. 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 R.A. and R. Leemans. 1992. *Comparing global vegetation maps with the Kappa statistic*. Ecological Modelling; 62: 275-293.

Numenta Inc. 2008. *Advanced Nupic Pogramming*. Available from: http://www.numenta.com/archives/education/nupic_prog_guide.pdf.

Perea A.J., Meroño J.E. and M.J. Aguilera. 2009. *Application of Numenta® Hierarchical Temporal Memory for land-use classification*. South Africal Journal of Science; 105(9-10): 370-375.

Rabatel G., Delenne C. and M. Deshayes. 2008. *A non-supervised approach using Gabor filters for vine plot detection in aerial images*. Computer and. Electronicic in Agriculture; 62(2):159-168.

Ranchin T., Naert B., Albuisson M., Boyer G. and P. Astrand. 2001. *An automatic method for vine detection in airborne imagery using wavelet transform and multiresolution analysis.* Photogrammetric Engineering of Remote Sensing; 67: 91-98.

Rodríguez-Pérez J.R., Álvarez-López C.J., Miranda D. and M.F. Álvarez. 2008 *Vineyard area estimation using medium spatial resolution satellite imagery*. Spanish Journal of Agricultural Research; 6(63): 441-452.

Soria-Ruiz J., Fernandez-Ordoñez Y. and I.H. Woodhouse. 2010. *Land-cover classification using radar and optical images: a case study in Central Mexico*. International Journal of Remote Sensing; 31(12): 3291-3305.

Vaudour E., Carey V.A. and J.M. Gilliot. 2010. *Digital zoning of South African viticultural terroirs using bootstrapped decision trees on* *morphometric data and multitemporal SPOT images*. Remote Sensing of Environment; 114(12): 2940-2950.

Warner T. and K. Steinmaus. 2005. *Spatial Classification of Orchards and Vineyards with High Spatial Resolution Panchromatic Imagery.* Photogrammetric Engineering of Remote Sensing; 71(2): 179-187.

Wassenaar T., Baret F., Robbez-Masson J.M. and P. Andrieux. 2001. *Sunlit soil surface extraction from remotely sensed imagery of perennial, discontinuous crop areas; the case of Mediterranean vineyards*. Agronomic for Sustainable Development; 21: 235-245.

Wassenaar T., Robbez Masson J.M., Andireux P. and F. Baret. 2002. *Vineyard identification and description of spatial crop structure by per-field frequency analysis*. International Journal of Remote Sensing; 23(17): 3311-3325.

Capítulo VII. Conclusiones Generales

«Los cerebros no son como las computadoras, a las cuales uno introduce símbolos, y luego sale de ellas algo diferente»

Jeff Hawkins (2004)

En esta tesis se ha realizado una evaluación de las nuevas técnicas de clasificación digital de imágenes de alta resolución espacial para la clasificación de usos del suelo. Para ello, se han desarrollado diversas metodologías para la actualización de bases de datos cartográficas de usos del suelo, basadas en la integración y el análisis de información cartográfica vectorial, ortoimágenes e información temática. Esta metodología se ha aplicado para su evaluación y discusión en un caso real. El primer punto a destacar de este trabajo consiste en la profundización en una nueva tendencia en el análisis de imágenes de la superficie terrestre para la actualización de bases de datos espaciales de usos del suelo mediante la integración de la información disponible de distintas fuentes y en distintos formatos, en lugar del empleo único o fundamental de imágenes. La combinación de distintos tipos de información lleva asociada la combinación de las técnicas y herramientas propias del tratamiento de esa información.

Las conclusiones que se derivan de la evaluación de los resultados de las diferentes metodologías de clasificación presentadas en los diferentes capítulos para la cuantificación de áreas agrícolas de las regiones del estudio, tanto en imágenes de satélite como en las imágenes aéreas digitales, son:

1º La combinación de bandas que mejores resultados proporciona en la clasificación supervisada es la imagen formada por los componentes principales y el NDVI.

2º El empleo de la clasificación orientada a objetos, el algoritmo de clasificación experta y el algoritmo basado en la Memoria Temporal Jerárquica han mejorado los resultados considerablemente con respecto a las técnicas de clasificación tradicionales.

Respecto a la clasificación orientada a objetos:

Cabe destacar que reduce en gran medida los problemas asociados al empleo de imágenes de alta resolución espacial, como es el caso del efecto «sal y pimienta».

La clasificación orientada a objetos utilizando el algoritmo de mínima distancia muestra una mejora considerable respecto a la clásica clasificación supervisada. Cabría esperar mejores resultados utilizando el algoritmo de máxima probabilidad.

La aplicación del algoritmo de clasificación experta ha supuesto las siguientes ventajas:

1º La clasificación experta realizada utilizando las clasificaciones previas: supervisada de la imagen formada por los componentes principales (C.P.), supervisada de la imagen formada por los C.P. y el NDVI y la clasificación orientada a objetos, ha presentado los mejores resultados.

2º Se ha podido incorporar las zonas de olivar y superficie edificada gracias a la información contenida en la imagen verdad-terreno. Estas clases presentaban unos valores de separabilidad bajos en la clasificación puramente espectral.

Por otro lado, las clases mejor discriminadas en las distintas clasificaciones han sido el monte/matorral y la retirada desnuda. Sin embargo, la alfalfa y la retirada cubierta han sido las clases con mayor dificultad de discriminación. La alfalfa presentaba bajos valores de fiabilidad del usuario debido a su baja separabilidad estadística con clases como cereal, proteaginosas o retirada cubierta.

En cuanto a la utilización multiespectral de fotografías aéreas, el NDVI, calculado a partir de los niveles digitales de la fotografía digital y no a partir de valores de reflectividad, ha mostrado ser de utilidad a la hora de hacer la clasificación supervisada, demostrando la eficacia de esta metodología usando esta información.

De la aplicación del Nuevo algoritmo basado en la Memoria Temporal Jerárquica podemos extraer las siguientes conclusiones:

1º Estamos ante una tecnología que supone un cambio en el paradigma de la inteligencia artificial en cuanto a programas de clasificación, ya que se basa en una teoría (teoría de memoria-predicción) que emula el funcionamiento del cerebro humano.

2º Una de sus mayores ventajas es que nos permite crear redes HTM en las que podemos diseñar cada uno de los parámetros de la misma, es decir, podemos construir aplicaciones de clasificación creando un diseño optimizado para el problema al que nos enfrentemos.

Se ha estudiado en profundidad la algorítmica y funciones de procesado de información que incorpora esta tecnología y es interesante reseñar ciertas diferencias novedosas frente a otras algorítmicas.

En primer lugar, en las redes MTJ, concretamente en sus nodos, no existe una algorítmica específica programada para la resolución de un problema específico como ocurre en muchos sistemas de clasificación en inteligencia artificial. Posee una algorítmica capaz de procesar cualquier tipo de patrones, realizando clasificación y predicciones sobre estos.

Un punto que también es novedoso es la arquitectura de los nodos, en la que se incorpora una función de análisis de información, es decir, un algoritmo que analiza la información a través del tiempo. Busca patrones comunes que se repiten a lo largo del proceso de entrenamiento.

Se ha desarrollado un experimento práctico con esta tecnología, concretamente en el reconocimiento y clasificación de cultivos en fotografías aéreas digitales, con el fin de evaluar la capacidad de esta tecnología como clasificador, obteniéndose una precisión del 96%.

Por tanto, estamos ante una tecnología novedosa que supone un cambio importante en el panorama de la inteligencia artificial y los clasificadores basados en esta ya que funciona bajo los preceptos de una teoría nueva, la cual queda plasmada en unos algoritmos específicos.

En resumen, según los resultados obtenidos en las distintas clasificaciones de la fotografía aérea digital, podemos decir que las imágenes de los sensores aéreos digitales pueden ser un efectivo reemplazo de las imágenes de satélite. Se han mejorado todos los resultados de las clasificaciones con respecto a los obtenidos en la imagen de satélite, lo cual nos demuestra la grandes ventajas que tiene la información espectral de estos sensores. Estos resultados caben que mejorarán en el futuro cuando se tenga la información sobre modelos que permitan obtener reflectividad a partir de los niveles digitales registrados en la imagen. Por otro lado, las nuevas técnicas de clasificación suponen un gran avance en el campo del control de ayudas por superficie. La calidad de los resultados obtenidos en la fotografía aérea digital unida al desarrollo de estas nuevas técnicas, posibilitan la reducción del número de visitas de campo para llevar a cabo los controles.