

niendo en cuenta que el máximo valor que puede tomar esta medida es la raíz cuadrada de 2, la distancia obtenida en este caso es relativamente baja y, por lo tanto, la discriminación entre ambas clases es claramente mejorable.

Además, elaborando un índice de confusión entre las clases olivar y suelo labrado que tome en consideración los píxeles de ambas clases clasificados erróneamente entre sí, el porcentaje de confusión toma un valor de 0.157, es decir, el 15.7% de los píxeles de ambas clases está clasificado erróneamente en la otra clase.

Según Duevas (1985), la inclusión de información espacial y de contexto puede mejorar los resultados de clasificación del olivar. En este trabajo estudiamos la utilización de información textural e índices de vegetación por tratarse de un desarrollo sencillo y de aplicación inmediata, mientras que los clasificadores de contexto, por su carácter más complejo desde el punto de vista metodológico y práctico, se abordan en un trabajo independiente (González et al., 1987).

EL AREA DE ESTUDIO. SELECCION DE LA IMAGEN

Como área de estudio para la realización de la experiencia se utilizó la misma zona en la que se había actuado en el trabajo previo de clasificación (González et al., 1987). Así, se aprovechó la existencia de la imagen libre de nubes, de trabajos de campo coetáneos con la fecha de obtención de la imagen, y se suprimieron labores previas como son el registro geográfico, la obtención de archivos de información temática y del archivo índice, y el tratamiento de la información gráfica derivada de la verdad de campo, incluyendo digitalización, calibración local, generación de máscaras, etc. Además, de esta manera se aseguraba la comparabilidad entre los resultados obtenidos en ambos trabajos.

El área de estudio está enclavada en la provincia de Toledo, como un transecto del valle del Tajo en su zona media. De Norte a Sur se extiende desde las estribaciones toledanas del Sistema Central, hasta los Montes de Toledo, incluyendo parte de las extensas comarcas agrarias de Torrijos, Talavera, La Jara y Montes de Navahermosa. Como consecuencia de ello, la diversidad de usos del suelo es muy amplia, abarcando desde los más forestales, como montes de cas-

taño y rebollo o repoblaciones de pino piñonero y pino pinaster, pastizales, matorrales y dehesas de quercineas, hasta cultivos de marcado carácter agrícola como son los cereales (trigo, cebada, avena), regadíos, y cultivos leñosos tales como el olivo, la vid y diversas especies de frutales.

Si bien el olivo se encuentra presente en cualquiera de las comarcas que componen la zona de estudio, su presencia es más notable en la comarca de Torrijos, situada al Noroeste del río Tajo, donde los olivares se asientan sobre arenas y arcosas, alternando con cereales y viñedos, y en la zona Sur, al pie de los Montes de Toledo, donde resulta predominante sobre el resto de cultivos, asentado sobre las rañas existentes en términos municipales como los de Navahermosa, Los Navalacrales, Los Navalvillos y Espinoso del Rey.

En cuanto a la imagen utilizada, se trata de una imagen MSS-Landsat 4, identificada como 4/202/032, obtenida en fecha 27/05/83.

En esta imagen se ha definido como área de estudio una ventana de la imagen comprendida entre las filas 950 y 2285 y las columnas 2491 y 3600 de la CCT.

SELECCION DE VARIABLES

Antes de diseñar una estrategia de actuación, cabe reflexionar acerca de la naturaleza de la respuesta espectral de un cultivo como el olivar. Esta respuesta no depende únicamente de la cantidad de biomasa verde existente o del vigor vegetativo de la cubierta, sino también del diámetro de las copas de los olivos, del marco de la plantación y del color del suelo. Estas propiedades no resultan suficientemente cuantificadas individualmente mediante las cuatro bandas del sensor MSS y, siendo quizás importantes a la hora de entrar en un proceso de clasificación supervisada, parece coherente utilizar un método que las tome en consideración, previamente a la clasificación.

Con objeto de realzar propiedades espectrales como la respuesta de la biomasa verde, el vigor vegetativo, o el brillo del suelo, cabe pensar en variables sintéticas, como los índices de vegetación, que son combinaciones lineales de las bandas visibles e infrarrojas de una imagen. Así pues, como variables sintéticas se ensayó un conjunto

de índices de vegetación relacionados con propiedades como la cantidad de biomasa verde (Rouse et al., 1973), el vigor vegetativo (Deering, 1978) y la importancia del color del suelo y del verdor de la cubierta vegetal (Kauth y Thomas, 1976).

Los índices de vegetación utilizados fueron los siguientes:

$$IV1 = \sqrt{B4 - B2 + 0.5}$$

$$IV2 = \sqrt{B3 - B2 + 0.5}$$

$$IV3 = (B4 - B2) / (B4 + B2)$$

$$IV4 = .43 * B1 + .63 * B2 + .59 * B3 + .26 * B4$$

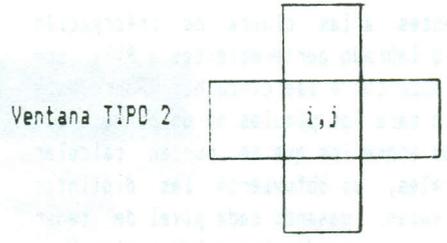
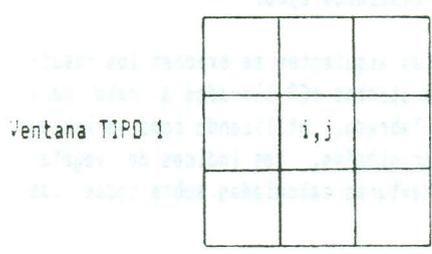
$$IV5 = -.29 * B1 - .56 * B2 + .60 * B3 + .49 * B4$$

Siendo B1, B2, B3 y B4 las cuatro bandas del sensor MSS.

Sin embargo, existen propiedades particulares del olivar que este tipo de variables no recogen; son todas aquellas concernientes a las relaciones espaciales. La textura, sin embargo, aporta información sobre la distribución espacial de las variaciones tonales dentro de una imagen.

Las medidas de la textura son muy numerosas aunque, en general, las más sencillas asignan a cada píxel un valor en función de los valores espectrales de los píxeles que le rodean. El tamaño de la orla o ventana, y la función de asignación empleada, dan origen a los distintos valores de la textura de una imagen.

En este estudio se han empleado dos tipos de ventana:



Siendo, en cada caso, i y j la fila y la columna del píxel central de la ventana.

Para la generación de los valores texturales, se han tomado en consideración los criterios de Haralick (1979) y Gordon y Philipson (1986). Las variables usadas fueron las siguientes:

$$TEXTURA = \sqrt{\frac{\sum_{\delta_i, \delta_j} (P_{i, j+\delta_i, j+\delta_j} - P_{i,j})^2}{n}}$$

$\delta_i = -1, 0, +1$
 $\delta_j = -1, 0, +1$

$$RUGOSIDAD 1 = \sum_{\delta_i, \delta_j} |P_{i,j} - P_{i+\delta_i, j+\delta_j}|$$

$$RUGOSIDAD 2 = \frac{1}{\sum_{\delta_i, \delta_j} P_{i+\delta_i, j+\delta_j}} \sum_{\delta_i, \delta_j} |P_{i,j} - P_{i+\delta_i, j+\delta_j}|$$

Una vez elegidas las herramientas de trabajo, se plantea la necesidad de elegir cual combinación de variables podría ser la más adecuada para entrar en el proceso de clasificación.

En primer lugar, hay que obtener los valores de cada uno de los píxeles de la imagen después de transformados. Para ello, se aprovechó una utilidad del sistema ERAFIS, que es capaz de generar un archivo en el que, para cada píxel, se indica la clase informacional a la que pertenece, la clase espectral, el Area de Verdad Terreno a la que pertenece, la Parcela de Verdad Terreno, si el píxel es o no de borde, la fila y la columna en donde está situado el píxel dentro de la imagen, y los cuatro valores espectrales de las cuatro bandas del sensor MSS correspondientes a cada píxel.

Este archivo se genera después de hacer la extracción de los píxeles pertenecientes a cada una de las clases de información seleccionadas. Para llevar a cabo la selección se utilizó un archivo depurado en el que se incluían únicamente los píxe-

les pertenecientes a las clases de información olivar y suelo labrado pertenecientes a PVTs con una superficie superior a las cinco ha. Sobre este archivo, y sólo para los píxeles no de borde, que son los únicos sobre los que se pueden calcular valores texturales, se obtuvieron las distintas variables derivadas, pasando cada pixel de tener cuatro valores, correspondientes a las cuatro bandas espectrales, a tener un total de 63 valores distintos.

Si bien el sistema ERAFIS está parametrizado y puede clasificar hasta con doce variables, es evidente que, dentro de la batería de variables calculadas, existe mucha información redundante o poco significativa, por lo que resulta necesario aplicar algún método de selección de las variables más útiles. Para ello se utilizaron dos técnicas estadísticas multivariantes clásicas: El Análisis Factorial Discriminante y el Análisis en Componentes Principales.

Para el Análisis Factorial Discriminante se utilizó el programa BMDP7M, que realiza un AFD no paramétrico sobre un conjunto de grupos, en este caso olivar y suelo labrado. En el cálculo de las funciones de discriminación, las variables son elegidas sucesivamente, según su poder discriminante, utilizando el método de la F para dicha extracción. En cada caso se introduce la variable que más contribuye a la separación entre grupos.

Los resultados de estos análisis son los siguientes:

Usando todas las variables:

	Ventana 1	Ventana 2
%c.c. olivar	92.5	89.2
%c.c. suelo labrado	73.8	62.5

Variables discriminantes:

TX(B2)	TX(B1)
TX(IV1)	TX(B3)
TX(IV4)	TX(IV1)
RU1(B1)	TX(IV2)
RU1(B2)	RU1(B2)
RU1(IV2)	RU1(B4)
RU1(IV5)	RU2(B4)
RU2(B1)	
RU2(B2)	
RU2(IV1)	

Cuando se realizan los análisis separando textura y rugosidades, los resultados (% de clasificación correcta) son siempre superiores para la textura, por lo que a continuación se expone el análisis en el caso de utilizar las variables texturales:

	Ventana 1	Ventana 2
%c.c. olivar	91.1	87.1
%c.c. suelo labrado	60.0	60.5

Variables discriminantes:

TX(B1)	TX(B1)
TX(B2)	TX(B2)
TX(IV1)	TX(B3)
TX(IV4)	TX(IV2)

A la vista de estos resultados, pueden ya deducirse algunas cuestiones:

- La variable Textura, es decir, la desviación típica de los valores espectrales de la ventana, discrimina más que cualquiera de las dos rugosidades.
- En todos los casos, para el olivar, la ventana tipo 1, ventana 3x3, es más discriminante que la ventana tipo 2.

Posteriormente a la realización del Análisis Factorial Discriminante, se pasó un Análisis en Componentes Principales. El objetivo de este análisis no era la obtención de nuevas variables derivadas, sino que se trataba de observar como se distribuían los coeficientes de correlación de las variables con los ejes factoriales, de forma que pudiéramos estudiar de qué forma se agrupaban y con qué peso en los distintos ejes.

En las tablas siguientes se exponen los resultados de los distintos ACP llevados a cabo para olivar y suelo labrado, utilizando como variables las bandas MSS originales, los índices de vegetación, y las texturas calculadas sobre todas las anteriores.

Hay que significar que los Fi representan los ejes factoriales, y que en lugar de exponer los coeficientes de correlación correspondientes a cada variable en cada eje, se muestran las variables con

coeficientes más altos, con su signo respectivo, dentro de cada eje factorial. Los valores escritos al final de cada columna corresponden al porcentaje de la inercia acumulada absorbida por cada eje.

VENTANA 1

Olivar

F1	F2	F3	F4
IV4	IV2	TX(IV1)	TX(IV4)
B2	IV5	TX(IV5)	TX(B3)
B1	IV3	TX(IV2)	TX(B2)
B3			TX(B1)
-IV3			
41%	61%	72%	87%

Suelo labrado

F1	F2	F3
IV2	TX(IV2)	IV4
IV5	TX(IV1)	B1
IV1	TX(IV5)	
IV3	TX(B4)	
B4	TX(B2)	
-B2		
-B1		
51%	74%	87%

VENTANA 2

Olivar

F1	F2	F3	F4
IV4	IV2	TX(IV2)	TX(IV4)
B2	IV5	TX(IV1)	TX(B3)
B1	IV3	TX(IV5)	
B3			
B4			
-IV3			
38%	59%	71%	81%

Suelo labrado

F1	F2	F3
IV2	TX(IV2)	IV4
IV5	TX(IV5)	B1
IV3	TX(IV1)	
IV1		
B4		
47%	71%	84%

Centrándonos en el caso del olivar, podemos observar como, claramente, dos de los ejes son, por decirlo de alguna manera, espectrales (ejes 1 y 2), y los otros dos son texturales (ejes 3 y 4). Además, dentro de cada grupo, uno de los ejes refleja la influencia del brillo (ejes 1 y 4) y el otro la influencia del verdor o la biomasa (ejes 2 y 3).

Así pues, las variables con las que se decidió experimentar en el posterior proceso de clasificación fueron las siguientes:

- Banda MSS 1
- Banda MSS 2
- Banda MSS 3
- Banda MSS 4
- IV3
- IV4
- Textura sobre Banda MSS 1
- Textura sobre Banda MSS 2
- Textura sobre Banda MSS 3
- Textura sobre Banda MSS 4
- Textura sobre IV1
- Textura sobre IV4

Es decir, se trata de doce variables, seis de ellas texturales y seis de ellas espectrales simples o derivadas, existiendo dentro de cada grupo variables de brillo y variables de verdor.

EL PROCESO DE CLASIFICACION

Como ya se comentó anteriormente, ERAFIS permite clasificar hasta con doce variables. Sin embargo, dado que lo que se persigue en este estudio es obtener resultados comparables con el trabajo de González et al. (1987), en el que se emprendía el proceso de clasificación con cuatro variables (las cuatro bandas MSS) y, además, se trataba de no aumentar el esfuerzo de cálculo en la clasifica-

ción, se decidió seguir utilizando cuatro variables aunque, eso sí, elegidas de entre las doce seleccionadas anteriormente.

Para ello se construyó un nuevo programa llamado CREAMZCT, que se integró dentro del sistema ERAFIS. Como input utilizaba el archivo CCTINIA, que no es más que el archivo con los datos MSS correspondientes a la ventana de estudio, pero reformateado para que cumpla las exigencias del sistema ERAFIS (Moro et al., 1986), y en el que, a través de una serie de menús, se podían seleccionar cualquier grupo de cuatro variables de entre las doce anteriores. Como output de CREAMZCT se obtenía un nuevo archivo, llamado ZCTINIA (López S., 1986), idéntico en cuanto a formato a CCTINIA aunque conteniendo los valores de las variables seleccionadas, y con el que se podía continuar directamente con el proceso de clasificación supervisada implementado en ERAFIS.

Si bien este sistema de elección de grupos ofrecía la posibilidad de seleccionar variables de una forma casi combinatorial, la clasificación se llevó a cabo eligiendo los grupos de variables que, según los resultados de los análisis efectuados en el apartado anterior, ofrecían mayor consistencia metodológica.

Se ensayaron seis grupos de variables, presentándose a continuación las dos combinaciones con las que se obtuvieron los resultados más satisfactorios:

COMBINACION 1

Banda 1
Banda 4
Textura sobre IV4
IV3

COMBINACION 2

IV4
IV3
Textura sobre IV4
Textura sobre IV1

Con cada una de estas dos combinaciones de cuatro variables, se inició un proceso de clasificación en el que se siguieron los mismos pasos que en el estudio de González et al. (1987).

Los resultados de la clasificación muestral fueron, para cada una de las dos combinaciones de variables, y comparados con los resultados del estudio de referencia (González et al., 1987), los siguientes:

RESULTADOS DE CLASIFICACION

	Referencia	Comb1	Comb2
1. Cereales	45.65	41.35	32.89
2. Leg. regadío	89.62	85.61	83.21
3. OLIVAR	81.67	83.22	81.47
4. Encinar	59.94	54.10	59.25
5. S. LABRADO	43.74	50.60	45.10
6. Dehencpast	80.62	79.23	79.90
7. Dehencsula	64.34	57.06	76.61
8. Pinea	67.71	73.11	74.50
9. Castanea	70.08	62.79	74.21
10. Repopinast	38.77	24.56	22.79
11. Jaral	55.43	63.79	75.26
12. Matomezcla	29.66	27.03	21.64
13. Embalses	99.25	100	100

Asimismo, las distancias de Jeffries-Matusita que se producen para la pareja de clases de información olivar/suelo labrado, son las siguientes:

DISTANCIAS J-M OLIVAR/SUELO LABRADO

Referencia	Comb1	Comb2
.9206	1.0023	.9929

CONCLUSIONES

Como consecuencia de los análisis efectuados, pueden deducirse las siguientes conclusiones:

1. La utilización de variables derivadas sean del tipo índices de vegetación, o sean del tipo textural, mejora los resultados de clasificación de las clases olivar y suelo labrado. Así, mediante la combinación 1 se alcanzaron porcentajes de clasificación correcta del 83.22% y del 50.60% respectivamente, mientras que con la combinación 2 la clase olivar descendió en dos décimas y la clase suelo labrado aumentó en dos puntos hasta el 45.10%. En suma se aprecian ciertas mejoras, aunque en una las combinaciones el olivar descendiera ligeramente, sobretudo en los porcentajes de confusión existentes entre ambas clases. Dado que aumenta la distancia J-M entre ambas clases (con la combinación 1 se alcanza un aumento del 9%), mejora la discriminación entre los píxeles de ambas clases.

2. Si bien la selección de variables puede requerir

un cierto esfuerzo de cálculo, la generación de imágenes derivadas apenas supone un 5% más de tiempo, dentro del tiempo total de clasificación. Así pues, si bien existen otros métodos de incorporación de información espacial, como pueden ser los clasificadores de contexto (González et al., 1987), estos requieren un esfuerzo de cálculo mucho mayor, haciéndolos rentables únicamente en determinados casos, mientras que la utilización de imágenes derivadas con datos de textura introduce rapidez y sencillez en tiempos y en cálculos.

3. Una de las causas por las que los resultados de la clasificación experimentan incrementos no demasiado altos, puede deberse a la propia naturaleza de las imágenes MSS. De hecho, debido al tamaño del pixel MSS, que es mucho mayor que los objetos que contiene, la respuesta es muy integradora, produciéndose, en general, imágenes con una pequeña varianza local (Woodcock y Strahler, 1987). La utilización de sensores de alta resolución, especialmente para clasificar clases forestales o con cultivos arbolados, o clases urbanas y suburbanas, produce imágenes con una alta varianza local que, si bien utilizadas directamente pueden disminuir la exactitud de la clasificación cuando se utilizan clasificadores espectrales, mediante la generación de imágenes derivadas en las que se incorpore información espacial, podrá mejorarse notablemente su calidad para clasificar determinados usos del suelo como el olivar.

BIBLIOGRAFIA

Deering, D.W., 1978. Rangeland characteristics measured by aircraft and spacecraft sensors. Ph. D. Texas University College Station. 338 pp.

Cuevas, J.M., 1985. La teledetección espacial agraria. Presente y perspectivas en España. INIA, nota técnica No. 27. Madrid.

González, F., Cuevas, J.M., Moro, J., 1987. Aplicación del sistema informático ERAFIS al reconocimiento y discriminación de usos del suelo en la provincia de Toledo a partir de imágenes digitales MSS-Landsat 4. Publicaciones INIA (en prensa).

González, F., López S., Llop, R., Cuevas, J.M., 1987. Utilización de la información espacial de contexto en la clasificación de imágenes de satélite. Comunicación presentada en la segunda Reunión Nacional del GTT. Valencia, 1987.

Gordon, D.K., Philipson W.R., 1986. A texture-enhancement for separating orchard from forest in Thematic Mapper data. International Journal of Remote Sensing 1986, Vol.7, No.2, pp.301-304.

Haralick R.M., 1979. Statistical and structural approaches to texture. Proc. IEEE, 67, 786.

Kauth, R.Y., Thomas G.S., 1976. The tasseled cap description of the spectral-temporal development of agricultural crops as seen by Landsat. Proc. Symp. on Machine Processing of Remotely Sensed Data. Purdue University, pp.41-51.

López Soria, S., 1986. Detección de cambios en el monte de El Pardo por regresión de imágenes de satélite Landsat. Proyecto Fin de Carrera. Escuela Técnica Superior de Ingenieros de Montes. Madrid.

Moro, J., González, F., Cuevas, J.M., 1986. Introducción del sistema ERAFIS (Estimación de Recursos Agrícolas y Forestales mediante Imágenes de Satélite). Monografías INIA, No.54, Madrid.

Woodcock, C.S., Strahler, A.H., 1987. The factor of scale in Remote Sensing. Remote Sensing of the Environment. Vol.21, No.3.