

CLASIFICACIÓN ESPECTRAL AUTOMÁTICA VS. CLASIFICACIÓN VISUAL: Un ejemplo al sur de la Ciudad de México*

José Luis Palacio Prieto **
Laura Luna González **

Resumen

Se realizaron dos clasificaciones multiespectrales, una supervisada y una no supervisada, con base en una imagen Landsat TM utilizando para ello un algoritmo de máxima similitud. En el primer caso se obtuvieron 29 clases a partir de unas 41 muestras y en el caso de la clasificación no supervisada se obtuvieron 27 clases. En ambos casos, el número final de clases se redujo una vez agrupadas las clases espectrales en clases de información, resultando seis en total. Por otra parte, se elaboraron diferentes compuestos en color para realizar una interpretación visual.

Los tres productos fueron comparados en un ambiente SIG contra una base de referencia de verdad en campo, consistente en una malla de puntos equidistantes a 1 km², totalizando 560 sitios de control.

Los resultados de la comparación permiten apreciar que los mejores valores de exactitud corresponden a la clasificación supervisada (82.32%) seguidos de la clasificación visual (78.72%) y la no supervisada (73.18%). Estos valores fueron obtenidos una vez agrupadas las clases afines.

Summary

Using a Maximum Likelihood algorithm a Landsat TM image was classified by both supervised and non-supervised approaches. In the first case, 12 classes were obtained based on 30 samples; the non-supervised procedure yielded 30 classes. Once grouped, both classifications considered 6 classes. Additionally, color composites were prepared and visually interpreted.

The three products were compared in a GIS environment using a regularly distributed network of points referring the field truth.

The results show that the lowest error correspond to the supervised classification (82.32% exactitude), followed by the visual interpretation (78.72%) and the non-supervised procedure (73.18%). These figures were obtained after grouping the classes according to their similarities.

* Recibido: 10 de noviembre de 1994.

** Instituto de Geografía, Universidad Nacional Autónoma de México, Apdo. Postal 20-850, 01000 México, D. F.

Introducción

Entre las aplicaciones más frecuentes del procesamiento digital de imágenes multiespectrales de satélite se encuentra la generación de cartografía relacionada con la identificación y estado de los diferentes tipos de cubierta del terreno. Éstos son diferenciados con base en su respuesta en distintas porciones del espectro electromagnético, y clasificados a partir de sus afinidades. De esta manera, se definen clases espectrales para las cuales existe una correspondiente clase de información o atributo, que identifica a él o los objetos contenidos en la mencionada clase espectral. Si bien desde el punto de vista teórico el proceso parece sencillo, existen varios problemas en la práctica.

Uno de los problemas comunes del procesamiento de imágenes multiespectrales con fines de clasificación de tipos de cobertura es la confusión de respuestas espectrales y de atributos. Así, objetos diferentes pueden poseer la misma o muy parecida respuesta, con lo cual pasan a formar parte de la misma clase espectral; igualmente, un objeto puede tener un comportamiento espectral diverso, con lo cual una clase de información individual, o atributo, puede corresponder a diferentes clases espectrales. Lo anterior introduce, finalmente, incertidumbre en los resultados obtenidos en una clasificación espectral, sea ésta supervisada o no supervisada.

La confusión espectral interclases es causante de que las clasificaciones automáticas posean cierto error, que en todo trabajo debe ser evaluado. Dicho error depende de variables tales como la resolución de la imagen utilizada, tipo de la misma, algoritmos de clasificación empleados, la variabilidad espacial de las coberturas de terreno y el número y naturaleza de las clases a determinar, entre otros aspectos.

Por otra parte, la interpretación visual de imágenes mejoradas (p. e. compuestos de color, imágenes obtenidas de procedimientos estadísticos como análisis de componentes principales o bien cocientes de bandas) por parte de personal experimentado y con el indispensable conocimiento del terreno, permite igualmente la definición de clases al involucrar en el análisis la evaluación y diferenciación de patrones espaciales de los atributos, mismos que en lo espectral muestran semejanzas tales que no pueden ser individualizados. La **tabla 1** refiere algunos resultados referentes a la exactitud de clasificaciones automáticas y visuales en los que puede apreciarse la variación de los valores de exactitud en función del tipo de imagen, resolución de la misma y clases consideradas.

El presente trabajo tiene como objetivo principal el de comparar algunos métodos automáticos de clasificación espectral (supervisada y no supervisada, a partir de un algoritmo de máxima similitud) y la interpretación visual para la definición de tipos de cubierta del terreno, en un área del centro de México.

Características generales de la zona estudiada

La zona en estudio se localiza al suroeste de la Ciudad de México y cubre cerca de 600 km² (**figura 1**). Sus coordenadas extremas son los 19°00' y 19°15' de latitud norte y los 99°10' y 99°20' de longitud oeste. Destacan como los poblados más importantes San Miguel Ajusco, al norte, y Tres Marías, al sur. Las elevaciones fluctúan entre los 2 700 y los 4 000 msnm; las máximas elevaciones están representadas en los volcanes Ajusco, al norte; El Pelado, al centro; Chichinautzin, al oriente, y el complejo de la Sierra de Zempoala hacia el suroeste.

Tabla 1. Exactitud de clasificaciones con base en diferentes tipos de imágenes. En la mayoría de los casos se utilizaron algoritmos de máxima similitud. (*) Interpretación visual de imágenes

Autor	Tipo de imagen	Atributos	Exactitud (%)
Conese y Maselli (1991)	Landsat TM	Varias coberturas	58
Gomasasca <i>et al.</i> (1993)	Landsat TM	Ciudad y entorno	80
Nageswara y Mohankumar (1994)	Landsat MSS, TM y IRS-1A/LISS-II	Cultivos varios	de 56.25 a 97
Kartikeyan <i>et al.</i> (1994)	IRS-1A/LISS-II	Varias coberturas	de 77 a 98
Kai-Yi y Mausel (1993)	Video multiespectral	Cultivos varios	82.23
Franklin y McDermid (1993)	SPOT	Bosque de Pino	59
Nelson y Horning (1993)	NOAA-AVHRR	Bosque-y no Bosque	74.7
Gastellu-Etchegorry <i>et al.</i> (1993)	NOAA-AVHRR	Bosque tropical	71
Nezry <i>et al.</i> (1993)	SAR espacial	Vegetación tropical	89.9
Joshi y Sahai (1993)*	Landsat MSS Landsat TM	Terrenos salinos	74* 90*

En buena medida, la distribución del uso del suelo y la vegetación guarda una estrecha relación con las características geológicas y geomorfológicas de la zona. La agricultura se encuentra restringida a las depresiones aluviales desarrolladas entre los derrames lávicos cuaternarios y a algunos piedemontes lávicos cubiertos por materiales piroclásticos. Por otra parte, los derrames lávicos recientes, frecuentemente carentes de un suelo desarrollado, se encuentran cubiertos por bosques de pino y zacatonal; son importantes las extensiones

cubiertas de bosques de oyamel en las faldas del Ajusco y la Sierra de Zempoala y sus alrededores.



Figura 1. Localización de la zona estudiada.

Materiales y métodos

Se contó con una imagen Landsat TM (1 125 líneas por 889 columnas, bandas 1, 2, 3, 4, 5 y 7) con fecha de octubre de 1991. Como material auxiliar se dispuso de las cartas de Uso del Suelo y Vegetación, Geológica y Topográfica (Hoja Milpa Alta, E14-A49, escala 1:50 000) editadas por el Instituto Nacional de Estadística, Geografía e Informática (INEGI) en 1972, 1976 y 1983. Igualmente, se contó con la carta topográfica editada en 1986 por Sistemas de Información Geográfica, S. A. (SIGSA) en escala 1:20 000; así también, se utilizaron mapas temáticos relacionados con vegetación y paisajes naturales elaborados en diferentes escalas por Velázquez (1993).

La imagen fue georreferida con base en las coordenadas UTM de 25 puntos de control extraídos de la cartografía topográfica. El valor de sigma para la georreferencia fue de 0.7 pixeles de error en la ubicación de coordenadas.

Se generaron imágenes correspondientes a los tres primeros componentes principales, los cuales fueron derivados de las seis bandas disponibles. Los valores correspondientes a los eigenvectores permitieron seleccionar las bandas más representativas para realizar las clasificaciones espectrales supervisada y no supervisada.

Con el fin de realizar el muestreo para la clasificación supervisada se prepararon diferentes compuestos de color generados a partir de imágenes mejoradas por filtrados pasa-altas (realce de bordes). Una vez seleccionados los compuestos, se realizó un muestreo dirigido con base en la interpretación visual, en la información auxiliar y en reconocimientos en campo. Las clases definidas sirvieron de base para la clasificación supervisada de la imagen, misma que fue elaborada a partir de un algoritmo de máxima similitud.

Para la clasificación no supervisada se tomó como base una matriz de seis columnas por cuatro líneas para toda la escena con objeto de definir los "clusters" o cúmulos iniciales; se estableció como máximo un número de 30 clases espectrales probables resultantes. Se usó igualmente un algoritmo de máxima similitud para obtener la imagen clasificada.

Los compuestos en color (453, 752 y 432) fueron interpretados visualmente por personal experimentado.

Posteriormente se realizó un control de la exactitud de las imágenes clasificadas automática y visualmente a partir de la verificación de información en puntos distribuidos de manera uniforme en una matriz de 1 km²; se usaron un máximo de 560 puntos de verificación en total, de los cuales se eliminaron aquéllos que por su ubicación en áreas limítrofes entre diferentes usos del suelo y/o vegetación introdujeran la posibilidad de confusión.

Finalmente, se hizo una evaluación de resultados para cada uno de los procedimientos utilizados. El procesamiento de las imágenes y la comparación con los sitios de verificación fue llevado a cabo en un sistema basado en computadora personal (ILWIS, 1993).

Resultados y discusión

Los primeros tres componentes principales obtenidos contienen el 98.44% del total de la varianza original de la escena (76.5, 18.6 y 3.3%, respectivamente). En la **tabla 2** se presenta la matriz de eigenvalores de la escena estudiada.

Tabla 2. Eigenvalores correspondientes a los tres primeros componentes principales de la escena

	Comp. 1	Comp. 2	Comp. 3
B1	0.187	-0.108	0.748
B2	0.148	-0.085	0.280
B3	0.188	-0.190	0.342
B4	0.519	0.837	0.079
B5	0.731	-0.344	-0.462
B7	0.323	-0.355	0.157

El primer componente presenta un alto contenido de información espectral de las bandas infrarrojas 4, 5 y 7, como puede observarse en la matriz. El segundo componente revela una importante aportación de información de la banda 4 y el tercero contiene información en su mayor parte de la banda 1 del visible. Con base en los eigenvalores, las bandas seleccionadas para realizar la clasificación multiespectral fueron, por tanto, 1, 4 y 5.

De los compuestos de color preparados, la combinación 453 permite distinguir claramente las formas del relieve, los bosques y los cultivos. Los cultivos de la zona de forrajeros (avena), de tal manera que contar con un compuesto que distinga, por ejemplo, cultivos forrajeros de pastizales inducidos y, a la vez, bosques de coníferas entre sí (pino-oyamel), es favorable para el muestreo de campos de entrenamiento, a diferencia de otros compuestos de color.

- Clasificación no supervisada

El proceso generó 27 clases considerando como clase el valor cero (píxeles no clasificados). Una vez reagrupadas, el número de clases finales fue de siete como se muestra en la **tabla 3**, que reúne las clases finales de las clasificaciones, los porcentajes para cada una de ellas y la diferencia entre ambas.

- Clasificación supervisada

A partir de 41 muestras se definieron 29 clases para toda la escena. Al igual que en la clasificación no supervisada, se generó una nueva imagen producto de la reagrupación de las clases resultantes, con el fin de establecer las comparaciones con la otra clasificación (**Tabla 3**).

Tabla 3. Resultados generales de las clasificaciones supervisada y no supervisada

	Clases finales	NO SUP %	SUP %	DIF %
0	sin clasificar	20.6	33.6	13.0
1	bosque de oyamel	6.2	9.6	3.4
2	bosque de pino	33.4	22.6	10.8
3	bosque de pino abierto/pastizal	20.9	16.3	4.6
4	bosque de pino/encino	0.0	5.3	5.3
5	pastizal	12.8	4.1	8.7
6	cultivo/pastizal	6.1	5.8	0.3
7	cultivo	0.0	2.3	2.3
	TOTAL	100.0	99.6	-

Puede apreciarse, en general, que el porcentaje de incertidumbre representado por el número de píxeles no clasificados (valor cero) es mayor en la clasificación supervisada, lo cual puede denotar falta de muestreo. En la clasificación no supervisada, por otra parte, no fueron diferenciados los cultivos, los cuerpos de agua y el bosque de pino-encino, que al ser integrados a otras clases explican las diferencias en los porcentajes obtenidos en ambos casos.

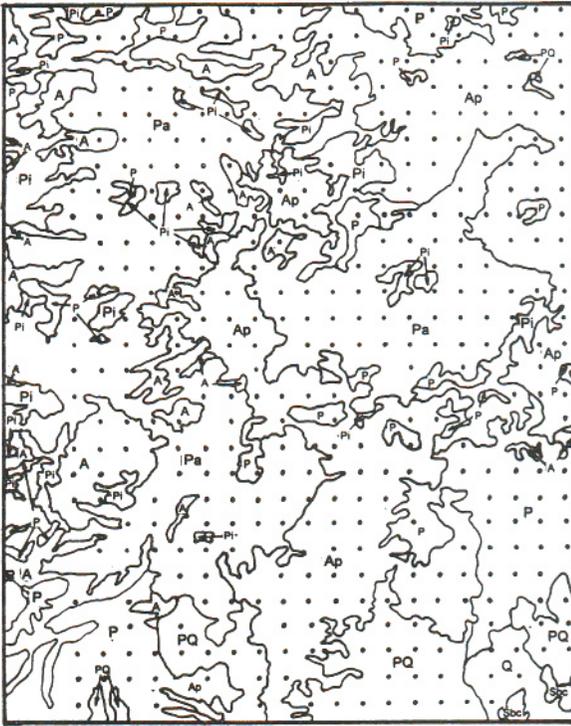
- Clasificación visual

Los compuestos de color fueron interpretados por un técnico experto. Las clases definidas (originalmente 9) fueron agrupadas para hacerlas compatibles con las obtenidas por procesos automáticos. La agrupación de clases fue hecha por los autores de este escrito a partir de trabajo de campo y material cartográfico auxiliar. La **figura 2** muestra el resultado de la clasificación visual, y proporciona una aproximación acerca de los diferentes usos del suelo y vegetación en la zona.

- Evaluación de la exactitud de las imágenes clasificadas por medios automáticos y visuales

Con objeto de evaluar la exactitud de las clasificaciones se reagruparon las clases en un total de cinco. Las clases resultantes fueron comparadas con objeto de obtener una matriz de error para cada uno de los tres mapas finales (**Tabla 4**).

Los resultados referidos en la tabla anterior permiten apreciar la dificultad de diferenciar algunas clases debido a su comportamiento espectral y apariencia visual. En los tres procedimientos, la respuesta espectral y la apariencia permiten individualizar aceptablemente al bosque de pino, con exactitud superior al 80%; sin embargo, el bosque de pino/oyamel difícilmente puede ser diferenciado, al igual que el bosque de pino/encino, en



LEYENDA	
A	BOSQUE DE OYAMEL
P	BOSQUE DE PINO
Pa	BOSQUE DE PINO ABIERTO
PQ	BOSQUE DE PINO-ENCINO
Q	BOSQUE DE ENCINO
Pi	PASTIZAL INDUCIDO
Sbc	SELVA BAJA CADUCIFOLIA
Ap	AGRICULTURA DE TEMPORAL

Figura 2. Uso del suelo (resultado de la interpretación visual).

donde la exactitud llega a ser de cero en el caso de la clasificación no supervisada. La agricultura tiende a confundirse con pastizales e incluso es probable que parte de la confusión espectral y visual con el bosque de pino se derive de los pastos que se intercalan en los bosques abiertos.

Tabla 4. Matrices de error de las clasificaciones supervisada, no supervisada y visual

Clasificación supervisada	Bosque de Oyamel	Bosque de Pino	Bosque de Pino/Encino	Agricultura/Pastizal	Bosque de Pino/Oyamel
Bosque de Oyamel	71.4	28.6	0	0	0
Bosque de Pino	1.2	92.7	1.2	5.9	0
Bosque de Pino/Encino	9.3	46.8	34.3	9.4	0
Agricultura/Pastizal	0	19.0	15.8	65.1	0
Bosque de Pino/Oyamel	16.7	83.3	0	0	0

Clasificación no supervisada	Bosque de Oyamel	Bosque de Pino	Bosque de Pino/Encino	Agricultura/Pastizal	Bosque de Pino/Oyamel
Bosque de Oyamel	39.1	60.9	0	0	0
Bosque de Pino	1.1	90.5	0	8.4	0
Bosque de Pino/Encino	8.6	71.8	0	0	21.8
Agricultura/Pastizal	0	28.6	0	71.4	0
Bosque de Pino/Oyamel	3.4	89.7	0	6.9	0

Clasificación visual	Bosque de Oyamel	Bosque de Pino	Bosque de Pino/Encino	Agricultura/Pastizal	Bosque de Pino/Oyamel
Bosque de Oyamel	66.7	29.2	0	4.1	0
Bosque de Pino	5.9	80.4	1.0	12.7	0
Bosque de Pino/Encino	2.2	32.6	50.0	15.2	0
Agricultura/Pastizal	2.1	23.6	1.4	72.9	0
Bosque de Pino/Oyamel	0	90.0	0	10.0	0

Los resultados de exactitud global de las tres clasificaciones se refieren en la **tabla 5**. Como puede apreciarse, los valores originalmente obtenidos refieren exactitudes entre 40 y 54%, lo cual es bajo. Sin embargo, si se reagrupan las clases afines en tres grupos (véase **figura 3**), los resultados son considerablemente mejores y coherentes con lo esperado en una zona con cubiertas más o menos homogéneas y relativamente poco impactadas por acción humana,

como sucede en el área estudiada (Tabla 5). Las tres clases finales referidas en la figura 3 permiten, si bien con menor diferenciación, una exactitud superior a 73% y hasta 82% en el caso de la clasificación supervisada. En el caso de la clasificación visual los resultados del 78.72% obtenidos son aceptables, si bien las expectativas eran en un principio mayores; sin embargo, la clasificación visual tuvo que sacrificar uno de sus mejores potenciales, que fue el de llegar a diferenciar entre clases (como entre agricultura y pastizales) a partir de la interpretación de patrones del uso del suelo. Igualmente, la clasificación visual tuvo que generalizar clases que eran detectables por la apariencia más no por su respuesta espectral, como el caso de ambientes semi-urbanos y algunos cuerpos de agua de escasas dimensiones, los cuales no son discernibles en lo espectral.

Tabla 5. Exactitud global de las clasificaciones

	Núm. de ptos. de verificación	% de exactitud global sin agrupar	% de exactitud global agrupada
supervisada	222	52.7	82.32
no supervisada	259	40.18	73.18
visual	342	54.0	78.72

Conclusiones

Los mejores resultados de exactitud fueron logrados a partir de un procedimiento supervisado de clasificación espectral, seguidos por la interpretación visual y, por último, por medios no supervisados.

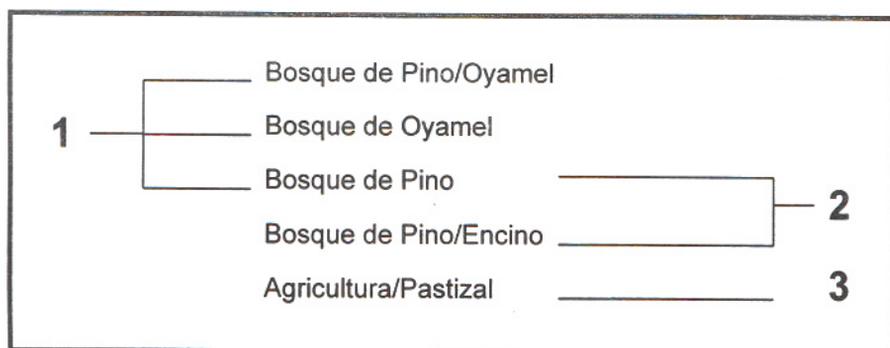


Figura 3. Reagrupación de clases afines para obtener tres clases finales.

Con el fin de lograr resultados aceptables, las clases debieron ser agrupadas en un total de tres, lo cual refleja la limitación de este tipo de imágenes en la diferenciación de clases espectrales muy cercanas, que de ser individualizadas incrementan el error de manera considerable, en cualquiera de los tres métodos seguidos.

No obstante lo anterior, y aunque los resultados sitúan al procedimiento visual en segundo sitio, se debe destacar que el procedimiento rinde resultados perfectamente comparables con los automáticos, considerados tradicionalmente como más precisos por su falsa apariencia derivada de procesos aparentemente objetivos, y que muy probablemente algún intérprete con amplia experiencia y, sobre todo, gran conocimiento de campo, puede llegar a superarlos.

Reconocimiento

La interpretación visual fue realizada por el Biól. Cutberto Carrillo, a quien se agradece su colaboración.

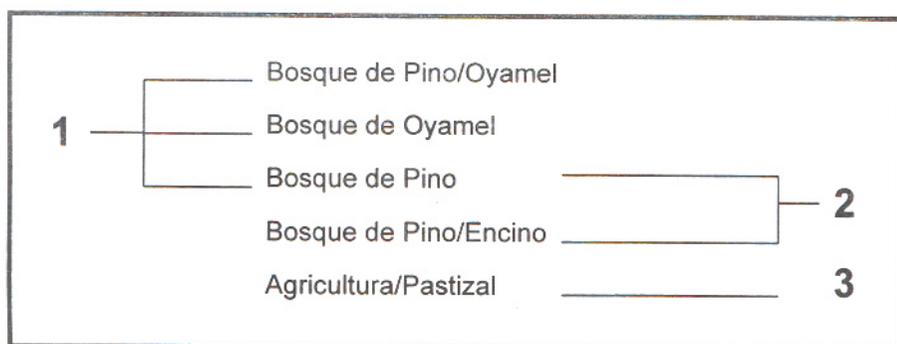


Figura 3. Reagrupación de clases afines para obtener tres clases finales.

Con el fin de lograr resultados aceptables, las clases debieron ser agrupadas en un total de tres, lo cual refleja la limitación de este tipo de imágenes en la diferenciación de clases espectrales muy cercanas, que de ser individualizadas incrementan el error de manera considerable, en cualquiera de los tres métodos seguidos.

No obstante lo anterior, y aunque los resultados sitúan al procedimiento visual en segundo sitio, se debe destacar que el procedimiento rinde resultados perfectamente comparables con los automáticos, considerados tradicionalmente como más precisos por su falsa apariencia derivada de procesos aparentemente objetivos, y que muy probablemente algún intérprete con amplia experiencia y, sobre todo, gran conocimiento de campo, puede llegar a superarlos.

Reconocimiento

La interpretación visual fue realizada por el Biól. Cutberto Carrillo, a quien se agradece su colaboración.

Cartografía consultada:

INEGI (1972), Carta de Uso del Suelo, Milpa Alta, E14-A49, esc. 1:50 000.

INEGI (1976), Carta Geológica, Milpa Alta, E14-A49, esc. 1:50 000.

INEGI (1983), Carta Topográfica, Milpa Alta, E14-A49, esc. 1:50 000.

SIGSA (1986), Carta Topográfica, esc. 1:20 000.