

## Identificación automática de zonas de manejo en lotes productivos agrícolas

Romina Galarza<sup>1</sup>, M. Nicolás Mastaglia<sup>1</sup>, Enrique M. Albornoz<sup>1,3</sup> y César Martínez<sup>1,2</sup>

<sup>1</sup> Centro de Investigación en Señales, Sistemas e Inteligencia Computacional  
(SINC(i))

Dpto. Informática, Facultad de Ingeniería - Universidad Nacional del Litoral  
CC217, Ciudad Universitaria, Paraje El Pozo, S3000, Santa Fe, Argentina

<sup>2</sup> Laboratorio de Cibernética, Fac. de Ingeniería-Universidad Nacional de Entre Ríos  
<sup>3</sup> CONICET, Argentina

rominavgalarza@gmail.com, nmastaglia@gmail.com,  
emalbornoz@fich.unl.edu.ar y cmartinez@fich.unl.edu.ar

**Resumen** La Agricultura de Precisión provee un conjunto de principios y herramientas que permiten manejar la variabilidad espacio-temporal en la producción agrícola. Dentro de sus incumbencias está la identificación de zonas de manejo en un lote productivo. Éstas son subregiones de un lote, que poseen características homogéneas y para las cuales una proporción de insumos resulta apropiada. El método desarrollado consiste en un proceso integrador que comprende toda la cadena de identificación y clasificación de zonas de manejo. En primer lugar, permite fusionar en una estructura común, datos de entrada georreferenciados provenientes de distintos sensores y con diferentes resoluciones. Luego, un algoritmo de agrupamiento de lógica difusa permite identificar y agrupar diferentes regiones. Mediante procesamiento de imágenes se mejoran los resultados a fin de identificar los bordes de las zonas de manejo. Finalmente, el método provee un archivo *shape* que es utilizado en las máquinas agrícolas con aplicación de dosis variable.

**Palabras claves:** Zonas de manejo, clustering, procesamiento de imágenes, segmentación, archivos *shape*.

### 1. Introducción

El cultivo sembrado en un campo registra un comportamiento dispar a lo largo de toda su extensión debido a múltiples factores que influyen de diferente manera en el crecimiento y posterior rendimiento del mismo. Identificar las áreas que dentro de un mismo lote se comportan de manera homogénea posibilita un adecuado manejo y un uso eficiente de los insumos (fertilizantes, semillas, tierra, tiempo, etc). De esta manera, el productor puede obtener mayores rendimientos y/o maximizar el beneficio económico al contar con datos que permitan realizar un manejo diferenciado, poniendo más énfasis en las zonas que poseen un mayor potencial productivo [1].

El continuo avance y perfeccionamiento tecnológico ha posibilitado que cualquier sembradora o pulverizadora disponible en el mercado cuente con un sistema tecnológico de dosis variable que posibilita, mediante una computadora incorporada, indicar cuánto fertilizante usar y dónde se lo debe aplicar. Empresas como John Deere y Case/New Holland implementan en sus maquinarias los sistemas integrales AMS y AFS respectivamente, que utilizan métodos para la clasificación que requieren un número preestablecido de ambientes. Para esto se necesita tener un conocimiento previo de las características del lote, lo que puede incurrir en el error de sobre o subestimar la cantidad de zonas [2,3].

En el mercado informático existe software comercial de manejo de información agrícola, como el caso de SSToolbox<sup>4</sup> y Farm Works<sup>5</sup>. Este último muestra detalladamente la representación gráfica de las distintas variables analizadas, pero carece de la posibilidad de realizar un análisis en profundidad sobre las posibles zonas de manejo que se encuentran dentro de un lote [4]. El SSToolbox posee diversas aplicaciones agronómicas, entre ellas generación, procesamiento y manejo de información georreferenciada, análisis de la misma y generación de recomendaciones de manejo agronómico. Es un software complejo que delimita las zonas de manejo únicamente por posición de paisaje [5].

En la Estación Experimental Agropecuaria Paraná del INTA (Entre Ríos), el proceso de delimitación de zonas de manejo se lleva a cabo actualmente en varias etapas. Algunas de las mismas constituyen un laborioso trabajo manual y se realiza con el apoyo de software diverso, entre los cuales podemos nombrar: MZA (Management Zone Analyst) [6], gvSIG, procesador de textos, procesador de cálculos. El software gratuito MZA sólo realiza la clasificación e identificación de las sub-regiones, lo cual implica el uso de otros programas para el pre y post procesamiento de los resultados. Esto conlleva a una compleja y tediosa tarea para arribar al resultado final.

En este trabajo se presenta el desarrollo de un método que automatiza las etapas mencionadas, desde el ingreso de los datos georreferenciados hasta la identificación de los distintos ambientes presentes en un lote. Se pretende con esta herramienta unificar o minimizar el uso de diferentes programas a lo largo de todo el trabajo. Además, se desea aportar al sector agrícola y a los usuarios de la Agricultura de Precisión una herramienta sencilla y eficiente que posibilite el uso de los datos georreferenciados de manera inteligente, sugiriendo la aplicación justa de insumos en el lugar exacto. De esta manera se podría maximizar la productividad y la rentabilidad de cada ambiente con sustentabilidad, mejorando la gestión del campo y el cuidado de los recursos naturales y el medio ambiente.

El resto del trabajo se organiza como se detalla a continuación. La Sección 2 describe el diseño y desarrollo del método propuesto. La Sección 3 presenta los experimentos y resultados obtenidos con un lote en particular. Finalmente, en la Sección 4 se encuentran las conclusiones y trabajos futuros.

---

<sup>4</sup> <http://www.sstsoftware.com/ssttoolbox.htm>

<sup>5</sup> <http://www.farmworks.com>

## 2. Método propuesto

El método se divide en varias etapas consecutivas. Se utilizan conjuntos de datos multiparamétricos que son obtenidos a través de diversos sensores, tales como imágenes satelitales o variables medidas en el campo (conductividad eléctrica, altimetría, rendimiento, etc.). Esto obliga a fusionar la información de manera tal que se adecue a una estructura común mediante técnicas de interpolación, las cuales establecen los valores que no han sido obtenidos experimentalmente.

El método busca identificar y agrupar los elementos que poseen características similares a través de la utilización del algoritmo de clustering de lógica difusa *Fuzzy C-Means* [7]. El mismo será validado mediante indicadores que midan la calidad de la clasificación realizada.

Mediante técnicas de procesamiento digital de imágenes se filtran los resultados obtenidos en el clustering y se eliminan las zonas que posean un tamaño menor a una determinada superficie. Luego se detectan los bordes de cada zona, para finalmente crear un archivo *shape*, formato estándar que reconocen las máquinas agrícolas dedicadas a aplicación de dosis variable en un campo.

Se utiliza el lenguaje de programación C++ para la impletación del método y además la librería QT<sup>6</sup> para el desarrollo de la interfaz gráfica del prototipo.

### 2.1. Definición de los datos de entrada

Los datos de entrada del método se pueden obtener utilizando diversas tecnologías tales como la teledetección (por ejemplo, imágenes tomadas por satélites) o medición de muestras en el lote (por ejemplo, rendimiento, altimetría, conductividad eléctrica, u otra). Estas mediciones conducen a conjuntos de datos heterogéneos con diferentes representaciones digitales y resoluciones, las cuales deben ser combinadas para crear un único modelo espacial de la grilla de puntos en estudio.

Un inconveniente en el manejo de datos es que cada fabricante de máquina agropecuaria incorpora diversos programas propietarios que originan archivos de formatos particulares. Además de esto, se debe sumar el inconveniente de que algunos programas también suelen cambiar con los años y esto podría generar problemas de compatibilidad. Estos programas permiten exportar la información en diversos formatos para que pueda ser utilizada por otras personas y sistemas. Entre los formatos estándares más usados se encuentran *txt*, *dat* y *csv*.

Por lo tanto, se definen dos tipos de datos de entrada. Por un lado, las variables que fueron obtenidas mediante la medición de puntos, generalmente distribuidos irregularmente en el espacio. Por otro lado, las imágenes GeoTiff que se basan en una representación raster del terreno con una distribución regular de los puntos (píxeles).

---

<sup>6</sup> <http://qt-project.org/>

### **Formato de los archivos de entrada.**

Se utilizan como datos de entrada los archivos *txt*, *dat* y *csv* que en su interior contengan el valor de la coordenada (latitud y longitud con Datum WGS84) y el valor de la variable en cada punto, separados por coma.

Cada uno de estos archivos contiene las muestras de una variable en particular y se puede ingresar tantos archivos como variables se tenga. Los datos dentro de cada archivo se ordenan de la siguiente manera: *Longitud, Latitud, Valor*.

En cuanto al procesamiento de imágenes, se trabaja únicamente con imágenes georeferenciadas debido a la necesidad de contar con información de la localización espacial de los píxeles. Este tipo de imágenes se representa mediante el formato *tiff* (GeoTiff, es decir *tiff* con coordenadas).

### **2.2. Transformación de coordenadas**

Se cuenta con dos grupos diferentes de datos de entrada, por un lado están las imágenes georeferenciadas con formato GeoTiff representadas mediante coordenadas planas y por otro los datos extraídos de los GPS, que se encuentran representados en coordenadas geográficas.

Aunque la utilización de las coordenadas geográficas resulta de gran utilidad en geodesia, su uso resulta incómodo para emplearlo en determinadas aplicaciones. Teniendo en cuenta que estas coordenadas se expresan en unidades angulares (grados, minutos y segundos de latitud y longitud), cuando se pretende determinar distancias entre puntos o direcciones definidas por éstos, se presentan limitaciones [8].

En efecto, las coordenadas geográficas necesitan una transformación a coordenadas planas para que todos los datos de entrada queden expresados en un mismo sistema geográfico. Esto permite operar rápida y fácilmente sobre la distancia de los puntos.

Para pasar de coordenadas geográficas a coordenadas planas se elige la proyección Gauss-Kruger [9]. En Argentina se la adoptó a partir de 1925 [10], dado que en territorios extendidos en dirección Norte-Sur son adecuadas las proyecciones cilíndricas transversales.

La rutina desarrollada toma como entrada los archivos descritos en la sección anterior, transforma las coordenadas (latitud-longitud) y obtiene como salida un archivo que posee las coordenadas planas (norte-este en metros) junto al valor del punto muestreado.

### **2.3. Fusión de los datos de entrada**

Esta etapa tiene como objetivo unificar el mallado de todas las variables que intervienen en el proceso. Para ello, se interpolan sus valores de manera que se tenga una sola grilla espacial común y que cada punto contenga el valor de todas las variables.

Se crea una grilla espacial con una distribución equiespaciada de sus puntos. La densidad de los mismos queda determinada con el ingreso de un parámetro

por parte del usuario. Luego se procede a interpolar los nuevos puntos a partir de la generación de una red irregular de triángulos cuyos vértices coinciden con los puntos muestreados. Dicha malla se obtiene mediante la triangulación de Delaunay. La interpolación de puntos dentro de la topología obtenida se realiza suponiendo que dichos puntos pertenecen a la superficie plana de primer orden que se apoya en los vértices de cada triángulo [11]. Sólo se calculan aquellos puntos del mallado regular que caen dentro de la envolvente convexa que forma cada variable. Para interpolar se puede elegir entre los métodos de interpolación que proporciona la librería CGAL [12]: interpolación Lineal, Cuadrática, Farin, Sibson y Sibson con raíz cuadrada.

La salida de esta etapa consiste en un archivo (formato separados por coma) que contiene todos los puntos del mallado regular. Cada uno consta de su respectiva ubicación espacial  $(x,y)$  y los valores que posee cada variable, obtenidos en la interpolación.

#### 2.4. Clasificación de las variables

Una vez obtenido el archivo interpolado se comienza con el proceso de clasificación e identificación de las zonas de manejo presentes en el lote productivo analizado.

El análisis de agrupamiento o *clustering* se define como la tarea de agrupar un conjunto de objetos en grupos, de modo que los objetos en el mismo grupo son más similares entre sí (en algún sentido) que al resto de los grupos.

Fuzzy C-Means (FCM) es un método de agrupamiento del estado del arte en reconocimiento de patrones [13,14] que ha demostrado su eficacia en este tipo de tareas, actualmente utilizado en programas como MZA. El mismo permite que un dato pertenezca a dos o más grupos y se basa en la minimización de la siguiente función objetivo:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, \quad (1)$$

donde  $1 < m < \infty$  es un índice de ponderación difuso que determina la falta de claridad de los grupos,  $x_i$  es el  $i$ -ésimo dato del conjunto medido,  $u_{ij}^m$  es el grado de pertenencia de  $x_i$  en el grupo  $j$ ,  $c_j$  es el centro del grupo  $d$ -dimensional,  $N$  el número de puntos del conjunto de datos,  $C$  la cantidad de grupos y  $\|\cdot\|$  es cualquier norma que exprese la similitud entre los datos medidos y el centro.

FCM se lleva a cabo a través de un proceso iterativo de optimización de la función objetivo (1), con la actualización de la matriz de pertenencia  $u_{ij}$  y los centros del grupo  $c_j$  por:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \quad (2)$$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}. \quad (3)$$

El proceso se detiene cuando  $\max_{ij} \{|u_{ij}^{k+1} - u_{ij}^k|\} < \varepsilon$ , donde  $\varepsilon$  es un criterio de finalización entre 0 y 1, mientras que  $k$  es la iteración. Este procedimiento converge a un mínimo local.

Todo el proceso de clasificación se realiza configurando los siguientes parámetros: exponente difuso, criterio de convergencia, máximo número de iteraciones y cantidad mínima y máxima de grupos. Además, se calculan tres índices que permiten al usuario realizar una evaluación acerca de la calidad de cada grupo. Estos se pueden dividir en dos categorías, la primera utiliza sólo los valores de pertenencia de la matriz  $u_{ij}$  de la partición de datos difusa ( $PC$  y  $PE$ ). El otro grupo implica tanto la matriz  $u_{ij}$  como el conjunto de datos ( $XB$ ), los cuales se explican a continuación.

El coeficiente de partición  $PC$ , se define como [15]:

$$PC = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C u_{ij}^2, \quad (4)$$

donde los agrupamientos encontrados tienen mayor separación a medida que  $PC \rightarrow 1$ .

El coeficiente de entropía de la partición  $PE$  es otro índice de esta categoría. Se define como [15]:

$$PE = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C u_{ij} \cdot \log_a u_{ij}, \quad (5)$$

donde el grupo obtenido es más definido cuando  $PE \rightarrow 0$ .

Sea  $\sum \sigma_i$  la *variación total* del conjunto de datos, donde  $\sigma_i$  es la varianza del  $i$ -ésimo conjunto, dada por:

$$\sigma_i = \sum_{j=1}^N d_{ij}^2, \quad (6)$$

donde  $d_{ij}$  es

$$d_{ij} = u_{ij} \|x_j - c_i\|, \quad (7)$$

se define el índice  $XB$  como [16]:

$$XB = \frac{\pi}{(d_{min})^2}, \quad (8)$$

donde

$$d_{min} = \min \|c_i - c_j\| \quad (9)$$

y la cantidad  $\pi = \frac{\sigma}{N}$  se llama compactación del conjunto de datos. Aquí, valores pequeños de  $XB$  son esperados para grupos compactos y bien separados.

Todos los resultados obtenidos en esta etapa (resultados de agrupaciones e índices de validación) son guardados en un archivo. Para continuar con la etapa siguiente es necesario que el usuario indique qué cantidad de zonas elige para diferenciar su lote.

## 2.5. Procesamiento de imágenes

En esta etapa se realiza el procesamiento de los resultados del clustering para obtener zonas bien definidas y con una superficie mayor que un determinado valor. Se trabaja con técnicas de procesamiento digital de imágenes que contienen diferentes métodos y algoritmos que manipulan y transforman una imagen en una señal con información de utilidad.

### Filtrado de imágenes

Se utilizan filtros estadísticos de orden: filtros espaciales no lineales cuya respuesta se basa en el ordenamiento (ranking) de los píxeles contenidos en una porción de la imagen. Las máscaras utilizadas tienen tamaño de  $3 \times 3$ ,  $5 \times 5$  ó  $7 \times 7$  píxeles y, a continuación, se sustituye el valor del píxel central con el valor que resulte según el criterio de clasificación [17].

Sea  $g(x, y)$  la imagen a procesar y  $\hat{f}(x, y)$  la imagen resultante, se aplican los siguientes procesos considerando los píxeles  $(s, t)$  de una vecindad  $S_{xy}$  centrada sobre cada píxel  $(x, y)$  original:

- Filtro de mediana: reemplaza el valor del píxel por la mediana estadística de los valores de intensidad en una vecindad del píxel (incluyendo el valor original del píxel en el cómputo), según

$$\hat{f}(x, y) = \underset{(s,t) \in S_{xy}}{\text{mediana}}\{g(s, t)\}. \quad (10)$$

- Filtro de moda: es el valor más frecuente de la zona, calculado como

$$\hat{f}(x, y) = \underset{(s,t) \in S_{xy}}{\text{moda}}\{g(s, t)\}. \quad (11)$$

Estos filtros son particularmente efectivos en presencia de ruido impulsivo.

### Eliminación de superficies pequeñas

La capacidad de cambiar la dosis que aplica la máquina agrícola no es instantánea. Por ello, no se justifica tener grupos con superficies pequeñas que no lleguen a cubrir una determinada superficie.

Para identificar aquellos sectores pequeños se utiliza el etiquetado por crecimiento de regiones, un proceso que consiste en agrupar píxeles o subregiones en regiones más grandes, basándose en un criterio predefinido  $P$ . La técnica comienza a partir de píxeles iniciales llamados “semillas” y se van incorporando

los píxeles vecinos que cumplen con una propiedad  $P$  seleccionada. El criterio es comprobado sobre una vecindad de 8 vecinos. El proceso termina cuando no existen más píxeles que cumplan con  $P$ . De esta manera, se logra identificar a cada zona junto a la cantidad de píxeles que la conforma, resultando directo el cálculo de la superficie ya que se conoce a priori el tamaño físico que representa cada píxel.

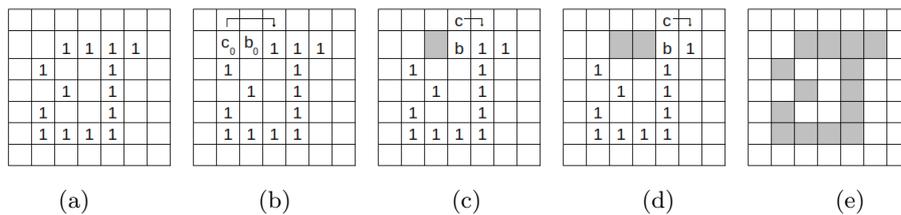
**2.6. Determinación de los polígonos de las zonas de manejo**

Para crear el archivo *shape* se necesita que los puntos que conforman el borde de cada zona estén ordenados. Para encontrar el trazado del contorno de cada zona y así determinar el polígono, se desarrolla una rutina basada en el algoritmo de Moore [17].

Este algoritmo comienza su funcionamiento a partir de una región binaria  $R$  o su contorno, y se ejecutan los siguientes pasos:

1. Tomar el punto inicial  $b_0$ , que es el punto superior izquierdo de la imagen con valor 1. Designamos  $c_0$  al vecino oeste de  $b_0$ . Claramente,  $c_0$  siempre es un punto del fondo. Se examinan los 8 vecinos de  $b_0$  comenzando con  $c_0$  y prosiguiendo en sentido horario. Se designa  $b_1$  al primer vecino encontrado con valor 1, y  $c_1$  es el punto (del fondo) inmediatamente precedente a  $b_1$  en la secuencia. Se almacenan las ubicaciones de  $b_0$  y  $b_1$  para usarse en el paso 5.
2. Tomar  $b = b_1$  y  $c = c_1$ .
3. Los 8 vecinos de  $b$ , comenzando por  $c$  y prosiguiendo en el sentido horario, se denominan  $n_1, n_2, \dots, n_8$ . Encontrar el primer  $n_k$  con valor 1.
4. Tomar  $b = n_k$  y  $c = n_{k-1}$ .
5. Repetir los pasos 3 y 4 hasta que  $b = b_0$  y el siguiente punto de contorno sea  $b_1$ . La secuencia de puntos  $b$  encontrada cuando el algoritmo se detiene constituye el conjunto ordenado de puntos del contorno.

La Figura 1 ilustra los pasos del algoritmo descrito anteriormente.



**Figura 1.** Ilustración del algoritmo de trazado de contorno: (a) sección de imagen original analizada; (b) paso 1; (c) paso 2; (d) paso 4; (e) fin del paso 5.

### Creación del archivo *shape*

Para crear un archivo *shape* una vez obtenidos los polígonos que delimitan las zonas de manejo, se utiliza la librería Shapefile C Library [18] con licencia LGPL. La misma ofrece la posibilidad de leer, escribir y actualizar el Shapefile y el archivo de atributo asociado (*dbf*).

De manera resumida, el proceso consiste en crear un objeto del tipo *Shape Polygon* al cual pasarle como parámetro las coordenadas  $(x,y)$  de los puntos que forman los polígonos que delimitan las zonas. Luego se crea el archivo *dbf* mediante una función de la librería a la cual se le pasa por parámetro los atributos que tiene cada polígono. Estos datos representan la dosificación variable que se aplicará a cada zona del lote.

## 3. Experimentos y resultados

Para realizar las pruebas se utiliza un conjunto de datos de un lote perteneciente a la Estación Experimental Agropecuaria Paraná del INTA (Entre Ríos). El lote analizado posee una superficie de 110 hectáreas y en la campaña analizada estaba sembrado con trigo. Se cuenta con una imagen satelital de  $74 \times 69$  píxeles, un mapa de rendimiento con 120000 muestras y la altimetría con 4687 valores, ambos distribuidos irregularmente sobre el lote.

Como primer paso se ingresan los archivos de rendimiento y altimetría para proceder a realizar la transformación de las coordenadas. Para crear la malla regular se elige una separación de 10 metros entre punto y punto, luego se procede a realizar la interpolación con el método de interpolación cuadrática, con el que se obtuvieron mejores resultados en pruebas preliminares.

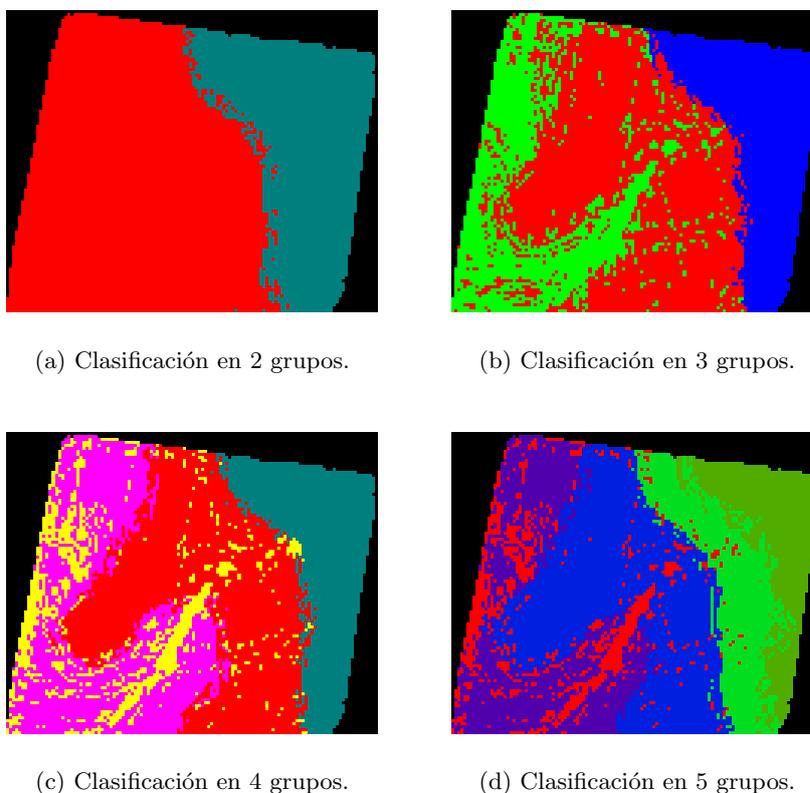
En la configuración de la etapa de clasificación se fijan los siguientes valores:

- exponente difuso = 1,7;
- criterio de convergencia = 0,0001;
- máximo número de iteraciones = 300;
- cantidad mínima de grupos = 2 y
- cantidad máxima de grupos = 5.

La Figura 2 muestra los resultados de clasificación en 2, 3, 4 y 5 grupos, respectivamente. Cada imagen muestra, con un mismo color, los puntos clasificados como pertenecientes al mismo agrupamiento.

Los índices aportan objetivamente una idea más clara acerca de cuál podría ser la clasificación óptima, aunque la selección final de la cantidad de grupos debe seguir una relación de compromiso entre lo sugerido por los índices y lo realmente practicable por la maquinaria de dosificación variable, la que que no tiene posibilidad de alcanzar cambios instantáneos.

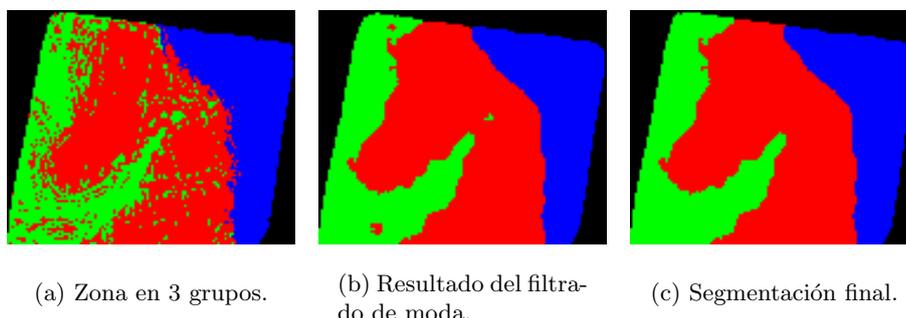
La Tabla 1 muestra los resultados de los diferentes índices para cada cantidad de agrupamientos. Para que la clasificación sea de calidad se busca que los índices *XB* y *PE* sean mínimos y *PC* máximo. Para observar de una manera más clara cuál es el número óptimo de grupos, se modifican los resultados de los



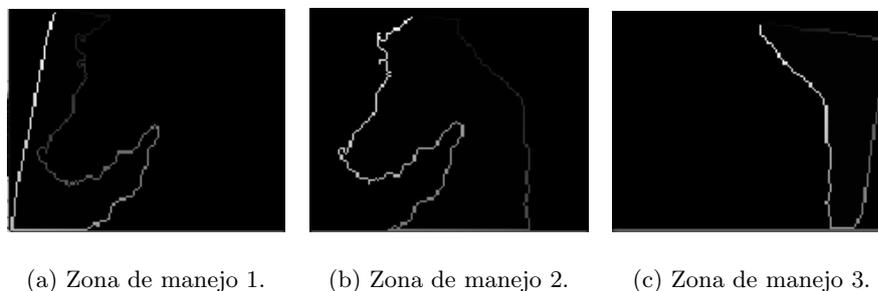
**Figura 2.** Resultado de la clasificación del lote en distinto número de grupos.

índices. Se recalcula  $PC_2 = 1/PC$  de manera que el valor mínimo en los tres índices represente la mejor elección (Tabla 1). Además se divide cada uno por el máximo para normalizar todos los valores entre 0 y 1. Se calcula la distancia euclídea ( $\sqrt{XB^2 + PE^2 + PC_2^2}$ ) y se selecciona la menor. Todos estos resultados se observan en la Tabla 2. De esta manera, se llega a la conclusión que la clasificación realizanda en 3 grupos es la que mejor se ajusta a lo buscado y por consiguiente es la elegida para continuar en la etapa siguiente.

Debido a que la clasificación que produce el FCM se realiza sobre los valores de las variables sin tener en cuenta su ubicación espacial, se pueden obtener zonas no bien definidas, con sectores no bien delimitados y compactos (Figura 3-a). Mediante el filtro de moda con máscara de  $7 \times 7$  se busca suplir estas imperfecciones (Figura 3-b). Además, existen zonas con una superficie pequeña que el filtrado de imágenes no es capaz de eliminar, por lo tanto se considera apropiado utilizar un parámetro de superficie mínima establecido en media hectárea ( $5000m^2$ ) (Figura 3-c).



**Figura 3.** Resultado del filtrado y segmentación de las zonas de manejo.



**Figura 4.** Resultado de la detección de borde en las zonas de manejo.

El siguiente paso es identificar el borde de cada zona. Partiendo de la imagen filtrada, se procede a ejecutar la rutina que realiza la búsqueda con el algoritmo de Moore. En la Figura 4 se ha marcando el sentido encontrado del polígono con un aumento en la intensidad de gris a medida que se avanza a lo largo de la trayectoria. Se obtiene como salida un archivo que contiene los puntos ordenados que conforman cada polígono.

Por último, se crea el archivo shape (*shx*) a partir de los polígonos obtenidos anteriormente. También se crea el archivo de datos (*dbf*) con los atributos

**Tabla 1.** Resultados de la evaluación de la calidad de los agrupamientos. En negrita, los mejores resultados por índice.

Índices	2 grupos	3 grupos	4 grupos	5 grupos
<i>XB</i>	1,80	<b>0,54</b>	0,68	0,90
<i>PC(PC<sub>2</sub>)</i>	<b>0,79 (1,27)</b>	0,74 (1,35)	0,67 (1,49)	0,63 (1,59)
<i>PE</i>	<b>0,33</b>	0,47	0,62	0,73

**Tabla 2.** Resultados normalizados de la calidad del clustering. En negrita, el mejor resultado.

Índices normalizados	2 grupos	3 grupos	4 grupos	5 grupos
$XB$	1,00	0,30	0,38	0,50
$PC_2$	0,80	0,85	0,94	1,00
$PE$	0,45	0,64	0,85	1,00
Distancia euclídea	1,36	<b>1,11</b>	1,32	1,50

correspondientes a cada polígono, según los valores de dosificación variable que se aplicará a cada zona.

#### 4. Conclusiones

En este trabajo se ha presentado un método que aborda el problema de identificar zonas de manejo en un lote productivo. El sistema desarrollado integra las diferentes etapas de procesamiento: carga de archivos de imagen satelital y relevamiento de datos, georreferenciación, segmentación automática del terreno según los parámetros relevados y creación del archivo necesario para su utilización en maquinaria agrícola de precisión.

Se ha implementado un procedimiento que mediante técnicas de reconocimiento de patrones como *Fuzzy C-Means* y el uso de los índices de validación de la calidad de los grupos, permite la evaluación simultánea de una serie de propuestas de agrupamientos. Así, el usuario puede comparar y elegir la cantidad de zonas de manejo que mejor se adapte a sus necesidades, utilizando información subjetiva (visualización de gráficas sobre el terreno) y objetiva (índices de calidad). Los resultados demostraron que el método logra automatizar el proceso, obteniendo una herramienta más rápida y sencilla que brinda al experto una gran ventaja en la operatividad de las máquinas agrícolas. Cabe mencionar que la obtención de buenos resultados depende de la calidad de las variables ingresadas.

Como trabajo futuro se puede nombrar el desarrollo de un módulo de adaptación de formato de los datos de entrada para todas aquellas variables que poseen formatos particulares (no estándares).

#### Agradecimientos

Los autores desean agradecer el apoyo brindado por la ANPCyT, la UNL (bajo los proyectos CAI+D 012-72, PAE 37122, PAE-PICT-2007-00052) y el CONICET. También se agradece especialmente a la Ing. Alejandra Kemerer perteneciente a la Estación Experimental Agropecuaria Paraná del INTA por el asesoramiento brindado.

## Referencias

1. Bongiovanni, R., Mantovani, E., Best, S., Roel, A.: Agricultura de precisión: integrando conocimientos para una agricultura moderna y sustentable. PROCISUR/IICA (2006)
2. Flowers, M., Weisz, R., White., J.G.: Yield-based management zones and grid sampling strategies: Describing soil test and nutrient variability. *Agronomy Journal* **97** (2005) 968–982
3. Taylor, J.C., Wood, G.A., Earl, R., Godwin, R.J.: Soil factors and their influence on within-yield crop variability. part ii: Spatial analysis and determination of management zones biosystems engineering. *Agronomy Journal* **84** (2003) 441–453
4. Bengolea, M.E.: Demostración de utilización de Farm Works Farm Site. <http://www.agriculturadeprecision.org/articulos/software/Farm-Works-Prec.asp> (1999) [Accedido 20-12-2012].
5. Gotthold, T.: Demostración de utilización de SSToolbox y sus posibles aplicaciones. AgriMax S.A. <http://www.agriculturadeprecision.org/articulos/software/SSToolbox.asp> (2005) [Accedido 02-08-2012].
6. Fridgen, J.J., N. R. Kitchen, K.A.S., Drummond, S.T., Wiebold, W.J., Fraisse, C.W.: Software Management Zone Analyst (MZA): Software for Subfield Management Zone Delineation. *Agronomy Journal* **96** (2004) 101–107
7. Bezdek, J., James, C., Ehrlich, R., Full, W.: FCM: The Fuzzy C-Means clustering algorithm. *Computers & Geosciences* **10**(2) (1984) 191–203
8. Huerta, E., Mangiaterra, A., Noguera, G.: GPS Posicionamiento Satelital. UNR Editora – Universidad Nacional de Rosario (2005)
9. Miretti, R., Cerati, E., Coronel, L.: Cartografía Matemática. Universidad Nacional del Litoral, Ediciones UNL (2012)
10. Francois, A.: Sistemas de coordenadas y transformaciones. Bases para el trabajo en SIG. Universidad de la República – Uruguay (2000)
11. Torres, F.J.A., Torres, M.A.A., Vega, F.A., Ramírez, F.C., Salmerón, P.L.S.: Efectos de la morfología del terreno, densidad muestral y métodos de interpolación en la calidad de los modelos digitales de elevaciones. *Anales de la ingeniería gráfica* **17** (2005) 25–35
12. Alliez, P.: Computational Geometry Algorithms Library. <http://www.cgal.org> (1998) [Accedido 05-10-2012].
13. Dunn, J.C.: Well separated clusters and optimal fuzzy partitions. *Journal of Cybernetics* **4** (1974) 95–104
14. Bezdeck, J.C., Ehrlich, R., Full, W.: FCM: Fuzzy C-Means Algorithm. *Computers and Geoscience* **10** (1984) 191–203
15. Halkidi, M., Batistakis, Y., Vazirgiannis, M.: On clustering validation techniques. *Journal of Intelligent Information Systems* **17** (2001) 107–145
16. Xie, X.L., Beni, G.: A Validity Measure for Fuzzy Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13** (1991) 841–847
17. Gonzalez, R.C., Woods, R.: *Digital Image Processing*. Pearson Prentice Hall, Upper Saddle River – NJ (2008)
18. Warmerdam, F.: Shapefile C Library. <http://shapelib.maptools.org> [Accedido 08-11-2012].