

Interfaz Gráfica para Extraer Coordenadas xy de Datos Vectoriales en Formato Shapefile

Graphical User Interface to extract xy coordinates from vector shapefiles

¹Marco Antonio Fuentes-Jiménez, ²Pedro Moreno , ²Alfonso D'Granda-Trejo , ²Jesús del Carmen Peralta- Abarca , ²Juana Enríquez-Urbano

¹FCAel-UAEM, ²Postgrado en Ingeniería y Ciencias Aplicadas CIICAp-UAEM
Av. Universidad 1001 Col. Chamilpa, Cuernavaca, Morelos, México. C.P. 62209

E-mail: markfuentes2991@gmail.com, pmoreno@uaem.mx, dgranda@uaem.mx, juana.enriquez@uaem.mx, carmen.peralta@gmail.com

PALABRAS CLAVE:

SIG, shapefile, ASCII (text),
GDAL, VectorMapConverter,
VectorMapView

RESUMEN

El tratamiento de información geográfica a partir de un conjunto de valores, para su análisis numérico, es un problema importante, debido a que la mayoría de las herramientas de Sistemas de Información Geográfica, tratan la información en formatos de imagen digital. En este trabajo presentamos una Interfaz Gráfica de Usuario, que a partir de una metodología propuesta, extrae coordenadas xy de los vértices de datos vectoriales en formato shapefile, a un archivo de texto ASCII. Así mismo, permite visualizar los datos espaciales extraídos.

KEYWORDS:

SIG, shapefile, ASCII (text),
GDAL, VectorMapConverter,
VectorMapView

ABSTRACT

The treatment of geographical information from a set of values for numerical analysis is an important problem. This is because, a lot of geographical information system tools treats the information in digital formats and not in ASCII text files. In this work, a Graphic User Interface is developed using a proposed methodology to extract xy coordinates from vector data in shapefile format, to ASCII-text format.

Recibido: 15 de octubre de 2016 • Aceptado: 31 de enero de 2017 • Publicado en línea: 28 de febrero de 2017

1 INTRODUCCIÓN

En la actualidad, la representación de mapas digitales se realiza a través de imágenes por computadora. Los Sistemas de Información Geográfica (SIG), manejan datos geográficos, permitiendo emplearlos para algún fin, y llevar a cabo operaciones sobre ellos, para obtener información. La información es, por tanto, el resultado de un dato y una interpretación. Trabajar con datos es un proceso enfocado a obtener de estos toda la información posible [1].

Representar de la mejor forma posible la información que se extrae de una determinada zona geográfica para almacenarse y analizarse, es indispensable en un SIG. Esto debido a que los datos reales son infinitos, y para poder ser tratados por una herramienta SIG, deben ser finitos. Esta información suele ser tratada en distintos tipos de formatos. Los más utilizados para exportar imágenes son los esquemas de modelos de representación vectorial y ráster [2].

En el modelo vectorial, se recoge la variabilidad y las características de una zona mediante entidades geométricas. Para cada entidad, las características son constantes. Los vectores son composiciones de polígonos, líneas y puntos. Los polígonos representan áreas geográficas, las líneas caminos o ríos y los puntos pueden representar unidades habitacionales, cacetas o cementerios entre otros [3]. Estas composiciones geométricas están a la vez formados por vértices.

Cuando trabajamos con modelos vectoriales, los vectores son comprimidos en archivos llamados shapefiles. Los shapefiles guardan información de vértices de cada geometría, es decir, su posición xy. El shapefile se compone de tres archivos, el Main, el Index y el dBASE. El Main es el archivo principal que contiene el registro de todos los vértices. Los vértices están almacenados en una estructura de datos. Por tal motivo, es necesario utilizar una técnica de indexación para su guardado, así como para su lectura. En el Main también está presente un encabezado con información relacionada al archivo (tipo, tamaño, número de registros, etcétera). En el archivo Index se encuentran escritos los índices de la estructura que guardan a los vértices, la forma de un fichero Index es similar a la del Main, puesto que también cuenta con una cabecera que describe al archivo. El archivo dBASE, es la base de datos donde esta guardada la información

de los atributos del objeto geométrico que representa al shapefile. Para hacer la lectura del mapa en formato vectorial, se toman en cuenta los tres archivos, de lo contrario no se puede realizar la lectura.

En el modelo ráster, la zona de estudio se divide en una serie de unidades mínimas denominadas celdas, en la que cada una recoge la información que la describe. El formato ráster está compuesto de rectángulos pequeños, es decir, se forma una malla para representar una imagen [3]. Cada celda de la malla tiene un valor llamado profundidad de color (color depth), este valor tiene un tamaño de 24 bits y está dividido en 3 partes de 8 bits, cada conjunto de 8 bits representa un color primario aditivo (rojo, verde y azul) [4]. La profundidad de color en cada celda se asocia para dar forma a la imagen. De esta forma la estructura regular de la malla ráster, permite conocer las coordenadas de las celdas, sin necesidad de almacenarlas.

Los archivos ráster vienen en formato "dataset", el cual es un conjunto de datos que definen como serán guardados los píxeles, el número de filas y columnas. Algunos dataset están comprimidos en formato JPEG, BitMap y ESRI Grid, entre otros.

La interfaz que se presenta, permite leer y escribir ambos tipos de archivos (Vector y Ráster). Se ha diseñado una metodología para acceder a los datos que han sido comprimidos, para extraerlos y posteriormente procesarlos mediante la interfaz gráfica, que exporta las coordenadas geográfica x y, a un formato de texto ASCII.

Este trabajo se estructura de la siguiente manera. La sección 2 muestra la motivación y los trabajos relacionados. En la sección 3 se presenta la metodología. La sección 4 describe la implementación. La sección 5 discute los resultados y la sección 6 concluye el trabajo.

2 MOTIVACIÓN Y TRABAJOS RELACIONADOS

Proveer una herramienta la cual permita extraer los datos de una imagen digital a un formato de texto ASCII, con la finalidad de facilitar el uso de los datos para realizar distintas operaciones sobre ellos, así como implementar distintos tipos de algoritmos, que permitan realizar análisis numérico de zonas geográficas, utilizando datos topológicos, con instancias de entrada en formato de texto ASCII. Además la herramienta genera instancias para un tipo específico de problema de análisis

de información geográfica, así como archivos personalizados a partir de imágenes digitales. Así mismo, se propone una metodología para la carga de geometrías, así como un espacio de trabajo que provee al usuario la visualización de características y otra información relevante de mapas con los que se trabaja.

Actualmente, existe software dedicado al mapeo, procesamiento de datos y análisis de espacios. Algunos de los paquetes de software SIG más utilizados y conocidos son:

Grass GIS, es una herramienta de información geográfica que se utiliza para el manejo de datos topográficos, procesamiento de imágenes y visualización de mapas [5]. Cuenta con varios módulos para el procesamiento de archivos ráster y vectoriales, de los cuales se listan algunos de ellos:

- Raster Analysis
- Vector Analysis
- Visualización de mapas
- SQL-Support: Soporta (DBF, SQLite, PostgreSQL, MySQL, ODBC).

gvSIG, es un sistema de información geográfica que se caracteriza por disponer de una interfaz amigable, siendo capaz de acceder a los formatos más comunes, tanto vectoriales como ráster. Cuenta con un amplio número de herramientas para trabajar con información de naturaleza geográfica[6].

Dentro de las funcionalidades de gvSIG se listan:

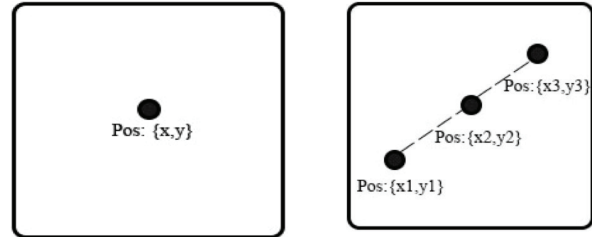
- Representación de ráster
- Representación de vector
- Geoprosesos
- Edición gráfica de mapas

ArcGIS ArcMap es un sistema de información geográfica, que permite visualizar y explorar los dataset SIG de su área de estudio [7].

Se utiliza para realizar muchas de las tareas habituales de SIG, como son:

- Trabajar con mapas.
 - Editar datasets.
 - Herramientas de visualización ráster.
 - Herramientas de conversión ráster.
 - Herramientas de visualización vector.
 - Herramientas de conversión vector
- Metodología

Es importante conocer cómo se encuentran organizados los vértices en el modelo vectorial. En la figura 1 se muestra la representación de un punto, una línea y un polígono respectivamente, como se ven en el plano.



a) Punto en el plano

b) Línea en el plano

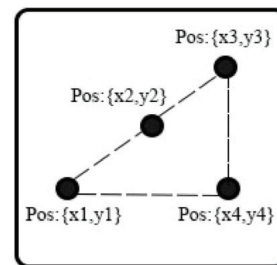
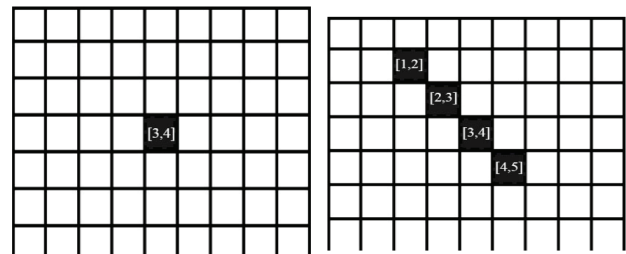


Fig. 1 c) Polígono en el plano.

Para las imágenes de tipo ráster, la representación de puntos, líneas y polígonos se hace iluminando la celda o el conjunto de celdas que corresponden a los puntos a mostrar. En la figura 2 se puede ver la representación de un punto, una línea y un polígono respectivamente, dentro de un modelo ráster.



a) Punto rasterizado

b) Línea rasterizada.

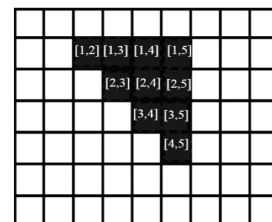


Fig. 2 c) Polígono rasterizado.

Tanto las geometrías de vectores como de ráster, tienen una posición en el espacio. Los valores de las coordenadas dentro de un vector está dada por metros, es decir un punto tiene una posición de xy metros alejados del origen. Mientras que en los ráster, los valores xy son sustituidos por la intersección entre la fila y la columna.

La estructura punto se compone de dos valores, cada uno capaz de representar valores reales y estos valores representan las posiciones tanto en x como en y, las cuales representan la latitud y longitud respectivamente. La figura 3 muestra los datos de la estructura.

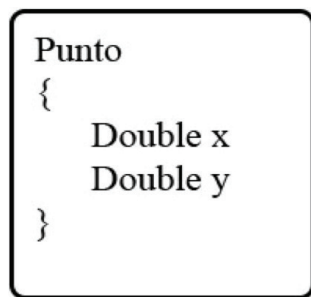


Fig 3. Estructura para guardar vértices

Para tratar y almacenar la información de polígonos y líneas, partimos del principio de que las líneas son una sucesión continua de puntos, mientras que los polígonos se componen de una cantidad finita de líneas y estos encierran una región en el plano. Visto desde otra perspectiva, tanto los polígonos como las líneas son una secuencia de puntos, por lo que la estructura que se muestra en la figura 4 muestra la forma de almacenar ambas geometrías.

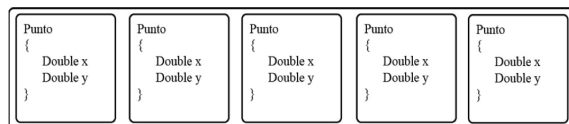


Fig. 4 Estructura de un arreglo el cual contiene elementos tipo estructura punto.

Es común que un archivo shapefile o ráster como esri grid, tenga más de una geometría guardada, es decir, contiene un conjunto de puntos representando una comunidad habitacional, un conjunto de líneas formando caminos y calles, o simplemente un conjunto de polígonos que conforman unas ciudades. Hasta este punto, las estructuras descritas anteriormente sirven para guardar una geometría de cada tipo (punto, línea o polígono). Y para guardar múltiples geometrías se utilizan tres estructuras de tipo arreglo. La estructura que se muestra en la figura 5 se utiliza para guardar múltiples

puntos, múltiples líneas y múltiples polígonos en un archivo.

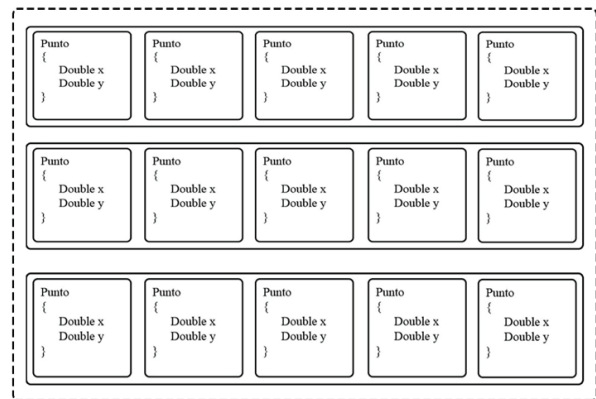


Fig. 5 Estructura para almacenar múltiples geometrías.

En la figura 6 se muestra el diagrama de flujo de la metodología utilizada para extraer coordenadas xy de datos vectoriales y ráster. Se aprecia como se leen y como se guarda la información en cada estructura, a partir de un archivo en formato shapefile o ráster.

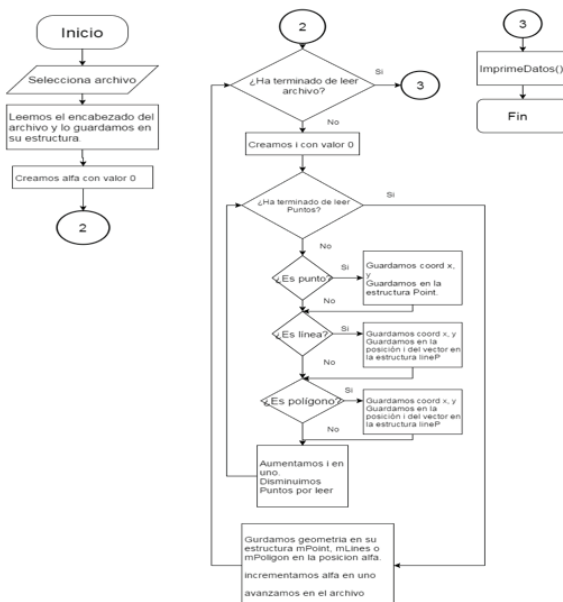


Fig. 6 Metodología para extraer información geográfica

Primero se selecciona el archivo que se va a leer, este puede ser en formato vectorial o ráster. Una vez seleccionado, se lee la parte del encabezado, en donde se obtiene información relevante al archivo el cual ayuda en la visualización. Antes de comenzar el ciclo iterativo de lectura de vértices, se declara una variable que funciona como un contador, el cual permite llevar el control sobre las geometrías leídas.

Dentro del ciclo iterativo, se leen repetidas ocasiones las geometrías registradas en el archivo shapefile o ráster. Es importante saber que un archivo puede tener n cantidad de elementos geométricos. Así es como se evalúa para todos y cada uno de los elementos geométricos registrados, en donde se leen todos los puntos o vértices que lo conforman, y dependiendo del tipo de archivo que se lee, se elegirá la estructura adecuada para almacenar la información extraída.

Cabe mencionar, que se invoca la función `imprimeDatos()` que aparece en la figura 6, para que a través de un proceso iterativo se realice la exportación o impresión de datos en un archivo de texto ASCII.

3 IMPLEMENTACIÓN

Se utilizó la librería de software libre GDAL [8] para tratar las posiciones de las celdas ([fila, columna]). Esta librería permite transformar las filas y columnas, a posiciones geográficas xy reales, en metros. Así se obtienen valores reales, evitando tener que implementar estructuras adicionales para guardar tipos ráster.

La interfaz gráfica para el manejo de información geográfica ha sido codificada en el lenguaje de alto nivel JAVA, por ser un lenguaje orientado a objetos que permite crear clases para las estructuras que guardan los datos de los vértices.

Además del manejo de las estructuras, JAVA permite definir objetos que actúan de manera independiente, permitiendo trabajar en un solo sistema, así como actualizar, deprecia e incorporar nuevos módulos de manera sencilla. Las librerías "ShapefileReader-1.0" y "OGR" son archivos .jar que se han integrado a través del entorno de desarrollo NetBeans, permitiendo importar ambas librerías de manera fácil. Además, la interfaz visual utiliza JAVA-swing y JOGL. Con JAVA-swing se crea la parte de interacción humano-aplicación, mientras que con JOGL se realiza la renderización en tiempo real del espacio de trabajo.

Resultados

VectorMapConverter es uno de tres módulos de la aplicación, el cual utiliza la metodología descrita en la sección anterior. VectorMapConverter se compone a su vez de otros módulos más específicos, de los cuales el primero se utiliza para la lectura de shapefiles, un segundo para la conversión a datos numéricos y carga en memoria, y el tercero se utiliza para la obtención de

archivos por medio de un cuadro de diálogo. La figura 7 detalla el cuadro de diálogo junto con un explorador de archivos

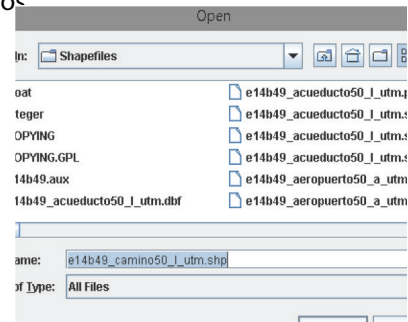


Fig. 7 Apertura de un archivo vector.

VectorMapView se encarga de generar la visualización en tiempo real en la interfaz gráfica. Utiliza dos módulos, uno para realizar renderizaciones activas en el espacio de trabajo y el otro para mostrar los elementos necesarios para la interfaz, como son: botones, áreas de texto y menús. En la figura 8 se puede apreciar la unión de los dos módulos

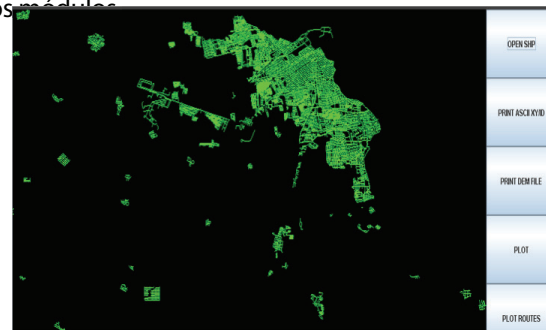


Fig. 8 A la izquierda se encuentra el área de renderización en tiempo real y a la derecha están los elementos de la interfaz gráfica.

ExportMap es el módulo que se encarga de la salida de archivos en nuevos formatos, en los cuales se añade información específica a los archivos de salida, a través de cuadros de diálogo que obtienen del usuario la información personalizada. La figura 9 muestra el cuadro de diálogo que permite añadir información personalizada y la figura 10 y la figura 11 muestran respectivamente



Fig. 9 Cuadro de diálogo especial para ingresar parámetros

Es importante personalizar la información de los datos de las capas topográficas de la zona de estudio. Por lo tanto se agrega un identificador (ID) a la información exportada que representa las coordenadas geográficas en texto plano. Así mismo, se agrega otro ID que identifica un dato que puede ser utilizado como penalización o accesibilidad a la capa de cobertura correspondiente a la topología utilizada. Estos datos personalizados son necesarios como parámetros de entrada para identificar las distintas capas de la cobertura de cada capa topológica, la cual puede ser utilizada en alguna aplicación con la finalidad de facilitar de forma sencilla e intuitiva la realización de operaciones de análisis utilizando lenguajes de alto nivel para resolver problemas complejos.

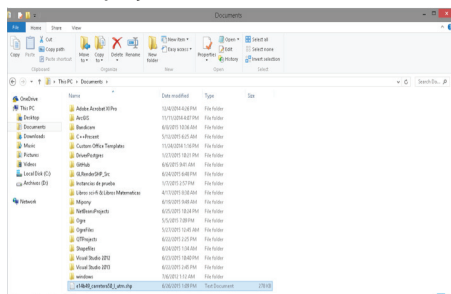


Fig 10. Archivo creado y guardado

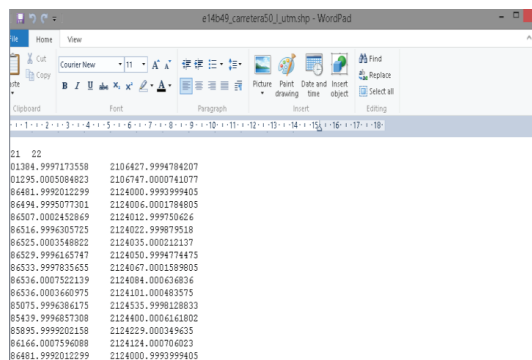


Fig. 11 Estructura de un mapa en formato texto ASCII, con los parámetros de identificación

4 CONCLUSIONES

Se presentó en este trabajo, una metodología para la lectura y extracción de formatos vectoriales y ráster. Así mismo, una interfaz gráfica que permite manejar información geográfica, al igual que muchos otros software SIG. La herramienta lee archivos de topografías en formato vectorial o ráster, lo que permite cargar en memoria vértices y así, hacer uso de ellos para generar una visualización o para crear un nuevo archivo, así como para facilitar el tratamiento de la información para

aplicaciones que requieran la realización de operaciones

de análisis numérico utilizando lenguajes de alto nivel para resolver problemas complejos.

Además permite crear archivos personalizados, así como exportarlos a un formato de uso genérico, para imprimir la información sobre la pantalla para su análisis.

REFERENCIAS

- [1] Olaya, Victor. (2012). Sistemas de Información Geográfica. España: CreateSpace Independent Publishing Platform.
- [2] Hearn D., Baker M. (1997) Computer Graphics, C version. Willard, Ohio: Pearson.
- [3] Greenberg I. (2007). Processing: Creative Coding and Computational Art. New York, NY: Springer.
- [4] Jeff Schewe. (2011). Home > Articles Real World Camera Raw: Digital Image Anatomy. 16/09/2015, de Pearson Sitio web: <http://www.peachpit.com/articles/article.aspx?p=1709190&seqNum=2>
- [5] The Open Source Geospatial Foundation (2015, February 24). General Overview <http://grass.osgeo.org/documentation/general-overview/>
- [6] The Open Source Geospatial Foundation (2015, July 31). gvSIG Info Sheet. <http://www.osgeo.org/gvsig>
- [7] ESRI (2015, July 31). ArcGIS Resource <http://resources.arcgis.com/es/home/>
- [8] GDAL - Geospatial Data Abstraction Library <http://www.gdal.org/>

Acerca de los autores



Marco Antonio Fuentes Jiménez. Es un joven programador, apasionado por la computación y programación gráfica, la misma que se aplica a videojuegos 2D/3D y simulaciones. Tiene un amplio conocimiento en tecnologías para el desarrollo de aplicaciones Desktop/Móvil así como aplicaciones de servidor y web, Actualmente está realizando tesis en el Centro de Investigación en Ingeniería y Ciencias Aplicadas (CIICAp) implementando una Interfaz gráfica de sistemas de ductos para transportar líquidos valiosos.



Alfonso D'Granda Trejo. Docente de la FCAel de la Universidad Autónoma del Estado de Morelos. Actualmente estudia un posgrado en el CIICAp de la Universidad Autónoma del Estado de Morelos.



Juana Enríquez Urbano. Docente de la FCQel de la Universidad Autónoma del Estado de Morelos. Actualmente estudia un posgrado en el CIICAp de la Universidad Autónoma del Estado de Morelos.



Pedro Moreno Bernal. Docente de la FCAel de la Universidad Autónoma del Estado de Morelos. Actualmente estudia un posgrado en el CIICAp de la Universidad Autónoma del Estado de Morelos.



Jesús del Carmen Peralta Abarca. Docente de la FCQel de la Universidad Autónoma del Estado de Morelos. Actualmente estudia un posgrado en el CIICAp de la Universidad Autónoma del Estado de Morelos.