

COMPARACIÓN DE MÉTODOS DE DIGITALIZACIÓN PARA EL INGRESO DE INFORMACIÓN ESPACIAL A LOS SISTEMAS DE INFORMACIÓN GEOGRÁFICA

Carlos E. Pacheco A.¹,
Alex G. Barrios R.², Juan I. López H.¹

RESUMEN

El objetivo del estudio es la comparación de métodos de digitalización (tableta, pantalla, semi-automático y automático) utilizados en la actualidad para introducir información espacial a los Sistemas de Información Geográfica (SIG). Con el fin de evaluar y comparar ventajas y desventajas, se aplicó una metodología que se sustenta en el tiempo neto efectivo para obtener un archivo vectorial y en la fidelidad de reproducción de los mismos. Este último aspecto se evaluó gráficamente mediante la visualización conjunta de curvas de nivel digitalizadas y las del mapa original, y calculándose el error medio cuadrático (RMSE) entre las coordenadas de puntos provenientes de los archivos vectoriales y las correspondientes en el mapa original.

Los resultados arrojaron que el método de digitalización semi-automático es superior en cuanto a su rapidez y reducido error medio cuadrático, no obstante, el método de digitalización manual en pantalla, aunque consume cerca de 25% más de tiempo que la digitalización en tableta, arroja errores medios cuadráticos cuatro veces menores y similares al obtenido con el método de digitalización semi-automático.

Palabras claves: métodos de digitalización, tableta, pantalla, semi-automático, digitalización automática, sistemas de información geográfica, raíz del error medio cuadrático.

¹ Laboratorio de Fotogrametría y Sensores Remotos. Facultad de Ciencias Forestales y Ambientales. Universidad de los Andes

² Centro de Estudios Forestales y Ambientales de Postgrado (CEFAP), Facultad de Ciencias Forestales y Ambientales. Universidad de los Andes. e-mail: carlosa@ula.ve, alexb@ula.ve y jlopez@ula.ve.

COMPARING DIGITALIZING METHODS FOR BRINGING SPATIAL INFORMATION INPUT TO GIS

Carlos E. Pacheco A.¹,
Alex G. Barrios R.², Juan I. López H.¹

ABSTRACT

The aim of the study is to compare digitalizing methods (tablet, on screen, semi-automatic and automatic) used nowadays to introduce information into GIS. In order to evaluate and compare advantages and disadvantages, an effective net time and fidelity of reproduction supplied method was applied for obtaining a vector file. Fidelity of reproduction was evaluated graphically by visualizing countour levels digitalized with the original map, and calculating root mean square error (RMSE) between control point coordinates from vector files and their correspondents in original map.

Results show that semi-automated methods have the best RMSE and are the fastest, while on screen manual digitalizing, with more than 25% of time than tablet digitalizing, shows RMSE four times lower and looks almost equal to the obtained with semi-automated digitalizing method.

Key words: digitalizing methods, tablet, screen, semi-automated, automated digitalizing, GIS, root mean square error.

¹ Laboratorio de Fotogrametría y Sensores Remotos. Facultad de Ciencias Forestales y Ambientales. Universidad de los Andes

² Centro de Estudios Forestales y Ambientales de Postgrado (CEFAP), Facultad de Ciencias Forestales y Ambientales. Universidad de los Andes. e-mail: carlosa@ula.ve, alexb@ula.ve y jlopez@ula.ve.

INTRODUCCIÓN

A pesar de la variabilidad de técnicas existentes para la producción de información espacial básica y temática utilizada en los Sistemas de Información Geográfica (SIG): levantamientos topográficos, Sistemas de Geoposicionamiento Global (GPS), interpretación de imágenes de sensores remotos, tales como, fotografías aéreas, imágenes satelitales y radar, entre otras, y finalmente, digitalización de la cartografía preexistente, la más utilizada es esta última (Bosque, 1992 y Longley et al. 1999), lo cual se explica por el elevado costo de los demás métodos, tanto en el trabajo de campo, como en la adquisición.

Esto ha obligado a estudiar las diferentes alternativas de digitalización para convertir la información analógica a información numérica digital: digitalización manual mediante el seguimiento de los elementos que conforman un mapa (punto, línea y polígono), a través de un cursor que envía señales magnéticas a una rejilla que está dentro de una tableta, estas señales son visualizadas en el monitor, registradas y almacenadas en la unidad central de procesamiento (Aumann et al. 1991, Carter 1988, y Nero y Cintra 2001); digitalización manual en pantalla, previo escaneo del mapa para obtener una imagen raster, mediante sensores ópticos de exploración (scanners) y, sobre esta imagen, generar capas vectoriales con la información que se desee del mapa (curvas de nivel, drenajes, vialidad, centros poblados, cobertura vegetal, entre otras), de manera manual con un cursor y el monitor (Nero y Cintra 2001 y Alfaro y Mosalvo 2006); o bien digitalización automatizada, utilizando una imagen raster procesada para generar los archivos vectoriales en forma semi-automática y automática, rastreando los píxeles de igual valor digital de manera interactiva o automática (Levachkine 2000, Decelis 2001, GPI 2002, Alfaro y Mosalvo, 2006 y Terrainmap, 2007).

Estos últimos son las más atractivos en cuanto a su rapidez y economía, sin embargo, existen una serie de elementos que lo limitan, puesto que se trabajan con imágenes escaneadas de documentos cartográficos que contienen información de carreteras, curvas de nivel, drenajes, toponimia, cuadrículas, entre otras; que son representadas digitalmente por píxeles que no tienen exactamente el mismo color, sino que aparecen bajo un efecto gradiente, “el color de los píxeles de los bordes de las líneas es más claro y menos saturado que el de los píxeles centrales, presentando aún tinte similar” (Fernández, López y Riveiro 2000); además, en dichos documentos no deben presentarse manchas, dobleces ni rotos, ya que estos se reflejan a la hora de ser

digitalizados bajo el “escáner” quedando registrados como información adicional que entorpecen el procedimiento.

Ante esta problemática, existen grupos de investigación que han generado programas especiales, tal es el caso del Grupo de Procesado de Imágenes (GPI) de la Universidad de Vigo (2002), que originó el Entorno de *Vectorización Automático* y Levachkine (2000). Ellos crearon una serie de algoritmos para el tratamiento digital de la imagen con fines de digitalización automatizada de mapas raster, así como también otros trabajos académicos referentes a la automatización de líneas sobre mapas raster (Decelis, 2001).

Todos con la finalidad de obtener un resultado ideal, es decir, generar una imagen binaria por cada capa de información, sin ruidos, donde el elemento que la conforma, como las curvas de nivel, estén realmente conectados entre sí y que no exista ningún otro tipo de elemento dentro de ella, de esta manera, cualquier algoritmo que se genere para la vectorización de mapas raster, bien sea de manera semi-automática o automática, siempre va a obtener buenos resultados. Sin embargo, no es fácil llegar a esto, puesto que los elementos antes señalados así lo indican.

En tal sentido, son pocos los estudios donde se han realizado comparaciones entre estos métodos de digitalización, para conocer las ventajas y desventajas que tienen en función al tiempo, precisión y exactitud, razón por la cual se plantea como objetivo central en esta investigación evaluar y comparar los diferentes métodos de digitalización: manual (tableta y pantalla) y automatizada, bajo ambiente SIG.

MATERIALES Y MÉTODOS

Área de estudio

Por su características topográficas de altas montañas se seleccionó la cuenca del Río Mucujún que se encuentra ubicada en Venezuela, entre los 08° 35' 39" y 08° 49' 27" de latitud Norte; 70° 58' 30" y 71° 09' 45" de longitud Oeste. Extendiéndose desde la desembocadura del Río Chama, en las cercanías de la ciudad de Mérida, a 1.400 msnm, hasta su nacimiento a 30,46 km, aguas arriba, a 4.400 msnm, ocupando una superficie de 191 km² (Figura 1).

Geográficamente, comprende una parte de Los Andes Centrales Venezolanos en el estado Mérida, específicamente denominada Cuenca del Río Mucujún, constituida por un valle longitudinal a la cordillera, muy cercana a la Ciudad de Mérida. Desde el punto de vista político-administrativo y según la Ley de División Político Territorial del Estado Mérida aprobada en 1.992, el área de estudio pertenece al Municipio Autónomo Libertador y ocupa la mayor parte del territorio de la Parroquia Gonzalo Picón Febres.

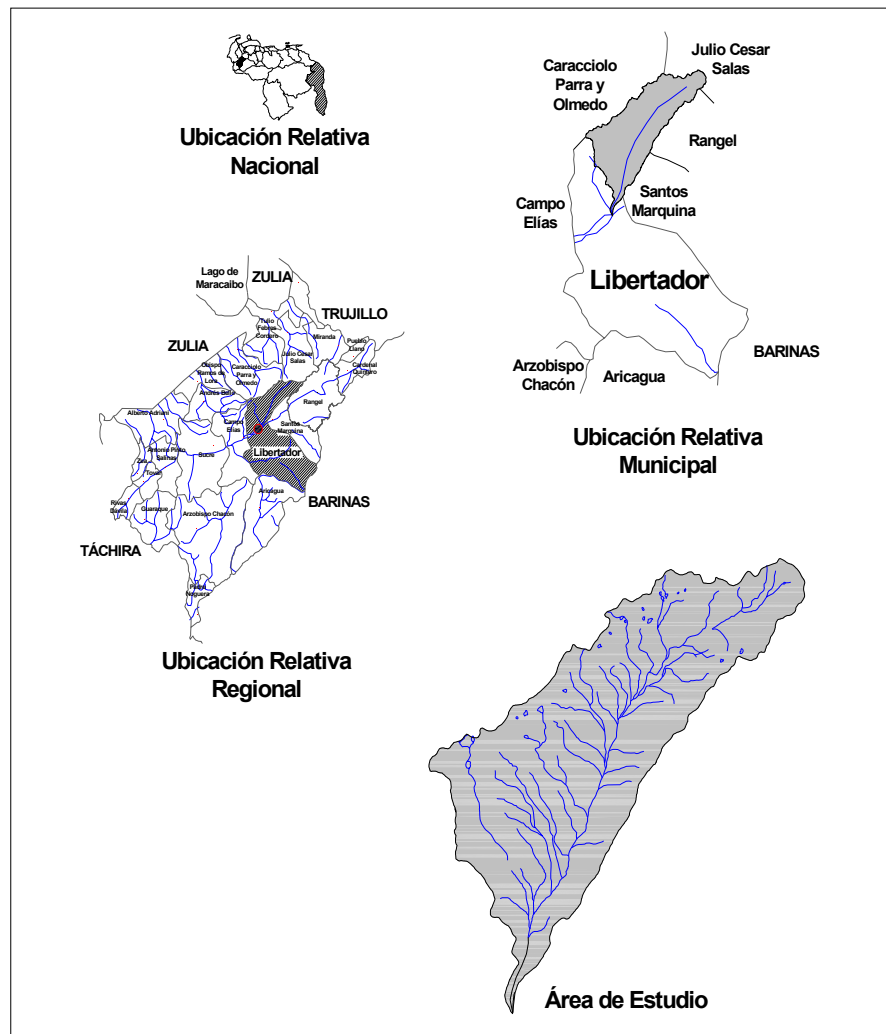


Figura 1. UBICACIÓN RELATIVA DEL ÁREA DE ESTUDIO

Métodos

Se utilizó una fracción de la carta topográfica a 1:100.000 del Instituto Geográfico Venezolano Simón Bolívar, número 6042, entre las coordenadas 958.400 y 976.000 Norte; y 280.000 y 287.150 Este; con un área de 136 km². Se definió la tolerancia usando ¼ de mm en el papel, que corresponde a 25 m en el terreno. Con este parámetro, se vectorizaron las curvas de nivel y los drenajes, mediante los siguientes métodos:

Método de Pantalla

Fue necesario obtener una imagen raster del documento original mediante escaneo digital, con una resolución de 300 puntos por pulgadas (dpi), la cual se georreferenció utilizando las coordenadas de 7 puntos de control de la imagen y del mapa, aplicando una “función lineal” y el “método del vecino más cercano”. Se realizó el remuestreo de la imagen, dejando la misma resolución espacial. Esto originó el RMSE para cada punto y el total (Cuadro 1).

Cuadro 1. COORDENADAS Y RMSE REGISTRADOS DURANTE LA GEORREFERENCIACIÓN DE LA IMAGEN

Puntos	Imagen		Mapa UTM		RMSE(m)
	x	y	X Este (m)	Y Norte (m)	
1	008,531812	1028,291833	280.000	976.000	2,144697
2	242,326546	1021,206395	284.000	976.000	0,478251
3	240,925000	0550,162400	284.000	972.000	0,792954
4	239,487440	0313,097021	284.000	968.000	0,733932
5	236,533307	0077,046053	288.000	964.000	1,119223
6	000,230970	0079,139504	284.000	960.000	0,549615
7	002,998279	0314,826434	280.000	964.000	2,373976
				Total	1,309675

Se usó la tolerancia calculada, para que el sistema automáticamente conecte los nodos colgantes cada 25 m, genere los vértices cada 10 m y seleccione arcos en un área no mayor de 50 m; de esta manera, se procedió a la creación de las

entidades gráficas, utilizando el zoom más apropiado dentro del área del monitor y la imagen raster del mapa (Figura 2.a).

Método de Tableta

Colocando la carta topográfica sobre la tabla digitalizadora se establecieron puntos de control para orientar la misma sobre su sistema de coordenadas. Esto se realizó en cada sesión de digitalización (3 en total) y originó una RMSE para cada punto y para el total (Cuadro 2), aceptados, ya que la tableta digitalizadora utilizada opera con una precisión de 0,005 pulg, que a la escala de trabajo representa 12,7 m de error mínimo.

Después de determinar el umbral mínimo de error de la tableta, se generaron las entidades gráficas (vector), y se editaron para obtener los archivos definitivos en formato digital (Figura 2.b).

Cuadro 2. COORDENADAS Y RMSE REGISTRADOS EN LA DIGITALIZACIÓN EN TABLETA

Puntos	Norte (m)	Este (m)	RMSE1 (m)	RMSE2 (m)	RMSE3 (m)
1	976.000	280.000	4,144697	5,342658	3,654328
2	976.000	284.000	2,478251	3,345689	1,235805
3	972.000	284.000	2,792954	3,801292	3,023489
4	968.000	284.000	2,733932	3,632879	2,253648
5	964.000	288.000	3,119223	4,009898	3,126453
6	960.000	284.000	2,731851	3,645004	4,256812
7	964.000	280.000	4,373976	5,234586	3,984652
		Totales	3,196412	4,144572	3,07645

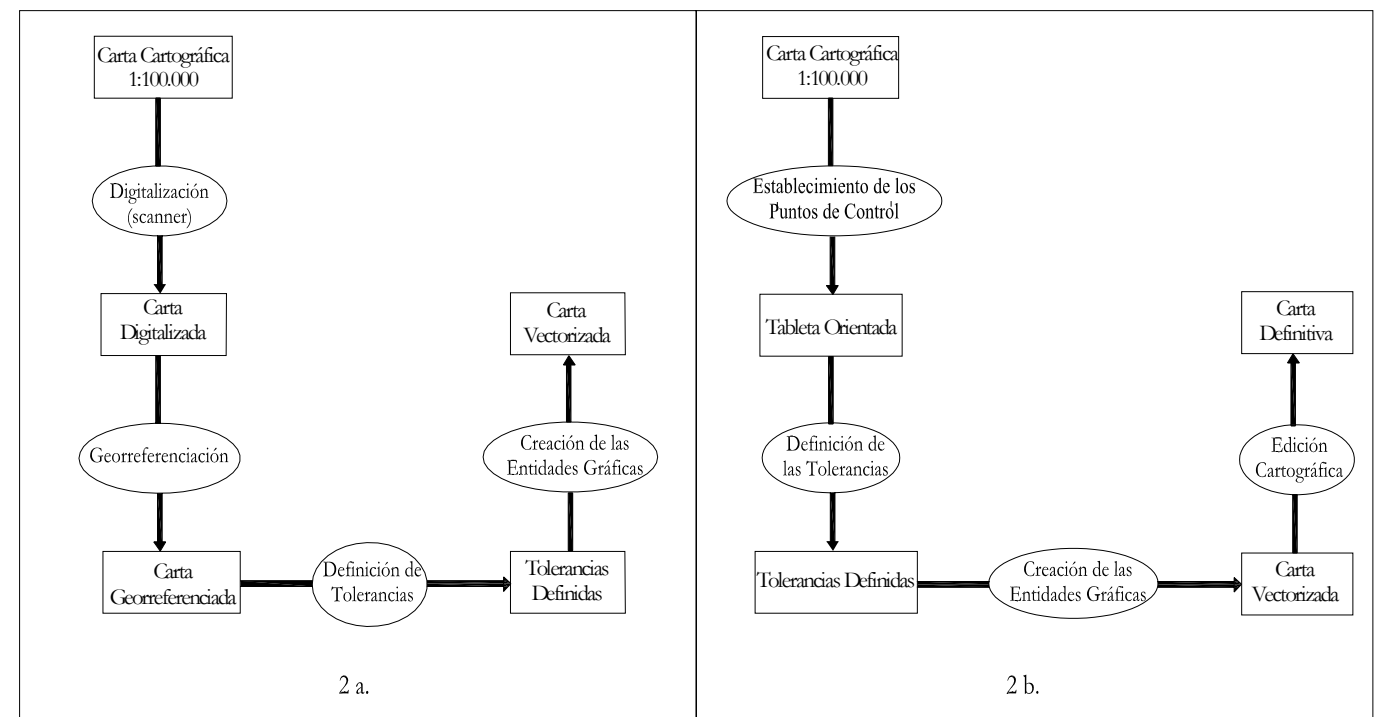


Figura 2. PROCEDIMIENTOS DE DIGITALIZACIÓN EN PANTALLA (2A) Y EN TABLETA (2B)

Método Semi-Automático

Se digitalizó la imagen con una resolución de 400 dpi para obtener mayor definición de los detalles de la carta, luego se georreferenció y se obtuvieron los resultados del Cuadro 3:

Posteriormente se procesó, transformándola a escala de grises, mejorando su contraste y brillo (ecualización). Se generó una imagen binaria para cada capa de interés, curvas de nivel y drenajes, las cuales se digitalizaron, seleccionando un píxel de arranque y remontando el vector de manera semi-automática. (Figura 3.a).

Cuadro 3. COORDENADAS Y RMSE REGISTRADOS EN LA DIGITALIZACIÓN SEMI-AUTOMÁTICA

Puntos	Imagen		UTM		RMSE (m)
	x	y	X Este (m)	Y Norte (m)	
1	002,07348	704,72045	280.000	976.000	1,025349
2	473,86823	711,86732	284.000	976.000	0,569217
3	944,17549	717,96409	284.000	972.000	0,649834
4	947,64374	481,20291	284.000	968.000	0,872291
5	950,53023	243,36050	288.000	964.000	1,064824
6	008,00006	231,22824	284.000	960.000	0,436872
7	005,96515	467,89031	280.000	964.000	2,587133
				Total	1,046341

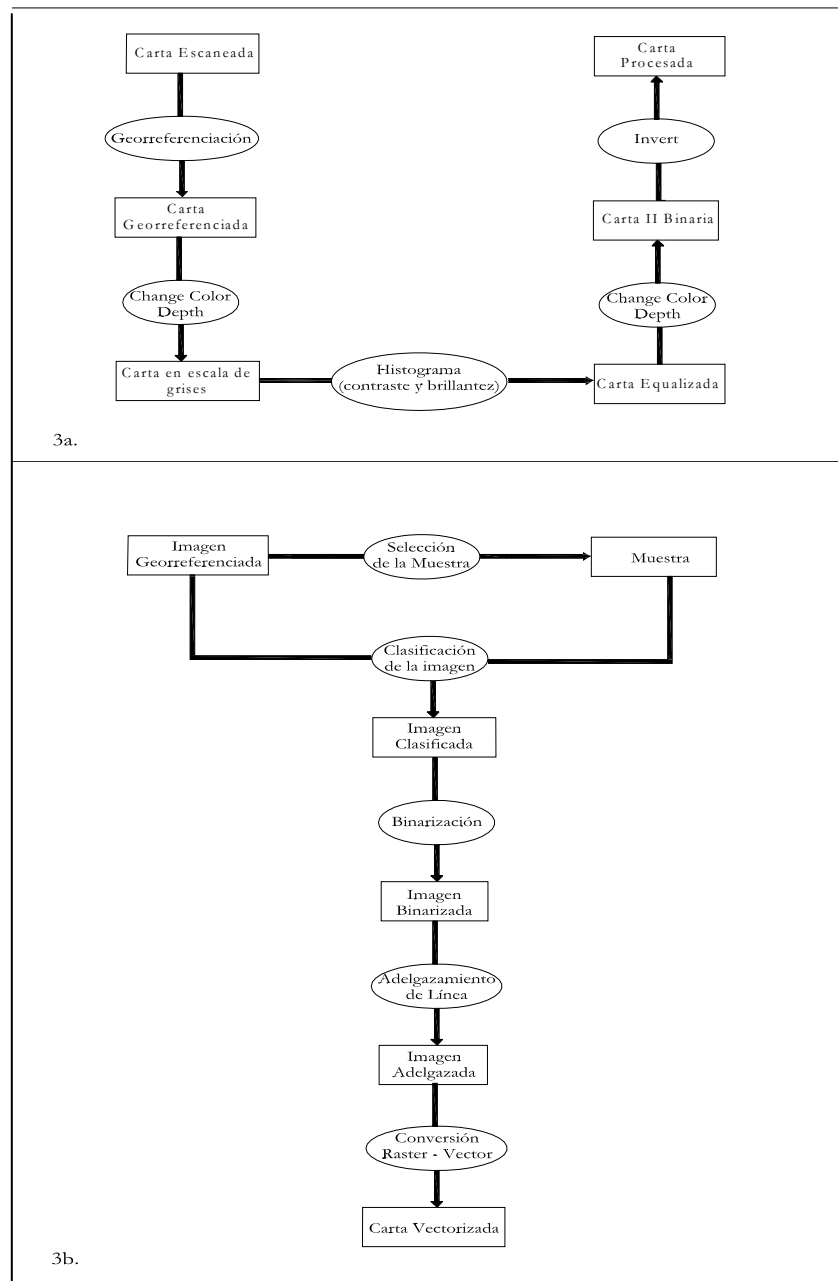


Figura 3. PROCEDIMIENTOS DE DIGITALIZACIÓN SEMI-AUTOMÁTICO (3A) Y AUTOMÁTICO (3B)

Método Automático

Este método es similar al semi-automático, pero con la salvedad de que el procesamiento digital de la imagen fue más riguroso y detallado para eliminar los ruidos presentes y obtener, de esta manera, una imagen binaria, ya que el proceso de la creación de las entidades gráficas no fue interactivo sino de manera automática, es decir, que el sistema mismo se encarga de generar el archivo vectorial de la imagen procesada (Figura 3.b).

Con el fin de obtener una imagen con los elementos de interés, se realizó una clasificación supervisada, seleccionando diferentes muestras: fondo de la imagen, curvas de nivel, drenaje y texto; después se obtuvieron de ella dos imágenes binarias: curvas de nivel y la red de drenaje, todo esto mediante el filtro de binarización; seguidamente, fue necesario aplicarles un filtro de adelgazamiento de líneas, ya que, tanto las curvas de nivel como la red de drenajes fueron muy gruesas; por último, se le aplicó el algoritmo de automatización de líneas para transformar la imagen raster en vectorial.

Criterios de comparación entre los métodos de digitalización

Se sustentó en el tiempo neto efectivo para obtener un archivo vectorial y, en la fidelidad, para la reproducción de las curvas y drenajes originales. Este último aspecto se evaluó gráficamente y estadísticamente.

Gráficamente, mediante la visualización conjunta de las curvas digitalizadas y las del mapa original, y estadísticamente, calculando el RMSE entre las coordenadas de los archivos vectoriales generados y los correspondientes en el mapa original, para ello, se seleccionaron puntos de control en las intersecciones de los drenajes y las curvas de nivel (Figura 4), tanto en la imagen digitalizada como en el mapa original.

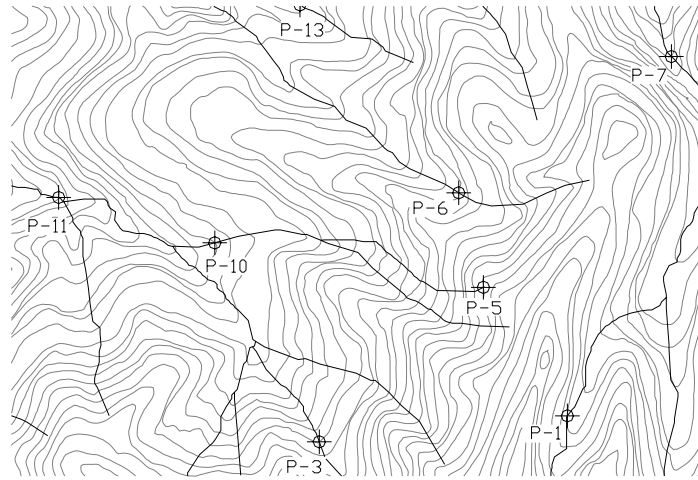


Figura 4. ALGUNOS PUNTOS DE CONTROL OBTENIDOS A TRAVÉS DE LA INTERSECCIÓN DE LAS CURVAS DE NIVEL Y LOS DRENAJES DIGITALIZADOS

Como la carta digitalizada tiene suficientes drenajes, se decidió obtener 6 puntos de control por cada cuadrícula de 4.000 m por lado, lo que originó 51 puntos de control, que se obtuvieron de la imagen raster georreferenciada, generada de la carta original.

Como la escala del mapa es 1:100.000, la precisión máxima utilizada fue de 1 mm, esto es 100 m del terreno como error máximo permitido. El procedimiento se realizó de manera automatizada, donde se obtuvieron en cada punto de control las coordenadas Norte (Y) y Este (X); de igual forma, se obtuvieron las coordenadas de los puntos homólogos dentro de los vectores generados por cada método. Con los puntos obtenidos anteriormente, se calculó el RMSE para cada vector digitalizado mediante los diferentes métodos:

$$RMSEy = \sqrt{\frac{\sum Ry^2}{n}} \quad RMSEt = \sqrt{(RMSEx)^2 + (RMSEy)^2}$$

$$RMSEx = \sqrt{\frac{\sum Rx^2}{n}} \quad Rx = X - X', \quad Ry = Y - Y'$$

Donde:

RMSEt = Raíz del error medio cuadrático total.

RMSEx = Raíz del error medio cuadrático en X.

RMSEy = Raíz del error medio cuadrático en Y.

Rx = Diferencia de coordenadas Este.

Ry = Diferencia de coordenadas Norte.

X = Coordenadas Este de la imagen raster georreferenciada.

Y = Coordenadas Norte de la imagen raster georreferenciada.

X' = Coordenadas Este de la intersección de las curvas y drenajes digitalizados.

Y' = Coordenadas Norte de la intersección de las curvas y drenajes digitalizados.

Materiales

Carta topográfica del Instituto Geográfico Venezolano Simón Bolívar (IGVSB) a 1:100.000 N° 6042.

Software básico: IDRISI Los Andes, TNT Mips 6.5, Cartalinx, AutoCad 2005, Raster Design 2005, EZYsurf, R2V y ArcView 3.2.

Hardware utilizado: PC Pentium IV Intel de 1.2 GHz, bajo Windows 2000, con 1 Gb en RAM y 120 Gb en disco duro, Scanners, Tableta digitalizadora e Impresora.

RESULTADOS Y ANÁLISIS

En los métodos semi-automático y automático se utilizaron varias alternativas que fueron exploradas, sin embargo, en todas se necesitó realizar un procesamiento previo a la imagen, esto trajo como consecuencia que el único método comparado fue el semi-automático, ya que no fue posible obtener resultados aceptables con el procedimiento automático, debido a que no hubo una correcta binarización de la imagen. Entre las causas estarían la baja calidad del material base escaneado, resolución inapropiada de la imagen raster, insuficiente saneamiento digital de la imagen.

Los resultados comparativos del tiempo neto al digitalizar un segmento de carta a escala 1:100.000, con equidistancia cada 40 m y con un área de 136 km²

mediante los métodos de pantalla, tableta y semi-automático se presentan en la Figura 5, se observa que tanto en drenajes como en curvas de nivel, el menor tiempo se obtuvo con el método semi-automático, seguido por los métodos de tableta y de pantalla con poca diferencia entre estos últimos. Con el método semi-automático, el tiempo de digitalización (vectorización) se reduce a la mitad o más del empleado con tableta y pantalla. A su vez, el método de pantalla es un 25% más largo que el de tableta.

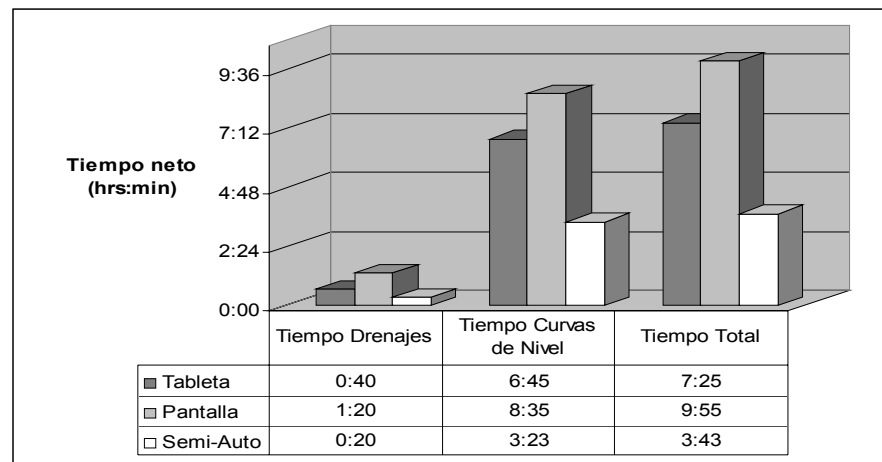


Figura 5. TIEMPO NETO REGISTRADO DURANTE EL PROCESO DE DIGITALIZACIÓN APLICANDO LOS MÉTODOS TABLETA, PANTALLA Y SEMI-AUTOMÁTICO

Con pantalla y semi-automático se obtuvieron los mejores ajustes visuales entre las curvas de nivel vectorizadas y los que subyacen en el mapa. El peor ajuste fue el logrado con la tableta (Figura 6). Esto se explica porque los errores humanos se evidencian más con este método, ya que, una vez que se realiza el seguimiento de la línea de manera manual, si se comete cualquier desviación durante el proceso, es en la etapa de edición donde se corregirá, pero las líneas corregidas no serán tan fidedignas con las originales, ya que la corrección se hace en pantalla y no en la tableta donde se encontraba el mapa original.

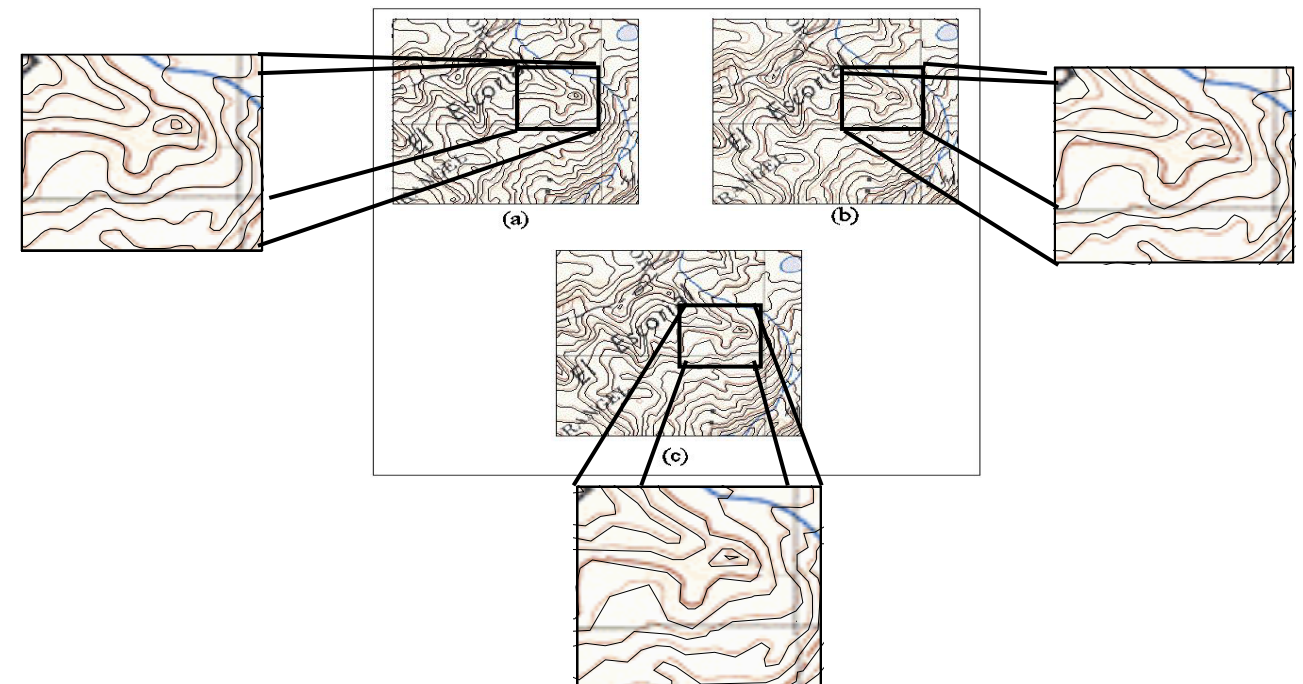


Figura 6. CURVAS DIGITALIZADAS SOBRE EL MAPA ORIGINAL: (A) PANTALLA, (B) TABLETA Y (C) SEMI-AUTOMATICO

En la Figura 7, se presenta el error medio cuadrático (RMSE) obtenido con cada método, además del error de calibración (tableta) o de georreferenciación (pantalla y semi-automático) de la imagen. Se observa que con la tableta se produjo un RMSE 300% mayor que con pantalla y semi-automático, siendo la diferencia entre estas últimas dos poco relevante (2%).

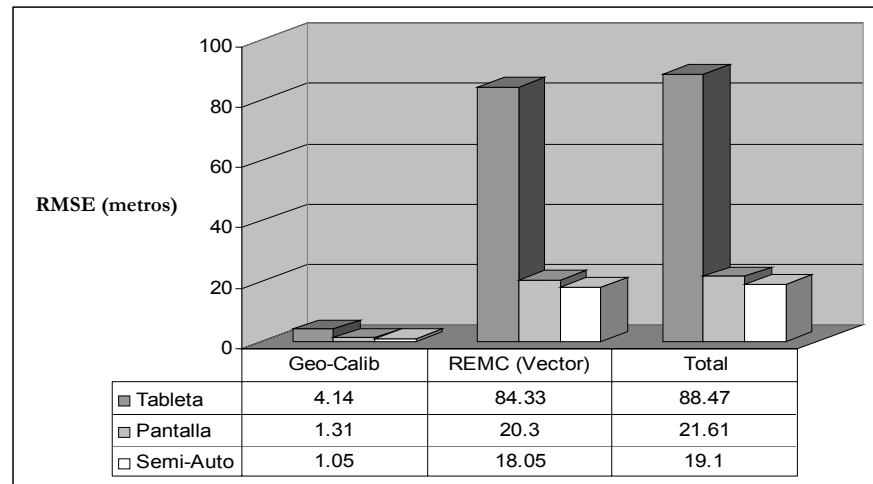


Figura 7. RMSE CALCULADO PARA LAS CURVAS DE NIVEL Y LOS DRENAJES ORIGINADOS MEDIANTE LOS MÉTODOS DE PANTALLA, TABLETA Y SEMI-AUTOMÁTICO

CONCLUSIONES Y RECOMENDACIONES

Por su rapidez y reducido error medio cuadrático, el método semi-automático luce superior a los de pantalla y tableta, siempre y cuando la imagen escaneada tenga regular calidad para su procesamiento preliminar (binarización de cada capa de información).

La digitalización en pantalla, aunque consume cerca del 25% de tiempo más que tableta, arroja errores medios cuadráticos cuatro veces menores (similar al obtenido con el método semi-automático).

REFERENCIAS BIBLIOGRÁFICAS

1. Alfaro, M. y Mosalvo, A. (2006). *Métodos de digitalización en el mapa forestal español [en línea] 2006*. Dirección URL: <http://www.nosolosig.com>. [Consulta: 06 Feb.2007].
2. Aumann, G., Ebner, H. y Tang, L. (1991). Automatic derivation of skeleton lines from digitized contours. *Journal of Photogrammetric and Remote Sensing*, 46(02), 259–268.
3. Bosque, S. (1992). *Sistemas de información geográfica*. Madrid: Ediciones RIALP, S.A.
4. Carter, J. (1988). Digital representations of topographic surfaces. *Photogrammetric Engineering and Remote Sensing*, 54(11), 1577–1580.
5. Decelis, J. (2001). *Digitalización automatizada de líneas en mapas Raster*. Recuperado el 12 de marzo de 2001 en <http://www.cic.inp.mx/ute/indiceute/tesistas/esistas/TesisMC/JOAQUINDECELIS/JOAQUINDECELISM2.htm>.
6. Fernández, X., López, O. Y Riveiro, F. (2000). *Separación de colores orientado a la vectorización de planos topográficos*. Recuperado el 6 de julio de 2005 en <http://www.gpi.tsc.uvigo.es/pub/reports/cap4>.
7. Grupo de Procesado de Imagen (2002). *Entorno de vectorización automático*. Recuperado el 6 de julio de 2002 en <http://www.gpi.tsc.uvigo.es/pub/papers/jipi.pdf>.
8. Levachkine, S. (2000). *Digitalización automatizada de mapas Raster*. Recuperado en 12 de marzo de 2006 en <http://www.revista.unam.mx/vol.0/art3/patrones.html>.
9. Longley, P., Goodchild, M., Maguire, D. y Rhind, D. (1999). *Geographical information systems* (Vol 1). USA: John Wiley and Sons, INC.

10. Nero, M. y Cintra, J. 2001. *Digitalización de mapas: estudio comparativo de metodologías*. Recuperado el 5 de junio de 2006 en <http://w.w.w.asovig.nstemp>.
11. TERRAINMAP.COM. (2007). *Digital Terrain Modeling and Mapping Journal*. Recuperado el 17 de abril de 2007 en <http://www.terrainmap.com>.