

APLICACIÓN DE ÍNDICES DE FORMA EN LA DETERMINACIÓN DE POLÍGONOS SLIVERS DERIVADOS DE LAS DIMENSIONES LINEALES DE UNA FIGURA Y SU ÁREA

APPLICATION OF SHAPE INDICES TO THE CHARACTERIZATION OF SLIVERS POLYGONS USING THE LINEAR DIMENSIONS OF A FIGURE AND ITS AREA

Miguel Sánchez Marco*, Áurea Gallego Salguero

Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría, Universidad Politécnica de Valencia, Camino de Vera s/n, Valencia 46022. msanchma@cgf.upv.es; augalsal@cgf.upv.es

Abstract:

Nowadays cartography is more and more required by users, companies and government agencies. A lot of tasks need digital, updated and quality cartography: Civil work design, planning, landscaping, environmental impact, conservation, etc. This need of cartography is mostly solved by spatial data infrastructures and map servers that make downloading and getting maps easy. Later, using Geographical Information Systems (GIS) it is quite easy to perform queries and geoprocessing operations in order to get valuable information which is difficult to get with other tools. However slivers polygons are a serious problem. They can alter the results in spatial analysis, queries and map portray. A map with sliver polygons is a map with false polygons and therefore with errors. We need to remove them. Removing sliver polygons is simple once they have been located. The real problem is to manage to locate them. Some sliver polygons are small, others are large but in both cases sliver polygons are difficult to locate and tell them apart from other polygons. Scientists have used formulas to calculate the shape of objects for many years. Today there are dozens of formulas available that could be used in the characterization and location of polygons slivers. In this article many of these shape formulas are analyzed and conclusions are drawn on whether they can be used in locating sliver polygons.

Key words: Geographical Information Systems, Sliver polygons, Form indices, Area, Perimeter, Linear dimensions

Resumen:

Hoy en día la cartografía es cada vez más demandada por usuarios, empresas y agencias gubernamentales. Muchas tareas necesitan cartografía digital, actualizada y de calidad: diseño de obra civil, planificación, paisajismo, impacto ambiental, conservación, etc. Esta necesidad de cartografía se resuelve principalmente mediante infraestructuras de datos espaciales y servidores de mapas que facilitan la descarga y la obtención de mapas. Utilizando los Sistemas de Información Geográfica (SIG) es muy fácil realizar consultas y operaciones de geoprosesamiento para obtener información valiosa que es difícil de obtener con otras herramientas. Sin embargo, los polígonos sliver son un problema serio, ya que pueden alterar los resultados en análisis espacial, consultas y representación de mapas. Un mapa con polígonos sliver es un mapa con polígonos falsos y, por lo tanto, con errores y tienen que ser eliminados. La eliminación de polígonos sliver es simple una vez que se han localizado, ya que el verdadero problema es lograr localizarlos. Algunos polígonos sliver son pequeños, otros son grandes, pero en ambos casos los polígonos sliver son difíciles de localizar y distinguirlos de otros polígonos. Los científicos han usado fórmulas para calcular la forma de los objetos durante muchos años. Hoy en día hay disponibles docenas de fórmulas que se podrían usar en la caracterización y localización de polígonos slivers. En este artículo se analizan muchas de esas fórmulas de forma y se obtienen conclusiones sobre si pueden ser utilizadas en la localización de polígonos sliver.

Palabras clave: Sistemas de información geográfica, polígonos slivers, índices de formas, área, perímetro, dimensiones lineales

1. Introducción

Los polígonos slivers son polígonos que se generan cuando los contornos de los polígonos no son coincidentes en su trazado. Estos polígonos se suelen generar normalmente cuando se superponen dos cartografías y los contornos de los polígonos no coinciden exactamente (Fig. 1). También se presentan cuando se digitaliza el mismo contorno dos veces para polígonos adyacentes y sus bordes no coinciden (Fig. 2).

Los polígonos slivers suponen un problema serio que afecta a los resultados en el análisis espacial, a las consultas en los Sistemas de Información Geográfica (SIG) o en la representación de mapas, siendo necesario eliminarlos de la base de datos.

La propagación de errores en el análisis de superposición ha sido estudiada por diversos autores (Shi *et al.* 2004; Veregin 1989). La aparición de estos polígonos origina que el usuario disponga de una cartografía con polígonos falsos y, por tanto, con

* Corresponding Author: Miguel Sánchez Marco, msanchma@cgf.upv.es

errores. Por lo tanto, es necesario que cualquier cartografía esté libre de estos polígonos erróneos (Maraş *et al.* 2010).

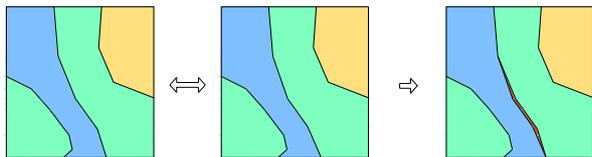


Figura 1: Superposición de elementos superficiales: generación de slivers.

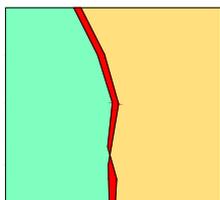


Figura 2: Digitalización: generación de slivers.

Para evitar que se generen polígonos slivers, durante la operación de superposición, existen técnicas que tratan los bordes de los polígonos como si fueran líneas fuzzy (las líneas tienen posición variable dentro de una tolerancia Épsilon (ϵ) prefijada, de modo que, si dos líneas caen dentro de una misma zona de tolerancia, ambas líneas se unen) (Klajnsek and Zalík 2005; Chrisman 1983).

Para evitar generar polígonos slivers durante la digitalización o compilación se debe emplear técnicas adecuadas. Por ejemplo, es recomendable hacer una digitalización topológica en vez de una digitalización en spaghetti, siendo necesario emplear técnicas de snap que permitan forzar nuestras coordenadas a puntos anteriormente digitalizados (Chrisman 1987).

Todas estas técnicas buscan evitar que aparezcan polígonos slivers en la cartografía. Pero en ocasiones esto es inevitable, bien porque nos proporcionan una cartografía que ya los tiene o, pese a utilizar tolerancias o anchuras epsilon, éstas no han sido suficientes. Aumentar el tamaño de la tolerancia no es la solución, ya que se simplifica peligrosamente la representación gráfica. Además, es difícil cifrar el tamaño correcto de esa tolerancia, ya que en un SIG podemos tener objetos de digitalización precisa (edificios, embalses, carreteras, etc.) junto a otros elementos cuyos bordes son menos precisos (bosques, tipos de suelo, etc.). Por todo lo visto anteriormente, suele ser bastante habitual que las bases de datos gráficas de los SIG tengan polígonos slivers, sobre todo las procedentes de la superposición de capas.

2. Características de los polígonos slivers

Los polígonos slivers son elementos especiales con formas muy variadas y formadas por diversos motivos. Se puede decir que, generalmente, los polígonos slivers presentan las siguientes características:

- El área suele ser pequeña, sobre todo con respecto a sus vecinos.
- Suelen presentar ángulos muy agudos.

- Los polígonos slivers suelen estar compuestos únicamente por dos arcos, mientras que esto no es habitual en el resto de polígonos.
- Los slivers suelen presentar cuatro arcos en un nodo, mientras que lo normal, en otros polígonos, es tres.
- Presentan correlación espacial. Un sliver suele tener polígonos slivers conectados y suelen hacerlo en ángulos agudos.
- Alternancia de atributos. (al superponer dos polígonos y los bordes no coinciden se generan polígonos sliver en cadena, perteneciendo un sliver a uno de los polígonos y el siguiente al otro de los polígonos).
- Suelen tener gran elongación (relación ancho/largo).
- Presentan mucho perímetro frente al área que generan.

3. Índices de forma derivados de las dimensiones lineales de una figura y su área en la determinación de polígonos slivers

Existe una gran variedad de métodos con sus respectivos índices para medir la forma.

Diversos autores han buscado en los índices de forma las siguientes características (Austin 1984; Moellering and Rayner 1981; Moellering and Rayner 1982; Griffith 1982):

- Regeneración de la forma original desde un índice de forma.
- Independencia de la escala.
- Independencia de la traslación de la figura.
- Independencia del punto de partida de la figura.
- Independencia de la rotación de la figura.
- Independencia del sistema de coordenadas empleado en la digitalización de la figura.
- Correspondencia uno a uno del índice resultante y la figura.

Stoddart (1965) añade las siguientes características:

- Medidas significativas de la forma en sí, más que de sus subpropiedades.
- Fácil de calcular.
- Que se pueda expresar en términos de un índice simple.

Este índice puede ser fácilmente reestablecido en términos de forma.

La proliferación de un número tan grande de índices de forma se debe a que es imposible definir completamente una forma con un simple índice, como lo demuestra Lee and Sallee (1970). Sin embargo, sí es posible cuantificar alguno de los aspectos de la forma (elongación, compactación). Tal como demuestran Blott and Pye (2008) y Krumbein (1941) es posible utilizar índices de

forma para describir, comparar y cuantificar características simples de las formas. En este sentido, los índices de forma pueden servir perfectamente para detectar polígonos slivers.

Son índices que utilizan conceptos sobradamente conocidos, como la anchura y la altura de las figuras. Según Frolov (1975), fue Vereschchagin (1930) quien propuso medir la elongación de una figura mediante el cociente del eje más largo por el eje más corto, ambos en direcciones perpendiculares, véase Ec. (1).

$$I_1 = \frac{l}{w} \quad (1)$$

Cuya imagen es $[1, \infty)$, 1 para el círculo, cuadrado o cualquier otra figura donde $w=l$, e ∞ para la recta. Otros autores, como Folk (1968), citan la inversa del anterior índice como un buen medidor de la elongación de partículas en petrología, véase Ec. (2):

$$I_2 = \frac{w}{l} \quad (2)$$

Donde ahora la imagen es $(0, 1]$, con 0 para la recta y 1 en las figuras donde $w = l$.

También existe una variante de la anterior, véase Ec. (3):

$$I_3 = \sqrt{I_2} = \sqrt{\frac{w}{l}} \quad (3)$$

Davis (2002) llama índice de forma de grano a la relación entre el perímetro de la figura y la longitud de su eje más largo, véase Ec. (4):

$$I_4 = \frac{P}{l} \quad (4)$$

Cuya imagen es $(2, \pi)$, donde 2 corresponde con un polígono en forma de recta y los valores grandes los consiguen los polígonos pequeños con mucho perímetro (muy enrollados).

En 1932 Horton utiliza el índice siguiente, véase Ec. (5):

$$I_5 = \frac{A}{l^2} \quad (5)$$

Cuya imagen es $(0, \frac{\pi}{4})$, donde 0 corresponde a la recta y $\frac{\pi}{4}$ al círculo.

En 1927 Pentland obtiene el siguiente índice, como resultado del cociente del área de la figura entre el área de un círculo con diámetro igual al eje más largo de la figura l , véase Ec. (6):

$$I_6 = \frac{4A}{\pi l^2} \quad (6)$$

Cuya imagen es $(0, 1]$, donde 0 corresponde a la recta y 1 al círculo.

En 1965 Haggett utiliza la fórmula de Pentland (1927), en la forma $\frac{1.27A}{l^2}$. Años antes, Schumm (1956), obtiene la siguiente fórmula, calculando el cociente del diámetro de un círculo que tiene la misma área de la figura de estudio, entre la longitud del eje más largo de la figura, véase Ec. (7):

$$I_7 = \frac{d_c}{l} = \frac{2\sqrt{A}}{l} \quad (7)$$

que es la fórmula de Pentland dentro de la raíz cuadrada y cuya imagen es también $(0, 1]$.

Son varios los autores los que argumentan la inutilidad de estos índices en la descripción de la forma (Folov 1975; MacEachren 1985; Austin 1984; Stoddart 1965). Se puede decir que los índices que utilizan la longitud o anchura de los polígonos a lo largo de ejes rectos, presentan un gran problema para detectar slivers, tanto los que no son rectilíneos como los que tienen formas normales. En la Figura 3 puede verse como dos slivers, con similares características en cuanto a área y perímetro, presentan valores muy distintos en cuanto a la longitud.

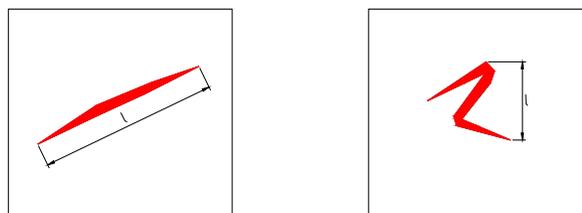


Figura 3: Slivers con la misma área y perímetro pero con distinta longitud.

En cambio, estos índices pueden jugar un buen papel cuando lo que buscamos son slivers con la forma clásica, alargada y rectilínea que muestra la figura anterior. Solo en este caso las conclusiones siguientes deben ser tenidas en cuenta. Cuando los slivers son rectilíneos y alargados, prácticamente todos los índices pueden ayudar a detectarlos. De hecho, Austin (1984) afirma que muchas de las diferencias que muestran muchos de estos índices no son más que un factor de escala.

4. Aplicación de los índices

En este apartado se va analizar el comportamiento de los índices de forma, derivados de las dimensiones lineales de una figura y su área, aplicados a una serie de polígonos con características y formas diferentes, siendo conocidas sus dimensiones lineales y áreas (Tabla 1).

Las nomenclaturas utilizadas para los diversos parámetros de los índices de estos polígonos son los siguientes:

- A: Área
- P: Perímetro
- l: Longitud eje largo figura
- w: Perpendicular

Las fórmulas utilizan la anchura y longitud, relacionadas entre ellas o con el área o el perímetro. Los índices empleados son los anteriormente definidos, van desde I_1 al I_7 .

Con respecto a los seis índices anteriores, se ha hecho una clasificación dependiendo de las imágenes, obteniendo cuatro casos.

En el primer caso corresponde al índice I_1 , que tiene una imagen con los valores de $[1, \infty)$ siendo el 1 para el círculo, cuadrado o cualquiera otra figura, e ∞ para la recta (Tabla 2).

Tabla 1: Características geométricas de los polígonos analizados.

	Área (m ²)	Perímetro (m)	l (m)	w (m)
	1567	844	275	125
	2.5	7.8	3.3	1.6
	58.9	236.6	117.8	0.5
	2400	616	300	8
	2400	616	88	76
	21955	2064	1560	408

Tabla 2 : Datos obtenidos en el cálculo de los índices derivados de las dimensiones lineales de una figura y su área. Para valores de [1, ∞).

	I_1
	2.2
	2.06
	235.6
	37.5
	1.158
	3.824

Cuanto más largo y estrecho sea el polígono mayor será el resultado, como ocurre en el tercer polígono a la vista de los resultados. En cambio, en los polígonos con formas más compactas, redondeados, incluso triangulares, doblados etc. el resultado se acerca más a 1, como ocurre en el segundo y quinto polígono.

En el segundo caso, la imagen tiene unos valores de (0, 1], donde el 1 corresponde al círculo y el 0 a la recta. Esta imagen se corresponde con los índices I_2, I_3, I_6, I_7 . En la Tabla 3 se observan los valores para estos índices.

Tabla 3: Datos obtenidos en el cálculo de los índices derivados de las dimensiones lineales de una figura y su área. Para valores de (0, 1].

	I_2	I_3	I_6	I_7
	0.454	0.674	0.026	0.162
	0.484	0.696	0.292	0.540
	0.004	0.065	0.005	0.073
	0.026	0.163	0.034	0.184
	0.864	0.929	0.395	0.628
	0.262	0.511	0.0114	0.107

Se puede observar que los valores más cercanos a 1 se corresponden con los polígonos que tienen cierta forma

triangular y están doblados, como es el caso del segundo y quinto polígono. Por lo tanto, podemos decir que este tipo de fórmulas presentan un buen rendimiento para los polígonos muy estrechos y alargados, como es el caso del tercer polígono.

En el tercero de los casos, la imagen que se utiliza es la correspondiente a $(2, \pi)$ para el índice I_4 , donde el 2 corresponde a la recta y π al círculo.

En la Tabla 4 se observa que el valor más cercano a 2 es el correspondiente al tercero de los polígonos, siguiendo a continuación el cuarto de ellos. El segundo de los polígonos se acerca también a este valor, debido a que su perímetro está muy cercano al doble de su longitud.

Tabla 4: Datos obtenidos para el cálculo de los índices derivados de las dimensiones lineales de una figura y su área. Para valores de $(2, \pi)$.

	I_4
	3.069
	2.363
	2.008
	2.053
	7
	1.323

En el cuarto caso se utiliza una imagen $(0, \pi/4)$ para el índice I_5 , donde el 0 corresponde a la recta y $\pi/4$ al círculo (Tabla 5).

Tabla 5: Datos obtenidos para el cálculo de los índices derivados de las dimensiones lineales de una figura y su área. Para valores de $(0, \pi/4)$.

	I_5
	0.0207
	0.229
	0.0003
	0.0266
	0.3099
	0.0092

En el índice I_5 el polígono más cercano a 0 es el tercero ya que tiene un área relativamente pequeña con respecto a su longitud.

Estas dos últimas fórmulas usan la relación entre el perímetro, área de la figura y su longitud referida al eje más largo.

5. Conclusiones de la aplicación de los índices

En el apartado anterior se han aplicado diversos índices a una serie de polígonos habituales en la cartografía vectorial. La respuesta de cada uno de los índices a cada polígono es diferente, como cabía esperar. No obstante, como se ha visto, existen casos particulares de polígonos en los cuales los índices arrojan valores un tanto inesperados.

A continuación, se presenta una tabla resumen con los resultados del estudio de los seis polígonos característicos utilizados en el apartado anterior. En ellas se indica para cada índice el resultado obtenido. Este resultado se ha codificado de tres formas, con la leyenda siguiente: B (índice adecuado), NR (no recomendable) y M (índice no adecuado). Cuando decimos que un índice es adecuado (B) para una forma particular de polígono, es debido a que el índice detecta con el valor numérico obtenido la forma (alargada, redondeada, etc.) que tiene el polígono.

Si un índice indica con el valor numérico obtenido, dentro de su imagen de valores, que el polígono es alargado cuando en realidad no lo es, calificamos ese índice para un polígono como no adecuado (M).

Por último, si el índice da un valor intermedio dentro de su imagen, que no nos permite decantarnos por ninguna forma en concreto, calificamos al índice para esa forma de polígono como no recomendable (NR).

Para los índices de forma derivados de las dimensiones lineales y su área obtenemos se obtienen los resultados mostrados en la Tabla 6.

No hay ningún índice de las tablas anteriores que obtenga una clasificación adecuada para todos los polígonos. Todos los índices tienen una calificación M en alguno de los polígonos, siendo el que presenta islas en su interior es el que peor calificación obtiene. Este tipo de polígonos se caracteriza por tener unas dimensiones, tanto en alto como ancho, considerables; su forma es redondeada, pero su área se ve reducida por la existencia de los polígonos islas que contiene en su interior. No obstante, su perímetro aumenta considerablemente, ya que debe computarse el perímetro exterior, así como el perímetro de las islas de

su interior. Sin embargo, se debe tener en cuenta que el propósito es ver si se puede usar alguno de estos índices para detectar polígonos slivers. Los polígonos slivers no suelen presentar polígonos con islas en su interior, por lo que se podría utilizar esta característica para diferenciar slivers de no slivers sin utilizar ningún índice de forma de los utilizados anteriormente.

(Tabla 6: Resultados de los índices de las dimensiones lineales y su área.

	I_1	I_2	I_3	I_6	I_7	I_4	I_5
	NR	B	NR	B	B	NR	B
	B	NR	B	NR	NR	NR	M
	B	B	B	B	B	B	B
	B	B	B	B	B	B	B
	M	M	M	B	NR	M	B
	B	M	NR	M	M	B	M

Todos los índices han dado un resultado positivo para los polígonos estrechos y alargados, que tienen poca área y mucho perímetro, siendo estas características propias de los polígonos slivers. Estos polígonos suelen tener gran elongación (relación ancho/largo) y además presentan mucho perímetro frente al área que generan. Esto no quiere decir que todos los polígonos que tengan esta forma sean slivers, por ejemplo un camino suele presentar mucho perímetro con respecto a su área, sin que deba de ser considerado un polígono sliver.

En futuros trabajos se plantea la utilización de otro tipo de índices que clasifique de forma más eficaz los polígonos en slivers o no slivers, para poder establecer estadísticamente unas cifras o umbrales, a partir de los cuales determinar qué polígonos son slivers con una probabilidad alta.

La detección de los polígonos slivers es un proceso complejo y necesario para dotar a la cartografía de consistencia topológica, necesaria para llevar a cabo determinadas operaciones de análisis espacial en SIG.

References

- AUSTIN, R.F., 1984. *Measuring and comparing two-dimensional shapes*, pages 293-312. Netherlands: Springer.
- BLOTT, S.J., and PYE, K., 2008. Particle shape: a review and new methods of characterization and classification. *Sedimentology*, **55**(1), pp. 31-63.
- CHRISMAN, N.R., 1983. Epsilon filtering: a technique for automated scale changing. *In 43 rd Annual Meeting of the American Congress on Surveying and Mapping*, pp. 322-331.
- CHRISMAN, N.R., 1987. Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. *International Journal of Geographical Information Science*, **1**(3), pp. 265-277.
- DAVIS, J.C., 2002. *Statistics and Data Analysis in Geology*. New York: (third ed.) John Wiley & Sons Inc. 638 pages.
- FOLK, R.L., 1968. *Petrology of sedimentary rocks*. Texas: Hemphill Publishing Company.
- FROLOV, Y.S., 1975. Measuring shape of geographical phenomena: A history of the issue. *Soviet geography: Review and Translation*, **16**(10), pp. 676-687.
- GRIFFITH, D.A., 1982. Geometry and Spatial Interaction. *Annals of the Association of American Geographers*, **72**(3), pp. 332-346.

- HAGGET, P., 1965. *Locational Analysis in Human Geography*. London: Arnold.
- HORTON, R., 1932. Drainage basin characteristics Transactions of the American. *Geophysical Union*, **13**, pp. 350-361.
- KLAJNSEK, G. and ZALIK, B., 2005. Merging polygons with uncertain boundaries. *Computers & Geosciences*, **31**(3), pp. 353-359.
- KRUMBEIN, W.C., 1941. Measurement and geological significance of shape and roundness of sedimentary particles. *Journal of Sedimentary Petrology*, **11**, pp. 64-72.
- LEE, D.R. and SALLEE, G.T., 1970. A Method of Measuring Shape. *Geographical Review*, **60**(4), pp. 555-563.
- MACEACHREN, A.M., 1985. Compactness of Geographic Shape: Comparison and Evaluation of Measures. *Geografiska Annaler. Series B, Human Geography*, **67**(1), pp. 53-67.
- MOELLERING, H. and RAYNER, J.N., 1981. The Harmonic Analysis of Spatial Shapes Using Dual Axis Fourier Shape Analysis (DAFSA). *Geographical Analysis*, **13**, pp. 64-77.
- MOELLERING, H. and RAYNER, J.N., 1982. The Dual Axis Fourier Shape Analysis of Closed Cartographic Forms. *The Cartographic Journal*, **19**, pp. 53-59.
- PENTLAND, A., 1927. A method of measuring the angularity of sands. *In Proceedings and Transactions of the Royal Society of Canada*, **21**.
- SCHUMM, S.A., 1956. Evolution of drainage systems and slopes in badlands at perth amboy. *Bulletin of the Geological Society of America*, **67**(5), pp. 597-646.
- SHI, W., CHEUNG, C.K. and TONG, X., 2004. Modelling error propagation in vector-based overlay analysis. *ISPRS Journal of Photogrametry and remote Sensing*, **59**(1), pp. 47-59.
- MARAŞ, H.H., AKTUĞ, B., MARAŞ, E.E., and YILDIZ, F., 2010. Topological error correction of GIS vector data. *International Journal of the Physical Sciencies*, **5**(5), pp. 476-483.
- STODDART, D.R., 1965. The shape of atolls. *Marine Geology*, **3**(5), pp. 369-383.
- VEREGIN, H., 1989. *Accuracy of Spatial Database*. London: Taylor & Francis, pp. 3-18.
- VERESHCHAGIN, G., 1930. *Methods of lake Morphometry*. Ekspeditsii, part II Geography.