Detección automática de edificios en imágenes aéreas mediante técnicas de teledetección y análisis contextual.

Hermosilla, T.; Ruiz, L.A., Fdez. Sarría, A.; Recio, J.A.
Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría
Universidad Politécnica de Valencia
Camino de Vera s/n
46022 Valencia, Spain.
txohergo@topo.upv.es

ABSTRACT

One of the most demanded applications in the field of digital image processing is the detection of buildings in urban and periurban areas for including them into cartographic databases. This paper proposes an automatic building detection method based on object oriented classification. These objects are obtained by employing a multirresolution segmentation method over satellite or aerial high resolution images, and their classification is combined with the study of the contextual relationships of the objects.

The first segmentation level is generated by using the cartographic limits provided by the cadastral subparcels. The second level is performed using the region growing segmentation method which automatically generates homogeneous sub-objects contained into the cartographic limits. Objects obtained from both segmentations are classified. The assessment of the method is performed by using two different points of view: quality in the detection of zones containing buildings (area or pixel level evaluation) and efficiency to find the buildings as independent units (object level evaluation).

RESUMEN

Una de las aplicaciones más demandadas en el ámbito del procesado digital de imágenes es la detección de edificios en áreas urbanas y periurbanas para su incorporación en las distintas bases de datos cartográficas. Se propone un método automático de detección de edificios basado en la clasificación orientada a objetos de polígonos obtenidos por segmentación multirresolución de imágenes aéreas o de satélite de alta resolución, combinados con el estudio de las relaciones de contexto de los objetos.

El primer nivel de segmentación se realiza utilizando los límites cartográficos de subparcelas catastrales. El segundo consiste en la generación automática de subobjetos, contenidos en los límites cartográficos por el procedimiento de crecimiento de regiones. Los objetos resultantes de ambas segmentaciones son clasificados posteriormente. La evaluación del método se realiza por medio de dos enfoques distintos: calidad en la detección de las zonas que contengan edificios (evaluación por área ó por píxel) y eficacia en encontrar los edificios como unidades (evaluación por objeto).

1. INTRODUCCIÓN.

Factores tales como el alto grado de dinamismo existente en nuestro territorio hace que las zonas periurbanas se hallen sometidas a una constante alteración de sus paisajes y usos característicos por la construcción de nuevas viviendas y edificaciones. Si a esto se añaden los casos referentes a la construcción de edificaciones de manera ilegal, se hace patente la necesidad de un permanente mantenimiento y actualización de los edificios en las bases de datos cartográficas y catastrales. Hasta el momento esta actualización se ha realizado mediante técnicas clásicas de fotointerpretación, con los costes en personal y tiempo que el empleo esta metodología conlleva. Con el fin de facilitar y automatizar en la medida de lo posible esta tarea, cada vez es más requerida la inclusión en el proceso de actualización de técnicas de tratamiento digital de imágenes.

La clasificación y detección de edificios, así como otros elementos propios de entornos urbanos y periurbanos es una tarea compleja de realizar mediante técnicas de teledetección. Valores con idéntica reflectancia espectral pueden corresponder a objetos distintos con distintas funciones. La variabilidad de

los edificios y su complejidad en cuanto a su disposición espacial de los edificios implican en muchos casos que los niveles digitales de los píxeles presenten la reflectancia de distintos materiales combinados. Debido a esto, con la utilización únicamente de variables derivadas de las características espectrales para la clasificación, ya sea esta orientada a píxeles (Knudsen y Nielsen, 2004) o a objetos (Hofmann et al., 2006) no se obtienen resultados consistentes, haciéndose necesaria la utilización de información geométrica y espacial que permita la contextualización de los elementos. La información geométrica ha sido definida hasta el momento mediante el estudio de las relaciones espaciales de las alineaciones obtenidas por diversos algoritmos de detección de bordes (Sohn y Downman, 2001) o por el empleo de la transformada de Hough (Wei et al., 2004). Para realizar la caracterización contextual de los edificios, diversos autores han analizado y utilizado las relaciones existentes de estos con sus sombras proyectadas, bien sea a partir de imágenes aéreas (Lin y Netavia, 1998) o de imágenes de satélite de alta resolución (Shackelford et al, 2004; Jin y Davis, 2005), de forma que esta información puede ser combinada con otras técnicas de clasificación.

Se presenta una metodología para automatizar la detección y localización de edificios en la que se combina la clasificación orientada a objetos de distintos niveles de resolución, con las relaciones contextuales existentes entre estos y las sombras presentes en la imagen. La metodología propuesta ha sido aplicada y evaluada sobre el polígono catastral número 19 del término municipal de Benicarló, en la provincia de Castellón. La evaluación de su funcionamiento se ha llevado a cabo utilizando dos enfoques distintos: evaluando por un lado la calidad en la detección de las zonas que contienen edificios (evaluación por área ó por píxel) y por el otro, comprobando la eficacia en la identificación de los edificios como unidades independientes (evaluación por objeto). Las pruebas se han llevado a cabo utilizando un mosaico en falso color infrarrojo –*Colour InfraRed* (CIR)–, con las bandas infrarroja (IR), roja (R) y verde (G), generado a partir de las ortofotos digitales independientes que adquiridas cámaras fotogramétricas DMC (*Digital Mapping Camera*). La resolución de la imagen final es de 0.5 m/píxel y ha sido obtenida mediante un proceso de fusión propio del sistema. La resolución espectral empleada es de 8 bits por píxel. Se ha utilizado, además, la cartografía de las subparcelas catastrales de la zona de estudio.

2. METODOLOGÍA.

La metodología que se propone para la detección y localización de edificios combina tres fuentes de información diferentes: una clasificación orientada a objetos realizada utilizando los límites de las subparcelas catastrales; una segunda clasificación orientada a objetos efectuada sobre los polígonos generados tras una segmentación automática de la imagen (subobjetos), y las relaciones espaciales de estos objetos con las sombras detectadas de la imagen. Así, mediante la clasificación de las subparcelas se restringen los objetos a estudiar a aquellos que han sido asignados a la clase *Edificios* o a alguna con la que esta puede ser confundida. Una vez definidas las clases sensibles a contener edificios, se aplica en paralelo una segmentación a la imagen, con el fin de tener subobjetos de menor tamaño y mayor detalle que los definidos por la cartografía catastral. Estos subobjetos serán nuevamente clasificados, y por último, se combinará la información de ambas clasificaciones con la posición de las sombras presentes en la imagen para la delimitación de los edificios. En la Figura 1 se muestra el esquema del proceso empleado para llevar a cabo la detección y localización de los edificios.

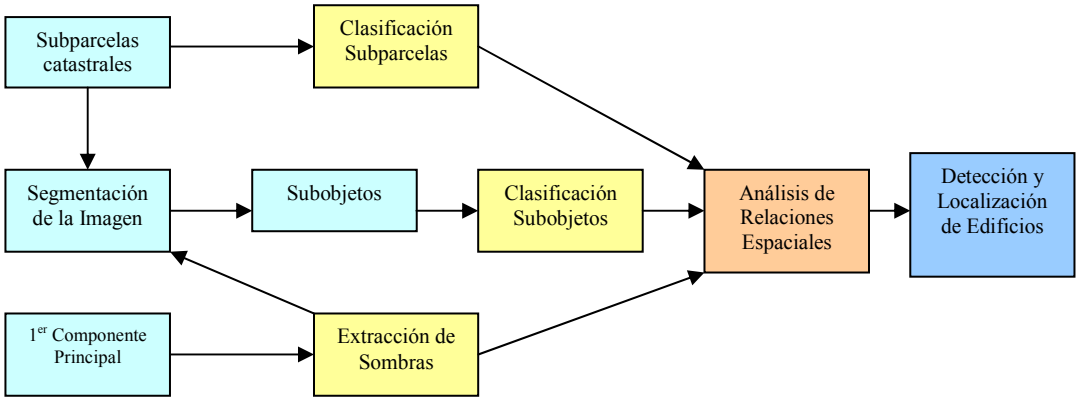


Figura 1.- Esquema de la metodología de detección de edificios propuesta.

2.1 Clasificación orientada a objetos cartográficos.

Como primer paso, se ha realizado una clasificación orientada a objetos, utilizando para la definición de éstos los límites de las subparcelas catastrales. De cada objeto (subparcelas) se han extraído características descriptoras de cuatro tipos: espectrales, de textura, estructurales y de forma. Las características espectrales informan a cerca de la respuesta espectral de los objetos, la cual depende de los tipos de coberturas existentes. Las características de textura aporta información acerca de la distribución espacial de los valores de intensidad en la imagen (Ruiz et al., 2004). Las características estructurales (Recio et al., 2006; Ruiz et al., 2007) están relacionadas con los patrones de distribución de los elementos vegetales de una parcela, principalmente los árboles. Las características de forma proporcionan información sobre propiedades como la complejidad de la forma o la compacidad de los objetos.

Previamente a la clasificación, se eliminaron aquellas características descriptoras redundantes mediante un proceso de selección. Para ello se emplearon las técnicas de análisis de componentes principales, y análisis discriminante descriptivo, se seleccionando las características más significativas para la separación de las distintas clases. Con estas características se realizó la clasificación utilizando la técnica de análisis discriminante predictivo, asignándose así a cada subparcela una de las clases. Estas fueron definidas tras el análisis de la zona de estudio, siendo finalmente once clases: *Cítricos, Cítricos jóvenes, Edificios, Frutales de Secano, Hortícolas, Invernaderos, Huerto familiar, Matorrales, Olivos, Vías de Comunicación,* y *Terrenos arables.*

Una vez clasificadas las subparcelas catastrales, el siguiente paso consiste en la búsqueda de aquellas subparcelas a las que se haya asignado una clase susceptible de contener edificios (*Edificios, Invernaderos, Huerto familiar*) o en las que se haya producido un alto grado de confusión con estas clases (*Terrenos arables, Vías de Comunicación*).

2.2 Detección y extracción de las sombras.

Pese a que las sombras han supuesto tradicionalmente una fuente de problemas al trabajar con imágenes de alta resolución en entornos urbanos, debido al ocultamiento de la información que estas producen (Madhavan et al., 2004), la metodología propuesta se beneficia de su presencia para la detección de los edificios.

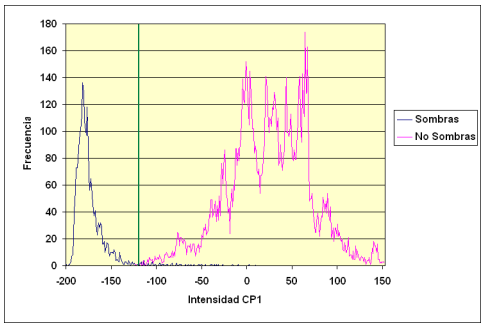
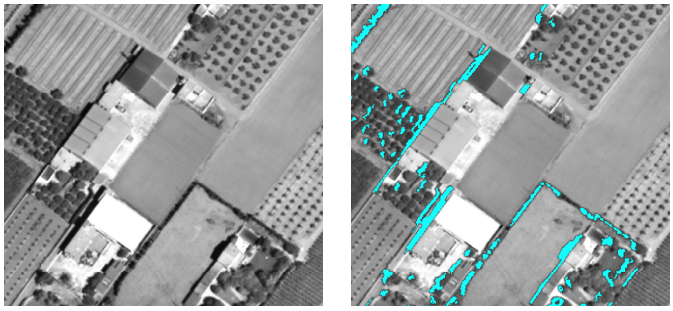


Figura 2.- Histogramas de píxeles con y sin sombra, y valor de umbralización utilizado.



a. Primer Componente Principal b. Sombras detectadas (en azul)

Figura 3.- Proceso de extracción de las sombras.

Para el proceso de extracción de las sombras de la imagen se ha efectuado en primer lugar el análisis de componentes principales (CP) de las tres bandas de la imagen. Posteriormente, y trabajando sobre el primero de los ellos, se busca el valor de intensidad que mejor permita aislar las sombras mediante una umbralización. Para ello, se seleccionan píxeles pertenecientes a zonas con sombras por un lado, y píxeles libres de estas por otro. Con los valores de ambos grupos se dibujan sus histogramas con los que se determina el valor de intensidad por el cual las sombras quedan aisladas (Figura 2).

2.3 Segmentación automática y clasificación de subobjetos.

Utilizando la segmentación automática se definen objetos de menor tamaño que las subparcelas catastrales. El algoritmo de segmentación empleado es el de multirresolución incluido en el programa *eCognition* de *Definiens Imaging*. Se trata de una técnica de crecimiento por regiones que comienza considerando objetos de un píxel. En pasos sucesivos, los objetos se mezclan con otros generando agrupaciones mayores, mientras se va calculando como varía la heterogeneidad, de forma que se unirán aquellos objetos adyacentes que minimicen esta modificación de la heterogeneidad. El factor de escala es

una medida que ha de ser especificada y representa el máximo cambio que puede producirse en la heterogeneidad cuando dos objetos son unidos. Con este factor se define indirectamente el tamaño de los objetos, ya que cuanto mayor cambio se permita, mayores serán los objetos resultantes. La selección de un valor del factor de escala óptimo es un problema común al trabajar con estas técnicas. Así, un factor de escala muy alto, podría suponer un exceso de simplificación en los subobjetos generados, de forma que los edificios no fueran definidos y aislados correctamente. Por el contrario, la utilización de un factor de escala muy bajo produciría que un mismo edificio contenga demasiados subobjetos.

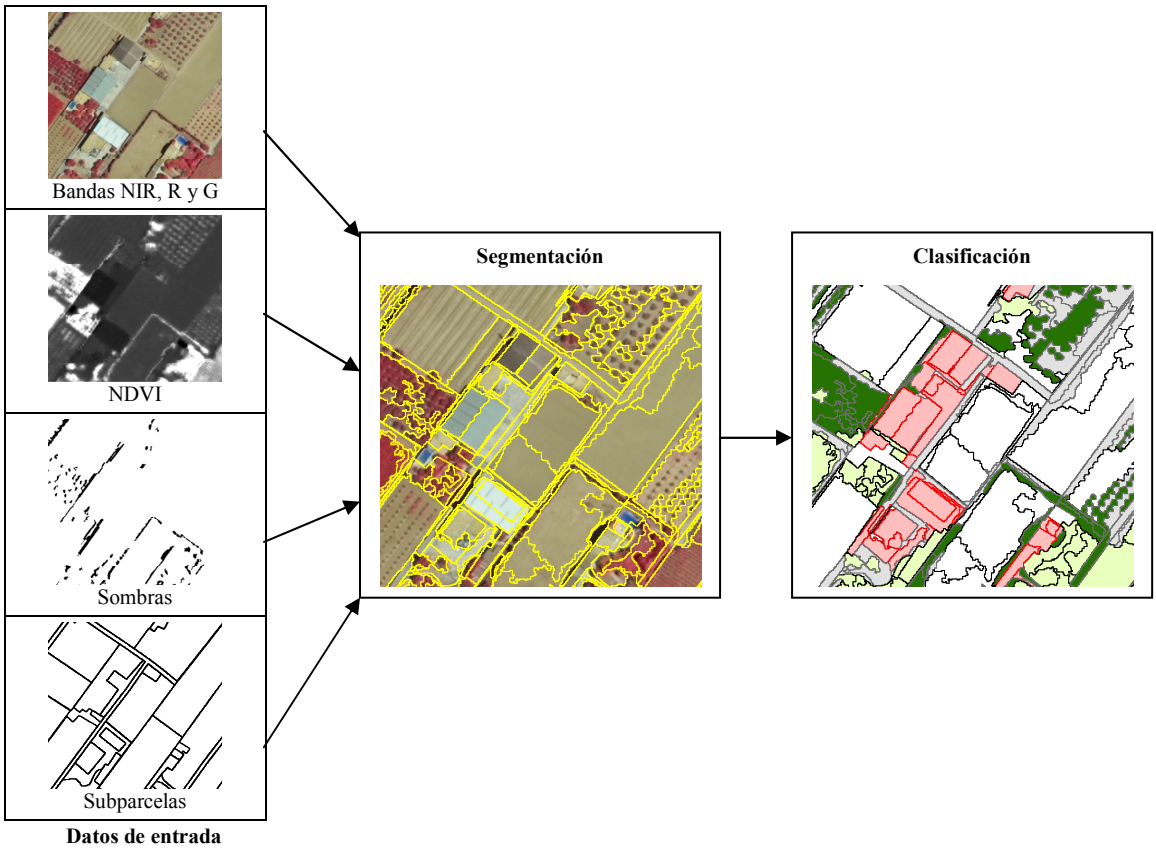


Figura 4.- Esquema del proceso de segmentación.

En la Figura 4 se esquematiza el proceso de segmentación para la obtención de los subobjetos. Como datos de partida se introducen las tres bandas espectrales disponibles (NIR, R, G), la imagen del NDVI y la imagen binaria que contiene las sombras detectadas, para que los polígonos generados se ajusten al máximo a la forma de las sombras. Además, se incluyen como superobjetos, los límites de las subparcelas catastrales, de forma que los polígonos que se generen se encuentren completamente contenidos dentro de los límites de la cartografía catastral. El último paso es la realización de una nueva clasificación para estos subobjetos empleando características espectrales, de textura, forma y estructurales. Las clases seleccionadas para esta clasificación se encuentran fundamentalmente enfocadas a la discriminación de los edificios

2.4 Análisis de las relaciones contextuales.

Utilizando la metodología descrita, las sombras se presentan como pilar fundamental para la localización de los edificios. Puesto que estas no son únicamente proyectadas por los edificios, se hace necesaria la comprobación de cada uno de los elementos de los que provienen las sombras. Para ello se ha definido una gramática que analiza el contexto y las relaciones espaciales de los elementos.

El proceso de análisis de las relaciones espaciales utiliza en primer lugar la dirección de incidencia solar. Así, a partir de una sombra se busca en la dirección opuesta a la incidencia solar aquellos subobjetos que (i) estén muy próximos a ésta, (ii) se encuentren contenidos en subparcelas catastrales definidas como susceptibles de incluir edificios, y (iii) hayan sido clasificados como *Vías de Comunicación* o *Edificios* en la clasificación realizada en la segmentación. En caso de cumplir estas condiciones se definiría a este subobjeto como edificio, siendo el siguiente paso estudiar la clase de los subobjetos vecinos a este, independientemente de la dirección de incidencia solar. Así, se añadirán como

edificios los polígonos que estén en contacto con aquéllos y que cumplan las condiciones gramaticales de pertenencia predefinidas.

3. EVALUACIÓN.

La metodología propuesta se ha evaluado desde dos perspectivas diferentes: una referida a la precisión en la detección de áreas (píxeles) y la otra a la precisión en la detección de los objetos (edificios). Para poder llevar a cabo la evaluación, los edificios presentes en la zona de ensayo fueron manualmente delineados y utilizados como referencia.

La evaluación analítica por medio de las áreas se ha realizado utilizando una serie de parámetros estadísticos ampliamente empleados para evaluar la detección de edificios (Lin y Netavia, 1998; Jin y Davis; 2005). Estos parámetros son el porcentaje de detección, el factor de exceso y el factor de defecto. El porcentaje de detección expresa el porcentaje de superficie de edificios reales correctamente detectado por el proceso automático. El factor de exceso es una medida del error por exceso, en el que el sistema detecta incorrectamente como edificios elementos que no lo son. El factor de defecto indica el error de omisión, esto es, edificios reales que no son detectados automáticamente como tales. Para la evaluación de la detección a nivel de objeto se han utilizado dos índices empleados por Shackelford et al. (2004). El índice de corrección (*correctness*) expresa el porcentaje de edificios detectados que se solapan, al menos en parte, con la posición de un edificio real presente en la imagen. Por su parte, el índice de compleción (*completeness*) hace referencia al porcentaje de edificios reales que son solapados por edificios automáticamente detectados.

Evaluación	Parámetro	Valor
por área	Factor de Exceso	1.25
	Factor de Defecto	0.22
	Porcentaje de Detección	82.2
por objeto	Corrección	67,9%
	Compleción	87.9%

Parámetro	Valor
Factor de Exceso	1.25
Factor de Defecto	0.22
Porcentaje de Detección	82.2

Tabla 1.- Resultados estadísticos de la evaluación a nivel de área.

	Detectados automáticamente	Referencia
Edificios correctamente extraídos	125	160
Total de edificios	184	182
	Corrección = 67,9%	Compleción = 87.9%

Tabla 2.- Resultados estadísticos de la evaluación a nivel de objetos.

Analizando los resultados obtenidos en la evaluación a nivel de las áreas detectadas (Tabla 1) destaca el alto valor que presenta el factor de exceso. Al buscar en la imagen las causas del elevado valor resultante, este se explica fundamentalmente por varias razones., aunque la más importante es debida a la identificación incorrecta de zonas asfaltadas como edificios, debido a la gran similitud espectral que estas presentan. El aspecto positivo es que las zonas asfaltadas se encuentran, por lo general, en el entorno de los edificios, por lo que pese a que produzcan una alta disparidad en el área detectada como edificio, forman una unidad contigua a los edificios reales. El factor de defecto obtenido es muy bajo, lo que indica que se deja de detectar poca superficie de edificios reales. Los errores que aquí se producen se deben fundamentalmente a una incorrecta clasificación de los objetos, bien a nivel de subparcela o de subobjeto. El alto porcentaje de detección obtenido, superior a un 80%, informa sobre la capacidad de la metodología propuesta para encontrar los edificios presentes en la imagen.

Figura 6.- Ejemplos de la detección de edificios.

A nivel de objetos (Tabla 2) los resultados de la evaluación muestran con el valor del factor de compleción que un alto porcentaje de los edificios reales son, aunque sea en parte, automáticamente detectados. Por otro lado, se vuelve a constatar, examinando el valor obtenido en el factor de corrección, la sobre-detección de edificios que se produce, ya que más del 30% de los edificios detectados no coincide con ningún edificio real.

4. CONCLUSIONES.

La constante evolución de los sensores conlleva una mayor disponibilidad de datos, lo que permite el desarrollo de métodos mejor adaptados, que deben ser incorporados de forma progresiva y en combinación con los procedimientos tradicionales. El reconocimiento automático de elementos y el establecimiento de relaciones entre ellos aportan un nivel más de información. No es aconsejable que recaiga casi todo el peso en la clasificación, por lo que es importante la integración de la información y de los métodos para llegar a soluciones más eficientes y completas. La utilización conjunta de clasificadores orientados a objetos y de elementos contextuales tales como las sombras, proporciona una mayor robustez a la detección y localización de edificios. De esta forma se evita la detección errónea de elementos que a nivel espectral son fácilmente confundibles con edificios, como son los terrenos improductivos, arables o las carreteras u otras vías de comunicación.

5. AGRADECIMIENTOS.

Los autores agradecen la financiación por parte del Ministerio de Educación y Ciencia en el marco del proyecto CTM2006-11767/TECNO. Así mismo al Instituto Cartográfico Valenciano por facilitar los datos necesarios para realizar este trabajo.

6. REFERENCIAS.

- Hofmann, P., Strobl, J.m Blaschke, T., Kux, H., 2006. Detecting informal settlements from QuickBird data in Rio de Janeiro using an object based approach. En: 1st International Conference on Object-based Image Analysis.
- Jin, X., Davis, C.H., 2005. Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual, and Spectral Information. EURASIP Journal on Applied Signal Processing 14, 2196-2206.
- Knudsen, T., Nielsen, A.A., 2004. Detection of buildings through multivariate analysis of spectral, textural, and shape based features. En: Geoscience and Remote Sensing Symposium, 5, 2830-2833.
- Lin, C., Netavia, R. 1998. Building detection and description from a single intensity image. Computer Vision and Image Understanding 72, 101-121.
- Madhavan, B.B., Tachibana, K., Sasagawa, T., Okada, H., Shimozuma, Y., 2004. Automatic Extraction of Shadow Regions in High-resolution Ads40p. En: XX ISPRS Congress, Istambul.
- Ruiz Fernández, L.A, Fernández-Sarría, A., Recio Recio, J.A., 2004. Texture feature extraction for classification of remote sensing data using wavelet decomposition: A comparative study. In: Proc. of the ISRPS 2004 Conference, International Archives of Photogrammetry and Remote Sensing. Vol. XXXV, part B, pp. 1109-1115, Istanbul, Turkey. 19- 23 July 2004.
- Recio, J.A., Ruiz,L.A., Fdez. Sarría, A., Hermosilla, T., 2006. Integration of multiple feature extraction and object oriented classification of aerial images for map updating. In: Procedings of the Second International Symposium on Recent Advances in Quantitative Remote Sensing, Torrent, Valencia, Spain, pp. 391-396.
- Ruiz, L.A., Recio, J.A., Hermosilla, T., 2007. Methods for automatic extraction of regularity patterns and its application to object-oriented image classification. In: Intl. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, 3/W49A, pp. 117-121. Photogrammetric Image Analysis (PIA07). Munich, September, 2007.
- Shackelford, A.K., Davis, C.H., Xiangyun Wang, 2004. Automated 2-D building footprint extraction from high-resolution satellite multispectral imagery. En: Geoscience and Remote Sensing Symposium, 3, 1996-1999.
- Sohn, G., Dowman, I., 2001. Extraction of buildings from high resolution satellite data. In: Baltsavias, E., Gruen, A., Van Gool, L. (Eds.), Automated Extraction of Man-Made Object from Aerial and Space Images (III). Balkema Publishers, Lisse, pp.345–355.
- Wei, Y., Zhao, Z., Song, J., 2004. Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection. Geoscience and Remote Sensing Symposium, 3, 2008-2010.