Building Verification from Disparity of Contour Points

Charles Beumier
Signal & Image Centre, Royal Military Academy, Brussels, Belgium
beumier@elec.rma.ac.be

Abstract

This paper presents building verification in the context of change detection for database update. Disparity values between the left and right images of a stereo couple are estimated from grey level correlation in the vicinity of edge pixels and harmonized along contours for better coherence and accuracy. This estimation is very fast as only edge pixels are considered and accurate thanks to grey level and geometrical constraints. Building evidence is derived from the presence of edge points with high disparity near or inside building polygons of the database. Verification results compared favourably with a previous work based on straight line and shadow detection. The system intends to assist geographers in the process of topographic database update for the building layer. It also aims at delivering useful by-products like building elevation (Z coordinate) and Digital Terrain Model (DTM).

Key Words- Remote sensing, building verification, stereo, disparity, topographic database, DTM

1. Introduction

Building detection is an important field of remote sensing since it relates to human activity and security. It is crucial in applications such as security management to identify potential targets in pre-crisis situations and in change detection for the evaluation of disaster impact in post-crisis situations. In the specific application of database update, the verification of a priori building information stored in database is useful to detect and resolve inconsistencies with recent data.

The Belgian National Geographic Institute (IGN Belgium) manages the topo-geographical inventory of Belgium and publishes topographic maps and databases. The current manual update of this information is expensive, time consuming and error prone. A semi-automatic system aims at highlighting changes to draw the attention of the geographer towards difficult areas where his/her expertise is useful. This focusing strategy becomes a necessity since the time elapsed between digital information updates has to drop in the very near future (e.g. navigation systems).

First techniques that have been used to localise buildings in 2-D images are corner detection, segment groupings and region detection [1]. They were often associated to models and combined with shadow evidence [2]. Recent trends for building extraction are to use 3-D data, possibly combined with 2-D features, to integrate a priori information or to fuse several cues and sensor data [3, 4].

This paper presents the continuation of research about building verification initiated in 2006 at the Signal and Image Centre, looking for fast automatic or semi-automatic solutions. First, image straight lines were used to score database polygons with a confidence measure about building presence [5]. Shadow cues were later combined in the polygon scores for better distinction between buildings and non buildings [6]. The difficulty with straight segments is that many of them do not refer to buildings (parking lots, road markings or edges, hedges). Even if database polygons were correctly labelled as building or non building in our experiments, candidate segments for new constructions were numerous and difficult to filter. In this contribution, the elevation of edge points is shown to provide more discriminative power to confirm building presence and to possibly detect new constructions.

2. Approach

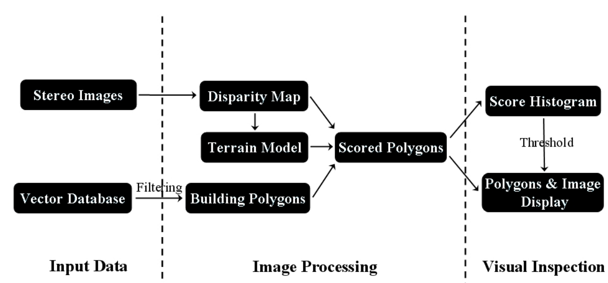

Figure 1: Approach overview

The overview of the approach for building verification, depicted in Figure 1, is separated into three parts. From the 'Input data' consisting of a stereo couple of images, and a vector database, building polygons of the database are scored automatically by 'Image Processing' tools. These estimate the elevation from the disparity of corresponding edge pixels in the left and right images. The semi-automatic aspect of the approach concerns the 'Visual Inspection' by the user of scored polygons superimposed on image data to let him/her check buildings insufficiently supported by elevation evidence.

In the 'Image Processing' part, the derivation of the disparity map from stereo images is the most important and difficult component. This is why the whole section 3 is devoted to disparity estimation. The terrain model and scoring of polygons exploit statistics of the disparity distribution and are described respectively in subsections 4.1 and 4.2. Subsection 4.3 presents the procedure of 'Visual inspection'.

Results about building verification are presented and discussed in section 5. Section 6 concludes the paper.

3. Disparity estimation

3.1. Computational stereo

Computational stereo refers to the determination of 3-D structure from two or more images taken from different viewpoints. Refer to [1, 7] for good surveys of methods. Having at one's disposal a stereo couple of images (Figure 2) from airborne or satellite sensors is more and more common in remote sensing applications.

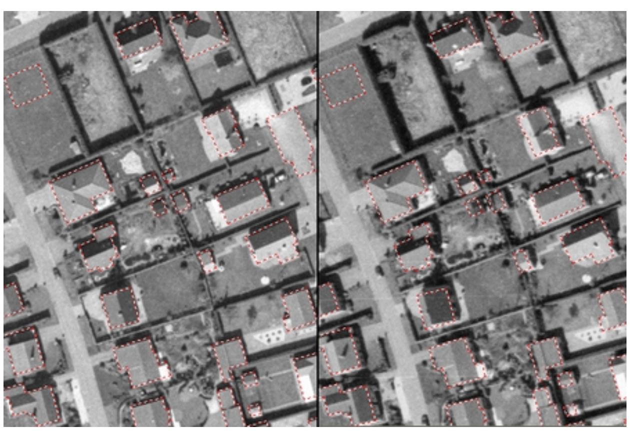

Figure 2. Part of a stereo couple of aerial images with superimposed database.

The change in point of view between images of a stereo couple induces a displacement of objects, called parallax or disparity, dependent on the 'object to sensor' distance. In many situations, stereo images are delivered in epipolar geometry meaning that parallax is horizontal, which simplifies object pairing in images to a one-dimensional search.

Disparity estimation is based on the correspondence of objects of the left and right images. Two main classes of algorithms exist: radiometric correspondence based on grey level correlation of any pixel and feature based correspondence that compares image primitives of interest like corners, contours, regions, etc.

Radiometric based algorithms generally lead to dense and robust disparity estimation but at the expense of many computations as sufficiently large windows have to be correlated for discrimination and all possible disparities have to be checked. On the contrary, feature pairing only considers matching of extracted primitives, which reduces the search space but which a priori leads to a sparse disparity map.

We have developed a mixed approach, combining the advantages of radiometric and feature pairing. Starting from the idea that local grey variations offer the cues for pairing, pixels with sufficient gradient magnitude are selected as primitive of interest. Primitives of the left and right images are matched thanks to local grey correlation leading to a confidence score for pairing. As edge pixels are accurately localized, grey correlation is limited to a small set of disparities.

3.2. Edge points detection

In order to improve the coherence between neighbouring pixels, images are first low-pass filtered. In smooth areas, a uniform 3x3 smoothing filter is applied. Everywhere else, smoothing (3x1) in the direction perpendicular to the gradient reduces noise while keeping edge strength.

After smoothing, the gradient magnitude and orientation is estimated at each pixel from vertical and horizontal differences. Edges are thinned by a non-maximal removal in the direction of the gradient, preserving edge connectivity.

The length and linearity of connected edge pixels can be additional constraints to focus at this stage on man-made structures.

3.3. Disparity estimation

Due to the availability of stereo images in epipolar geometry, only horizontal search for correspondence is considered here. A candidate edge pixel of the left image is matched against all candidate edges of the right image verifying the a priori disparity range and gradient orientation similarity (+/- 20°).

For each plausible pair of edges, the local grey level profiles are compared thanks to the standard deviation of the grey level difference for only 3 disparities (as detected edges have precise localization). The pair with minimal standard deviation is chosen to derive the disparity to be assigned to the left edge pixel. The quality of the match is used as confidence for the measured disparity.

In the case of horizontal edge segments, the disparity per pixel is ambiguous as no significant horizontal grey level variation exists. The whole horizontal segment is matched against a corresponding one in the other image and a disparity value is issued only if the segments lengths are comparable.

So far, pixels are processed individually, except for the special case of horizontal segments. A final pass is provided to harmonize the individual disparity values by low-pass filtering them along connected pixels to improve the global coherence and possibly increase the accuracy.

We prefer this pixel oriented processing instead of contour matching because image contours are not very stable with viewpoint due to occlusions, appearances and geometrical changes. Mention that disparities along a connected segment vary in the case of an object with varying elevation, but values are normally progressive along the contour. Low-pass post-processing reinforces this continuity.

Figure 3 shows the results of disparity estimation for a small part of the aerial image couple. Results are better interpreted when using false colours. Blue and violet (mid grey) contours correspond to the ground level. Elevated contours, likely to refer to building structures, are in green or yellow (bright). Database polygon outlines are drawn dotted (white/red). Notice the large proportion of ground pixels except in building areas where most edges correspond to high elements.

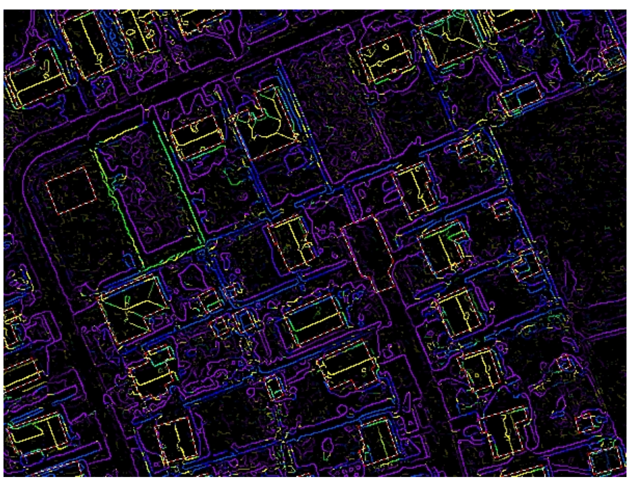

Figure 3: Disparity at edge pixels: low values in blue or violet (mid grey), high in green or yellow (bright). Database polygon outlines are dotted.

3.4. Disparity vs. elevation

Once disparities are estimated, corresponding elevation values (Z coordinate) can be derived from the knowledge of the sensor position of each image after a calibration phase or can be approximated thanks to known values at reference points.

However, for this work about building verification, disparity differences are sufficiently precise to offer a discriminative criterion so that real elevations are not computed.

4. Building verification

Disparity values extracted as explained in section 3 are exploited to derive the Digital Terrain Model (4.1) and to verify buildings (4.2, 4.3) for database update (4.4).

4.1. Digital Terrain Model

According to section 3.3, a sparse disparity map with values at pixels with sufficient gradient magnitude is obtained from the stereo pair of images. The estimated disparities related to ground objects are used for DTM estimation. As ground objects are numerous, with slowly varying elevation, and corresponding to smaller disparity values, we propose to derive the DTM from the analysis of disparity distribution thanks to local histograms.

DTM values are evaluated at pixels of a square lattice with 8-pixel spacing in each direction. For each pixel (x,y), disparity values of a centred 192x192 pixel square are accumulated in the bins of a histogram. The size of the square (about 60x60m), larger than most buildings, is chosen to ensure sufficient ground point representatives. We retain the percentile 20 as the approximated local ground disparity. This value, to be the DTM estimation at (x,y), is refined by picking the more populated bin distant of maximum 1 bin from the approximated value.

Considering the computational load of large square areas for statistics collection, the technique of ‘sliding histogram’ was implemented line by line. The histogram population is first initialized at the left of each image line. For each pixel of the square lattice, the local histogram is updated from the last histogram at the left, adding values of pixels which joined the area on the right and deleting those which left it at the left.

Since rectangular areas are large and largely overlap, the DTM is dense and continuous so that low-pass filtering and interpolation are not needed. The result on a pair of aerial images is presented in Figure 4. The range of disparity spans 10 pixels (about 10m) expressing that the terrain is flat in this region.

Mention that the DTM has been estimated so far without the knowledge of building presence. The use of the database could be a future refinement.

The DTM is an interesting by-product in itself. It helps analyzing flood conditions or assisting environmental planning. In our context, it is used as a basis to define the height of building pixels. Height is a stronger cue for building assessment than absolute elevation.

Figure 4: Extracted Digital Terrain Model

4.2. Scoring database polygons

In order to show the potential of 3D cues for building verification, the database polygons are assigned a score related to the presence of specific disparity values. The values of the DTM obtained in 4.1 are deducted from the disparity map (subsection 3.3) to normalize the disparities at edge pixels so that they indicate a local height more appropriate for building detection than an absolute disparity value.

The polygons of the database are first filtered to reject the roads, the very small structures (< 20 m^2) and to skip those not fully inside the image. The polygons, representing the footprint of buildings, are enlarged by 5 pixels (1.5 m) in each direction to account for image registration imprecision and parallax displacement of elevated structures.

For each polygon, the bins of a disparity histogram are incremented by the disparity confidence of edges contained in the enlarged polygon. The lowest three histogram bins are expected to refer to ground segments. Bins of higher values are summed up to represent the building score of the polygon. The building score is normalized by the polygon area to represent a density of elevated pixel edges. Scores are finally rescaled in the [0..100] range from the statistics of encountered polygon score values.

4.3. Visual inspection

The main objective of this work is to detect inconsistencies between the topographic (2-D) database of buildings and recent images in order to update the database. As presented in the introduction, delivering potential changes with a confidence measure to the geographer is a desired semi-automatic solution to let him/her focus on ambiguous cases.

Once polygon scores are obtained, the system displays a building verification map representing the confidence and the acceptance/rejection of a building hypothesis (Figure 5). Each building is drawn with intensity proportional to its score and in a colour depending on a decision threshold: green (red) if its score is higher (lower) than the threshold. The user typically sets the threshold to highlight a few red buildings as the number of database inconsistencies is usually small. An alternative is to present the histogram of polygon scores for threshold selection. The user then looks for visual confirmation of building in the images and possibly refines the threshold to look for other mismatches.

With this semi-automatic procedure, what will happen next depends on the application. If the user aims at fixing database errors, he will check the image for visual confirmation of red buildings and will adapt the database accordingly. In the case of damage assessment, he can plot the building verification map to report on detected changes.

4.4. Database update

The building verification map with threshold adaptation helps finding supported and unsupported buildings of the database. For new constructions, not referenced by the database, another procedure has to be considered, like the highlight of elevated areas for visual inspection. Preliminary results are given in subsection 5.2.

As a by-product, the Z coordinate which can be derived from the disparity associated to building elements can be integrated in the building description of the topographic database. IGN is indeed already enlarging its topographic database with elevation so that automatic estimates will be welcome, either to complete the coverage with Z estimates or for updates.

5. Results

5.1. Building verification

The procedure presented in section 4 for building verification has been applied on a stereo couple of images a small part of which is shown in Figure 2.

The full images have 4000x3000 pixels and cover 1 km^2. They contain isolated buildings (about 700 items) in residential areas. Their resolution is high (30 cm per pixel) but they are quite noisy, corresponding to the old generation of image capture at IGN Belgium (scanned photographs, no colour). The terrain is rather flat and we dispose of a layer of digital polygons describing building outlines and roads.

Figure 5. Building verification map. Dark polygons have low scores and should be rejected. The rectangle outlines the area of Figure 3. Blind alleys are marked 'x' and phantoms 'o'.

Figure 5 shows the building verification map of about 15% of the full image. Notice in red (dark) the few polygons not supported by 3D clues. They correspond to two special cases of the database which offered impostor tests. 'Blind alleys', marked x, are ends of the road network, outlined by a polygon. 'Phantoms', marked 'o', stand for buildings present in the database but not in the images because the latter are slightly older than the database. Phantoms also appear in the case of building demolition.

Table 1 gives the number of accepted blind alleys and phantoms and rejected buildings as a function of the score threshold. As the threshold increases, we notice the rapid reduction of the 13 imposters (blind alleys, phantoms) and the plateau before the number of falsely rejected buildings increases. Compared to the method based on linear segments and shadow cues [6], the current method appears more robust against impostors.

Table 1. Number of accepted blind alleys and phantom buildings and rejected genuine buildings versus the score threshold.

Scores	0	5	10	15	20	30	50	100
# Blind Alleys	7	1	0	0	0	0	0	0
# Phantoms	6	3	1	0	0	0	0	0
# Buildings	0	0	0	0	1	9	183	657

5.2. Building detection

Detecting buildings not referenced by the database was attempted with the same procedure as for verification. The density of elevated edges is estimated from the disparity histogram for an arbitrary 10x10m square polygon moved all over the image. In the resulting Figure 6, the threshold was set to the highest value that generated no error in verification. Notice the major overlap of green / light grey (elevated) areas by building polygons. Polygons not overlapping green areas correspond to road or impostors (phantoms, blind alleys). Most green areas uncovered by polygons correspond to wooded zones or hedges. Adding radiometric (colour, NDVI) and geometric (straightness) constraints will help handling elevated vegetation.

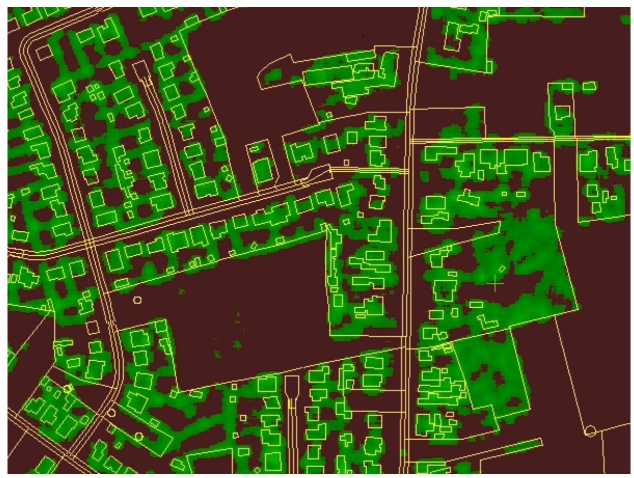

Figure 6. Potential areas for buildings in green (light grey).

5.3. Discussion

The proposed approach for building verification, simple in concept and implementation, was successfully applied to a stereo pair of images with 12 Mpixel (4000x3000) which is representative of a Belgian rural area. The algorithms are very fast: about 30s for disparity estimation, and 1s to derive the DTM and to prepare the building verification or detection map on a 2.33 GHz PC.

Compared to the previous developments based on straight segments and shadow cues, 3D is more discriminative, what appears a clear advantage for new building detection. Shadow cues and geometric constraints may be combined to add more discriminative power to the 3D approach.

Encountered limitations mainly concern what disturbs disparity estimation (occlusion, inversion of contrast or large geometrical changes between the left and right images). These conditions are more likely to be met with large Base/Height ratio of the stereo capture which improves disparity accuracy but at the expense of larger image discrepancies (in visibility, geometry and/or image intensity). Urban areas are usually more problematic than rural zones due to the higher density of elevated structures and occlusion. To a lesser extent, trees introduce elevated elements which may confuse building detection. For extended wooded areas, the derived DTM may be incorrect. Complications due to vegetation can be reduced thanks to multi-spectral data (colour, NDVI) and geometrical constraints ("roundness of trees" compared to straightness of buildings).

6. Conclusion

We have presented the verification of buildings described as polygons in a database thanks to 3D cues obtained from stereo disparity measures. Stereo correspondence has been achieved by a hybrid correlation and feature based approach which derives disparities from the best matching of edge pixels in the sense of local grey correlation.

Compared to previous verification results obtained on the same data with a combination of linear segments and shadow cues, the verification from 3D evidence is quicker, simpler and more discriminative. This last advantage enables the automatic detection of elevated areas in order to find new constructions, necessary to complete database updates.

The presented method also derives a DTM from the disparity map which can be a useful by-product.

Limitations of the approach are bound to the stereo correspondence problem. A low base/height ratio is to be preferred to reduce image discrepancies although a minimum is required for precision in disparity estimation. Urban areas generally contain more difficult cases due to occlusion. Trees introduce false positives and should be filtered by multi-spectral and/or geometrical constraints.

In the future, we intend to improve building detection by reducing ambiguity thanks to the integration of more views and additional information like shadow cues and multi-spectral data. Moreover, roof models will be adapted to gain further assessment of building presence and derive more precise z coordinates for the topographic database.

7. Acknowledgement

This study has been funded by the Belgian Ministry of Defence. Image and vector data have been gracefully obtained from IGN (Institut Géographique National) of Belgium.

8. References

[1] C. Baillard, "Analyse d'images aériennes stéréoscopiques pour la restitution 3D des milieux urbains. Détection et caractérisation du sursol" (In french), Thèse Signal et Images, TSI, Télécom Paris [ENST], 1997.

[2] J. Shufelt, "Performance Evaluation and Analysis of Monocular Building Extraction From Aerial Imagery", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, NO. 4, Apr 1999, pp 311-326.

[3] E. Baltsavias, "Object extraction and revision by Image Analysis Using Existing geospatial Data: State-of-the-art and Steps Towards Operational Systems", *Int. Archives of Photogrammetry Remote Sensing and Spatial information Sciences*, Vol 34, Part 2, 2002, pp 13-22.

[4] X. Jin and C. Davis, "Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual and Spectral Information", *in EURASIP Journal on Applied Signal Processing, 2005*:14, pp 2196-2206.

[5] C. Beumier, "Building verification from geometrical features", *27th Earsel Symposium - Geoinformation in Europe*, Bolzano Italy, 4-7 June 2007.

[6] C. Beumier, "Building verification from geometrical and photometric cues", *Applications of Digital Image Processing XXX*, Proceedings of SPIE Vol. 6696, San Diego, California, 26-30 Aug 2007.

[7] M. Brown, D. Burschka and G. Hager, *"Advances in Computational Stereo"*, IEEE PAMI, Vol. 25, NO 8, Aug 2003, pp 993-1008.