

EXTRACTING 3D FREE-FORM SURFACE BOUNDARIES OF MAN-MADE OBJECTS FROM MULTIPLE CALIBRATED IMAGES: A ROBUST, ACCURATE AND HIGH RESOLVING POWER EDGEL MATCHING AND CHAINING APPROACH

Franck Jung, Nicolas Paparoditis

Institut Géographique National, MATIS, 4 Avenue Pasteur 94165 StMandé Cedex, Franck.Jung@ign.fr, Nicolas.Paparoditis@ign.fr

Commission III, WG III/ 3

KEY WORDS: Photogrammetry, 3D Reconstruction, Curved 3D Lines, 3D Edgels, Multiple Images, Feature Matching, Robust LMS bundle adjustment.

ABSTRACT

This paper describes a purely geometrical approach for matching and chaining edgels, i.e. contours points with their tangent direction, to reconstruct in a general way 3D rectilinear or curved surface boundaries from multiple calibrated images. Results show that the 3D reconstructed edgels are highly reliable and accurate, and correctly reproduce the 3D limits of very small image structures thus demonstrating the high resolving power of this technique and enabling a finer description and understanding of urban scenes. 3D Edgels are a key feature for reconstructing man-made objects. Practical examples of edgels reconstruction are given on various type of 3D objects (roundabouts, zebra crossings, cars, classical and odd shape buildings, dormer windows) in surveys at a ground pixel size of 25 cm. Practical examples of edgels integration within a surface reconstruction or 3D pattern recognition and object extraction processes are also provided. Indeed 3D edgels can be used as characteristic and constraining lines for the surface topography reconstruction and used to build the limits of cartographic objects thus ensuring a natural geometrical and topological coherence between Digital Surface Models and Digital Object Models.

1 INTRODUCTION

Feature-based matching techniques from stereopairs to reconstruct 3D linear features corresponding to surface radiometric or geometric changes have been a very largely investigated subject within the computer vision and the photogrammetric communities in the three last decades (Medioni and Nevatia, 1985), (Serra and Berthod, 1995), (Baillard and Maitre, 1999). Indeed, surfaces reconstructed by purely area-based stereo-matching techniques do not render correctly the localisation and the morphology of surface changes and 3D objects (e.g. building discontinuities or slope breaks). Surface models and 3D objects can only be coherent if the 3D object limits or their corresponding image features are injected in some way in the surface reconstruction scheme.

Matching features within stereopairs is a geometrical incorrectly posed problem. Indeed, contour and segment features themselves are not stable i.e. they are not characterised in the same way in both images (e.g. due to occlusions). Moreover, any couple of image contours (resp. segments) could generate a 3D point (resp. segment) by intersection of their 3D corresponding lines (resp. planes) in object space. A way of reducing geometrical ambiguities is to search for invariant-based matching attributes (given an analysis area around the feature). These parameters are usually not very stable within urban areas due to non lambertian reflectance of surfaces and inversion contrast. And even when they are stable, considering the fact that these attributes are calculated on a neighbourhood, problems can occur when objects are small and features are close to one another and thus inevitably reduces the resolving power of the matching algorithm.

The acquisition of multiple views solves many of these problems. Actually, provided a sufficient overlapping exists, there will most of the time be more than two images among the set of images in which a relevant feature of the landscape is characterised in a similar way. Having a high redundancy of geometrically different feature observations will increase the reliability, the accuracy

and especially the resolving power of the reconstruction algorithm due to the fact the pure intrinsic geometrical information within the feature will be sufficient to match it.

Recent papers have shown that multiple views simplify and improve considerably the reconstruction of 3D segments (Noronha and Nevatia, 1997), (Heuel and Förstner, 2001), (Taillandier and Deriche, 2002). Although results shown in these papers are impressive they are not adapted for reconstructing small structures (in general smaller than 10 pixels in image space). In addition, even though 3D segments can describe efficiently a very large number of man-made structures many other man-made objects - which should appear in 3D city models or traditional high scale cartographic databases- have curved boundary limits: roundabouts, pavements, and more and more buildings with the new standards and fashions of architecture which are in general very complicated and time consuming to stereoplot by a manual operator in comparison with classical buildings.

We will thus present an edgel matching technique suited to reconstruct both straight and curved linear object boundaries. An edgel is an elementary point of contour (maxima of intensity gradients) with its related tangent direction.

2 EXTRACTING 2D EDGELS FROM IMAGES

Contours within each image of the set are detected by applying the Canny-Deriche edge detector and by extracting the local maxima in the gradient images. The local maxima are then located with sub-pixel accuracy by determining the peak value of the parabola that fits locally the gradient samples in the gradient direction. The sub-pixel contour points are thereafter chained and the 2D tangent direction for each contour point is calculated by fitting a straight line on the few closest chain samples on each side of it. This estimation can be sensitive to noise, change of direction, etc. thus, in order to get a robust estimation of the associated direction, we use an M-estimator (Rousseeuw and Leroy,

1987). We thus obtain chains of edgels: points with tangent direction. We now want to match each 2D edgel with the $N - 1$ other images.

3 IDENTIFYING EDGEL-MATCHING HYPOTHESES IN SLAVE IMAGES THROUGH OBJECT SPACE

First we choose a master image. For each edgel point \tilde{p} in this master image, we determine the corresponding bundle in object space (Figure 1). Having previously corrected the chains of contours of the image distortion, the epipolar curves i.e. the projection of the master bundle in all slave images, are straight lines. Thus the set of homologous 2D edgels hypotheses englobes all the 2D edgel segments intersecting the epipolar lines. A quadtree search is performed to reduce efficiently the combinatorial of the hypotheses determination. The 3D object space edgel point P corresponding to \tilde{p} is determined by an adjustment of all the bundles corresponding to all the set of hypotheses as explained in the following paragraph.

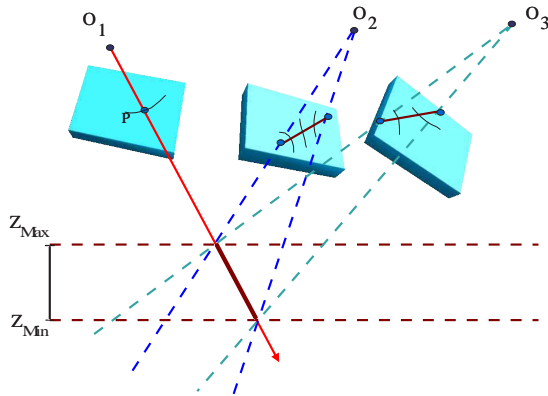


Figure 1: The master point \tilde{p} is reprojected in every slave image using an epipolar constraint and a bound on the altitude $z \in [Z_{min}, Z_{max}]$. Image i contains a set S_i of possible matches with \tilde{p} .

4 ROBUST 3D EDGEL ESTIMATION (POINT LOCATION WITH TANGENT DIRECTION)

The first step of our adjustment aims at finding a point matching algorithm related to the master point \tilde{p} . During this step, we use a RANSAC technique in order to select the best cluster of bundles close to the master bundle. The matching is described by the master point \tilde{p} , the set of associated slave points and a 3D location P . Each slave image contains at most one slave point. It is important to notice that we do not systematically find a matching for each slave image. First, we reject a matching candidate as soon as the distance between the bundle and P is larger than a given threshold. This threshold is only a function of the various $\frac{B}{H}$ parameters. Second, the whole matching is rejected as soon as it contains a poor number of points. The lower bound on the number of points depends on the robustness required by the application (in our application, we kept matchings containing at least four points). More details can be found in (Jung et al., 2002). Practically, we minimize the third residual. Bundles having a distance larger than the altimetric resolution (in our case $5m$) are automatically rejected. Our evaluation procedure (see next section) proved that this threshold was meaningful.

After the RANSAC location adjustment, we suppose that the remaining measures follow a gaussian error. Hence, it is now possible to refine the 3D location by a least square technique. After this step, the importance of the master bundle is reduced and we get a precise 3D location P associated to our matching.

To estimate a 3D tangent direction associated to P , we tailored a direction adjustment. The aim is to find a 3D direction as close as possible to every plane formed by a matching point, the associated 2D direction and the projection center of the associated camera. Intersecting all these planes two by two gives us a set of hypothesis for the 3D direction (Figure 2). We first use a RANSAC technique to get a robust estimation of the direction. We reject all measures that are too far from our solution. The threshold should only depend on the accuracy of the 2D edgels and on the geometry of the different point locations. A threshold of 10° was fixed in our applications. The accuracy evaluation performed were able to confirm us that this threshold was meaningful.

Then a least square technique is used considering the directions close to the RANSAC solution.

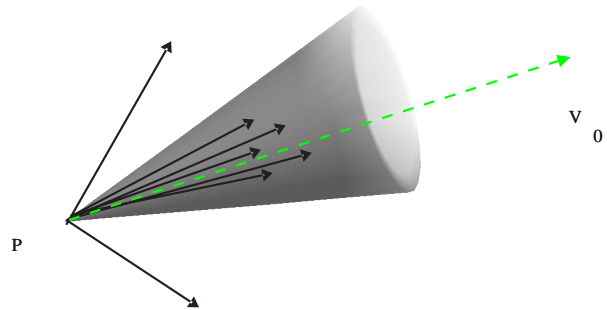


Figure 2: Direction Adjustment. The RANSAC step aims at eliminating outliers. The second step consists of a L_2 estimation in order to refine the direction.

We have separated the location adjustment from the directional adjustment. Because the location residuals are metric, while the directional residuals are angular, it is hard to compare these quantities. Besides, considering simultaneously position and direction estimation stays challenging due to the large number of potential matches to compare with a robust technique.

As soon as all master image edge locations are processed, we repeat the very same algorithm choosing a master image from the remaining set of images. This enables the detection of 3D edgels not significantly detected in the previously processed master images. Moreover, seeing that the object sampling changes in each view, repeating the process master images and accumulating the 3D edgels allows a super resolution object characterization.

Many applications take advantage of the use of robust external surface data : robust DSM computed by correlation techniques (Paparoditis et al., 2001), laser scanning surface, etc. Using a photogrammetric or a LASER DSM in our process reduces drastically the matching search-space combinatorial and consequently the matching ambiguities and erroneous edgel features.

Having removed most of the erroneous matches, it is possible to associate a geometric uncertainty to each reconstructed feature using Förstner and Heuel's technique of uncertainty propagation (Heuel and Förstner, 2001). This enables a thresholdless grouping and chaining of our low-level 3D features to reconstruct higher-level features and objects.

5 3D EDGEL GROUPING

Our algorithm provides a 3D description of an entire scene using edgels. Due to the extraction process, it is easy and natural to group these edgels using the 2D edgels topology. This algorithm proceeds in two steps: first, a refinement of the edgel position is performed, second a polygonal approximation is performed.

5.1 Edgel refinement

Using as many master images as available, we will finally get a redundant description of our scene. All these descriptions are not independent but they complete one each other. Thus, we use this redundancy in order to refine our description of the scene. To each 3D edgel we associate the master 2D edgel p_i used for the former described adjustment technique. We associate a new 3D edgel using a RANSAC technique in order to compute a new 3D line thanks to all the 3D neighboring edgels (Figure 3). More precisely, we select the 3D line D (which supports an edgel) which minimize a given residual of the distance of the set of edgels to D . In order to compute D we use all 3D edgel associated to every master image. If we get too few edgels inside our bounding box, we stop the refinement process for that given edgel.

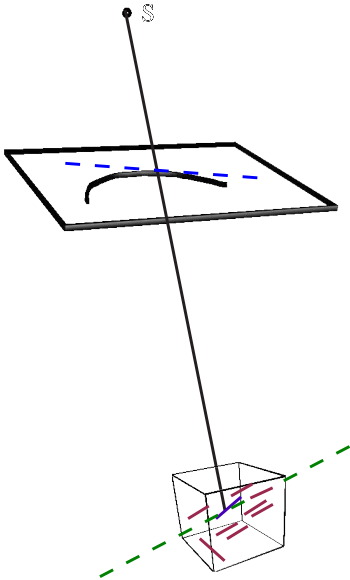


Figure 3: We estimate a 3D line candidate minimizing the distance to the neighbors edgels using a RANSAC algorithm. The dashed line represents the solution chosen by our algorithm.

We now associate a new 3D edgel P to p_i (Figure 4). It is important to notice that P does not necessarily belong to the bundle ($S p_i$).

5.2 3D Polygonal approximation

Using the chain structure of the 2D edgels contours in each image, we will be able to compute a 3D polygonal approximation with a classical Ramer-Douglas-Peucker algorithm. We are using a 2D threshold $0.5m$ and an altimetric threshold of $0.5m$ in order to readjust the location without changing the edgel orientation. Hence, we get a set of 3D edgels chains. We recursively cut in two pieces every chain as soon as the distance of two consecutive 3D edgels inside the chain is bigger than twice the image resolution. We must notice that this threshold is not critical for our applications (Figure 5).

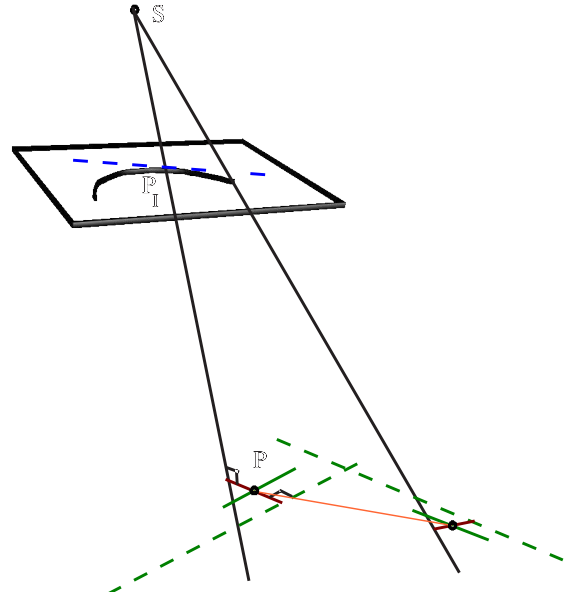


Figure 4: To each 2D edgel, we associate a 3D refined edgel. We take advantage of the chains of 2D edgel in order to provide a chaining of 3D edgels.

The described algorithm is simple and easy to implement. Nevertheless, there remains some major drawbacks associated to this technique. First, we do not necessarily get a unique description of the contours. A matching of the different chains describing the same 3D contour has still to be performed. This task can be very tricky but is not absolutely necessary for many scene interpretation tasks.

6 RESULTS

First, we must note that some general accuracy and detection evaluation experiments can be found in (Jung et al., 2002). In this section we illustrate the resolving power of our 3D edgels, first on synthetic data (Figure 6) and then on real data (Figure 7, 8). We can notice that very small objects (five to twenty image pixels) are well described. This makes us believe that 3D edgels are well suited for 3D object recognition tasks.

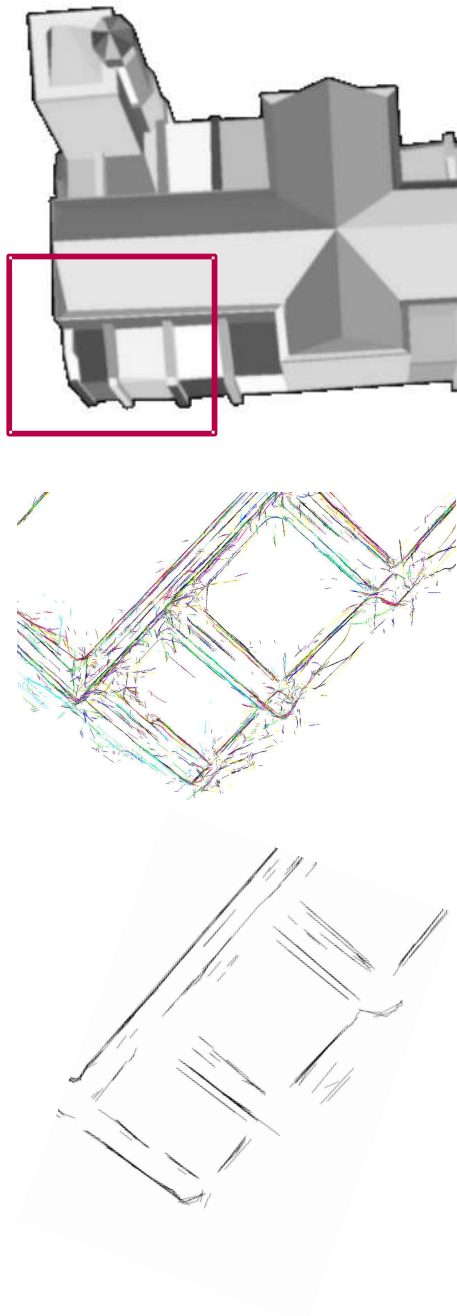


Figure 5: Example of a 3D polygon approximation based on a set of 3D edgels. The upper figure represents the results of our 3D extraction algorithm. The bottom figure provides the result of our chaining algorithm.

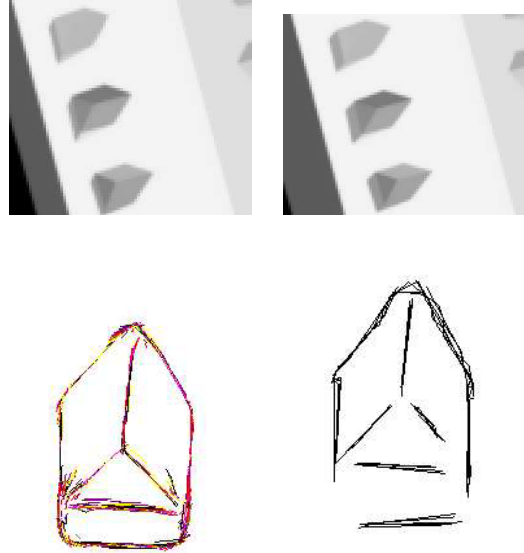


Figure 6: Example of 3D edgels of a dormer window reconstruction using edgels. The dormer window is composed by edges of size 10 to 15 pixels. The two upper images represent two of the 9 synthetic images used to extract the edgels.

Second, curved objects are well described (Figure 9). As a matter of fact, modern architecture provides numerous examples of very complex buildings which can not be modeled by straight lines. Urban areas will contain more and more of these curved shaped buildings. Automatic 3D building reconstruction schemes will have to integrate non-straight elements in order to reach a higher automatic performance.

Third, we give an example of application using 3D edgels in order to estimate an accurate 3D surface model (Figure 10). We performed a 2D constrained Delaunay tessellation (using the x and y coordinates of our 3D edgels) in order to get a description of a 3D scene with triangles. We used some prior information concerning the DTM in order to get a smooth ground surface. These information can be provided by a ground/above ground classification like in (Baillard and Maitre, 1999).

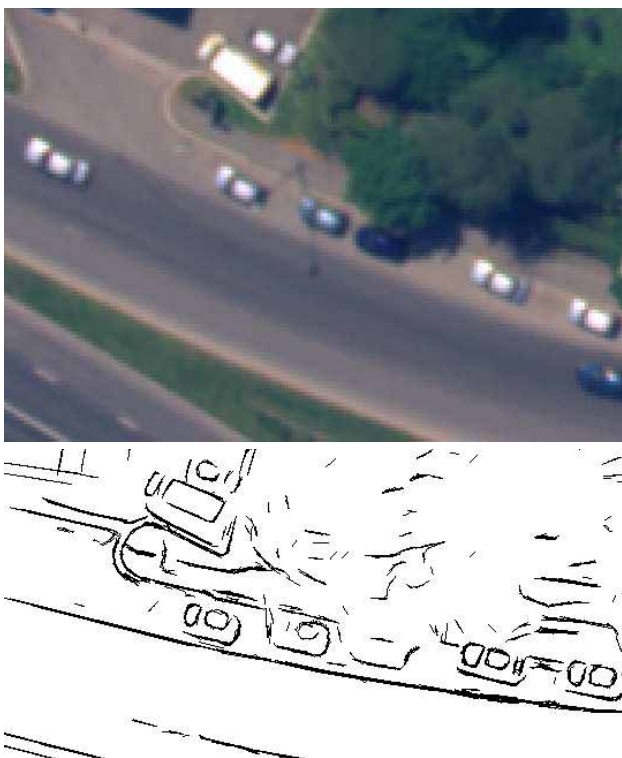


Figure 7: Example of car description using 3D edgels. We can notice that the main structure of all static cars are well described. Pixel size : 25cm. The image length of a car is approximately 10 to 15 pixels.



Figure 9: Example of 3D edgel extraction and chaining. Curved objects are well described.



Figure 8: Example of non-linear object extraction with edgels. The shape of small objects (cars, zebra crossing, etc.) are well reconstructed using edgels. The square (upper left and bottom left images) is used as a 1m x 1m metric reference.

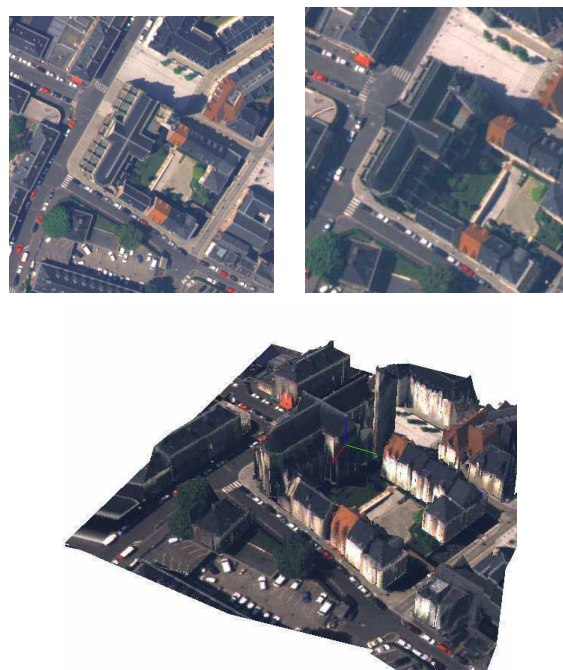


Figure 10: Example of a 3D surface reconstruction using edgels. We computed a 2D Delaunay tessellation combined to prior information of the DTM.

7 FUTURE WORK AND CONCLUSION

We have presented a technique to extract in 3D the smallest significant features within the images. Indeed the outputs of such a reconstruction is extremely rich but the information is rather complete and just needs to be sorted. The features reconstructed are robust, accurate (Jung et al., 2002) and the features come with their geometric uncertainty estimation. The quality of the reconstruction is also put onto the account of the digital frame camera (Thom and Souchon, 2001) used for the surveys which has a very high Signal to Noise Ratio thus guaranteeing accurate tangent directions.

As presented in the paper these features can have many applications in an cartographic context. These features with their geometric uncertainty can also be combined in order to build higher level information (Heuel and Förstner, 2001): e.g. to reconstruct planar faces of buildings by grouping 3D edgels with or without the help of other features e.g. 3D points provided by a multi-view template matcher (Paparoditis et al., 2001). 3D Edgels can also be injected as such as an input into 3D pattern recognition experts: e.g. car experts, zebra-crossing experts, etc. all to draw evidence of roads. Indeed, many pattern recognition algorithms (neural networks, decision trees, etc.) just need a "recoded" input image in order to take a decision. All these experts will participate in a holistic approach to a better general understanding and splitting of the scene in elementary mono-thematic focusing areas (Paparoditis et al., 2001): e.g. ground, roads, pavements, trees, vegetation, buildings, upon which object reconstruction experts will be applied.

REFERENCES

- Baillard, C. and Maitre, H., 1999. 3d reconstruction of urban scenes from aerial stereo imagery: a focusing strategy. *Computer Vision and Image Understanding* 13, pp. 244–258.
- Heuel, S. and Förstner, W., 2001. Matching, reconstructing and grouping 3d lines from multiple views using uncertain projective geometry. In: *CVPR '01*, IEEE.
- Jung, F., Tollu, V. and Paparoditis, N., 2002. Extracting 3d edgel hypotheses from multiple calibrated images: a step towards the reconstruction of curved and straight object boundary lines. In: *Proceedings of PCV02 (Vol B)*, Graz, pp. 100–104.
- Medioni, G. and Nevatia, R., 1985. Segment-based stereo matching. *CVGIP*.
- Noronha, S. and Nevatia, R., 1997. Detection and description of buildings from multiple aerial images. *Proceedings, CVPR*.
- Paparoditis, N., Maillat, G., Taillandier, F., Jibrini, H., Jung, F., Guigues, L. and Boldo, D., 2001. Multi-image 3D feature and DSM extraction for change detection and building reconstruction. In: B. B. Verlag (ed.), *Automatic Extraction of Man-Made objects from aerial and space images*, Ascona.
- Rousseeuw, P. and Leroy, A., 1987. *Robust Regression and Outlier Detection*. John Wiley & Sons, New-york.
- Serra, B. and Berthod, M., 1995. Optimal subpixel matching of contour chains and segments. In: *ICCV*, pp. 402–407.
- Taillandier, F. and Deriche, R., 2002. Reconstruction of 3d linear primitives from multiple views for urban areas modelisation. In: *Proceedings PCV02, Vol B*, pp. 267–272.
- Thom, C. and Souchon, J. P., 2001. Multi-head digital camera systems. *GIM International* 15(5), pp. 34–37.