3D BUILDING RECONSTRUCTION WITH PARAMETRIC ROOF SUPERSTRUCTURES

M. Brédif [12], *D. Boldo* [1], *M. Pierrot-Deseilligny* [1]

H. Maître [2]

[1] French Mapping Agency (IGN)
2/4, avenue Pasteur
94165 Saint-Mandé, Cedex, France

[2] GET - Telecom Paris
46, rue Barrault
75013 Paris, France

ABSTRACT

This paper describes an automatic method for the detection and detailed reconstruction of 3D building models that include roof superstructures such as dormer windows or chimneys from a very high resolution Digital Elevation Model. Buildings are reconstructed as a set of roof planes with a set of parametric shapes that model roof superstructures. The proposed model-based approach minimizes a Minimum Description Length energy.

Index Terms— 3D Modeling, Building superstructure, Parametric approach, Minimum Description Length, Digital Elevation Model.

1. INTRODUCTION

3D city models that include buildings with roof superstructures (detailed volumes present on the roofs) achieve a level of detail that finds applications in virtual or augmented reality and urban planning. This level of precision impacts directly on the degree of immersion of the visualization and on the user confidence in the model quality. Current 3D city modeling algorithms consider the roof superstructures as outliers or noise that will inevitably bias their results. Carefully taking into account their presence and characteristics will benefit in more robust reconstructions. Some applications, like natural hazard management or electromagnetic wave propagation may not require such a level of detail. Nevertheless, they would benefit from the increased robustness and accuracy of taking into account those superstructures, even if they are discarded in the end to keep only the main roof planes.

Generic methods have been introduced to produce free-form polyhedral building models from aerial imagery[1][2] but satellite data requires model-based approaches[3][4] due to their inferior resolution. The context of this paper is new because the 3D-reconstruction of buildings with superstructures using very high resolution aerial images presents two different subproblems. The data of the main roof planes is good enough to use a generic approach, but superstructures are much smaller features and their data quality can be compared to the one of entire buildings in satellite images.
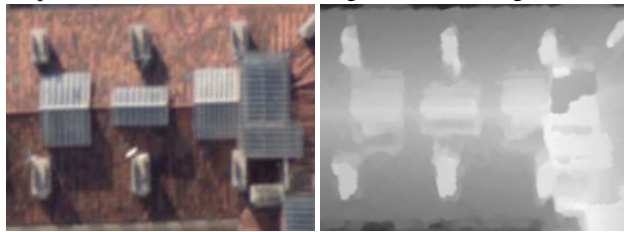
Fig. 1. Image and DEM crops featuring chimneys and glass roofs.

The proposed approach requires an initial building reconstruction algorithm such as [1][2], a collection of superstructure types (figure 2), and a Digital Elevation Model (DEM). The DEM (figure 1-right) is a height-map generated by the correlation [5] from multiple very high resolution (10cm) aerial images (figure 1-left). We aim at refining the geometry of the initial building, while adding a set of parametric superstructure objects to the roof planes.

First, the search space will be defined in Section 2 as the set of buildings modeled by a polyhedral roof base structure with a set of parametric superstructure objects. Then, the objective function is described in Section 3 as a Minimum Description Length (MDL) energy[6]. Finally, Section 4 presents the proposed algorithm along with results and analysis in Section 5.

2. HYBRID POLYHEDRAL/PARAMETRIC MODEL

The proposed hybrid model, similar to the architectural model in [7], allows the generic modeling of the main planes of the building with 3D-polygons, while parametric objects are used to model the smaller features that are the roof superstructures, where the DEM presents reconstruction artefacts that are incompatible with a generic polyhedral reconstruction.

The support $supp(m)$ of a 3D-model m is defined as the 2D-domain formed by its orthogonal projection on the horizontal plane, as illustrated in figure 3. We will assume that the roof planes do not overlap, that the superstructures do not overlap and that each superstructure is contained in a single roof plane, using the following definitions on 3D-models m_1 and m_2:
- m_1 is contained in m_2 if $supp(m_1) \subset supp(m_2)$
- m_1 and m_2 overlap if the area $\mathcal{A}(supp(m_1) \cap supp(m_2)) > 0$

A building model is a collection of roof planes defined by a 3D-polygon and a set of contained volumetric primitives corresponding to various roof superstructures such as chimneys or dormer windows. A building model is defined by $\mathcal{B} = (\mathcal{R}, \mathcal{S})$ where the roof plane set $\mathcal{R} = \{\mathcal{R}_i\}_{1 \leq i \leq n}$ defines the geometry of the 3D-polygon that models each of the n roof planes, and the superstructure set $\mathcal{S} = \{\mathcal{S}_{ij}\}_{1 \leq i \leq n}^{1 \leq j \leq n_i}$ defines, for each roof plane $\mathcal{R}_i$, its n_i contained superstructures $\mathcal{S}_i = \{\mathcal{S}_{ij}\}_{1 \leq j \leq n_i}$.

A superstructure is the parametric object $\mathcal{S}_{ij} = (t_{ij}, \theta_{ij}, \phi_{ij})$ where :
- t_{ij} is one of the superstructure types that are illustrated in figure 2.
- ϕ_{ij} is the set of parameters that are specific to the superstructure type, documented as the h values in figure 2.
- $\theta_{ij} = (x_0, y_0, x_1, y_1)$ is a parameterization of the minimum gutter-aligned rectangle that contains $supp(\mathcal{S}_{ij})$, where the gutter of roof is defined as the segment of the 3D-polygon of $\mathcal{R}_i$ that contains the lowest point and has the smallest slope. Note that the support and the rectangle defined by θ_{ij} are identical except for the dormer window.
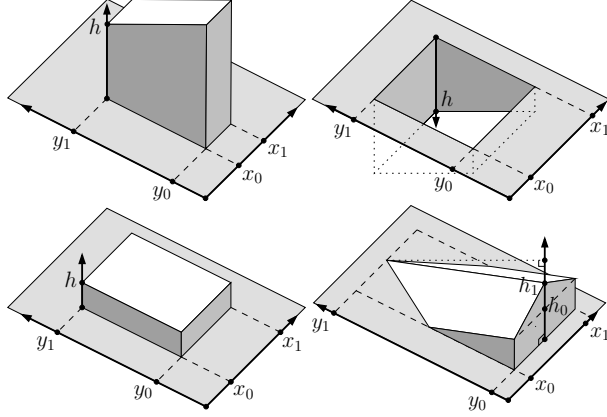
Fig. 2. The superstructure types : a chimney (4+1 parameters), a roof terrace (4+1), a glass roof (4+1) and a dormer window (4+2).
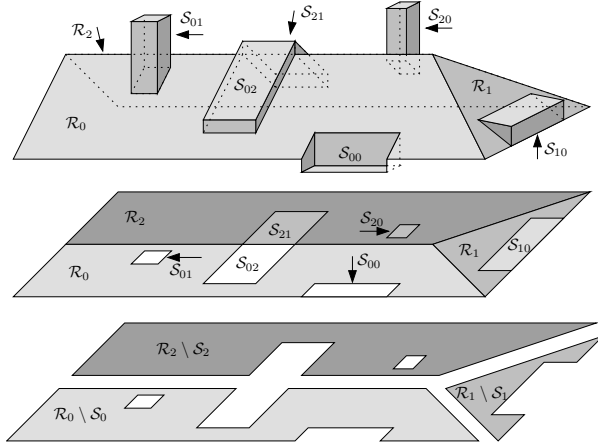
Fig. 3. A building model (top), its supports (middle) and the supports of its effective roof planes (bottom).

These 4 parameters θ_{ij}, common to all superstructure types, define the position and the horizontal extents of the superstructure.

A superstructure type t_{ij} defines **hard constraints** on the possible parameter values (θ_{ij}, ϕ_{ij}) of its instances. This introduces **loose priors** that guarantee that, for instance, dormer windows are convex ($h_0 \leq h1$) and leave enough space for a window on their front facet. The only tight hard constraint is the alignment of every superstructure with the gutter of its containing roof plane. This constraint does not seem to be restrictive and introduces robustness in the orientation of the superstructures. Finally, a **description length constant**[6] L_t encodes the complexity of each model type t.

The effective roof plane $\mathcal{R}_i \setminus \mathcal{S}_i$ is the subset of the roof plane $\mathcal{R}_i$, whose support $supp(\mathcal{R}_i \setminus \mathcal{S}_i) = supp(\mathcal{R}_i) \setminus \bigcup_{s \in \mathcal{S}_i} supp(s)$ is cut out by the supports of its superstructures $\mathcal{S}_i$ in the final building model $\mathcal{B}$, as illustrated in figure 3. Similarly, the **effective roof** $\mathcal{R} \setminus \mathcal{S}$ is defined as $\{\mathcal{R}_i \setminus \mathcal{S}_i\}_{1 \leq i \leq n}$. Note that, given our assumptions, $\mathcal{P}_i = \{supp(\mathcal{R}_i \setminus \mathcal{S}_i)\} \bigcup_{s \in \mathcal{S}_i} \{supp(s)\}$ forms a **partition** of $supp(\mathcal{R}_i)$ and that $\mathcal{P} = \bigcup_i \mathcal{P}_i$ is a partition of the support $supp(\mathcal{B})$ of the entire building. This partition $\mathcal{P}$ is used to decompose the DEM into regions of homogeneous geometry.

3. ENERGY FORMULATION

The energy $E_\lambda(\mathcal{B})$ of a building model $\mathcal{B}$ uses the MDL formulation:

$$E_\lambda(\mathcal{B}) = d_\mathcal{B}(\mathcal{B}) + \lambda.L(\mathcal{B}) \qquad (1)$$

$$d_{m_1}(m_2) = \int_{supp(m_1)} \|z_{DEM}(x,y) - z_{m_2}(x,y)\| \quad (2)$$

$$\lambda.L(\mathcal{B}) = \lambda.\sum_{ij} L(\mathcal{S}_{ij}|\mathcal{R}_i) \qquad (3)$$

The data term $d_\mathcal{B}(\mathcal{B})$ is evaluated with the Z-error distance defined in equation 2 where $\|.\|$ is a chosen metric, and m_1 and m_2 are 3D-models of superstructures or roof planes. Simpler models are favored using the additive description length term $\lambda.L(\mathcal{B})$ of equation 3, where the description length $L(\mathcal{S}_{ij}|\mathcal{R}_i) = L_{t_{ij}}$. Assuming that $L(\mathcal{R}_i)$ is constant for evey hypothesis, it is not taken into account. The λ factor may be set to zero when the data is of good quality, yielding a **parameter-less error-driven reconstruction**. Nevertheless, setting $\lambda > 0$ becomes useful to prevent over-fit with noisier data.

From the definition of $E_\lambda(\mathcal{B})$ in equation 1, the reformulation in equation 4 separates the Z-error of the effective roof $\mathcal{R} \setminus \mathcal{S}$ from the rest of the energy, whereas equation 5 exhibits the benefit $e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i)$ of adding each superstructure $\mathcal{S}_{ij}$ to the building reconstruction compared to the background hypothesis of the building $\mathcal{B}^0 = (\mathcal{R}, \emptyset)$ that has the roof planes $\mathcal{R}$ of the building $\mathcal{B}$ but no superstructures.

$$
\begin{aligned}
E_\lambda(\mathcal{B}) &= d_{\mathcal{R} \setminus \mathcal{S}}(\mathcal{R}) + d_\mathcal{S}(\mathcal{S}) + \lambda.L(\mathcal{B}) \qquad (4)\\
&= d_{\mathcal{R} \setminus \mathcal{S}}(\mathcal{R}) + d_\mathcal{S}(\mathcal{R})\\
&\quad - \sum_{ij} \underbrace{d_{\mathcal{S}_{ij}}(\mathcal{R}_i) - d_{\mathcal{S}_{ij}}(\mathcal{S}_{ij}) - \lambda.L(\mathcal{S}_{ij}|\mathcal{R}_i)}_{e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i)}\\
&= E_\lambda(\mathcal{B}^0) - \sum_{ij} e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i) \qquad (5)
\end{aligned}
$$

4. OPTIMIZATION

The MDL solution of this problem is the building that minimizes the energy $E_\lambda(\mathcal{B})$.

$$\mathcal{B}_{min} = \underset{\substack{\text{hard constraints}(\mathcal{S})\\ \mathcal{P}(\mathcal{B}) \text{ is a partition}}}{argmin} E_\lambda(\mathcal{B}) \qquad (6)$$

We will assume that the exact topology and an approximate initial geometry of the building main roof planes $\mathcal{R}$ have been computed using algorithms such as [2]. The optimal minimization of the roof planes geometry and of the superstructures is a difficult problem that is not tackled in this paper. Therefore, we have chosen a suboptimal method : the roof plane geometry and the set of superstructures are optimized alternatively.

Algorithm 1 Optimization algorithm overview

Require: a DEM, $\lambda \geq 0$ and a number of iteration n.
  $\mathcal{R} \Leftarrow$ InitialReconstruction(DEM)
  repeat n times
    for all Roof plane $\mathcal{R}_i$ **do**
      $\mathcal{S}_i \Leftarrow$ Superstructures($\mathcal{R}_i, \lambda$,DEM) $\longleftarrow$ section 4.2
      $\mathcal{R} \Leftarrow$ RefineRoofPlaneGeometry($\mathcal{R}, \mathcal{S}$,DEM) $\longleftarrow$ section 4.1
  return the optimized building $\mathcal{B} = (\mathcal{R}, \mathcal{S})$

4.1. Roof plane geometry refinement

Given a building model $\mathcal{B} = (\mathcal{R}, \mathcal{S})$, this optimization step minimizes the $d_{\mathcal{R}\backslash\mathcal{S}}(\mathcal{R})$ term of equation 4. This is a minimization of $E_\lambda(\mathcal{B})$ while fixing the geometries and supports of superstructures $\mathcal{S}$ and the number of roof planes. For each roof plane $\mathcal{R}_i$, its plane parameters are optimized according to the metric of E_λ over their effective support $supp(\mathcal{R}_i \backslash \mathcal{S}_i)$. Then, using the new plane parameters and the fixed topology, the 3D-polygons of $\mathcal{R}$ are updated. This step refines the geometry of the roof planes by taking out of the optimization the areas that are currently labeled as superstructures.

4.2. Superstructure detection and reconstruction

This complementary optimization step detects and reconstructs the superstructures with constant background roof planes $\mathcal{R}$. Using the reformulation of equation 5, the minimization of $E_\lambda(\mathcal{B})$ requires, if $E_\lambda(\mathcal{B}^0)$ is held constant, the maximization of $\sum_{ij} e_\lambda(\mathcal{S}_{ij}|\mathcal{R})$. This amounts to compute the Maximum Weighted Clique in the compatibility graph where each superstructure candidate $\mathcal{S}_{ij}$ that satisfies the hard constraints corresponds to a node with a weight $e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i)$. The compatibility graph edges link superstructures that do not overlap, guaranteeing the property that $\mathcal{P}(\mathcal{B})$ is a partition.

This discrete optimization requires a finite set of superstructure candidates. This discretization is achieved, for each roof plane separately, by sampling the DEM on a regular 2D-grid that is aligned with the gutter of this roof. The parameters $\theta = (x_0, y_0, x_1, y_1)$ (figure 2) that are common to all superstructure types are then quantized on this grid to yield, for each superstructure type, the finite set of plausible containing rectangles according to the hard constraints of its superstructure type.

Given the DEM, a roof plane $\mathcal{R}_i$ and the θ parameters, a maximization of $e_\lambda(s|\mathcal{R}_i)$ constrained by the hard constraints of its superstructure type allows the continuous computation of the best specific parameters ϕ of a superstructure s for each superstructure type with a given minimum containing rectangle parameterized by θ. This means, for instance, computing the best height of a chimney of a given support, or the best roof angles of a dormer window. The supports are constrained, so the number of possible discrete rectangles θ is not proportional to the square of the roof plane area, but only to the area of the roof plane times the number of possible discrete extents $(x_1 - x_0, y_1 - y_0)$ allowed by the superstructure type. To sum-up, for each discrete rectangle of each roof plane, each superstructure type presents its best candidate, if one satisfies the type constraints.

The list of superstructure candidates must be filtered for the Maximum Weighted Clique problem to become tractable. Firstly, candidates s with a negative contribution $e_\lambda(s|\mathcal{R}_i)$ can safely be discarded. Then, the candidates that are not a local maximum of e_λ for the adjacency relation $\sim_\theta$ are discarded, where $\mathcal{S}_{ij} \sim_\theta \mathcal{S}_{ik}$ is defined by $\|\theta_{ij} - \theta_{ik}\|_\infty \leq 1$ quantization step.

The filtered set of superstructure candidates $\mathcal{C}_i$ is finally optimized without approximation using a Branch and Bound method that will prevent the brute force exploration of the $2^{\#\mathcal{C}_i}$ leaves of the search tree. The specific structure of the graph with weighted nodes that is produced by this problem allows the design of an algorithm that performs far better than a generic NP-Hard Maximum Weighted Clique solver. The energy attainable with a non-overlapping subset $\mathcal{S}_i$ of a given set of candidate superstructures $\mathcal{C}_i$ can be bounded between 0 and the maximum energy of a simpler relaxed problem. The score of the initial problem (equation 7-left), can be upper-bounded by the optimization of the same objective function with a less restrictive condition on the support areas (equation 7-right) :

$$\underset{\substack{\mathcal{S}_i \subset \mathcal{C}_i \\ \mathcal{P}_i\text{ is a partition}}}{argmax} \sum_j e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i) \leq \underset{\substack{\mathcal{S}_i \subset \mathcal{C}_i \\ \sum_j \mathcal{A}(supp(\mathcal{S}_{ij})) \leq \mathcal{A}(\bigcup_{c \in \mathcal{C}_i} supp(c))}}{argmax} \sum_j e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i) \tag{7}$$

This condition on the areas is verified by sets of disjoint superstructures and is easy to upper-bound by sorting the candidates by efficiency $e_\lambda(\mathcal{S}_{ij}|\mathcal{R}_i)/\mathcal{A}(supp(\mathcal{S}_{ij}))$ and selecting greedily the most efficient compatible candidates. This bound will let us use backtracking to prevent the exploration of entire subtrees that provably do not contribute to the optimal solution.

The connected components of the union of the supports $supp(\mathcal{C}_i)$ are treated separately, because they are independent of each other. The order of the nodes is defined as the decreasing order of scores e_λ of the superstructures that are compatible with those already chosen. Thus, the solution that is evaluated first, is the solution of the greedy algorithm that uses a best-in approach. Its approximate solution is close enough to the optimal solution to speed-up greatly the optimization process. The order in which the presence of a node is tested in the search tree impacts the computing time because it will affect the order in which the search tree is explored. The algorithm could be further improved by examining earlier, nodes that are articulation points of the incompatibility graph.

5. RESULTS

The $\mathcal{L}^1$ and $\mathcal{L}^2$ metrics have been implemented. The $\mathcal{L}^1$ metric is more adapted than the $\mathcal{L}^2$ metric to the correlated, non-Gaussian noise typically present in correlation DEMs, but the computation and optimization of the latter is much faster. The computing time is dominated by the optimizations of the specific parameters ϕ. The typical roof plane of figure 4 contains 300 000 superstructure candidates with $e_\lambda > 0$ and 100 locally minimum superstructures according to $\sim_\theta$. Computing times range from a few seconds with the $\mathcal{L}^2$ metric to a few minutes with the $\mathcal{L}^1$ metric. Furthermore, setting the number of iterations $n > 2$ does not improve the reconstruction.
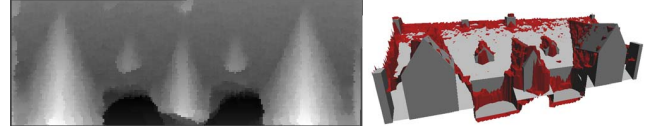
Fig. 4. A DEM (left) and its reconstruction with a 3D-triangulation that represents the DEM (right).
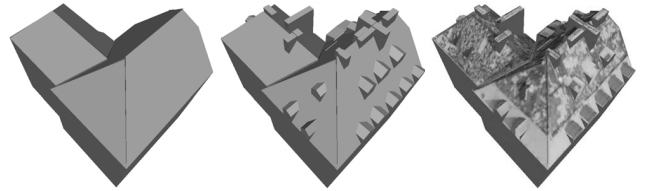

Fig. 5. The input model (left), the reconstructed building (center) and its textured version (right) where each polygon is textured by the most front facing aerial image.
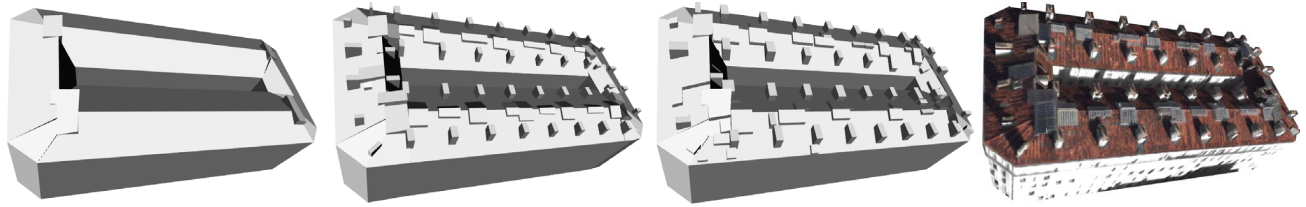
Fig. 6. The input model, the ground truth reconstructed manually, the reconstructed building and its textured version.

A ground truth has been generated by an operator using the images of the building of figure 6. This ground truth contains 46 chimneys and 40 glass roofs. The false alarm rate is of 11% (9 overdetected glass roofs) and the detection rate is of 85% (1 underdetected chimney and 12 glass roofs) without taking into account the labeling errors : 5 chimneys and 7 glass roofs have been assigned the wrong superstructure type. The true negative (TN) chimney is due to 2 true glass roofs and a true chimney that are detected as a single larger glass roof. The false positive (FP) glass roofs are due to the small volume of this object type and to the regularization of the DEM : small glass roofs may be detected near chimneys to take into account the smoothed discontinuities of the DEM. The false positive (FP) regions in figure 7 illustrate the general overestimation of the support, whereas the true negative (TN) regions are mainly caused by under-detection of superstructures. The under-detection is due to the difficulty to separate the imperfections of the DEM from the glass roof models which are allowed to have only a small height h.
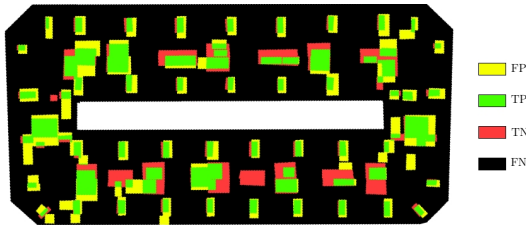
Fig. 7. Evaluation of the classification of the DEM pixels as pixels of the superstructure supports.

With a DEM resolution of 25cm, medium-sized and large superstructures are reconstructed correctly such as the 3 dormer windows of figure 8. But smaller structures are altered and hard to distinguish from noise. Coherent regions due to the regularization of the DEM are spuriously reconstructed as small superstructures. At these resolutions, the small superstructures should be disabled in the superstructure library or the MDL parameter λ has to be tuned up, because their signal is comparable to the imperfections of the DEM.

6. CONCLUSION

The proposed method achieves a reasonably fast detection and reconstruction of buildings with roof superstructures using only a DEM, an initial building model without superstructures and an easily extensible collection of parametric models defining the available superstructure types. This approach gives convincing results and is fully automatic with 10cm data using the parameters $n = 2$ and $\lambda = 0$. This method could be extended to laser scanned point clouds.

Improving the geometric accuracy would require the direct use of the images rather than processing only the DEM. Such algorithms could either fit DEM-produced models to the images as a post-process as in [8] or produce those models directly from the images. The energy could be reformulated as a Bayesian energy that handle interactions between the superstructures and the roof planes, to introduce stronger priors such as alignments or to model the imperfections of the DEM. Finally, the simultaneous optimization of the whole building model may be explored. It would require a broader framework, of which Reversible Jump Monte Carlo Markov Chain[4] seems to be a good candidate.
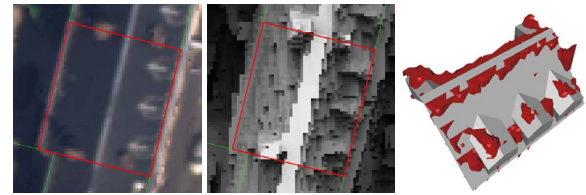
Fig. 8. An image with 25cm resolution, its shaded DEM (same resolution), and the reconstructed building with a DEM triangulation.

7. REFERENCES

[1] C. Baillard and A. Zisserman, "A plane-sweep strategy for the 3D reconstruction of buildings from multiple images," in *19th ISPRS Congress and Exhibition*, Amsterdam, The Netherlands, 2000.

[2] M. Durupt and F. Taillandier, "Automatic building reconstruction from a digital elevation model and cadastral data: an operational approach," in *Proc. PCV*, Bonn, Germany, 2006, ISPRS.

[3] R. Nevatia and K. Price, "Automatic and interactive modeling of buildings in urban environments from aerial images," in *Proc. ICIP*, New York, 2002.

[4] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny, "An automatic building reconstruction method : A structural approach using high resolution images," in *Proc. ICIP*, Atlanta, USA, 2006.

[5] S. Roy and I.J. Cox, "A maximum-flow formulation of the n-camera stereo correspondence problem," in *Proc. ICCV*, Bombay, 1998.

[6] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.

[7] A.R. Dick, P.H.S. Torr, and R. Cipolla, "Modelling and interpretation of architecture from several images," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 111–134, Nov. 2004.

[8] K. Schindler and J. Bauer, "A model-based method for building reconstruction," in *Proc. ICCV workshop on Higher-Level Knowledge in 3D Modeling and Motion (HLK'03)*, Nice, France, 2003.