AUTOMATIC PRODUCTION OF OCCLUSION-FREE RECTIFIED FAÇADE TEXTURES USING VEHICLE-BASED IMAGERY

Sébastien Bénitez, Eloïse Denis and Caroline Baillard
SIRADEL, 3 allée Adolphe Bobierre, CS 24343, 35043 Rennes Cedex, France
sbenitez@siradel.com, edenis@siradel.com, cbaillard@siradel.com

Commission III, WG III/4

KEY WORDS: Texture, Building, Multi-sensor, Close-range, Terrestrial, Registration

ABSTRACT

This paper presents a fully automatic method for computing occlusion-free rectified façade textures, using an input 3D model and data acquired with a mobile mapping vehicle. The purpose of this study is to enhance available building 3D models with realistic textures. The terrestrial data consist of geo-referenced terrestrial images and a laser point cloud. The method is three-folded. Firstly the input 3D model and the terrestrial data are registered, and each façade image is geometrically rectified. Secondly "predictable" and "non-predictable" occlusions are detected combining laser and image information. Finally the images are merged into rectified occlusion-free seamless façade textures. The main novelty of the system is the method for occlusion management.

1. INTRODUCTION

The demand for 3D realistic urban models has tremendously increased in the past few years. Public tools like Google Earth or Bing Maps have become very popular. More and more local authorities use digital 3D city models to study and present new urban projects. The visual realism of the 3D models is very dependent on the façade textures.

The first urban models consisted of grey building models with flat roofs extruded from 2D outlines. With the emergence of more powerful visualization tools, very efficient approaches based on generic textures have appeared (Parish and Müller, 2001, Wonka *et. al*, 2003). They are still necessary when no real image can be found, as in historical reconstitutions for example. Another kind of approach consists in using oblique aerial imagery (Frueh *et al*., 2004). This technique is very popular thanks to its ability to produce very large areas at a moderate cost. However the resulting models cannot be navigated at a street level because of the large difference in the acquisition and the navigation points of view. The texture resolution is poor and side effects can occur, like roller blinds or balconies projected down to the wall. Recently, new approaches based on mobile mapping systems have emerged in public laboratories (Bentrah *et a*l. 2004, Goulette *et al*. 2007), as well as in private companies (Hunter, 2009; Mrstik and Kusevic, 2009). The mapping vehicles are often equipped with accurate geo-localization systems based on GPS, IMU and DMI, close range optic cameras and high pulse laser scanners. With the development of mobile mapping systems, a lot of research has focused on the automatic extraction of façade textures from terrestrial imagery. A possibility consists in texturing a mesh derived from the laser point cloud (Brun *et al.* 2007). If an existing 3D building model is available, it can be textured with images of the whole façade using 2D/3D fitting techniques (Haala and Böhm, 2003, Colleu *et al*. 2007). However these techniques are not appropriate to very close range imagery. In (Kang *et al*., 2007), close-range images are first rectified using vanishing points, then combined via the adjustment of the exterior orientation parameters. The resulting mosaic is finally refined using horizontal lines and iterative corner detection. In (Hoegner and Stilla, 2009), building models are textured with infrared images.

A key issue is the detection and the removal of occlusions in the images. In (Ortin and Remondino, 2005) and (Böhm 2004), two methods for occlusion-free texture generation are described, but a minimum of 3 overlapping images is necessary, which requires a high acquisition rate or a very slow driving speed in narrow streets. In (Frueh *et al*. 2005) a method based on laser points is presented. The point cloud is used to generate a 3D mesh that is then classified as foreground or background in order to predict occlusions.

This paper focuses on the production of occlusion-free façade textures based on close-range imagery. SIRADEL has produced many 3D city models around the world using stereo aerial imagery. The purpose of this study is to enhance these 3D models with realistic textures acquired with a mobile mapping vehicle. The terrestrial data consist of geo-referenced terrestrial images and a laser point cloud. A façade texture generally consists of a mixture of 10 to 20 overlapping input views. An example of input images is shown in Figure 9.

The method consists of three processing stages. The first stage aims at registering the input 3D model and the terrestrial data: the 3D model is first refined using laser points and image correlation (Section 2.1), then images are rectified according to the corresponding façade (Section 2.2). The second stage of the method is the occlusion detection (Section 3). In the final stage, the images are merged into a rectified occlusion-free seamless façade texture (Section 4). Experimental results are presented and discussed in Section 5.

2. MODEL REGISTRATION AND IMAGE RECTIFICATION

2.1 3D Model registration

The 3D models derived from aerial imagery are in general less accurate than the 3D information derived from geo-referenced terrestrial data. Moreover, the roof outlines derived from aerial imagery often correspond to the gutter limit but not necessarily to the wall positions. As inaccurate wall positions can create serious artifacts in the texturing process, the input 3D models need to be refined according to the terrestrial data.

Laser point clouds are prioritarily used to refine the façade position and delineation. After a preliminary global registration using the ICP algorithm, each visible façade is individually processed and compared to a filtered point cloud containing only the façade points. More details about the method can be found in (Denis and Baillard, 2010).

If the refinement process based on laser data is not successful (no laser data available or not enough laser points describing the façade), then a second method based on image correlation is performed. The wall is translated along its normal direction, and each position hypothesis is assessed by a score. The score is computed by correlating the two images closest to the wall centre, using their corresponding orientation parameters given by the GPS+IMU system. Only the pixels with a strong gradient and belonging to the façade are taken into account. Besides, pixels located at the top or the bottom of the façade are excluded. From a practical point of view, a correlation score is computed every 20cm, starting at the original wall position and ending 3 meters inward. The best score determines the best wall location. Two additional translations of 10cm around the best location are scored to refine the result (see Figure 1). If the best score or the score dispersion is too low, the registration is considered as non-reliable.

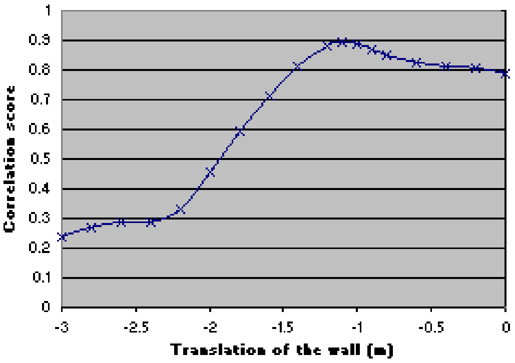


Figure 1. Example of correlation score as a function of the wall translation.

2.2 Image selection and rectification

Each façade is first associated to a set of images. The image selection is particularly important for large-scale façade mapping where thousands of images can be available. A wall-by-wall analysis based on 3D ray tracing is performed, which is a good compromise to achieve a relevant selection whilst limiting computation (Bénitez and Baillard, 2009).

The next step is to create geometrically rectified texture images. The original images are back-projected onto the wall plane. The resolution of the resulting images is predetermined using the image closest to the wall. Two cases are distinguished. If the 3D model has been successfully registered to the terrestrial data, the back-projected images are well registered, thanks to the rigid mounting of lasers and cameras. However, if the model registration has failed, an image-based correction based on SIFT points (Lowe, 1999) is applied to the back-projected images. The most central image is selected as the reference set. Then every other image is processed in turn. It is wrapped to the reference set using an homography estimated as follows. Let *Ic* be the current image and *Ir* the closest image in the reference set (camera axes intersecting the wall plane at the shortest distance). SIFT points are computed over *Ic* and *Ir*, then the n (five in our case) best matching pairs according to the distance between SIFT parameters are extracted. The computing time is reduced by sorting the points along the x-axis. Only symmetrical matches are accepted. Additionally, a threshold is applied on the correlation score computed over the candidate pairs. If several pairs meet the matching constraints, then the smallest SIFT distance is favoured. Finally, a threshold on the maximal spatial distance between matched points is introduced, in order to avoid false matches between repetitive objects (like façade windows for instance). If the number of matched points over the images is too small, the algorithm is iteratively applied with an increasing threshold on the SIFT distance. When enough matched points have been determined (20 points practically), the homography between the two images is estimated using RANSAC. The homography is applied to *Ic* and the resulting rectified image is added to the reference set. At the end of this stage, each image is geometrically rectified and can be superimposed in the façade plane.

3. OCCLUSION MANAGEMENT

Two kinds of occlusions are distinguished. The predictable occlusions are caused by buildings or walls present in the input 3D model. The non-predictable occlusions are caused by objects that were not modelled (pedestrians, vehicles, posts…), or not accurately enough (vegetation). The purpose of this stage is to associate an occlusion layer to each rectified image, depicting both kinds of occlusions.

3.1 Predictable Occlusions

A mask of predictable occlusions is computed for each rectified façade image. More precisely, each pixel of the rectified image is associated to a 3D ray passing through the camera center point. The 3D ray is tested with respect to all the walls compatible with the camera (wall plane facing the camera within a given distance). If the ray intersects another façade at a location closer than the current façade, then the original pixel is marked as an occlusion point. To compensate for possible positioning errors, the façade delineations are dilated from a small distance. The result is an occlusion layer describing slightly dilated occlusion areas. Thanks to the image overlap, the dilatation will not affect the final texture. Figure 2 shows an example of a predictable occlusion layer and its benefit on the final texture.

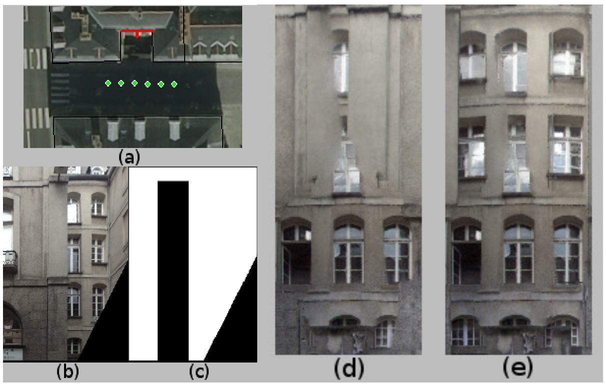


Figure 2. Example of occlusion layer benefits: (a) Study case: wall with a projection aisle (in thick red) and camera centre positions (in green); (b, c) one of the 12 rectified images and its occlusion layer; (d) final texture without using predictable occlusions; (e) final texture using predictable occlusions.

3.2 Non-predictable Occlusions

Most occlusions occurring in urban areas cannot be predicted using the input 3D model. They are generally caused by moving objects (cars, pedestrians), non-modeled objects (traffic lights, road signs, fences), or inaccurately modeled objects (vegetation). In order to complete the occlusions layers, a new process for detecting non-predictable occlusions has been designed. Occlusions are first detected using the laser point cloud acquired by the vehicle, then completed with an image-based correlation technique.

3.2.1 Laser-based occlusion detection

The 3D model has previously been registered to the laser point cloud (see Section 3.1). Each façade can therefore be associated to a laser acquisition time interval. All the laser points acquired during this time interval are extracted from the original cloud. Besides, all the points belonging to the façade or its background can be removed.

The points belonging to the ground are also extracted. A first set of ground points is detected within each vertical scan line as the lowest significant peak in the elevation histogram. These points are then used as seeds for a local surface growing algorithm applied to the whole cloud. The ground points are iteratively and chronologically stored into a small size queue, the acquisition order corresponding to the progression along the street. At each iteration, a mean square plane is computed over the stored points. All the points of the cloud belonging to this plane are marked as ground points. The last acquired ground points are used for updating the queue. Thus, the seed location moves along the street at each iteration, and follows the ground curvature.

The remaining points describe occluding objects related to the current façade. These points are projected onto the occlusion layers associated to the rectified images of the façade. As laser points only provide a spatial sampling of the objects, the points are replaced by squares corresponding to the base of the camera beam. The base height is derived from the laser vertical resolution and the distance to the camera. The base width is derived from the vehicle displacement between two laser scan lines. As in (Frueh et al. 2005), it would be interesting to involve the acquisition angle to refine the base width.

Figure 3 shows three laser-based occlusion layers superimposed with the corresponding rectified images. Figure 4a and Figure 4b show a fourth rectified image and its associated laser-based occlusion layer. The car has been correctly detected but not the pedestrian. A false detection can also be observed just above the car, caused by another pedestrian not visible in the image. Moving objects are particularly difficult to handle because laser data and images are not acquired exactly at the same time. The cameras are triggered every n meter whereas laser data are continuously collected according to scan lines. Two cases of failures are distinguished:

- False-positive: an occluding object was detected in the laser cloud but it is not visible in the image.
- False-negative: no occlusion was detected in the laser data although a mobile object is visible in the image.

These two cases are handled using image information, as explained below.

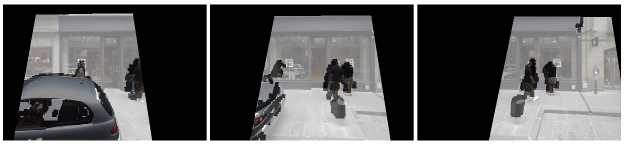


Figure 3. Example of laser-based occlusion layers superimposed with the corresponding rectified images.

3.2.2 Image-based occlusion refinement

In order to solve the false-negative cases, the laser-based occlusion detection is completed with an image-based technique based on (Böhm, 2004). Occlusions are detected with a background estimation technique. Each façade point is projected onto the various images, and the corresponding pixels are clustered in a RGB space. The cluster containing most pixels is assumed to describe the background, and the other pixels are marked as image-based occlusions in the corresponding occlusion layer. A dilatation and erosion are subsequently applied to remove small regions.

Figure 4c shows a result of the image-based occlusion detection. Figure 4d shows the occlusion layer obtained by combining the laser-based and image-based detections. The mobile pedestrian has been almost entirely detected. The residual false detections have no effect on the final texture as the radiometry can be taken from another image.

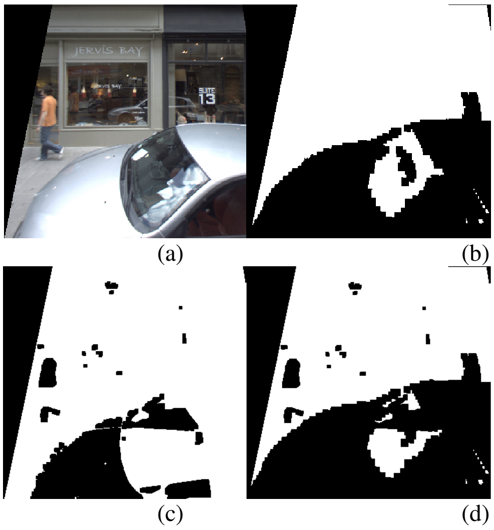

(a) (b)
(c) (d)

Figure 4. (a) Rectified image; (b) Laser-based occlusion layer; (c) Image-based occlusion layer; (d) Combination of laser-based and image-based occlusion detection.

The false-positive laser-based detection is solved by a similar technique. First the occluding laser points are grouped into connected components describing potential occluding objects. The laser points associated to each occluding object are then projected onto the various images, and the corresponding pixels are clustered in a RGB space. Small clutter dispersion reinforces the presence of a static object at the location indicated by the laser point. However, high clutter dispersion indicates that the detected occluding object might actually be mobile and should not be taken into account at this particular location in the occlusion layers. Hence, if a majority of points are associated to a high dispersion, then the corresponding occluding object is discarded. It is a case of mobile occlusion that should be handled using image-based occlusion detection as explained previously.

The method is illustrated in Figure 5. Figure 5a shows the dispersion scores computed for each laser point. Figure 5b shows the thresholded scores and Figure 5c shows the classification of the occluding objects as valid (black, low dispersion) or discarded (white, high dispersion). Figure 6 shows an example of valid occlusion laser points, colored using image RGB information. Figure 7 shows an example of final texture computed with and without occlusion detection. Most occlusions have been detected and replaced. Two errors can still be observed. The car windscreen has not been removed because it was not scanned by the laser nor detected by image-based clustering. One of the pedestrians standing in front of the window has not completely disappeared either, because he stands very close to the wall.

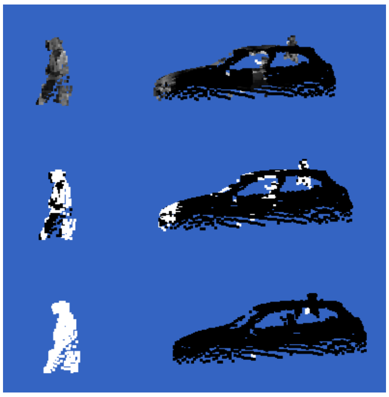


Figure 5. Example of foreground objects in front of a building (a) Dispersion scores (white = low similarity) (b) thresholded dispersion scores (c) Object classification

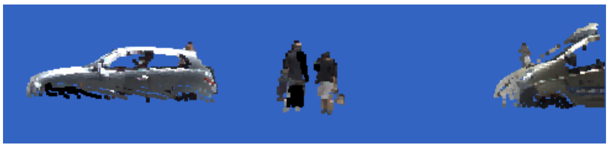


Figure 6. Coloured occlusion points

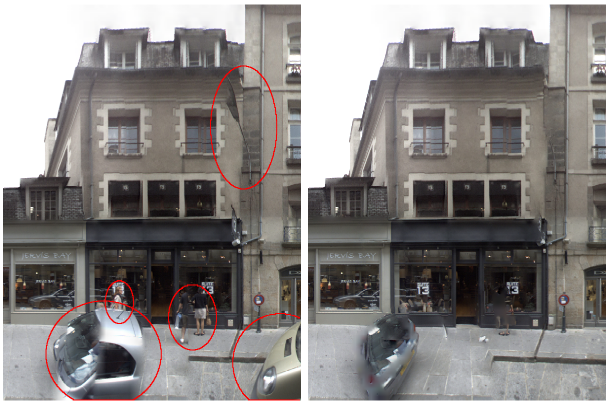


Figure 7. Left image: result without occlusion detection; Right image: result with combined occlusions detection.

4. IMAGE MERGING

The purpose of this final stage is to produce seamless mosaics from a set of input rectified images and the corresponding occlusions layers. It consists of three different steps. The radiometry of each rectified image is first corrected by gain correction. Then the images are merged together by multi-band blending. Finally the occluded areas are filled in using inpainting techniques.

4.1 Gain correction

The image brightness of a camera can significantly vary within a façade. The variations are even more important between two different cameras. An example of luminance variations over a façade can be seen in Figure 9.

A standard method for correcting luminance consists in estimating a gain correction for each image. Given a reference image, the standard gain correction of the current image is given by:

$$GainCor_{Cur} = \frac{Lumi_{Ref}}{Lumi_{Cur}} \times GainCor_{Ref},$$

where $Lumi_{Ref}$ and $Lumi_{Cur}$ are respectively the luminance of the reference and the current image computed over their overlapping area, and $GainCor_{Ref}$ is the gain correction of the reference image.

The method was extended to the simultaneous management of top/bottom and left/right image pairs. A weighted cross-correction favouring the up/bottom correction has been defined:

$$CrossGainCor_{Cur} = \frac{GainCor_{left/right} + n \times gainCor_{up/bottom}}{n+1},$$

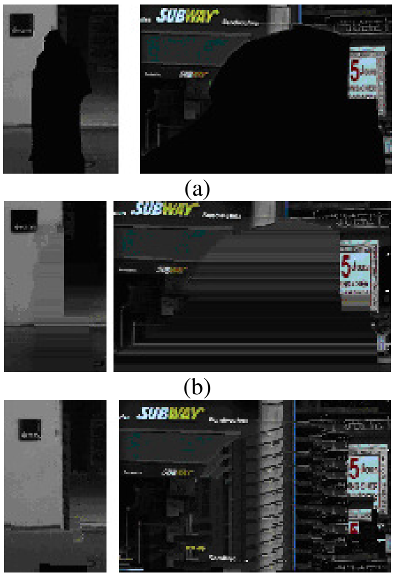
where $GainCor_{left/right}$ and $GainCor_{up/bottom}$ are the standard gain corrections computed horizontally and vertically, and n is a integer >1 depending on the camera setting (aperture…). The weight n can vary between each mission and needs to be estimated on one or two buildings before use on the full set. The cross-correction is first performed on the most central image then propagated on its neighbours. As bright occluding objects can mislead the gain estimation (a white van parked in front of a building for instance), it is important to locate large occluding objects before computing the gain. Figure 11b shows the result of the gain correction on the example of figure 9. In the future, an adaptive gain correction will be studied, where the correction can vary inside the image.

4.2 Multi-band blending

Standard methods based on averaging or p-norm are simple and quick but they often introduce ghosts in image transitions. Another kind of approach consists in finding the best seamlines separating two images (Mallick, 2002, Gracias et al, 2003). The chosen method is a multi-band approach (Burt et al., 1983) based on Laplacian pyramids associated to weight images. It combines the techniques described in (Brown and Lowe, 2007) and (Wang *et al*., 2002). The weight images take into account an occlusion layer and an obliqueness layer. The occlusion layer is the combination of the predictable occlusions and non-predictable occlusions presented in Section 4. The obliqueness layer describes the view angle of the imaged pixel, giving priority to the pixels close to the center (near the principal point). This method has proved to be a good compromise between quality and computing time.

4.3 Inpainting

Large objects or objects near the façade can create occluded areas that cannot be seen in any of the camera views (see example of Figure 8a). The pixel values can be interpolated from the surrounding pixels only if the area is small with a homogeneous radiometry, otherwise it does not give satisfying results (see Figure 8b). A method based on (Rasmussen and Korah, 2005) has been implemented. Pixels along the border of the empty area with a strong gradient are processed first. A patch is defined around the pixel, and the most similar patch around this position is searched. The best match is used to replace the current pixel. The process is iterated until all empty pixels are processed. This method gives better results (see Figure 8c) but requires more computing time. It could be improved by introducing a confidence term propagated with the colour information (Criminisi *et al*, 2004) or by introducing a façade grammar (Konushin and Vezhevets 2007, Korah and Rasmussen 2008).


(a)

(b)

(c)

Figure 8. Inpainting techniques. (a) Input images with unknown radiometry at occluded areas; (b) Inpainting result using interpolation; (c) Inpainting result using patch matching.

5. EXPERIMENTATION

5.1 Test data set

The method has been tested on a small area of an historical center. The path is of about 1.1 km long (see Figure 10).

The input 3D model is derived from aerial geo-referenced images, and they have a metric accuracy. The terrestrial mobile mapping system is equipped with cameras, lateral laser scanners, and a precise geo-referencing system composed of a Global Positioning System (GPS), an Inertial Measurement Unit (IMU) and an odometer. The data collection was performed at a normal speed. The images were captured every 2 meters. A façade texture generally consists of a mixture of 10 to 20 overlapping input views (see example of Figure 9). The data set consists of about 2300 images with 400 buildings. It is completed with a georeferenced laser point cloud describing the façades. Laser data collection was continuous along this track. All the facades are assumed to be planar.

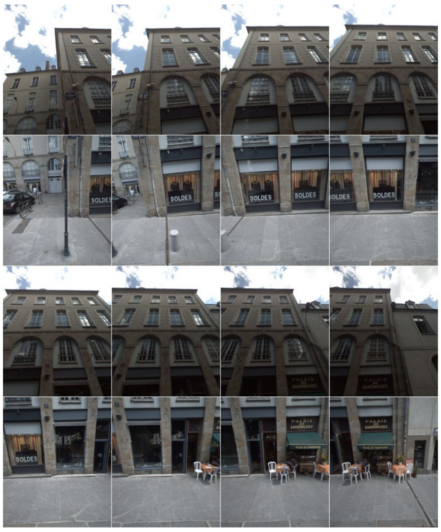


Figure 9. Example of input images

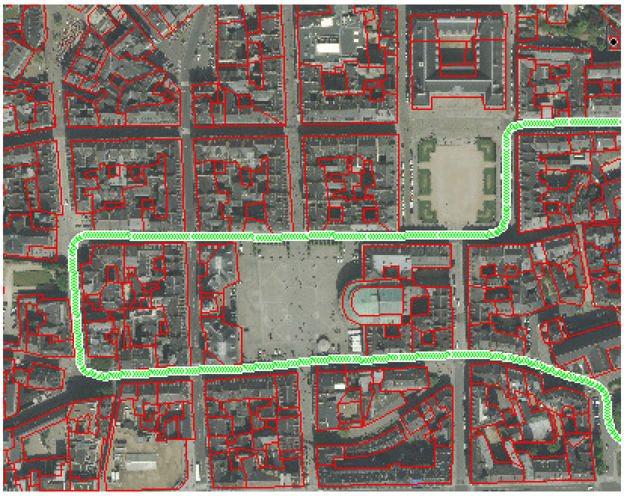
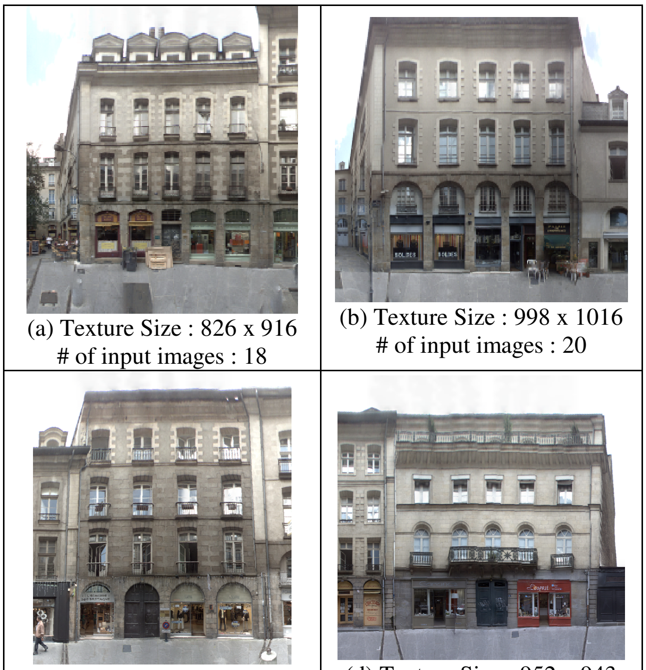


Figure 10. Test Area: building polygons and vehicle track

5.2 Results

Texture images have been generated for 46 façades, at a typical resolution of 2cm per pixel. Figure 11 shows various examples of automatically computed façade textures. Figure 12 shows a view of the final textured model, where the texture images were automatically snapped to the model. The computing time was around 4 minutes per façade on a standard computer dual core 1.86GHz.

The images have been rectified and merged into homogeneous textures. The luminance is corrected and quite stable. Some occluding objects have not been not removed (bins, tables, signs) because they are very close to the wall; in this case it is better to leave them in the texture rather than using inpainting techniques.


(a) Texture Size : 826 x 916
# of input images : 18

(b) Texture Size : 998 x 1016
# of input images : 20

(c) Texture Size : 871 x 991
# of input images : 18

(d) Texture Size : 952 x 943
# of input images : 22

Figure 11. Some resulting facades

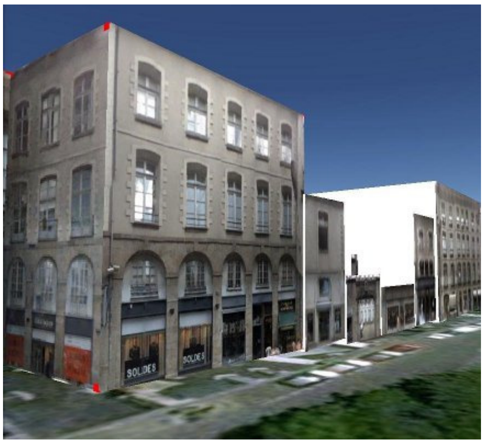


Figure 12. Textured 3D model

6. CONCLUSION AND PERSPECTIVES

The method presented in this paper is a major step towards a fully automatic system for texturing existing 3D building models. The terrestrial images are automatically corrected and merged into rectified façade textures. Special attention has been paid to occlusion management, and most occluding objects are removed from the final textures. The results show the feasibility of a fully automatic approach based on vehicle-based data. The resulting textured 3D models are realistic and can be navigated from the street level. In the future, the image merging stage will be improved, especially the inpainting technique. Importantly, a method for automatically registering the façade textures to the façade model is necessary in order to propose a robust and fully automated texturing system.

REFERENCES

Benitez, S. and Baillard, C., 2009, Automated Selection of Terrestrial Images from Sequences for the Texture Mapping of 3D City Models. *IAPRS & SIS, vol. XXXVIII (Part 3/W4), Paris, France pp. 97–102*.

Bentrah, O., Paparoditis, N., Pierrot-Deseilligny, M., 2004. Stereopolis : an Image Based Urban Environments Modeling System. *In International Symposium on Mobile Mapping Technology (MMT)*, Kunming, China, March 2004.

Bohm, J., 2004. Multi-Image Fusion for Occlusion-Free Façade Texturing. *IAPRS&SIS, vol. XXXV (Part B5), Istanbul, Turkey pp. 867-872.*

Brown, M. and Lowe, D., 2007, Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision, vol. 74 (1), pp. 59-73, August 2007.*

Brun, X., Deschaud, J.E. and Goulette, F., 2007, On-the-way City Mobile Mapping Using Laser Range Scanner and Fisheye Camera. *IAPRS & SIS, vol. XXXVI (Part 5/C55), Padua, Italy pp. 29-31*.

Burt, P. J. and Adelson E. H., 1983, A Multiresolution Spline with Application to Image Mosaics. *ACM Transactions on Graphics, Volume 2, Issue 4, October 1983*

Colleu, T., Sourimant, G., and Morin, L., 2007, “Une méthode d’initialisation automatique pour le recalage de données SIG et vidéo,” *CORESA*, 2007.

Criminisi, A., Pérez, P., and Toyama, K., 2004, Region Filling and Object Removal by Examplar-Based Image Inpainting. *IEEE Trans on Image Processing, Vol. 13, n°9, Sept.2004.*

Denis, E. and Baillard, C., 2010, Refining existing 3D building models with terrestrial laser points acquired from a mobile mapping vehicle. *IAPRS & SIS, Vol. XXXIX (Part 5), Newcastle upon Tyne, England.*

Frueh, C., Sammon, R., Zahcor, A., 2004, Automated Texture Mapping of 3D City Models with Oblique Aerial Imagery. *3D Data Processing, Visualization and Transmission, 3DPVT 2004, pp. 396-403, 2004*.

Frueh, C., Jain, S. and Zakhor, A., 2005. “Data Processing Algorithms for Generating Textured 3D Building Facade Meshes from Laser Scans and Camera Images,” *International Journal of Computer Vision, vol. 61, Issue 2, p. 159-184, January 2005.*

Goulette, F., Nashashibi, F., Abuhadrous, I., Ammoun, S. and Laurgeau, C., 2007. An Integrated On-board Laser Range Sensing System for On-the-way City and Road Modeling. *Revue française de photogrammétrie et de télédétection, vol. 185, p. 78.*

Gracias, N., Mahoor, M., Negahdaripour, S. and Gleason, A., 2009, Fast Image Blending using Watersheds and Graph Cuts. *Image and Vision Computing, Volume 27, Issue 5, pp. 597-607.*

Haala, N. and Böhm, J., 2003, “A multi-sensor system for positioning in urban environments,” *ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58 (1-2), pp. 31-42, 2003*

Hoegner, L. and Stilla U., 2009. Thermal leakage detection on building facades using infrared textures generated by mobile mapping. *In 2009 Urban Remote Sensing Joint Event, Shangai, China.*

Hunter, G, 2009. Streetmapper mobile mapping system and applications in urban environments. In *ASPRS Annual Conference, Baltimore, USA.*

Kang, Z., Zhang, L., and Zlatanova, S., 2007, “An automatic mosaicking method for building facade texture mapping,” *IAPRS & SIS, vol. XXXVI (Part 4/W45), Stuttgart, Germany pp. 1-9.*

Konushin, V. and Vezhnevets, V., 2007, "Automatic building texture completion" *Proc. of Graphicon'2007, pp. 174-177, Moscow, Russia.*

Korah, T. and Rasmussen, C., 2008, “Analysis of Building Textures for Reconstructing Partially Occluded Facades”, *Proceedings of the 10th European Conference on Computer Vision, pp.359-372, Marseille, France, 2008.*

Lowe, D. G., 1999. Object Recognition from Local Scale-invariant Features. *In Proc. of the International Conference on Computer Vision ICCV 1999*, pp. 1150-1157.

Mallick, S. P., 2002, Feature based Image Mosaicing. *Technical Report*, University of California, San Diego, Aug. 2002.

Mrstik, P., and Kusevic, K., 2009. Real Time 3D Fusion of Imagery and Mobile Lidar, *ASPRS Annual Conf*, Baltimore, USA.

Ortin, D. and Remondino, F., 2005, “Occlusion-free Image Generation for Realistic Texture Mapping”, *IAPRS & SIS, vol. XXXVI (Part 5/W17), Venice, Italy, on CD-ROM*.

Parish, Y.I.H. and Müller, P., 2001, “Procedural Modeling of Cities,” *SIGGRAPH*, 2001, pp. 301-308.

Rasmussen, C. and Korah, T., 2005, Spatiotemporal Inpainting for Recovering Texture Maps of Partially Occluded Building Facades. *IEEE International Conference on Image Processing, Volume 3, pp. 125-128, September 2005.*

Wang, X., Totaro, S., Taillandier, F., Hanson., A. R., Teller, S., 2002. Recovering Facade Texture and Microstructure from Real-World Images, *Proc. 2nd International Workshop on Texture Analysis and Synthesis, pp. 145-149, Copenhague, Denmark, June 2002*

Wonka, P., Wimmer, M., Sillion, F. and Ribarsky, W., 2003, Instant Architecture, *ACM Transations on Graphics, vol. 22 (3), pp. 669-677.*