DETECTION OF WINDOWS IN BUILDING TEXTURES FROM AIRBORNE AND TERRESTRIAL INFRARED IMAGE SEQUENCES

D. Iwaszczuk, L. Hoegner, U. Stilla

Technische Universitaet Muenchen (TUM), Photogrammetry & Remote Sensing, Munich, Germany - (iwaszczuk, hoegner, stilla)@bv.tum.de

Commission VI, WG VI/4

KEY WORDS: Infrared, Image Sequences, Texture Mapping, Structure Detection

ABSTRACT:

Infrared (IR) images depict the thermal radiation the surface. Imaging the building facades and roofs with a IR camera thermal inspection of the buildings can be carried out. In such inspections a spatial correspondence between IR-images and existing 3D building models can be helpful. Texturing 3D building models with IR images this spatial correspondence can be created. It textures heat leakages can be detected. However an automatic extraction of leakages is disrupted by windows. Therefore the windows should be marked out before detecting leakages.

Most common algorithms for window detection were thought for images in visual band. In this paper, a new algorithm for windows detection in textures extracted from terrestrial IR images is proposed. In the first step small objects have to be removed by scaling down the image (texture). A segmentation of this scaled image is done using a local dynamic threshold. Morphological operations are used to remove false detections and unify substructures of windows. For every extracted region, which is a candidate for a window, the center of gravity is calculated. The airborne images have lower resolution, so the scaling is usually not necessary. However, in these images most heat leakages are not visible. Therefore the airborne images can be used as complementation of the terrestrial images for windows detection. The segmentation is done in the same way as for downscaled terrestrial images.

It is assumed that windows on facades are ordered in regular rows and columns. This assumption allows doing a least squares adjustment for the centers of gravity, so that they build a regular grid. This set of points is used as center points of windows for the estimation of the geometric shape of the windows. For this purpose, rectangular masks are matched in a local window around every center point. It is assumed that, if all windows are ordered in a regular grid, they have the same geometric appearance. This assumption allows combining all parameters of all windows in an adjustment to find the optimal mask for all windows. This mask is then used for every center point to mark out the windows from the texture. The remaining texture now contains only the thermal signature of the façade and no more errors caused by mirroring effects. This would allow the localization of thermal leakages on the façade.

First experiments are done with data of IR image sequences of test area “Technische Universitaet Muenchen” in Munich, where one set is taken from a moving car and one set is taken from a helicopter. The recorded facades can be divided in two groups. Group one are facades with regular structures and uniform thermal behavior. The second group contains façades with some irregular structures and different thermal behavior for different parts of the façade. First results show that the combination of terrestrial and airborne based IR textures reduces the number of false detections of windows and increases the completeness of the window detection. The restriction of a regular grid of windows given in the adjustment process limits the usability to regular and approximately known structures. For irregular arrangements of windows the assumption of regular grid cannot be used and thus the adjustment of the centers of gravity and the estimation of additional positions of windows are not possible. In this case, the only improvement of window detection is the combination of textures from terrestrial and airborne IR images.

1. INTRODUCTION

Mobile mapping systems are used mainly for a fast acquisition of special data. A mobile mapping system usually consists of a moving platform (vehicle or aircraft), a navigation sensor and a mapping sensor (Li, 1997). Typically for navigation Global Positioning System (GPS) combined with Inertial Navigation System (INS) as well as with other navigation sensors (e.g. vehicle wheel sensors) are used. On the one hand GPS provides rough knowledge of position and orientation. On the other hand, an INS yields good short-term accuracy, but in a longer time a systematic drift occurs. However, the combination of GPS and INS allows to avoid the INS drift and to bridge any loss of satellite signal by GPS (Yastikli & Jacobsen, 2005).

Using the information from GPS/INS the images can be geo-referenced directly and matched with existing 3D building models (Frueh et al., 2004; Eugster & Nebiker, 2009; Avbelj et al., 2010). Images combined with a 3D city model can be used in variety applications, e.g. visualization, urban planning, navigation.

1.1 Motivation

In this paper we present an application of mobile mapping system for thermal inspection of buildings. Due to increasing energy costs and climate changes, energy efficiency of buildings became in recent years an important topic. Much effort is required for reducing the energy loss. Sustainable building as well as thermal inspection and monitoring of old buildings contribute to further development saving energy. Thermal radiation of the building hull can be captured by infrared (IR) thermal cameras. In these images weak spots can be detected. In such inspections a spatial correspondence between IR-images and existing 3D building models can be helpful. Usage of the 3D building models enables processing of image sequences taken by a mobile platform capturing whole districts or cities. The correspondence can be created by geo-referencing of the images. This can be done directly using GPS/INS data from the devise mounted with the camera on the mobile platform. After system calibration the 3D model can be projected into the image and for each surface of the model a region of the image can be selected for texture. Using terrestrial image sequences from a camera mounted in a vehicle frontal faces can be captured, while airborne image sequences can be taken for roofs and inner yards. The physical behavior of the IR spectrum causes camera systems with lower resolution and smaller viewing angle than normal video or photo cameras. Therefore, on the one hand, in terrestrial images it is not possible record a complete building in dense urban areas with narrow streets in one image. Because of this, textures have to be combined from multiple images. On the other hand, in airborne images the resolution is low and not all details can be seen.

In these textures with 3D spatial reference certain structures can be detected, which automatically get the spatial reference and belonging to the building from the texture. These structures can represent e.g. windows, heat leakages and other objects in walls and roofs.

Windows in terrestrial IR images appear similarly to the heat leakages and usually disturb leakage extraction. Hence, to distinguish heat leakages, which are most interesting for thermal inspection, an additional algorithm for windows detection is necessary. Actually, the appearance of a window in an IR image is not representing its temperature, but it is a reflection caused by sun or objects on the opposite side. Infrared images are usually blurred, so that the windows edges cannot be easily detected. Furthermore, due to the reflections, windows can appear with irregular shape, thus it is difficult to fit rectangles for windows detection.

1.2 Related work

Existing algorithms for façade reconstruction were developed for optical textures and point clouds extracted both from image sequences and laser. Dick et al. (2004) introduce a texture prior generated from training data to identify image areas as supposable façade elements. A group of histograms is generated for every texture using horizontal and vertical wavelet transforms at for different scale levels. A likelihood functions evaluates possible structures with their priors. Ripperda (2008) and Mayer and Reznik (2006) propose reversible jump Markov chain Monte Carlo (rjMCMC) (Green 1995) for the estimation of optimal parameters for windows. Ripperda (2008) is using a formal grammar to describe the behavior of windows. Becker (2009) uses cell decomposition to extract geometry from images as well as from LiDAR point clouds. Formal grammars are used to predict invisible parts and to enhance the window detection. A problem of grammars is the complexity of big scenes and the limitations of the rules defined for the grammars. Reznik and Mayer (2007) are using implicit shape models (Leibe and Schiele 2004) to define a set of windows to check with the given image. Werner and Zisserman (2002) use regular structure primitives like vanishing points or symmetry.

The low optical resolution and smooth edges in infrared images limit the usability of most approaches. The description of both implicit shape models and explicit geometry are limited to well-known optical behaviors whereas windows in infrared show a big variety of optical behavior. Edge points are hard to determine exactly. Window frames can appear lighter or darker than the façade. The façade often contains neither homogeneous nor regular structure due to temperature differences. Approaches dealing with more general geometric primitives should be preferred. In Sirmacek et al. (2011) L-shape structures are searched in textures and combined in a bottom up approach. In Hoegner and Stilla (2009) morphological operators are used to mask window areas and distinguish them from other objects like heat leakages.

1.3 Overview

In this paper we present a new method for window detection in textures extracted from IR images. Our method, which is described in Section 2, consists of three steps. In the first step a segmentation of the image using local dynamic threshold is carried out and morphological operations are used to improve the results of the segmentation (Section 2.1). Further geometric constrains are applied to achieve a regular grid of the windows (Section 2.2). Finally, a local correlation is carried out to improve the final position of the window (Section 2.3). The presented method was applied for a test data set "Technische Universität München" (Section 3). First results of the experiment are presented in Section 4. Results and further work are discussed in Section 5.

2. METHOD

2.1 Segmentation

Terrestrial images of building facades captured from a vehicle usually are taken from small distance. Hence the resolution of the textures extracted from these images is relatively high. In these textures many details can be seen, the most meaningful for the thermal inspection of buildings are heat leakages. In textures it is difficult to detect heat leakages, because they appear similarly to the windows. Thus, before detecting thermal leakages, localization of the windows should be defined. Most windows are rectangular and build a grid on the façade. Using the knowledge about can help to detect them. For windows detection the small objects have to be removed by scaling down the image (texture).

In the next step the texture is segmented using a local dynamic threshold. The local dynamic threshold is used because the windows, according to different reflections, appear differently even in the same texture and no global threshold can be found.

After segmentation some windows are fused into one object. It is related to the heat leakages, which appear similarly to windows and often connect few windows (see Fig. 1 and Fig. 2). In such cases two or more windows together with the leakage build one object. To separate the windows morphological operations are used. These operations also help to remove false detections (small objects). Then, for every extracted region, which is a candidate for a window, a rectangular bounding box and its center of gravity is calculated.

2.2 Geometric constrains

For façade reconstruction a pre-knowledge about the regularity of windows can be used. In most buildings the windows build regular grid. The gaps between the columns and between rows are often the same size. However the gravity centers extracted from IR textures in the previous step usually don't build a regular grid of windows. It is related to the fact that the IR images depict distribution of thermal radiation which does not have sharp borders, as well as to the still included thermal leakages which could not be removed through the morphological operations. Hence, in the next step of our method the extracted gravity points are adjusted to regular grid. The gravity centers which are close to each other are merged.

2.3 Standardized masked correlation

To refine the location of windows a standardized masked correlation is applied. Locally around the point (gravity center) the correlation with a priori mask is applied to find the best position of the window.

The idea of masked correlation consists in a priori knowledge about the geometry of a searched object. According to this geometry a binary raster is created and the areas which are not relevant for the matching are labeled as "don't-care-areas".

3. EXPERIMENTS

The method presented in the Section 2 was tested with an exemplary dataset consisting of airborne and terrestrial image sequences of a test area "Technische Universitaet Muenchen".

3.1 Terrestrial image sequences

Current IR cameras cannot reach the optical resolution of video cameras or even digital cameras. The camera used for the acquisition of the test sequences offers an optical resolution of 320x240 pixels with a field of view (FOV) of only 20°. The FLIR SC3000 camera is recording in the thermal infrared (8 - 12 μm). On the top of a van, the camera was mounted on a platform which can be rotated and shifted. Like in the visible spectrum, the sun affects infrared records. In long-wave infrared the sun's influence appears only indirect, as the sun is not sending in the long wave spectrum, but of course is affecting the surface temperature of the building.

Caused by the small field of view and the low optical resolution it was necessary to record the scene in oblique view to be able to record the complete facades of the building from the floor to the roof and an acceptable texture resolution. The image sequences were recorded with a frequency of 50 frames per second. To minimize holes in the textures due to occlusion caused by the oblique view, every façade was recorded with a view forward looking and a view backward looking. The viewing angle related to the along track axis of the van was constant. An example of a recorded sequence is shown in Fig.1. The position of the camera was recorded with GPS and, for quality measurements from tachymeter measurements from ground control points.

3.2 Textures

The terrestrial textures used for the experiment where extracted from terrestrial image sequences applying the method which is described in detail by Hoegner et al. (2007). In first step the image sequences were relatively oriented using 5 point algorithm. Further, the relative path was matched with the GPS path registered during the measurement and 3D building model. Finally, the sequences where projected on the 3D building model and combine into the textures.

Figure 1. Example images from one sequence along a building.

3.3 Segmentation

Segmentation was carried out in a commercial program MVTec HALCON. The texture was scaled down with factor 8. The image was segmented using local dynamic threshold, then closing and opening operation was applied. Further from the segmented area the separate regions were extracted and for every region a rectangular bounding box was created. Finally gravity centers for every bounding box where calculated and taken for further processing.

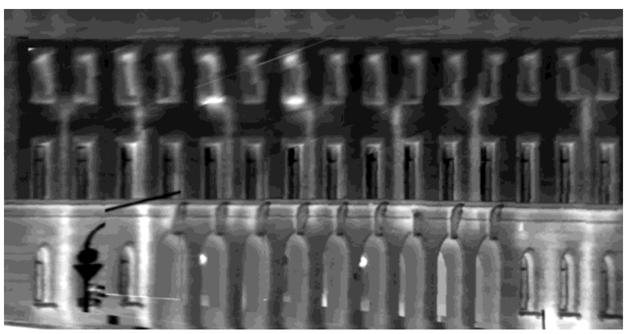


Figure 2. An IR texture extracted from terrestrial data

3.4 Grid and correlation

In our test area there are up to four floors in the buildings. Meanwhile in the ground floor the windows have often untypical form and irregular distribution. Hence adjustment in rows only was possible. We used the pre-knowledge about the number of floors. We sorted the points (gravity centers) into groups. The number of groups was equal to the number of floors. For every group the mean height was calculated and all points of one group were shifted on one level. Finally the precise placing of the window is carried out using a priori mask which is depicted in the Fig. 3. The a priori mask consists of a binary image and so called “don’t-care-area”. The idea of this correlation is that parts of the mask which are not interesting for correlation are masked out. In our case we created a mask with “don’t-care-areas” inside. Thanks to this reflections in windows can excluded from correlation. Also the area around the expected edge of the window was masked out. Thereby, the blurred in IR image edge cannot influence the result.

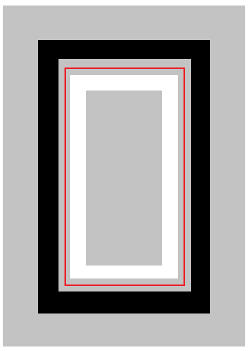


Figure 3. Mask used for correlation: red – expected shape of the window; black & white – binary mask; grey – don’t-care-areas

3.5 Airborne textures

We tested our method also with textures extracted from airborne IR images. Airborne images usually have lower resolution than textures from terrestrial images. Therefore scaling down is often not necessary. Using our test data the windows have size of few pixels. Thus, creating a proper mask and matching are difficult.

Our experience shows that better results can be achieved applying the segmentation and adjustment in rows only.

4. RESULTS

Our first results on segmentation are shown in the Fig. 4. In Fig. 5 we see results on regions and gravity centers extraction, which show that many extracted gravity centers correspond to the approximated position of windows, however in two top rows only. In the lowest row (ground floor) windows could not be extracted. It is related to the irregularity of the windows and entrances in this floor.

a)

b)

Figure 4. a) Extraction of candidates for windows from IR texture using local dynamic threshold; b) separation of the regions and removing false detections using morphological operations

a)

b)

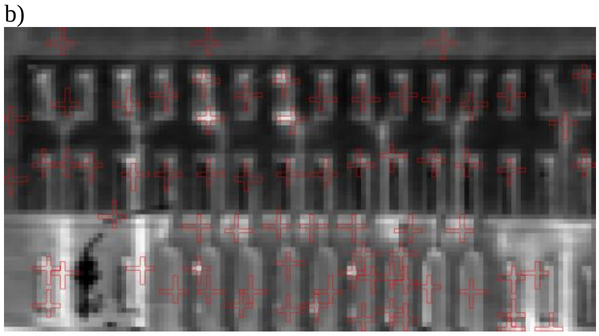


Figure 5. a) Rectangular bounding boxes for the extracted regions; b) gravity centers

In the Fig. 6 results on window extraction after masked correlation are shown. In the two top rows some windows could be detected. In few cases the detection was false. These false detections were strongly related to striking leakages which connect two windows together. The windows from the lowest row were not detected according to low correlation coefficient.

An example of window extraction from airborne textures is presented in Fig.7. One row of windows only could be extracted completely. This is related (1) to the fact that the windows in the top row are smaller than other windows and (2) to occlusion with trees in other rows.

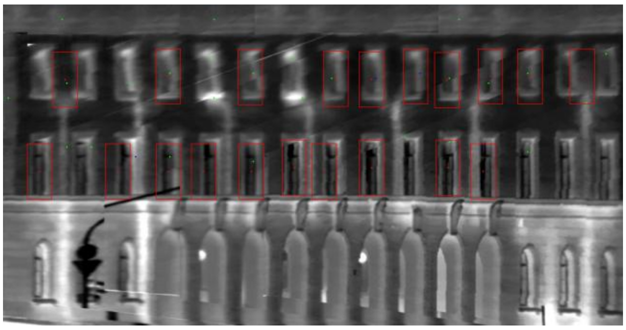
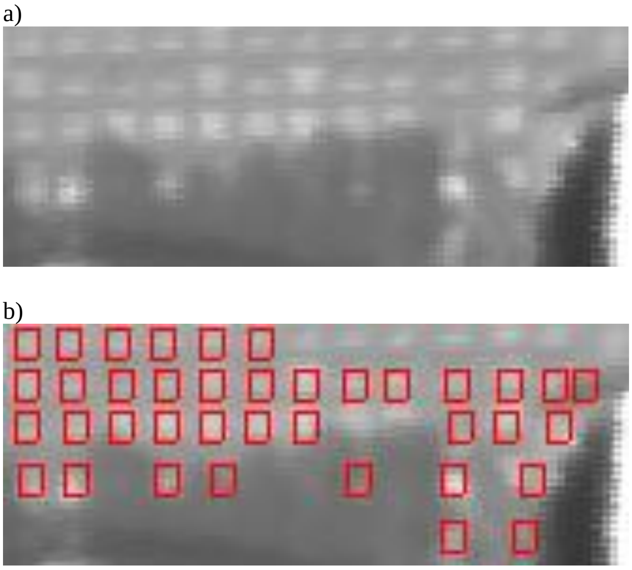


Figure 6. Windows extracted from terrestrial textures

a)



b)

Figure 7. a) An example of a texture extracted from airborne image sequences; b) windows detected in an airborne texture

5. DISCUSSION AND FUTURE WORK

In IR images the glass reflects the temperature of the sky and surrounding. It is the reason why windows should be excluded from thermal inspection of buildings. But also these reflections can be used for window detection. In IR images windows usually strongly differentiate from the rest of the façade. However, every window can appear differently, because of different objects which are reflected. According to this fact we applied a local dynamic threshold to segment the image. Thanks to differently appearing regions (candidates for windows) could be extracted. Thanks to geometric constrains many windows could be found on their correct position. Using masked correlation the precise location of the window in the façade could be corrected.

However, still few windows couldn't be detected or were detected on a wrong position. First of all a reason for this can be the strong bright reflections, for example Fig. 6, 5th and 7th window from right in the top row. In this case the reflections are located on the top and bottom edge. Hence, the area around the horizontal edges is bright and around the vertical edges is dark. But the mask was created for a case that inside the window is bright and outside dark or vice versa. So, for further development of the method other masks should be developed and tested. Furthermore the achieved grid of windows should be completed.

Further problem in window detection are occlusions by trees, traffic signs and lights as well as by other buildings. Thus, in the future should be researched on occlusion free texture extraction.

Finally the combination of terrestrial and airborne textures should be considered. Fusing regions (candidates for windows) extracted from both textures double detection can get a higher probability to be a window.

REFERENCES

Avbelj J, Iwaszczuk D, Stilla U (2010) Matching of 3D wire-frame building models with image features from infrared video sequences taken by helicopters. PCV 2010 - Photogrammetric Computer Vision and Image Analysis. International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 38(3B): 149-154

Becker S (2009) Generation and application of rules for quality dependent facade reconstruction, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 64, Issue 6, November 2009, Pages 640-653

Dick A, Torr P, Cipolla R, Ribarsky W (2004) Modelling and interpretation of architecture from several images. International Journal of Computer Vision 60(2), pp. 111–134

Eugster H., Nebiker S. (2009) Real-time georegistration of video streams from mini or micro UAS using digital 3D city models. Proceedings of 6th International Symposium on Mobile Mapping Technology, Presidente Prudente, São Paulo, Brazil, July 21-24, 2009

Frueh C, Sammon R, Zakhor A (2004) Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery, Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04).

Green P (1995) Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. Biometrika 82, pp. 711–732.

Hoegner L, Kumke H, Meng L, Stilla U (2007) Automatic extraction of textures from infrared image sequences and database integration for 3D building models. PFG Photogrammetrie Fernerkundung Geoinformation. Stuttgart: Schweizerbartsche Verlagsbuchhandlung. 2007(6): 459-468

Hoegner L, Stilla U (2009) Thermal leakage detection on building facades using infrared textures generated by mobile mapping. Joint Urban Remote Sensing Event (JURSE 2009). IEEE

Leibe B, Schiele B (2004) Combined Object Categorization and Segmentation with an Implicit Shape Model. In: ECCV'04 Workshop on Statistical Learning in Computer Vision, pp. 1–15.

Li R (1997) Mobile Mapping - An Emerging Technology for Spatial Data Acquisition. Journal of Photogrammetric Engineering and Remote Sensing, Vol.63, No.9, pp.1085-1092.

Reznik S, Mayer H (2007) Implicit Shape Models, Model Selection, and Plane Sweeping for 3D Facade Interpretation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (3/W49A), S. 173-178

Ripperda N (2008) Grammar Based Facade Reconstruction using RjMCMC. PFG Photogrammetrie Fernerkundung Geoinformation. Stuttgart: Schweizerbartsche Verlagsbuchhandlung. 2008(2): 83-92

Sirmacek B, Hoegner L, Stilla U (2011) Detection of windows and doors from thermal images by grouping geometrical features. In: Stilla U, Gamba P, Juergens C, Maktav D (eds) JURSE 2011 - Joint Urban Remote Sensing Event, 133-136

Stilla U, Kolecki J, Hoegner L (2009) Texture mapping of 3D building models with oblique direct geo-referenced airborne IR image sequences. ISPRS Hannover Workshop 2009: High-resolution earth Imaging for geospatial information. 38(1-4-7/W5)

Werner T, Zisserman A (20029 New Techniques for Automated Architectural Reconstruction from Photographs. In: Seventh European Conference on Computer Vision, Vol. II, pp. 541–555.

Yastikli N, Jacobsen K (2005), Direct sensor orientation for large scale mapping – potentials, problems, solutions. The Photogrammetric Record 20(111), September 2005: 274-284