Co-registration of Time-of-Flight (TOF) camera generated 3d point clouds and thermal infrared images (IR)

LUDWIG HOEGNER [1], MARTIN WEINMANN [2], BORIS JUTZI [2], STEFAN HINZ [2] & UWE STILLA [1]

Abstract: This article presents an investigation on the co-registration of a 3d point cloud and intensity image of a near-infrared (NIR) time-of-flight (TOF) camera system and a thermal infrared (TIR) camera. The TIR-camera is relatively oriented to the TOF-camera. As the radiometric behavior of the TIR-camera and the NIR-image of the TOF-camera is different for most pixels, corners are detected in both images and co-registered. The intensity values of TIR-camera are projected into the TOF point cloud and from there in the TOF-camera image plane. The quality of the projection is evaluated comparing the projected TIR image from the view of the TOF-camera and the TOF NIR image.

Zusammenfassung: Gegenstand dieses Artikels ist die Koregistrierung einer 3d Punktwolke und eines Nah-Infrarot (NIR) Bildes einer Time-of-Flight (TOF) Kamera mit einer thermischen Infrarot (TIR) Kamera. Die TIR Kamera wird relativ zur TOF Kamera orientiert. Da sich die radiometrischen Eigenschaften der TIR Kamera und des NIR Bildes der TOF Kamera für die meisten Pixel unterscheiden, werden Ecken in beiden Bildern detektiert und koregistriert. Die Intensitätswerte der TIR Kamera werden in die Punktwolke der TOF Kamera projiziert und von dort in die Bildebene der TOF Kamera. Die Qualität der Projektion wird durch Überlagerung des in die TOF Kamera projizierten TIR Bildes und des NIR Bildes der TOF Kamera überprüft.

1 Motivation

Deriving an appropriate 3D description of man-made and natural environments is of great interest in Computer Vision, Photogrammetry and Remote Sensing. Most of the current approaches are based on the use of image and/or range data (HARTLEY & ZISSERMAN, 2008). Sets of images and image sequences can be used to simultaneously calculate the relative orientation of the images and generate 3d points clouds (POLLEFEYS et al, 2008; MAYER, 2007). Image-based methods of deriving 3d point clouds allow the observation and reconstruction of moving objects and include color information for the 3d points. These methods are nevertheless limited to textures scenes where corresponding homologues image points can be found. Range measuring sensors like laserscanners overcome this limitation but are in general not capable of capturing moving objects. The combination of laserscanners and a camera allows the generation of a colored 3d point cloud.

Simultaneously capturing intensity information of high quality as well as range information by images with a single measurement, new active time-of-flight (TOF) sensors seem to be well-suited for combining the advantages of range measurement and images (WEINMANN et al., 2011). However, the acquired intensity typically represents information of the visual domain and hence, only radiometric and geometric surface properties of observed objects are captured

1) Photogrammetrie & Fernerkundung,Technische Universität München (TUM)
Arcisstraße 21, 80333 München, www.pf.bv.tum.de
2) Institut für Photogrammetrie und Fernerkundung, Karlsruhe Institute of Technology (KIT)
Englerstr. 7, 76131 Karlsruhe, http://www.ipf.kit.edu

which might not always be sufficient. In contrast to this, infrared (IR) cameras sense thermal radiation in the infrared spectrum which is emitted by objects in the scene and not visible in visual images. Thus, the captured images offer a different look on objects and the extraction of additional information like temperature and material of observed objects. Especially in building observation (IWASZCZUK et al., 2011; HOEGNER et al., 2007), the extraction of textures on facades of buildings allows a reconstruction of the surface temperature and a look into the interior behavior of a wall. Due to different radiometric behavior in different bands, well-known features points like SIFT are not applicable for the co-registration of thermal infrared and visible band images, but a combination of these different bands can enhance features in both kinds of images and even reveal new features that might not be present in either IR or visual images (CHEN & LEUNG, 2009; BAI et al., 2011). The registration of such image data representing information in different spectral bands has for instance been investigated using a segment based approach (COIRAS et al., 2000) or an approach involving normalized mutual information (PARK et al., 2008). This includes that an image transformation is required for mapping the IR texture onto intensity images. For almost planar scenes, the transformation model of a homography was investigated (WEINMANN et al., 2012). Another solution for co-registration is the use of ground control points, that can be identified in both images that are to be matched. In case of a TOF camera, the ground control points are directly related to image coordinates of the TOF camera and with this coordinates directly to 3d points of the TOF point cloud. The alignment of such point clouds is commonly referred to as point cloud registration, where standard techniques such as the Iterative Closest Point (ICP) algorithm (BESL et al., 1992) or Least Squares 3D Surface Matching (LS3D) (GRUEN & AKCA, 2005) only exploit information of the spatial 3D geometry. The ICP algorithm relies on iteratively minimizing the difference between two point clouds which can be very time-consuming for large point clouds and requires a good initial alignment of the respective 3D point clouds. In contrast, the LS3D approach minimizes the distance between matched surfaces which also requires a good initial alignment of the point clouds. If unreliable 3D/3D correspondences might be present in the set of all 3D/3D correspondences, a RANSAC-based filtering (FISCHLER & BOLLES, 1981) scheme can be applied in order to increase the robustness of the estimation (SEO et al., 2005).

In this paper, the matching of both a TOF camera with intensity image and 3d point cloud and a thermal infrared camera with an intensity image is investigated for scenes without the restriction of planarity. At first, a relative orientation of the IR camera to the TOF camera is done using homologues points. From these points, 3d coordinates in a relative coordinate system are calculated. These homologues points are directly related to 3d points of the TOF camera which can now be enriched with an additional thermal intensity value. Additionally, ground control points in an object coordinate system are used for verification of the results.

2 Methodology

The method investigated in this paper deals with two main aspects: the relative orientation of the cameras and the co-registration of the TOF 3d point cloud and the thermal infrared intensity image. The relative orientation is based on 5 parameters, namely the displacement in axis Y and Z for a fixed displacement in X and the 3 angles around the axes. It is assumed that the cameras are already calibrated geometrically (ALBERTZ & WIGGENHAGEN, 2009). These 5 parameters are classically determined using homologues feature points in two images like Förstner or SIFT. Due to the different radiometry of the TOF camera intensity (near infrared) and the thermal camera (thermal infrared), radiometric features cannot be compared with sufficient certainty. Thus, in both images, only the coordinates of features are used. If the relative orientation of the cameras is roughly known, the assignment of the features only from their image coordinates is possible. From this set of homologues points, the relative orientation of the thermal camera to the TOF camera is calculated.

In the second step, the model coordinate system of the intensity images of the two cameras with the origin in the TOF camera has to be transferred into the model coordinate system of the TOF camera 3d point cloud. This is necessary because the relative orientation of the intensity images does not include a scalar factor and additionally refines the orientation of the cameras. This transformation is done by a 3D/3D matching. From the relative orientation, 3d points in the model coordinate system of the camera are calculated for the homologues points. In the TOF camera, these points are connected to a depth value of the point cloud with the same image coordinate. So corresponding 3D coordinates in both model coordinate systems are known.
If reliable 3D/3D correspondences are already known, there is no need to estimate the closest points as should be done before an ICP matching. Further assuming that the 3D/3D correspondences represent physically almost identical 3D points, there is no need to estimate and align the respective object surfaces. Hence, the straight forward approach consists of applying a standard rigid transformation which minimizes the difference between the point clouds in the standard Least Squares sense.
For aligning 3D point clouds in a common coordinate frame, the respective transformation between them has to be estimated. This transformation consists of a rotation described with the rotation matrix **R** and a translation represented as translation vector **t**. The standard rigid transformation minimizes the error between the point clouds by minimizing the energy function

$$E = \sum_i \left\| \mathbf{X}'_i - \left(\mathbf{R}\,\mathbf{X}_i + \mathbf{t} \right) \right\|^2$$

and thus the sum of squared distances between respective 3D points. For estimating such a standard rigid transformation, at least three 3D/3D correspondences are required, but the presence of more correspondences increases the accuracy of the estimated transformation.

The 3d point cloud of the TOF camera is now projected into the thermal infared image to interpolated thermal intensities for every TOF 3d point.

3 Experimental results

3.1 Camera system

For obtaining an infrared-textured 3D model of a scene, thermal information about the local environment has to be captured with a thermal infrared device (InfraTec VarioCAM hr) and the respective 3D information has to be captured with a time-of-flight (ToF) camera (PMD[vision] Camcube 2.0). Furthermore, the used scanning device should also be suited for recording intensity information which is co-registered to the spatial measurements. These devices have been mounted on a sensor platform as shown in Figure 1. Thus, the devices are coupled and a fixed relative orientation between them is preserved.

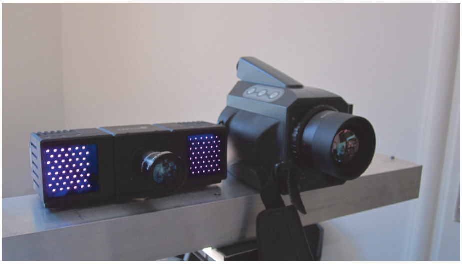

Figure 1: sensor configuration: left: TOF camera, right: thermal infrared camera

A PMD[vision] CamCube 2.0 simultaneously captures geometric information as well as radiometric information and thus various types of data by images with a single shot and a frame-rate of 25 frames per second. In contrast to the classical stereo observation techniques, the monostatic sensor configuration of the PMD[vision] CamCube 2.0 preserves information on a discrete 2D grid without the need of a co-registration of the captured data. For each point on this discrete grid, three features are measured, namely the respective range, the active intensity and the passive intensity. The active intensity depends on the illumination emitted by the sensor, whereas the passive intensity depends on the background illumination arising from the sun or other external light sources. Using the grid information, the different types of data can be represented as images. These images have a size of 204x204 pixels which corresponds to a field of view of 40°x40° and hence, the device provides measurements with an angular resolution of approximately 0.2°. However, the non-ambiguity range which is also called unique range is less than 10 m and depends on the tunable modulation frequency. In order to overcome this range measurement restriction, image- or hardware-based unwrapping procedures have recently been introduced (JUTZI, 2009; JUTZI, 2012).
The used infrared camera is a bolometer-based VarioCAM hr from InfraTec. Its sensor records in the wavelength interval from 7.5 to 14 µm with a radiometric resolution of 0.05 K. The image representation of the captured thermal information has a size of 384x288 pixels and, considering an angular resolution of approximately 0.16°, this corresponds to a field of view of approximately 61° to 46°. As the frame rate is 25 fps, this device can also be applied for observing dynamic scenes. Both cameras have been calibrated geometrically before to obtain their interior orientation and distortion.

3.2 Scenario

The test scenario is an indoor scene with different depth layers of a static scene and moving objects and persons. Both cameras record synchronized images with 25 fps. Figure 2 shows an exemplary image pair of the scene. In the color-coded thermal image of the infrared camera, the cube that is held by the person is rarely visible, but the person can easily be separated from the cold background. The edges and corners of the shelves of the board and the door can be seen. In the TOF passive intensity image, the borders and edges of the shelves and the door also appear clearly, but show different contrast and behavior in their local image neighborhood. In contrast to the thermal image, where all the books show their temperature, in the TOF intensity image, they show almost their visible spectrum intensity.

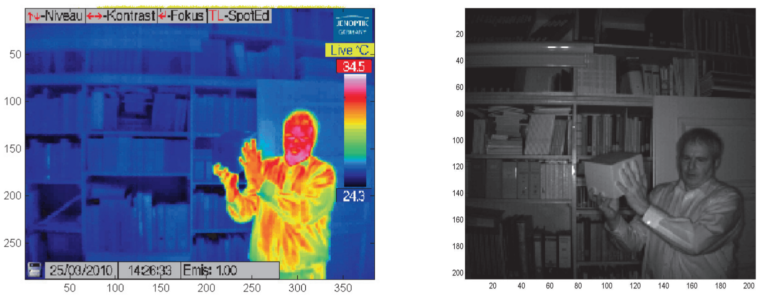

Figure 2: synchronized image pair: left: thermal infrared image with color-coded temperature, right: active intensity image of the TOF camera. Both images before distortion correction.

The distance image of the TOF camera (Fig. 3 left) shows that the cube and the person can easily be separated from the background, but the cube and person can hardly be separated. This is eary to see in the TOF point cloud generated from the depth image (Fig. 3 right). A combination of both the thermal and depth information allows to segment foreground and background (TOF depth image) and cube and person (thermal image).

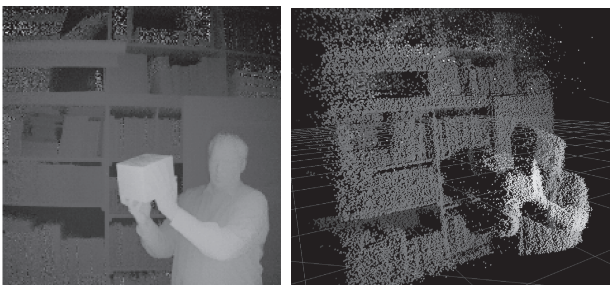

Figure 3: Left: Distance image of the TOF camera. Right: Point cloud generated from the depth image. Bright values have a smaller distance to the sensor than dark values.

3.3 Relative Orientation

Due to the similarity of the observed scene it is assumed, that geometric features like edges and corners appear in both intensity images in the same geometric order. Features are automatically selected from their reliability. This takes into account the position and orientation of the feature in both images, the certainty of the depth value, and the distribution in the images (Fig. 4). Points are found on the shelves, in the shelves and at the door. The cube and the person are ignored here. The person shows no straight lines and the outline of the person has unsure depth values. The cube shows no edges in the thermal image.

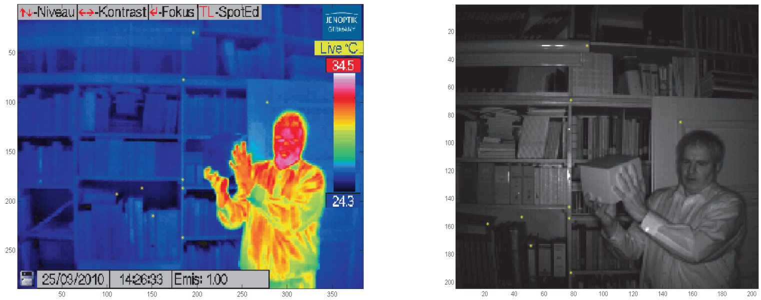

Figure 4: Same image pair as Fig. 2. Corresponding homologues points are extracted for the determination of the relative orientation.

The 5 parameters of the relative orientation are no determined in a bundle adjustment with the TOF camera in the origin and the baseline b_x with length 1 fixed. Using the relative orientation, 3d points for the homologues points are computed and matched with the corresponding TOF depth points to calculated the scale factor of the relative orientation to the coordinate system of the TOF camera. The baseline b_x was manually measured with about 18 cm. After the matching of the relative oriented point cloud onto the TOF point cloud, the calculated baseline is given with 16,6 cm.

In the last step, the TOF 3d points can now be projected into the image plane of the thermal camera to interpolate thermal intensity values for every TOF 3d point (Fig. 5). In the point cloud, there is now a cold background representing the shelves with a big gap, where the shelves are hidden by the person. The person and the cube are the second group of 3d points that is clearly before the shelves. The shape of the person is clearly visible in the 3d point coordinates and the thermal colors of the points which indicate a good fit of the TOF point cloud and the thermal image. In front of the gap in the shelves, there are a few single 3d points with warmer color. These are points at the outline of the hand of the person, where the depth value is unsure. At the right edge of the scene, a few points show a temperature color distribution. This is the projection of the caption of the original infrared image.

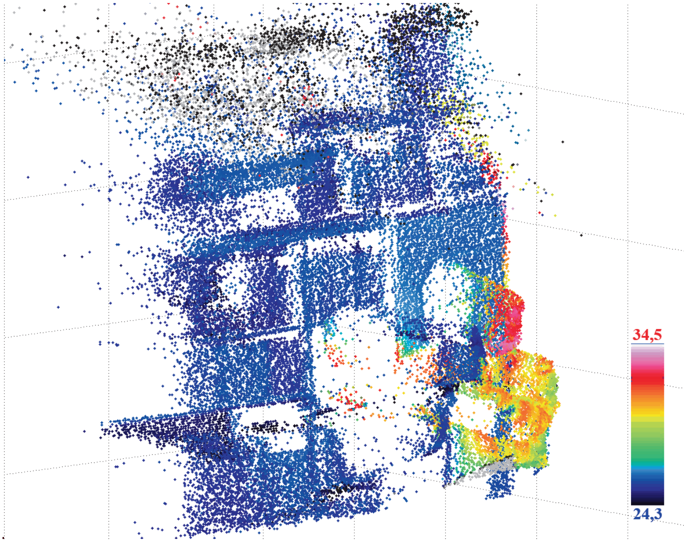

Figure 5: TOF 3d points with thermal intensities.

4 Conclusion

The co-registration of the TOF and thermal IR camera using relative orientation and the TOF 3d point cloud shows more flexibility and accuracy compared to the homography approach (WEINMANN et al., 2012). For scenes with different depth layers of moving objects, the new approach allows a better matching of TOF and thermal images. The approach depends on the possibility to find corresponding, non moving geometric elements in both images. For scenes without structures elements, further investigation are necessary on the extraction of feature pairs in different spectral domains. The experiments so far, are working also on image sequences, but there are so far no investigations on how the sequence itself could be used to improve the orientation. The relative orientation is determined once in one image pair and then used for the hole sequence.

5 Literature

ALBERTZ, J. & WIGGENHAGEN, M., 2009:Taschenbuch der Photogrammetrie und Fernerkundung, Wichmann, Heidelberg. ISBN 978-3-87907-384-9

BAI, X., ZHOU, F. & XUE, B., 2011: Fusion of infrared and visual images through region extraction by using multi scale centersurround top-hat transform. Optics Express 19(9), pp. 8444– 8457.

BESL, P. J. & MCKAY, N. D., 1992: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2), pp. 239-256.

CHEN, S. & LEUNG, H., 2009: An EM-CI based approach to fusion of IR and visual Images. 12th International Conference on Information Fusion, pp. 1325–1330.

COIRAS, E., SANTAMARIA, J. & MIRAVET, C., 2000: A segmentbased registration technique for visual-IR images. Optical Engineering 39, pp. 282–289.

FISCHLER, M. A. & BOLLES, R. C., 1981: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6), pp. 381–395.

GRUEN, A. & AKCA, D., 2005: Least squares 3D surface and curve matching. ISPRS Journal of Photogrammetry and Remote Sensing 59(3), pp. 151-174.

HARTLEY, R. I. & ZISSERMAN, A., 2008: Multiple view geometry in computer vision. University Press, Cambridge.

HOEGNER, L., KUMKE, H., MENG, L. & STILLA, U., 2007a: Automatic extraction of textures from infrared image sequences and database integration for 3D building models. PFG Photogrammetrie, Fernerkundung, Geoinformation 6 / 2007, pp. 459–468.

IWASZCZUK, D., HOEGNER, L. & STILLA, U., 2011: Detection of windows in IR building textures using masked correlation. In: Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B., Butenuth, M. (Eds.), Photogrammetric Image Analysis, ISPRS Conference - Proceedings. Lecture Notes in Computer Science, Vol. 6952, Springer, pp. 133–146.

JUTZI, B., 2009: Investigations on ambiguity unwrapping of range images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3 / W8), pp. 265–270.

JUTZI, B., 2012: Extending the range measurement capabilities of modulated range imaging devices by time-frequency multiplexing. AVN - Allgemeine Vermessungs-Nachrichten 2 / 2012.

MAYER, H., 2007: 3D Reconstruction and Visualization of Urban Scenes from Uncalibrated Wide-Baseline Image Sequences. In: Photogrammetrie – Fernerkundung – Geoinformation 2007(3), pp. 167–176.

PARK, C., BAE, K.-H., CHOI, S. & JUNG, J.-H., 2008: Image fusion in infrared image and visual image using normalized mutual information. Proceedings of SPIE, Vol. 6968, pp. 69681Q–1–9.

POLLEFEYS, M., NISTÉR, D., FRAHM, J. M., AKBARZADEH, A., MORDOHAI, P., CLIPP, B., ENGELS, C., GALLUP, D., KIM, S. J., MERRELL, P., SALMI, C., SINHA, S., TALTON, B., WANG, L., YANG, Q., STEWÉNIUS, H., YANG, R., WELCH, G. & TOWLES, H., 2008: Detailed real-time urban 3D reconstruction from video. In: International Journal of Computer Vision (IJCV), 78(2-3), 143-167

SEO, J. K., SHARP, G. C. & LEE, S. W., 2005: Range data registration using photometric features. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2, 1140-1145.

WEINMANN, M., WURSTHORN, S. & JUTZI, B., 2011: Semi-automatic image-based co-registration of range imaging data with different characteristics. PIA11 - Photogrammetric Image Analysis. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science. 38(3) / W22, S. 119-124.

WEINMANN, M., HOEGNER, L., LEITLOFF, J., STILLA, U., HINZ, S. & JUTZI, B., 2012: Fusing Passive and Active Sensed Images to Gain Infrared-textured 3D Models. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 39(B1), 71-76

Publikationen der Deutschen
Gesellschaft für Photogrammetrie,
Fernerkundung und Geoinformation e.V.

DGPF

Band 22 **2013**

Vorträge

felis

33. Wissenschaftlich-Technische Jahrestagung der DGPF

27. Februar – 1. März 2013
in Freiburg i. B.

Dreiländertagung D - A - CH

ISSN 0942-2870

Eckhardt Seyfert, Hrsg.