A Combined Approach for Estimating Patchlets from PMD Depth Images and Stereo Intensity Images

Christian Beder, Bogumil Bartczak and Reinhard Koch

Computer Science Department
Universiy of Kiel, Germany
{beder,bartczak,rk}@mip.informatik.uni-kiel.de

Abstract. Real-time active 3D range cameras based on time-of-flight technology using the Photonic Mixer Device (PMD) can be considered as a complementary technique for stereo-vision based depth estimation. Since those systems directly yield 3D measurements, they can also be used for initializing vision based approaches, especially in highly dynamic environments. Fusion of PMD depth images with passive intensity-based stereo is a promising approach for obtaining reliable surface reconstructions even in weakly textured surface regions.
In this work a PMD-stereo fusion algorithm for the estimation of patchlets from a combined PMD-stereo camera rig will be presented. As patchlet we define an oriented small planar 3d patch with associated surface normal. Least-squares estimation schemes for estimating patchlets from PMD range images as well as from a pair of stereo images are derived. It is shown, how those two approaches can be fused into one single estimation, that yields results even if either of the two single approaches fails.

1 Introduction

Vision-based passive stereo systems [1] and active systems such as structured light or laser scanners [2] are complementary methods to measure 3D depth of a scene. However, the application domain of all this systems is restricted. For instance the algorithmic complexity of stereo systems is quite high and they are not applicable in case of weakly textured surfaces. Laser scanners and structured light approaches on the other hand cannot cope with moving objects, because capturing is not instantaneous for this systems.

A new promising development for the area of surface reconstruction, that is able to cope with those caveats of the existing techniques, is the Photonic Mixer Device (PMD), which measures distances directly for a two dimensional field of pixels based on the time of flight of incoherent, modulated infrared light. Recently PMD cameras have been developed that are capable of capturing reliable depth images directly in real-time. Those cameras are compact and affordable, which makes them attractive for versatile applications including surveillance and

computer vision [3]. The successful application of this technology in Structure from Motion [4], motion capturing [5] and face tracking[6] have been demonstrated.

In [7] the complementary nature of stereo vision based systems and PMD cameras is discussed qualitatively and a simple method for fusing the information gathered from those two system is proposed. Yet neither a quantitave comparison of both systems nor a statistically optimal fusion is done. Recently, a systematic and quantitative comparison of both approaches was investigated in [8]. The evaluation showed that a combination of both approaches, either by initialisation of stereo with PMD, or by fusion of both methods, could prove beneficial. However, no systematic analysis of a statistically optimal fusion of both modalities exists.

The main contribution of this work is therefere the development of a statistically optimal fusion of both systems based on the estimation of patchlets [9], i.e. small planar surface patches with an associated surface normal, being a very useful surface representation for tasks such as segmentation [10] and visualization [11].

First a short introduction to the technology underlying the PMD image formation will be given in section 2 in order to provide some background information.

Then in section 3 two least-squares estimation schemes for the PMD images as well as for the stereo images will be presented and it will be shown, how optimal estimates for the patchlets together with their covariance matrices can be obtained. While the PMD camera provides direct geometric measurements, which are used in the estimation of the patchlet, the stereo matching is based on estimating a local homography between the images [12–14], which optimally aligns the image intensities between the stereo image pair [15–17]. Because both estimation schemes have the same structure, a fused estimation using input data from both sources is possible.

In section 4 a brief quantitative analysis of the three approaches on synthetic images for different noise levels will be presented and some surface reconstructions from real images will be shown.

2 The PMD-Camera

We will first give some background information on the Photonic Mixer Device (PMD), which is a semiconductor structure based on CCD- or CMOS-technology [18]. Integrated in an image sensor array it is capable of modulating the current that is generated by received light intensities in every single pixel. This can be utilized to build affordable cameras that are able to measure depth with high precision. One such camera is shown in figure 3. It mainly consists of a camera with the PMD sensor and light emitter arrays that are used to send out modulated light. The light is reflected by 3D scene points and received by the PMD image sensor.

The depth measurement performed with a detector using the Photonic Mixer Device is based on the time of flight principle. Different approaches for measuring the time-of-flight with light exist [19]. One method suitable for the use with the PMD is to modulate the emitted light intensity with a periodic pattern. Depending on the "time of flight" a phaseshift of the periodic pattern is observable. The PMD-Camera is able to extract this phase shift in every pixel.

Though different intensity modulations using square waves or pseudo noise coding are possible, the use of a sinusoidal signal is technically well realizable [20]. Generally an intensity wave $I(x,t) = I_0 + I_A \cos(2\pi\nu_m(t + \frac{x}{c}) + \varphi_0)$ with modulation frequency ν_m, propagation speed c and initial phase φ_0 is sent out. At two points x_0 and x_1 of the wave the phase shift

$$\Delta\varphi = 2\pi\nu_m(x_1 - x_0)/c \tag{1}$$

is observable. Extracting this shift from the wave therefore delivers the distance between x_0 and x_1. Due to the repetitive nature of the wave the non-ambiguous wavelength of the measurement is $\lambda_{max} = c/\nu_m$. The effective measurement range is $\frac{\lambda_{max}}{2}$ because light wave has to return to its source to be detected. Typically a modulation frequency of $\nu_m = 20$MHz is used which gives the camera 7.5 meters of unambiguouse depth range. Reflections from distances beyond this range might cause measurement errors due to phase wrapping, however usually the reflected light intensity is too small to cause such errors. For very small distances below 2m, the reflected strong light intensity might cause nonlinear saturation effects which might limit the accuracy and cause bias [3]. Therefore, the usable range was chosen between $2 - 7.5$ meters.

The phase difference is measured by cross correlation between the sent and received modulated signal by the PMD chip. Since the resolution of the phase difference measurement is independent from distance, the achievable depth resolution is independent from scene depth. This is in contrast to stereo triangulation where depth accuracy is proportional to inverse depth.

After taking depth calibration and lens distortions [21, 3] into account, the model of a central perspective projection for the geometric description of the PMD-Camera measurements can be utilized.

3 Estimation of patchlets

In the following two patchlet estimation schemes for PMD images as well as stereo images based on the Gauss-Markoff-Model [22] will be presented. Because the structure of both estimations is identical, both approaches can easily be fused, which will be shown in section 3.3.

3.1 Estimation of patchlets from PMD images

The PMD-camera determines for each ray direction corresponding to pixel $\mathbf{x}$ its distance λ to the optical center. If the camera geometry is given by a projection matrix [23, p.141f] as

$$\mathsf{P}_3 = \mathsf{K}_3(\mathsf{R}_3|\boldsymbol{t}_3) \tag{2}$$

then the corresponding 3d point is obtained directly from the distance λ as

$$\boldsymbol{X} = \frac{\lambda R_3^T \mathsf{K}_3^{-1}\mathbf{x}}{\sqrt{\mathbf{x}^T \mathsf{K}_3^{-T}\mathsf{K}_3^{-1}\mathbf{x}}} - R_3^T \boldsymbol{t}_3 \tag{3}$$

This 3d point lies on the plane $(\boldsymbol{n}^T, d)^T$, if

$$\boldsymbol{X}^T \boldsymbol{n} + d = 0 \tag{4}$$

or equivalently

$$\left(\sqrt{\mathbf{x}^T \mathsf{K}_3^{-T}\mathsf{K}_3^{-1}\mathbf{x}}\boldsymbol{t}_3^T R_3 - \lambda \mathbf{x}^T \mathsf{K}_3^{-T} R_3\right)\boldsymbol{n}' = \sqrt{\mathbf{x}^T \mathsf{K}_3^{-T}\mathsf{K}_3^{-1}\mathbf{x}} \tag{5}$$

using the substitution $\boldsymbol{n}' = \frac{\boldsymbol{n}}{d}$ to parameterize the plane. This expression is linear in the plane parameters $\boldsymbol{n}'$ so that initial values are easily computed. Solving this expression for the unknown depth λ yields

$$\lambda = \sqrt{\mathbf{x}^T \mathsf{K}_3^{-T}\mathsf{K}_3^{-1}\mathbf{x}}\frac{\boldsymbol{t}_3^T R_3 \boldsymbol{n}' - 1}{\mathbf{x}^T \mathsf{K}_3^{-T} R_3 \boldsymbol{n}'} \tag{6}$$

Now using the Jacobian

$$\boldsymbol{a}^T(\mathbf{x}, \lambda, \boldsymbol{n}_0') = \frac{\partial \lambda}{\partial \boldsymbol{n}'} \tag{7}$$

$$= \frac{\sqrt{\mathbf{x}^T \mathsf{K}_3^{-T}\mathsf{K}_3^{-1}\mathbf{x}}}{\left(\mathbf{x}^T \mathsf{K}_3^{-T} R_3 \boldsymbol{n}'\right)^2}\left(\mathbf{x}^T \mathsf{K}_3^{-T} R_3 \boldsymbol{n}' \boldsymbol{t}_3^T R_3 - \left(\boldsymbol{t}_3^T R_3 \boldsymbol{n}' - 1\right)\mathbf{x}^T \mathsf{K}_3^{-T} R_3\right) \tag{8}$$

the Taylor expansion of this expression then yields for every point on the plane

$$\underbrace{\boldsymbol{a}^T(\mathbf{x}, \lambda, \boldsymbol{n}_0')}_{\boldsymbol{a}_i^T}\underbrace{(\boldsymbol{n}' - \boldsymbol{n}_0')}_{\Delta \boldsymbol{n}'} \approx \underbrace{\lambda - \sqrt{\mathbf{x}^T \mathsf{K}_3^{-T}\mathsf{K}_3^{-1}\mathbf{x}}\frac{\boldsymbol{t}_3^T R_3 \boldsymbol{n}_0' - 1}{\mathbf{x}^T \mathsf{K}_3^{-T} R_3 \boldsymbol{n}_0'}}_{\Delta l_i} \tag{9}$$

In the following section a similar expression will be derived for the stereo system.

3.2 Estimation of patchlets from stereo images

Given a calibrated stereo system with the first camera being

$$\mathsf{P}_1 = \mathsf{K}_1(I_3|\mathbf{0}_3) \tag{10}$$

and the second camera being

$$\mathsf{P}_2 = \mathsf{K}_2(R_2|\boldsymbol{t}_2) \tag{11}$$

with the known rotation matrix R and translation vector t, the homography relating points on the plane $(\boldsymbol{n}^T, d)^T$ from the first into the second camera is given by [23, p.314]

$$\mathsf{H} = \mathsf{K}_2 \left(R_2 - \boldsymbol{t}_2 \frac{\boldsymbol{n}^T}{d} \right) \mathsf{K}_1^{-1} \tag{12}$$

which is linear in the vector $\boldsymbol{n}' = \frac{\boldsymbol{n}}{d}$. Hence, points on the plane are transformed according to

$$\mathbf{x}_2 = \mathsf{H}(\boldsymbol{n}')\mathbf{x}_1 = \mathsf{K}_2 R_2 \mathsf{K}_1^{-1}\mathbf{x}_1 - ((\mathbf{x}_1^T \mathsf{K}_1^{-T}) \otimes (\mathsf{K}_2 \boldsymbol{t}_2))\boldsymbol{n}' \tag{13}$$

We now assume, that the grey value of corresponding points is equal in the two images. Using the Euclidean normalization function

$$\boldsymbol{h}(\mathbf{x}) = \boldsymbol{h}\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \frac{1}{w}\begin{pmatrix} u \\ v \end{pmatrix} \tag{14}$$

this is expressed in terms of the two images I_1 and I_2 as

$$I_1(\boldsymbol{h}(\mathbf{x}_1)) = I_2(\boldsymbol{h}(\mathbf{x}_2)) \tag{15}$$

or substituting equation (13)

$$I_1(\boldsymbol{h}(\mathbf{x}_1)) = I_2(\boldsymbol{h}(\mathsf{K}_2 R_2 \mathsf{K}_1^{-1}\mathbf{x}_1 - ((\mathbf{x}_1^T \mathsf{K}_1^{-T}) \otimes (\mathsf{K}_2 \boldsymbol{t}_2))\boldsymbol{n}')) \tag{16}$$

Applying chain rule, the partial derivatives of this expression are given by

$$\boldsymbol{b}^T(\mathbf{x}_1, \boldsymbol{n}') = \frac{\partial}{\partial \boldsymbol{n}'} I_2(\boldsymbol{h}(\mathsf{K}_2 R \mathsf{K}_1^{-1}\mathbf{x}_1 - ((\mathbf{x}_1^T \mathsf{K}_1^{-T}) \otimes (\mathsf{K}_2 \boldsymbol{t}))\boldsymbol{n}')) \tag{17}$$

$$= -(\nabla I_2)(\boldsymbol{h}(\mathsf{H}(\boldsymbol{n}')\mathbf{x}_1))\mathsf{J}(\mathsf{H}(\boldsymbol{n}')\mathbf{x}_1)\mathsf{K}_2 \boldsymbol{t}_2 \mathbf{x}_1^T \mathsf{K}_1^{-T} \tag{18}$$

with the Jacobian of the normalization function being

$$\mathsf{J} = \frac{\partial}{\partial \mathbf{x}} \boldsymbol{h} = \begin{pmatrix} \frac{1}{w} & 0 & -\frac{u}{w^2} \\ 0 & \frac{1}{w} & -\frac{v}{w^2} \end{pmatrix} \tag{19}$$

Hence, the Taylor expansion of equation (16) yields for every point on the plane

$$\underbrace{\boldsymbol{b}^T(\mathbf{x}_1, \boldsymbol{n}'_0)}_{\boldsymbol{b}_j^T} \underbrace{(\boldsymbol{n}' - \boldsymbol{n}'_0)}_{\Delta \boldsymbol{n}'} \approx \underbrace{I_1(\boldsymbol{h}(\mathbf{x}_1)) - I_2(\boldsymbol{h}(\mathsf{H}(\boldsymbol{n}'_0)\mathbf{x}_1))}_{\Delta m_j} \tag{20}$$

This equation is compatible with equation (9), so that a fused best linear unbiased estimation is possible as shown in the next section.

3.3 The fused estimation

Using a window containing N points from the PMD depth image and M points from the stereo intensity images on the patchlet, the plane parameter updates may be estimated iteratively as [22]

$$\widehat{\Delta \boldsymbol{n}'} = \left(\sum_{i=1}^{N} \sigma_{l_i}^{-2} \boldsymbol{a}_i \boldsymbol{a}_i^T + \sum_{j=1}^{M} \sigma_{m_j}^{-2} \boldsymbol{b}_j \boldsymbol{b}_j^T \right)^{-1} \left(\sum_{i=1}^{N} \sigma_{l_i}^{-2} \boldsymbol{a}_i \Delta l_i + \sum_{j=1}^{M} \sigma_{m_j}^{-2} \boldsymbol{b}_j \Delta m_j \right) \tag{21}$$

having the expected covariance matrix

$$\boldsymbol{\Sigma}_{\widehat{\boldsymbol{n}'}\widehat{\boldsymbol{n}'}} = \left(\sum_{i=1}^{N} \sigma_{l_i}^{-2} \boldsymbol{a}_i \boldsymbol{a}_i^T + \sum_{j=1}^{M} \sigma_{m_j}^{-2} \boldsymbol{b}_j \boldsymbol{b}_j^T \right)^{-1} \tag{22}$$

where $\sigma_{m_j}^2$ is twice the variance of the image noise for stereo pixels and $\sigma_{l_i}^2$ is the variance of the distance uncertainty of the PMD camera. Note, that those quantities need only be specified up to scale so that only the relative weighting between the stereo system and the PMD system is required.

To specify this relative weighting the variance factor can be estimated from the residuals as

$$\widehat{\sigma}_0^2 = \frac{1}{N + M - 3} \left(\sum_{i=1}^{N} \sigma_{l_i}^{-2} (\Delta l_i - \boldsymbol{a}_i^T \Delta \boldsymbol{n}')^2 + \sum_{j=1}^{M} \sigma_{m_j}^{-2} (\Delta m_j - \boldsymbol{b}_j^T \Delta \boldsymbol{n}')^2 \right) \tag{23}$$

By looking at the variance factors resulting from estimations with $N = 0$ and $M = 0$ respectively the relative weights can be determined.

Putting everything together we obtain three different patchlet estimation algorithms, namely using only the depth images, using the depth image for initialization (cf. equation (5)) and the stereo images for the estimation and finally a fused approach using all available data. In the following section those three alternatives will be compared.

4 Results

For evaluating the performance of the fused estimation scheme a stereo-rig was used, which is shown in figure 3. It consists of two Sony cameras which deliver images with a resolution of 1024×768 pixels and a field of view of $40° \times 32°$. The PMD camera in the middle is PMDtech's model 3K-s with a resolution of 64×48 pixels over a viewing angle of $22° \times 17°$. Hence, we used windows of 20×20 pixels in the intensity images and windows of 3×3 pixels in the depth images in order to cover approximately the same viewing angle with both systems. The rig was calibrated using a calibration pattern so that the internal parameters of each of the three camera as well as the relative poses of all cameras with respect to the

left stereo camera are known. The stereo system had a baseline of approximately 30 cm and the orientation was close to standard stereo geometry.

We started by generating synthetic data using the calibration parameters of the real rig to produce three images of a well-textured plane $3m$ in front of the camera. We then added white noise of different standard deviation to the three images and estimated 100 patchlets on the surface. The distance of the estimated patchlets from the ground truth as well as the expected standard deviation of the distance is plotted against the added image noise in figure 1. The expected accuracy plotted on the left hand side is worst for the approach depending solely on the PMD images due to the low resolution of the depth image. The accuracy is better for the estimation initialized with the depth image and optimized using the intensity images. Best results are expected for the completely fused approach. However, as the patchlets are estimated at equally distant sample-points rather than only at positions of high intensity gradient, the mean distance from the ground truth is worse for the texture-dependent intensity based estimation than for the other two approaches and the variation is high over the set of patchlets. On the average the purely depth based approach performed best and the fused approach, being an average over both, lies somewhere in between. This is depicted on the right hand side of figure 1.

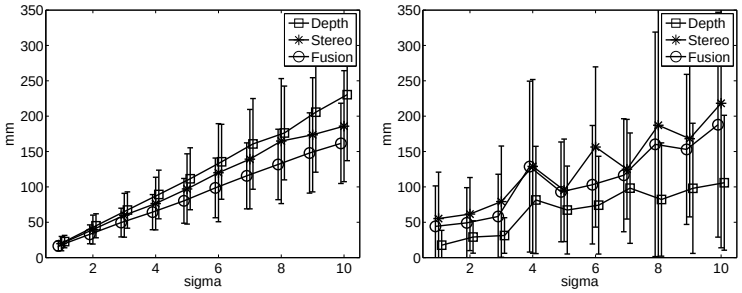

Fig. 1. *Left:* Expected standard deviation of estimated patchlet distance against image noise standard deviation. *Right:* Average distance of the estimated patchlets from ground truth against image noise standard deviation.

Figure 2 shows the same analysis for the normal angle. On the right hand side the angular difference to the ground truth is plotted against the image noise while on the left hand side the expected accuracies are plotted. Here the purely depth map based method is inferior to the intensity based estimation while the fused approach yields the best results.

Now we will present some results on real data. We used the rig depicted on the left hand side of figure 3 and took a picture of the scene shown on the right hand side. The resulting patchlets are depicted in figure 4. We removed patchlets, where the angular accuracy was below a common threshold of 10° in order to demonstrate the capabilities of the different methods. On the left hand side of

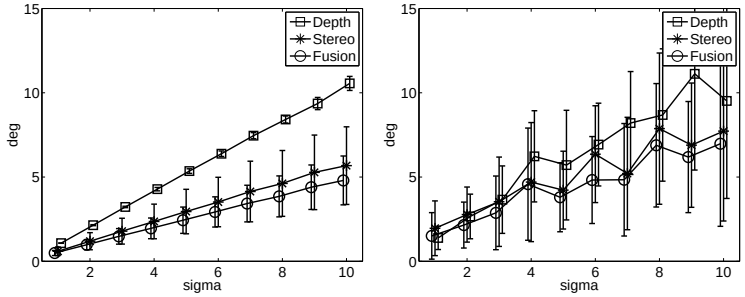

Fig. 2. *Left:* Expected standard deviation of estimated patchlet angle against image noise standard deviation. *Right:* Average angular distance of the estimated patchlets from ground truth against image noise standard deviation.

figure 4 the remaining patchlets for the purely depth image based estimation are shown. It can be seen, that the angular accuracy of the patchlet estimation increases with distance for the depth image, because the surface covered increases and the distance measurement accuracy is equal over the whole image. Further observe, that the dark regions are measured slightly off the plane. The middle picture shows the remaining patchlets for the intensity based estimation with depth initialization. As expected only the textured regions of the image yield good patchlets. Finally the fused estimate yields the patchlets depicted on the right hand side of figure 4. As expected, the best accuracy is achievable using the statistically optimal fused method.

Fig. 3. *Left:* The Rig used to obtain the results. It consists of two color cameras that frame the PMD-Camera. *Right:* One of the stereo images.

5 Conclusion

We have presented a fused estimation scheme for patchlet based surface reconstruction from stereo images as well as PMD range images. It has been shown, how the two systems can be integrated yielding more accurate surface reconstructions than either system alone.

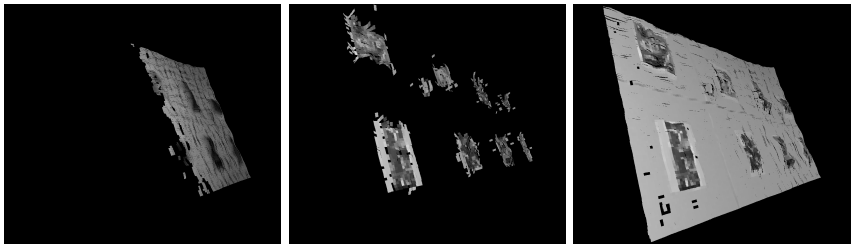

Fig. 4. *Left:* Patchlets from purely depth based estimation. *Middle:* Patchlets from stereo based estimation using depth initialization. *Right:* Patchlets from fused estimation.

The PMD camera yields an accurate direct distance measurement for each pixel and is therefore required for initializing each of the three proposed algorithms. However, the low resolution of current PMD cameras is the major factor limiting the stand-alone applicability of such a system.

Stereo intensity based systems on the other hand have a much higher resolution but their depth accuracy is depending on texture and object distance. Furthermore some initialization is required for those systems to work robustly.

Hence, both systems can be considered complementary in terms of resolution, depth accuracy and scene coverage. The proposed fusion of both approaches therefore constitutes a method for obtaining accurate and robust scene reconstructions including surface normals using a camera rig such as the one depicted in figure 3.

Acknowledgements

The PMD camera used in the experiments is courtesy of Alexander Prusak and Hubert Roth, University of Siegen, Germany.

This work was supported by the German Research Foundation (DFG), KO-2044/3-1.

References

1. Scharstein, D., Szeliski, R., Zabih, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms (2001)
2. Hoppe, H., DeRose, T., Duchamp, T., Halstead, M., Jin, H., McDonald, J., Schweitzer, J., Stuetzle, W.: Piecewise smooth surface reconstruction. Computer Graphics **28**(Annual Conference Series) (1994) 295–302
3. Lindner, M., Kolb, A.: Lateral and depth calibration of pmd-distance sensors. In: International Symposium on Visual Computing (ISVC06). Volume 2., Springer (2006) 524–533
4. Streckel, B., Bartczak, B., Koch, R., Kolb, A.: Supporting structure from motion with a 3d-range-camera. In: Scandinavian Conference on Image Analysis (SCIA07). (June 2007)

5. Grest, D., Koch, R.: Single view motion tracking by depth and silhouette information. In: Scandinavian Conference on Image Analysis (SCIA07). (2007)
6. Gokturk, S., Tomasi, C.: 3d head tracking based on recognition and interpolation using a time-of-flight depth sensor. In: Proc. CVPR. (2004) 211–217
7. Kuhnert, K., Stommel, M.: Fusion of stereo-camera and pmd-camera data for real-time suited precise 3d environment reconstruction. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). (October 2006)
8. for review., B.: Blank for review.
9. Murray, D.R.: Patchlets: a method of interpreting correlation stereo 3D data. PhD thesis, The Univeristy of British Columbia, Vancouver, Canada (2004)
10. Murray, D., Little, J.J.: Segmenting correlation stereo range images using surface elements. In: 3DPVT '04: Proceedings of the 3D Data Processing, Visualization, and Transmission, 2nd International Symposium on (3DPVT'04), Washington, DC, USA, IEEE Computer Society (2004) 656–663
11. Szeliski, R., Tonnesen, D.: Surface modeling with oriented particle systems. Computer Graphics **26**(2) (1992) 185–194
12. Molton, N.D., Davison, A.J., Reid, I.D.: Locally planar patch features for real-time structure from motion. In: Proc. British Machine Vision Conference, BMVC (September 2004)
13. Pietzsch, T., Grossmann, A.: A method of estimating oriented surface elements from stereo images. In: Proc. British Machine Vision Conference. (2005)
14. Pritchett, P., Zisserman, A.: Wide baseline stereo matching. In: Proceedings of the International Conference on Computer Vision. (1998) 754–760
15. Baker, S., Matthews, I.: Lucas-kanade 20 years on: A unifying framework: Part (2002)
16. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: IJCAI'81. (1981) 674–679
17. Triggs, B.: Detecting keypoints with stable position, orientation, and scale under illumination changes. In: Proceedings of European Conference on Computer Vision. (2004) 100–113
18. Xu, Z., Schwarte, R., Heinol, H., Buxbaum, B., Ringbeck, T.: Smart pixels - photonic mixer device (pmd). In: Mechatronics and Machine Vision in Practice. (1998) 259 – 264
19. Lange, R., Seitz, P., Biber, A., Schwarte, R.: Time-of-flight range imaging with a custom solid state image sensor. Proc. SPIE Vol. 3823, p. 180-191, Laser Metrology and Inspection, Hans J. Tiziani; Pramod K. Rastogi; Eds. **1999**(Tiziani, H. J. and Rastogi, P. K.) (sep)
20. Zhang, Z.: Untersuchung und Charakterisierung von Photomischdetektor-Strukturen und ihren Grundschaltungen. PhD thesis, Department of Electrical Engineering And Computer Science at Univeristy of Siegen (December 2003)
21. Kahlmann, T., Remondino, F., Ingensand, H.: Calibration for increased accuracy of the range imaging camera swissrangertm. In: IEVM06. (2006)
22. Förstner, W., Wrobel, B.: Mathematical concepts in photogrammetry. In J.C.McGlone, E.M.Mikhail, J.Bethel, eds.: Manual of Photogrammetry. ASPRS (2004) 15–180
23. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN: 0521623049 (2000)