21^{st} *Computer Vision Winter Workshop*
Luka Čehovin, Rok Mandeljc, Vitomir Štruc (eds.)
Rimske Toplice, Slovenia, February 3–5, 2016

A Novel Concept for Smart Camera Image Stitching

Majid Banaeyan* , Hanna Huber* , Walter G. Kropatsch* and Raphael Barth+
Vienna University of Technology
*Pattern Recognition and Image Processing group
{majid,hanna,krw}@prip.tuwien.ac.at
+Indiecam
raphael@indiecam.com

Abstract. *As panoramic images are widely used in many applications, efficient image stitching methods that provide visually pleasant image mosaics are needed. In this paper we discuss a novel concept for smart camera image stitching based on graph pyramids. For a multi-camera system, the images have to be aligned accordingly to create an image mosaic. Instead of calculating the corresponding transformations centrally, we aim at enabling each camera to individually calculate the transformation of the image it takes. Graph pyramids used for image segmentation provide information about the segmentation process. We analyze how this information can be used to calculate the transformations for image alignment.*

1. Introduction

Panoramic views form the basis of many applications including augmented reality applications. Producing video content with high quality seamless and artefact-free 360° of coverage is challenging per se and even more challenging if all related processing, especially seamless stitching, has to work automatically and in real-time for live productions.

A suitable approach has to solve a system conflict between omnidirectional simultaneous video capture on one hand, which cannot be done from the nodal point due to mechanical collision problems, and parallax-free stitching of panoramas without any parallax, ghosting and distortion artefacts on the other hand.

Using high on-board computing power of smart cameras and a dedicated communication network between cameras could be used to integrate the entire image processing for automatic real-time stitching into the cameras themselves, avoiding the need for further peripheral processing devices.

Common image stitching techniques take images taken from different views and align them using image registration in overlapping regions. So far, all images are collected and aligned centrally, which suffers from high computational cost. Thus, we aim at parallelizing parts of this process by developping smart cameras that are able to perform some of the image transformations themselves.

The camera systems we consider use fish-eye lenses. General camera models such as the pinhole model cannot be applied to these lenses, because they do not conform to the perspective projection due to their large field-of-view. Simple models are given for different projections of ideal fish-eye lenses. They provide a formula for the radius r which is the distance between an image point and the principal point. The principal point is the point where the optical axis intersects the image plane. In case of the equidistant projection the radius is given by

$$r = f\theta, \tag{1}$$

where f is the focal length and θ is the the incident angle of the ray from an object point. However, this formula does not reflect the behavior of real lenses. Instead, extended models are developped which take into account the high level of distortion. Parameter values are estimated using calibration, defining a final model for a particular camera [16]. In fish-eye lenses, both radial as well as tangential distortion is present. While radial distortion reduces the spatial resolution towards the periphery of an image and distorts rectilinear objects, geometric shifts are the result of tangential distortion [15].

In this paper, we define our concept of smart camera image stitching and present ideas how to realize it. We will first give an overview of related work in the fields of fish-eye lenses, image stitching and smart cameras in section 2. After declaring our functional goal in section 3, we discuss open problems that we aim to solve in order to realize it and present our ideas for possible solutions including novel approaches in section 4. Finally, in section 5, we conclude our paper.

2. Related Work

In this section we present a selection of state-of-the-art techniques in fields that are related to smart camera image stitching.

2.1. Smart Cameras

The name of smart camera goes back to the middle of 1970s [29] when Ron Schneidermann applied it in developing systems for controlling the shutter. Then in 1981 the optical mouse was invented by Richard Lyon [24, 25] which was the first realized smart camera including an imaging device and embedded processing unit as a compact system. "Smart camera is a label which refers to cameras that have the ability to not only take pictures but also more importantly make sense of what is happening in the image." [4, Chapter 2, page 21] Smart cameras employ various concepts of computer vision and machine vision which can extract useful information from images resulting in special decisions based on that information. Smart cameras can be classified into three main categories including integrated, compact-system and distributed smart cameras [4, chapter 2]. Integrated smart cameras can be further subdivided into three types including single-chip [2, 11], embedded [20] and stand-alone smart cameras. Distributed smart cameras involve some sort of networking and have recently attracted significant interest in academic and industries fields [28]. Indeed, some problems such as depth information in foreground detection and occlusion are difficult to be solved by single smart cameras. In this case, using multiple cameras with a powerful computing platform is an advantage. However, we encounter some physical limitations of the acquisition hardware. Although current professional cameras capture images at a horizontal resolution of about 4k to 5k [27], they are insufficient for large scales and wide-angle viewpoints. Additionally, there are some noticeable distortions caused by wide-angle optics such as distortion in the border regions of images in fish-eye lenses which result in additional loss in image resolution.

Smart camera networks have a wide range of applications in various areas including surveillance systems, security monitoring, traffic control and telemedicine [1]. For instance, Kawamura et. al [17] proposed a reliable surveillance system for railway stations. Their system tracks suspicious behavior by applying multiple camera fields of view. Smart sensors communicate with each other over a wireless mesh network. Moreover, as an application for the airport, Shirmohammadi et. al [31] introduced a decentralized target tracking scheme. Smart camera nodes automatically identify neighboring sensors with overlapping fields and produce a communication graph which reflects how the nodes will interact to fuse measurements in the network.

2.2. Multi-View Setups and Image Stitching

Moreover, numerous publications deal with panoramic images. For image registration, feature-based methods which use distinct image points are generally favored over area-based techniques which compare images window by window [39]. Lowe et al. [6, 21, 7] introduced scale-invariant feature points (SIFT) which have been widely used since. They use a 128-dimensional feature vector. Ke and Sukthankar [18] adopted this approach, but reduced the dimension of the descriptor to 36. Alternatively, Bay et al. [3] presented a faster method based on Haar wavelets using speeded up robust features (SURF).

All these features work well with standard perspective projection since they are invariant to affine transformations and provide a sufficient number of corresponding points to recover the parameters of the homography. Multi-view images taken from cameras at different positions lead to parallax errors. These errors cannot be fully eliminated. Still, these effects can be reduced. Global image transformations that are calculated by fitting a homography to matched feature points cannot handle parallax well. Zhang and Liu [38] address this problem by combining the transformation using a homography with local content-preserving warping. The homography is no longer chosen as the best fir for all feature point pairs, but considers only neighboring feature points. Additionally, they use a tolerant fitting threshold. Per-

azzi et al. [27] describe an algorithm for generating videos from unstructured camera arrays. They apply the basic concept of local warping to remove the parallax and define a new error measure with increased sensitivity to stitching artifacts. Their method tries to smooth out the blurring, ghosting and some other distortions caused usually when videos which feed from unstructured camera arrays are combined to create a single panoramic video. Deen et al. [9] create image mosaics for scientific purposes. Thus, they focus on correct rather than visually pleasant results. Parallax errors are reduced by performing pointing correction.

Existing tools for panoramic image stitching as well as camera calibration include Hugin [1] and PTGui [2], which are both based on Panorama Tools [3] by Dersch.

German et al.[14] investigate the application of different map projections to panoramic images including projections of fish-eye lens images. Multi-view setups are addressed by Sturm et al.[33] who develop a multi-view geometry model for central and non-central cameras based on structure-from-motion and by Luo et al.[23] who focus on saliency detection in multi-camera setups.

2.3. Fish-Eye Lenses

Schwalbe [30] develops a geometric model for fish-eye lens cameras based on the approximately linear relation between the incident angle of the ray from an object point and the distance from the corresponding image point to the principal point. Distortion is accounted for by using conventional distortion polynomials. Alternatively, Kannala and Brand [16] present a flexible camera model which is applicable for fish-eye as well as narrow-angle lenses. They use a polynomial imaging function as well as two additional terms for radial and tangential distortion, respectively. The final camera model includes 23 parameters. It provides both a forward as well as a backward model. Moreover, Luhmann et al. [22] deal with the correction of chromatic aberration in fish-eye images. Standard distortion correction methods use odd polynomial models as used by Mallon and Whelan [26]. These models describe the distorted radius r_d as a polynomial function of the undistorted radius r_u, using only odd terms. For high level distortion, however, fewer terms are needed with the division model [10]. Based on this approach, Aleman-Flores et al. [12] formulate a one-parameter model.

In order to determine lens parameters, various calibration procedures [37, 19, 32, 36, 34] have been developped. In many cases, they extract features such as lines or corners from the image of a calibration pattern for which the world coordinates are known [15]. A self-calibration method based on circle-fitting which does not require information about the objects' world coordinates is presented by Bräuer-Burchardt and Voss [5]. However, the distortion of an image needs to exactly fit the chosen distortion model. Aleman-Flores et al. [12] determine the distortion parameter automatically by introducing it into Hough space and detecting distorted lines.

3. Our Goal: a 360° Image Mosaic

We consider a multi-camera system of small high-quality cameras, in order to create a 360° image mosaic. The system consists of six fish-eye lens cameras. At this point we use the *indieGS2K* model produced by Indiecam[4]. Two adjoining cameras share an overlapping region, respectively. Position and optical parameters can be chosen arbitrarily, but will be fixed for a specific system. Each camera creates an image using fish-eye projection. Additionally, it holds the information about the other cameras' settings. In the end, an image mosaic using equirectangular projection is created. This means that the horizon is a straight line in the middle of the image and vertical lines in real world are vertical lines in the image [14]. Before the actual stitching can be performed, the respective images have to be transformed accordingly.

Eventually, our goal is to develop the respective coordinate transformation model. For any two images I_j and I_{j+1} from the six cameras with overlapping view, a function $F_j : I_j \rightarrow I_{j+1}$ has to be found such that $F(p_j) = p_{j+1}$ for all corresponding pixels (p_j, p_{j+1}) in the overlapping region with $p_j \in I_j$, $p_{j+1} \in I_{j+1}$. The resulting algorithm should take the image of one camera as well as the settings (position, optics) of the other as input.The output will be the accordingly transformed image.

[1] http://hugin.sourceforge.net/

[2] https://www.ptgui.com/

[3] http://panotools.sourceforge.net/

[4] www.indiecam.com

4. A Novel Concept for Image Alignment

At this point, the following problems have to be solved in order to determine the image transformation:

1. Calibrate the fish-eye lens and determine the distortion.
2. Calculate the transformation from the fish-eye projection to the equirectangular projection.
3. Perform a geometrical classification of possible setups. Considering two cameras C_1 and C_2, calculate critical points and distances in order to distinguish between the following classes:
   - region in which points can only be seen by C_1
   - region in which points can only be seen by C_2
   - closer part of the overlapping region with visible parallax errors
   - part of the overlapping region with negligible parallax errors
4. Calculate the coordinate transformation

In order to solve these problems described in the previous section, we consider the following approaches.

4.1. Lens Calibration and Image Alignment using Graph Pyramids

While traditional lens and distortion models have been studied extensively, we follow a different approach. Our goal is to extract the distortion information using graph pyramids.

4.1.1 Overview

Traditionally, lens calibration is based on a geometric model depending on parameters. The respective parameter values are determined during the calibration procedure. This is a characteristic that previous lens calibration methods have in common, even though different models and procedures have been developped. By establishing a model for which the parameters are specified, these methods already make fundamental assumptions about the structure of the distortion. On the contrary, we propose a calibration method that determines the distortion including its structure. In a graph pyramid as used for structurally consistent image segmentation (SCIS) [8], the information about the segmentation process is stored. As this process is performed based on the structure of the underlying image, it also contains information about the distortion. A target coordinate system is defined by the continuous curves of a checkerboard pattern which follow the isolines of the coordinate system. By applying the segmentation to this pattern and storing the details of the segmentation process, the distortion information of the coordinate system is retrieved.

4.1.2 Features of the SCIS Algorithm

The SCIS algorithm segments an image based on Local Binary Patterns and the Combinatorial Pyramid. It works on the local structure of the image and preserves structural correctness [8, Chapter 4, page 39] and topology of an image. For this purpose, five topological classes based on Local Binary Patterns of regions are applied which by combination with the dual graph are able to remove redundant structural information. As a result, by using this approach the image graph will be simplified and connected regions will be merged without introducing structural errors [8].

The SCIS algorithm performs image segmentation using a graph-based image representation. It provides the image at any level of segmentation as well as the information about the segmentation process up to that level. The latter contains information about the distortion structure.

Initially, each pixel corresponds to a vertex and each edge to a neighborhood relation in the graph, which represents the base level of the combinatorial pyramid. Subsequently, pixels are merged to regions which are in turn merged to larger regions based on their intensity values. On higher levels, each vertex corresponds to an image region. Merging corresponds to edge contraction and removal. The SCIS algorithm creates the entire pyramid as well as the contraction history. The latter is represented by the *contraction kernels*. Thus, it is able to reconstruct the segmented image at any level. An example of a combinatorial pyramid is shown in Figure 1.

An evaluation study of stereo matching by Joanneum Research [13] shows that the SCIS algorithm achieves the highest matching quality compared to different compression methods.

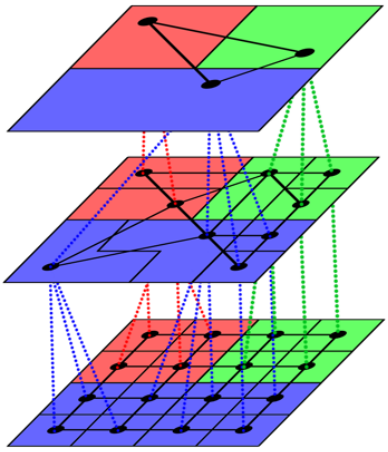

Figure 1. Example of a Combinatorial Pyramid. Image taken from [8]

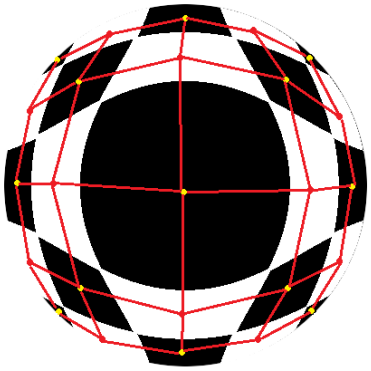

Figure 2. Distorted checkerboard pattern with corresponding primal graph at the top of the pyramid. Each vertex (yellow) corresponds to a patch. Vertices of the adjacent patches are connected by an edge (red).

4.1.3 Calibration Procedure

The canonical representation of the combinatorial pyramid stores it as a single array. The elements in this array are half-edges, called *darts*. They are ordered according to the contraction history. In order to extract information about the distortion from the combinatorial pyramid, we consider the image of a checkerboard pattern, where each patch is assigned an absolute coordinate. At the top level of the pyramid, each vertex corresponds to a single patch (see Figure 2).

As a result we get the contraction history. The top level delivers a single vertex for every patch of the checkerboard with its adjacency. All contracted edges of a patch form a spanning tree of the corresponding region in the primal graph. We do not know anything about the contraction kernels inside the homogeneous regions.

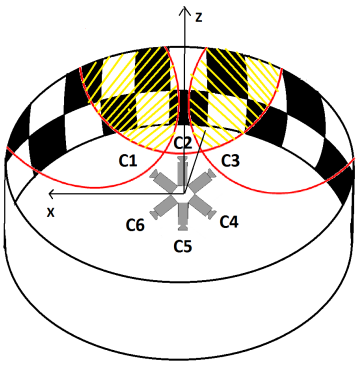

Figure 3. Multi-camera calibration setup for six cameras $C1$ - $C6$.

Since they have all the same value it cannot be said which edges are contracted or which are removed. For making the process more precise we can consider two solutions. One is to apply geometry of target coordinates and perform linear interpolation. However, this approach has the drawback that we do not know the size of the distorted patch, which is particularly problematic in our case where we expect severe deformation. The second approach is to shift the checkerboard pattern and create a new image from a different viewpoint. By iteratively applying this process, the regions inside the patches will be refined. For instance, we can take M captures with different offsets. Next, the idea is to freeze only the boundaries of which we are sure that they are precisely delineated. Indeed, by taking two different positions (randomly) and overlapping with the two contraction kernels, both boundaries should be preserved. Therefore, the random space of patches will be smaller and smaller as the process is used more and more.

There are two ways for applying this strategy. On the one hand, it can be performed sequentially by freezing the contraction kernels corresponding to the boundaries from the previous iteration. On the other hand, it can be performed randomly. Given the contraction kernels at every point and knowing the position of a boundary, we can integrate the contraction kernels using high weights at boundaries and low weights in between. For homogeneous regions, the contraction kernels provided by the shifting approach will converge towards the proper kernel.

With the contraction kernels provided, the information about the distortion is stored implicitly, allowing us to apply it to any new image. Conveniently, the canonical representation stores this in-

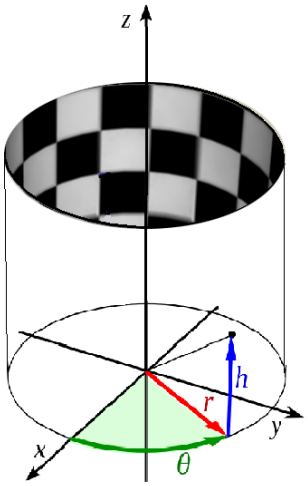

Figure 4. Calibration pattern using a cylindrical target coordinate system with radius r, azimuthal angle θ and height h.

formation in an ordered array. Thus, the calibrated kernels which have to be applied to get to a particular level of the pyramid can be re-used.

The calibration setup for a multi-camera system is illustrated in Figure 3.

4.1.4 Advantages of Calibration using Graph Pyramids

Apart from the fact that the graph-based approach does not make any assumptions about the structure of the distortion, it yields other advantages compared to previous calibration methods. Accuracy can be increased simply and reached to the resolution of original images by increasing the number of shifts. Additionally, we do not need a global model of the geometric projection for calibration, which is needed for many estimation methods of the parameters. Finally, our method does not depend on a particular coordinate system. Instead, any target coordinate system can be chosen. It is defined by the checkerboard pattern where the continuous curves correspond to the the isolines of a target coordinate system. Thus, various geometries can be used for this approach such as cylindrical (see Figure 4) or spherical (see Figure 5) coordinate systems. In particular, the coordinate system of the final mosaic can be chosen as target coordinate system. In this case, the transformation provided by the calibration method does not only consider lens distortion, but also includes remapping to equirectangular projection as well as image alignment, and this simultaneously for all six cameras.

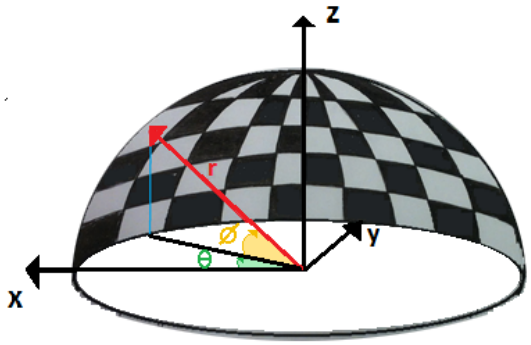

Figure 5. Calibration pattern using a spherical target coordinate system with radius r, azimuthal angle θ and elevation angle ϕ.

4.2. Projection remapping

The remapping from fish-eye to equirectangular projection can also be handled by the graph-based calibration method presented in the previous section. For comparison, it can be addressed individually following German et al. [14]. Information about the camera's roll, which is the rotation angle about the optical axis, and pitch, which is the elevation angle from the horizontal axis, allows the remapping from a fish-eye to an equirectangular projection. Roll and pitch can be determined manually or by using horizontal or vertical control lines.

4.3. Setup Classification

The classification of the setup with regard to parallax errors can be performed using partial edge contours as used by Wang et al. [35]. The edge contour of an obstacle is mapped from one image to the other. The parallax is then calculated as the transverse distance between corresponding edge contour pixels.

4.4. Image Transformation

Similar to the projection remapping, the image tranformation used for image alignment can be determined by the graph-based approach. For comparison, the calibration of the multi-camera system can be performed using feature extraction and matching. For this purpose, SIFT features [21] will be used. In order to reduce parallax errors, the image transformation will be calculated following the parallax-tolerant approach used by Zhang and Liu [38].

5. Conclusion

We presented a novel concept for the smart camera image stitching. It aims at reducing the cost of the stitching process by enabling each camera of a multi-camera system to align the image that takes individually. Lens calibration can be performed using graph pyramids, which yields several advantages compared to traditional lens calibration methods. Additionally, the same method can be used to directly determine the image transformation required for image alignment. Currently, the work is in progress, but in near future we are planning to experimentally prove the applicability of the proposed ideas.

Acknowledgment

Hanna Huber and Majid Banaeyan are supported by Vienna Business Agency with application number ID:1360780 and Image Processing (PRIP) group via PRIP-Club respectively.

References

[1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury. A survey on wireless multimedia sensor networks. *Computer Networks*, 51:921–960, 2007. 2

[2] L. Albani, P. Chiesa, D. Covi, G. Pedegani, A. Sartori, and M.Vatteroni. VISoc: A smart camera SoC. In *Proceedings of the 28th European Solid-State Circuits Conference*, pages 367–370, 2002. 2

[3] H. Bay, T. Tuytelaars, and L. V. Gool. Surf: Speeded up robust features. In *In ECCV*, pages 404–417, 2006. 2

[4] A. N. Belbachir, editor. *Smart Cameras*. Springer, 2010. 2

[5] C. Bräuer-Burchardt and K. Voss. A new algorithm to correct fish-eye- and strong wide-angle-lens-distortion from single images. *Proceedings 2001 International Conference on Image Processing 2001, Vol.1, pp.225-228*, 2001. 3

[6] M. Brown and D. G. Lowe. Recognising panoramas. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, ICCV '03, pages 1218–1225, Washington, DC, USA, 2003. IEEE Computer Society. 2

[7] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vision*, 74(1):59–73, Aug. 2007. 2

[8] M. Cerman. Structurally correct image segmentation using local binary patterns and the combinatorial pyramid, 2015. Wien, Techn. Univ., Dipl.-Arb., 2015, Technical Report 133. 4

[9] B. Deen. In-Situ Mosaic Production at JPL/MIPL. Pasadena, CA : Jet Propulsion Laboratory, National Aeronautics and Space Administration, 2012. Planetary Data: A Workshop for Users and Software Developers 2012, JPL TRS 1992+. 3

[10] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2001, Vol.1, pp.I-I*. 3

[11] B. Flinchbaugh. Smart cameras systems technology roadmap. In B. Kisacanin, V. Pavlovic, and T. Huang, editors, *Real-Time Vision for Human-Computer Interaction*, pages 285–297, 2005. 2

[12] M. A. Flores, L. Á. León, L. G. Déniz, and D. E. S. Cedrés. Automatic Lens Distortion Correction Using One-Parameter Division Models. *IPOL Image Processing OnLine (Special Issue on Lens Distortion Models)*, 4:327–343, 2014. 3

[13] B. Froehlich and M. P. Caballo-Perucha. Evaluation of image compression algorithms version 1.0, issue D1, 2015. Joanneum Research. 4

[14] D. M. German, P. d'Angelo, M. Gross, and B. Postle. New Methods to Project Panoramas for Practical and Aesthetic Purposes. In D. W. Cunningham, G. Meyer, and L. Neumann, editors, *Computational Aesthetics in Graphics, Visualization, and Imaging*. The Eurographics Association, 2007. 3, 6

[15] C. Hughes, M. Glavin, E. Jones, and P. Denny. Review of geometric distortion compensation in fish-eye cameras. In *IET Irish Signals and Systems Conference, 208. (ISSC 2008)*, 2008. 1, 3

[16] J. Kannala and S. S. Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE TRANS. PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 28:1335–1340, 2006. 1, 3

[17] A. Kawamura, Y. Yoshimitsu, K. Kajitani, T. Naito, K. Fujimura, and S. Kamijo. Smart camera network system for use in railway stations. In *SMC*, pages 85–90. IEEE, 2011. 2

[18] Y. Ke and R. Sukthankar. PCA-SIFT: A more distinctive representation for local image descriptors. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR'04, pages 506–513, Washington, DC, USA, 2004. IEEE Computer Society. 2

[19] M. Kedzierski and A. Fryskowska. Precise method of fisheye lens calibration. In *Proceedings of the ISPRS-Congress*, pages 765–768. International Society for Photogrammetry and Remote Sensing, 2008. 3

[20] B. Kisacanin, S. Bhattacharyya, and S. Chai. *Embedded Computer Vision*. Springer, 2007. 2

[21] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, Nov. 2004. 2, 6

[22] T. Luhmann, H. Hastedt, and W. Tecklenburg. Modelling of chromatic aberration for high precision photogrammetry. *Remote Sensing and Spatial Information Sciences*, 36 (Part 5):173–178. 3

[23] Y. Luo, M. Jiang, Y. Wong, and Q. Zhao. Multi-camera saliency. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(10):2057–2070, 2015. 3

[24] R. Lyon. The optical mouse, and architectural methodology for smart digital sensors. In H.T.Kung, B.Sproull, and G.Steele, editors, *Computer Science Press. Invited Paper, CMU Conference on VLSI structures and Computations*, 1981. 2

[25] R. Lyon. Apparatus for controlling movement of a curser in computer display system, 1983. European Patent. 2

[26] J. Mallon and P. Whelan. Precise radial un-distortion of images. *Proceedings of the 17th International Conference on Pattern Recognition, 2004, Vol.1, pp.18-21*. 3

[27] F. Perazzi, A. Sorkine-Hornung, H. Zimmer, P. Kaufmann, O. Wang, S. Watson, and M. Gross. Panoramic video from unstructured camera arrays. In *Proc. Eurographics 2015*, volume 34, 2015. 2, 3

[28] B. Rinner and W. Wolf. An introduction to distributed smart cameras. In *Proceedings of the IEEE*, volume 96, pages 1565–1575, 2008. 2

[29] R. Schneidermann. Smart cameras clicking with electronic functions. *Electronics*, 48:74–81, 1975. 2

[30] E. Schwalbe. Geometric modelling and calibration of fisheye lens camera systems. In *Proceedings 2nd Panoramic Photogrammetry Workshop, Int. Archives of Photogrammetry and Remote Sensing*, pages 5–8, 2005. 3

[31] B. Shirmohammadi and C. J. Taylor. Distributed target tracking using self localizing smart camera networks. In *Proceedings of the Fourth ACM/IEEE International Conference on Distributed Smart Cameras*, pages 17–24, New York, NY, USA, 2010. 2

[32] P. Srestasathiern and N. Soontranon. A novel camera calibration method for fish-eye lenses using line features. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 327–332, Aug. 2014. 3

[33] P. Sturm, S. Ramalingam, and S. Lodha. On calibration, structure-from-motion and multi-view geometry for general camera models. In R. Reulke and U. Knauer, editors, *2nd ISPRS Panoramic Photogrammetry Workshop,*, Berlin, Allemagne, 2005. ISPRS. Published in the Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI-5/W8. 3

[34] S. Urban, J. Leitloff, and S. Hinz. Improved wide-angle, fisheye and omnidirectional camera calibration. {*ISPRS*} *Journal of Photogrammetry and Remote Sensing*, 108:72 – 79, 2015. 3

[35] X.-H. Wang, W.-P. Fu, and W. Chen. Detection of obstacle based on nocular vision. *2010 International Conference on Intelligent Computation Technology and Automation, May 2010, Vol.2, pp.71-74*. 6

[36] Z. Wang, H. Liang, X. W. andYipeng Zhao, B. Cai, C. Tao, Z. Zhang, Y. Wang, S. Li, F. Huang, S. Fu, and F. Zhang. A practical distortion correcting method from fisheye image to perspective projection image. In *Information and Automation, 2015 IEEE International Conference on*, pages 1178 – 1183, 2015. 3

[37] X. Ying, Z. Hu, and H. Zha. Fisheye lenses calibration using straight-line spherical perspective projection constraint. In P. J. Narayanan, S. K. Nayar, and H.-Y. Shum, editors, *ACCV (2)*, volume 3852 of *Lecture Notes in Computer Science*, pages 61–70. Springer, 2006. 3

[38] F. Zhang and F. Liu. Parallax-tolerant image stitching. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '14, pages 3262–3269, Washington, DC, USA, 2014. IEEE Computer Society. 2, 6

[39] B. Zitov and J. Flusser. Image registration methods: a survey. *Image and Vision Computing, 2003, Vol.21(11), pp.977-1000*. 2