CAMERA ARRAY FOR MULTISPECTRAL IMAGING

03CS6903 Seminar I

13/MCS/2020 CHN20MT013 Shivangi.M
M. Tech. Computer Science & Engineering
(Image Processing)

Department of Computer Engineering
College of Engineering Chengannur
Alappuzha 689121
Phone: +91.479.2165706
http://www.ceconline.edu
hodcs@ceconline.edu
College of Engineering Chengannur
Dept. of Computer Engineering

C E R T I F I C A T E

This is to certify that, this report titled ***Camera array for multispectral imaging*** is a bonafide record of the 03CS6903 Seminar I presented on March 17 2021 by

13/MCS/2020 CHN20MT013 SHIVANGI M.

First Semester M. Tech. Computer Science & Engineering (Image Processing) scholar, under our guidance and supervision, in partial fulfillment of the requirements for the award of the degree, M. Tech. Computer Science & Engineering (Image Processing) of APJ Abdul Kalam Technological University.

Guide

Ms. Sreelekshmi KR
Asst. Professor in
Computer Engineering

Coordinator

Mr. Ahammed Siraj K K
Associate Professor in
Computer Engineering

Head of the Department

March 17, 2021

Dr. Smitha Dharan
Professor
in Computer Engineering

Acknowledgments

Primarily, I thank Lord Almighty for his eternal support throughout my Seminar work.

I express my sincere thanks to Dr. Jacob Thomas V., Principal, College of Engineering Chengannur for extending all the facilities required for doing my seminar. My heartfelt words of gratitude to Dr. Smitha Dharan, Professor and Head of Department of Computer Engineering, for providing constant support.

Now I express my gratitude to my seminar co-ordinator Mr. Ahammed Siraj K K, Associate Professor in Computer Engineering and my seminar guide Ms. Sreelekshmi , Assistant Professor in Computer Engineering who played a great role for valuable suggestions and expert guidance.

Abstract

Despite the growing demand, the acquisition of multi-spectral data is still very complicated. Often, expensive, inflexible, or low resolution acquisition setups are only obtainable for specific professional applications. To overcome these limitations, a novel camera array for multispectral imaging is presented for generating consistent multi-spectral videos. As differing spectral images are acquired at various viewpoints, a geometrically constrained multi-camera sensor layout is introduced, which enables the formulation of novel registration and reconstruction algorithms to globally set up robust models.On average, the novel acquisition approach achieves a gain of 2.5dB PSNR compared to recently published multispectral filter array imaging systems. Moreover, depth information is generated, so that 3D imaging applications, e.g., for augmented or virtual reality, become possible. The proposed camera array for multispectral imaging can be set up using off-the-shelf hardware,which allows for a compact design and employment in, e.g., mobile devices or drones, while being cost effective.

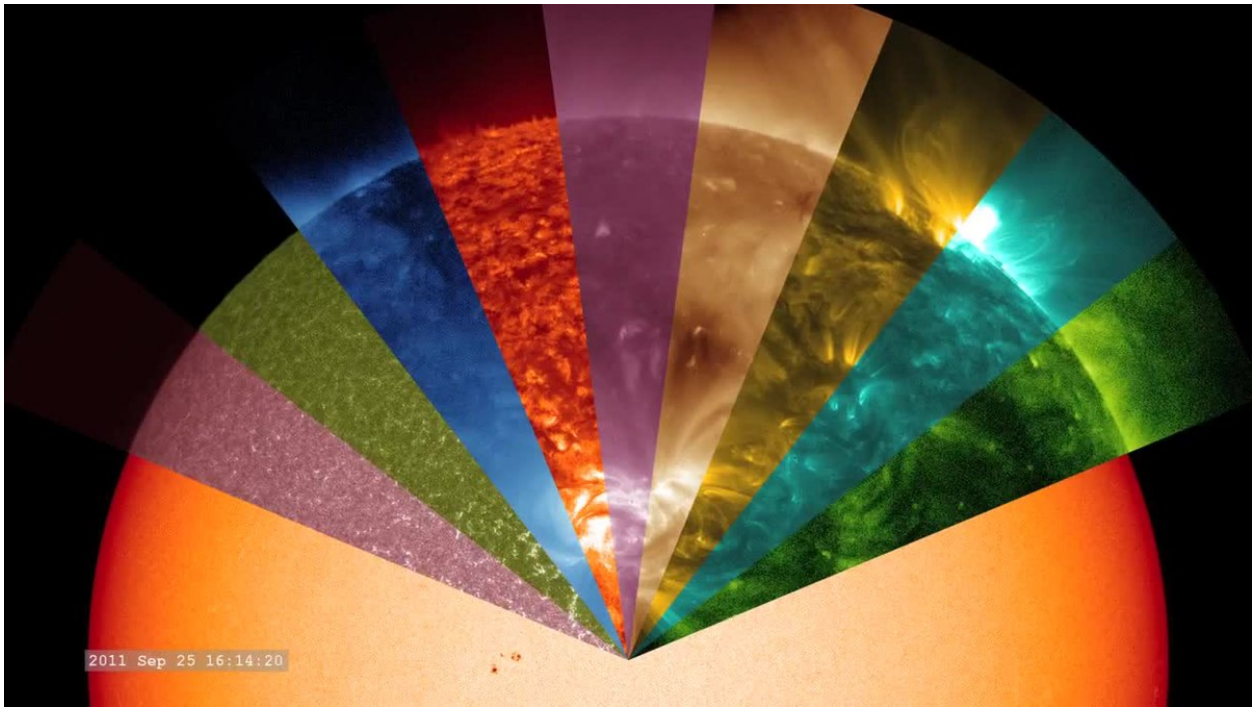
Contents

Chapter 1

Introduction

Multispectral imaging means methods for spectral imaging where one obtains images corresponding to at least a couple of spectral channels.The used spectral regions are often partially outside visible spectral range,covering parts of the infrared and ultraviolet region.ie in simple terms multispectral imaging gives us additional information that human eye cannot see.Multispectral imaging measures light in a small number (typically 3 to 15) of spectral bands.Multispectral imaging means methods for spectral imaging where one obtains images corresponding to at least a couple of spectral channels.The used spectral regions are often partially outside visible spectral range,covering parts of the infrared and ultraviolet region.A multispectral camera captures multiple images in visible ,uv and IR regions.Each of image taken with a multispectral camera is passed through a filter to restrict light to a specific colour or wavelength.

Despite growing demand the acquisition of multispectral data is still very complicated. Often expensive, inflexible or low resolution acquisition setups are only obtainable for specific applications.There are many disadvantages in the existing techniques.The breakthrough of multi-spectral imaging is impeded by the different disadvantages of existing acquisition systems.



To overcome the existing disadvantages a novel camera array for multispectral imaging CAMSI is introduced.As differing spectral images are acquired at various viewpoints, a geometrically constrained multi-camera sensor layout is introduced.On average, the novel acquisition approach achieves a gain of 2.5dB PSNR compared to recently published multispectral filter array imaging systems. At the same time, the proposed acquisition system ensures not only superior spatial, but also a high spectral,and temporal resolution.While filters are flexibly exchangeable by the user depending on the application.The proposed camera array for multispectral imaging can be set up using off the-shelf hardware,which allows for a compact design and employment in, e.g., mobile devices or drones, while being cost effective.Moreover, the system must be price-efficient and suited for the consumer market, as well as inexpensive for small batch sizes.

Chapter 2

Literature survey

1. **S. Mendis, S. E. Kemeny, and E. R. Fossum, "CMOS active pixel image sensor,"** IEEE Trans. Electron Devices, vol. 41, no. 3, pp. 452–453,Mar. 1994.

Most recent image sensor arrays have not been able to operate at frame rates that exceed tens to hundreds of thousands of frames per second.The main bottleneck preventing imaging at higher frame rates is the time required to access the array, convert the image data from analog to digital, and transmit the data off the image sensor chip.Most significant source of delay, mainly due to the limited number of input and output ports available on the chip.This work allows for a significant increase in image capture rate by separating the image acquisition phase from the conversion and readout phase.This was done by capturing eight frames at a high capture rate and temporarily storing the multiple frames into analog memory units that are incorporated inside the pixel. The design was implemented in a deep-submicron CMOS 130 nm technology that allows for high-speed operation.The pixel can achieve an electrical image capture rate of 1.25 billion fps.The sensor used can achieve a speed of 1 million fps. For higher frame rates, we recommend using a high sensitivity light sensor such as a single photon detector with the pixel.

2 **H. Park and K. B. Crozier, "Multispectral imaging with vertical silicon nanowires**," IEEE Transaction on image processing., vol. 3, no. 1, pp. 1–2, Aug. 2013.

Multispectral imaging systems generally are expensive and bulky, and multiple exposures are needed. Here, demonstration of a compact multispectral imaging system Iis introduced that uses vertical silicon nanowires to realize a filter array.

Multiple filter functions covering visible to near-infrared (NIR) wavelengths are simultaneously defined in a single lithography step using a single material (silicon). Nanowires are then stretched and embedded into polydimethylsiloxane (PDMS), thereby realizing a device with eight filter functions.By attaching it to a monochrome silicon image sensor, we successfully realize an all-silicon multispectral imaging system. In our approach, an array of PDMS-embedded vertical silicon nanowires is attached to a monochrome image sensor . The unit cell of the array contains eight different spectral filters and a transparent window in its center. This configuration is analogous to the dye-based filter arrays (red/green/blue) used in colour image sensors, but with eight filterfunctions that span visible to NIR wavelengths, rather than three visible-wavelength filter functions. The device permits an image to be acquired from each spectral channel in a single exposure.

3. **L. Miao, H. Qi, R. Ramanath, and W. E. Snyder, "Binary tree-based generic demosaicking algorithm for multispectral filter arrays**," IEEE Trans. Image Process., vol. 15, no. 11, pp. 3550–3558, Nov. 2006

In this paper, the idea of using mosaicked color filter array (CFA) in color imaging, which has been widely adopted in the digital color camera industry, to the use of multispectral filter array (MSFA) in multispectral imaging. The filter array technique can help reduce the cost, achieve exact registration,and improve the robustness of the imaging system. However,the extension from CFA to MSFA is not straightforward. First,most CFAs only deal with a few bands (3 or 4) within the narrow visual spectral region, while the design of MSFA needs to handle the arrangement of multiple bands (more than 3) across a much wider spectral range. Second, most existing CFA demosaicing algorithms assume the fixed Bayer CFA and are confined to properties

only existed in the color domain. Therefore, they cannot be directly applied to multispectral demosaicking. The main challenges faced in multispectral demosaicking is how to design a generic algorithm that can handle the more diversified MSFA patterns, and how to improve performance with a coarser spatial resolution and a less degree of spectral correlation. In this paper,they present a binary tree based generic demosaicing method. Two metrics are used to evaluate the generic algorithm, including the root mean-square error (RMSE) for reconstruction performance and the classification accuracy for target discrimination performance.Experimental results show that the demosaicked images present low RMSE (less than 7) and comparable classification performance as original images. These results support that MSFA technique can be applied to multispectral imaging with unique advantages.

4 **Y. Monno, D. Kiku, M. Tanaka, and M. Okutomi, "Adaptive residual interpolation for color and multispectral image demosaicking,**" Sensors, vol. 17, no. 12, pp. 2787–2811, Dec. 2017.

Color image demosaicking for the Bayer color filter array is an essential image processing operation for acquiring high-quality color images. Recently, residual interpolation (RI)-based algorithms have demonstrated superior demosaicking performance over conventional color difference interpolation-based algorithms. In this paper, adaptive residual interpolation (ARI) is introduced that improves existing RI-based algorithms by adaptively combining two RI-based algorithms and selecting a suitable iteration number at each pixel. These are performed based on a unified criterion that evaluates the validity of an RI-based algorithm. Experimental comparisons using standard color image datasets demonstrate that ARI can improve existing RI-based algorithms by more than 0.6 dB in the color peak signal-to-noise ratio and can outperform state-of-the-art algorithms based on training images.And further extend ARI for a multispectral filter array, in which more than three spectral bands are arrayed, and demonstrate that ARI can achieve state-of-the-art performance also for the task of multispectral image demosaicking.

5 **P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry,"** Comput. Vis. Image Understand., vol. 78, no. 1, pp. 138–156, Apr. 2000

A new robust estimator, MLESAC, which can be used in a wide variety of estimation tasks.In particular, MLESAC is well suited to estimating complex surfaces or more general manifolds from point data. It is applied here to the estimation of several of the multiple view relations that exist between images related by rigid motions.
One reason is that the minimal point set initially selected MLESAC is known to provide a good estimate of the image relation (because there is a lot of support for this solution). Hence the initial estimate of the point basis provided by MLESAC is quite close to the true solution and consequently the nonlinear minimization typically avoids local minima.Second, the parameterization is consistent, which means that during the gradient descent phase only image relations that might actually arise are searched for It has been observed that the MLESAC method of robust fitting is good for initializing the parameter estimation when the data are corrupted by outliers.In this case there are just two classes to which a datum might belong, inliers or outliers. The MLESAC method may be generalized to the case when the data has arisen from a more general mixture model involving several classes, such as in clustering problems.

6 **A. Geiger, F. Moosmann, O. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot,**" in Proc. IEEE Int.Conf. Robot. Autom., May 2012, pp. 3936–3943
Here a toolbox with a web interface for fully automatic camera to-camera and camera-to-range calibration is introduced. The system is easy to setup and recovers intrinsic and extrinsic camera parameters as well as the transformation between cameras and range sensors within one minute.In contrast to existing calibration

approaches, which often require user intervention, the proposed method is robust to varying imaging conditions, fully automatic, and easy to use since a single image and range scan proves sufficient for most calibration scenarios. Experimentally, demonstrate that the proposed checkerboard corner detector significantly outperforms current state-of-the-art. Furthermore, the proposed camera-to-range registration method is able to discover multiple solutions in the case of ambiguities.

Chapter 3

CAMSI
(camera array for multispectral imaging)

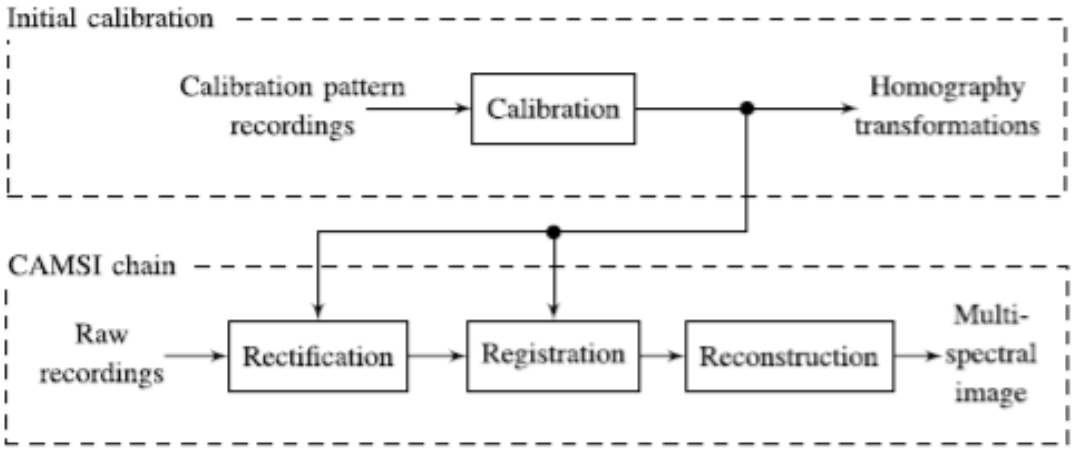
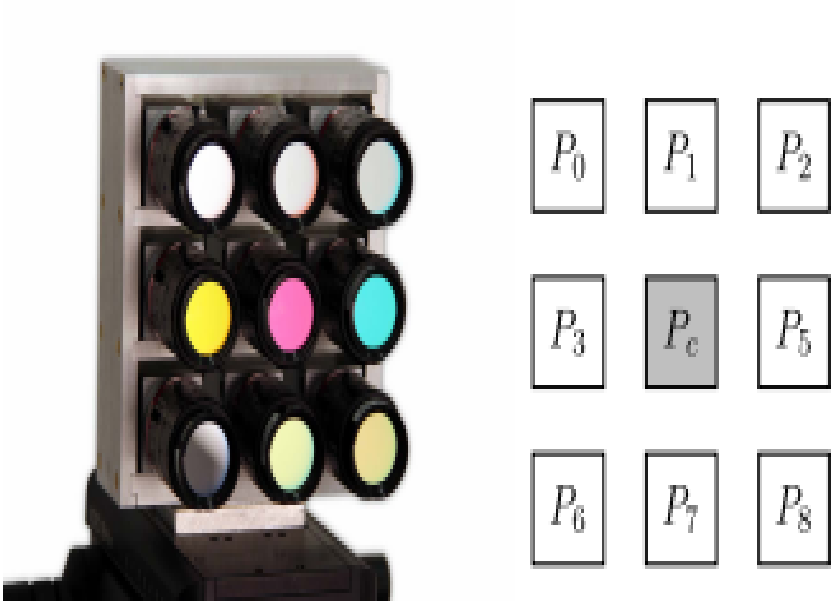


Fig. 4. Overview of the proposed CAMSI chain. The calibration of the camera array is carried out only once. Afterwards, the estimated homography transformations are used to calculate multi-spectral images.

3.1 HARDWARE DESIGN.

The proposed CAMSI system consists of K = 9 cameras,with K being selected flexibly depending on the required number of channels. The different camera positions are denoted as Pk allowing index k in the range 0 to K − 1 with center camera position Pc and index c = 4. Cameras are aligned on a 3×3 grid, so that the displacement from a peripheral camera to the center is always horizontal, vertical, or diagonal . The camera array's size must be chosen quadratic and the array length must be odd, e.g.,

3×3 or 5×5, to align all peripheral views around the center camera. The cameras are mounted in a solid aluminum enclosure to ensure a consistent calibration and a good thermal conduction The baseline between each camera measures 4 cm, so that the distance between each peripheral and the center camera is 4 cm and $4\sqrt{2} \approx 5.7$ cm, respectively. The professional monochrome cameras that record images up to a resolution of 1600 × 1200 pixels at 60 frames per second and a bit depth of 12, while the sensor pixel size measures 4.5 μm. The cameras are connected to a PC via Gigabit Ethernet while being synchronized using the camera trigger feature to ensure that all cameras take the pictures at the same time and still as well as moving content is recordable.The cameras allow for an easy mounting of different lenses,so that a wide range of problems can be tackled with this setup. Professional image processing color filters were used together with steep 50 nm band pass filters from stock to record color, ultra-violet, and near infrared images



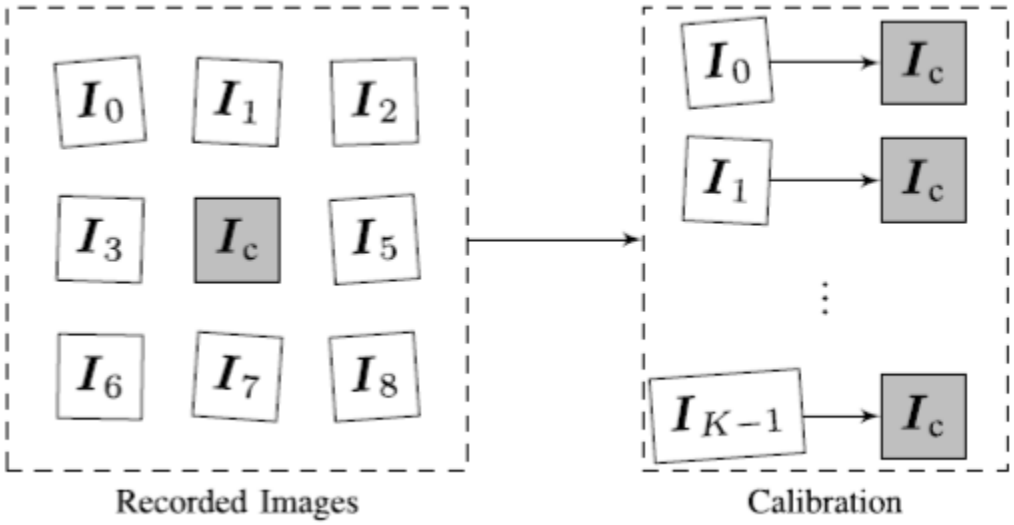
3.2 CALIBRATION AND RECTIFICATION

All cameras exhibit arbitrary displacements to each other. On the one side, it is not possible to perfectly align the cameras onto the array grid. On the other side, the image sensor may not be installed precisely enough in the camera body itself. All cameras in the array must be calibrated with respect to the central camera before a registered multi-spectral image can be calculated.

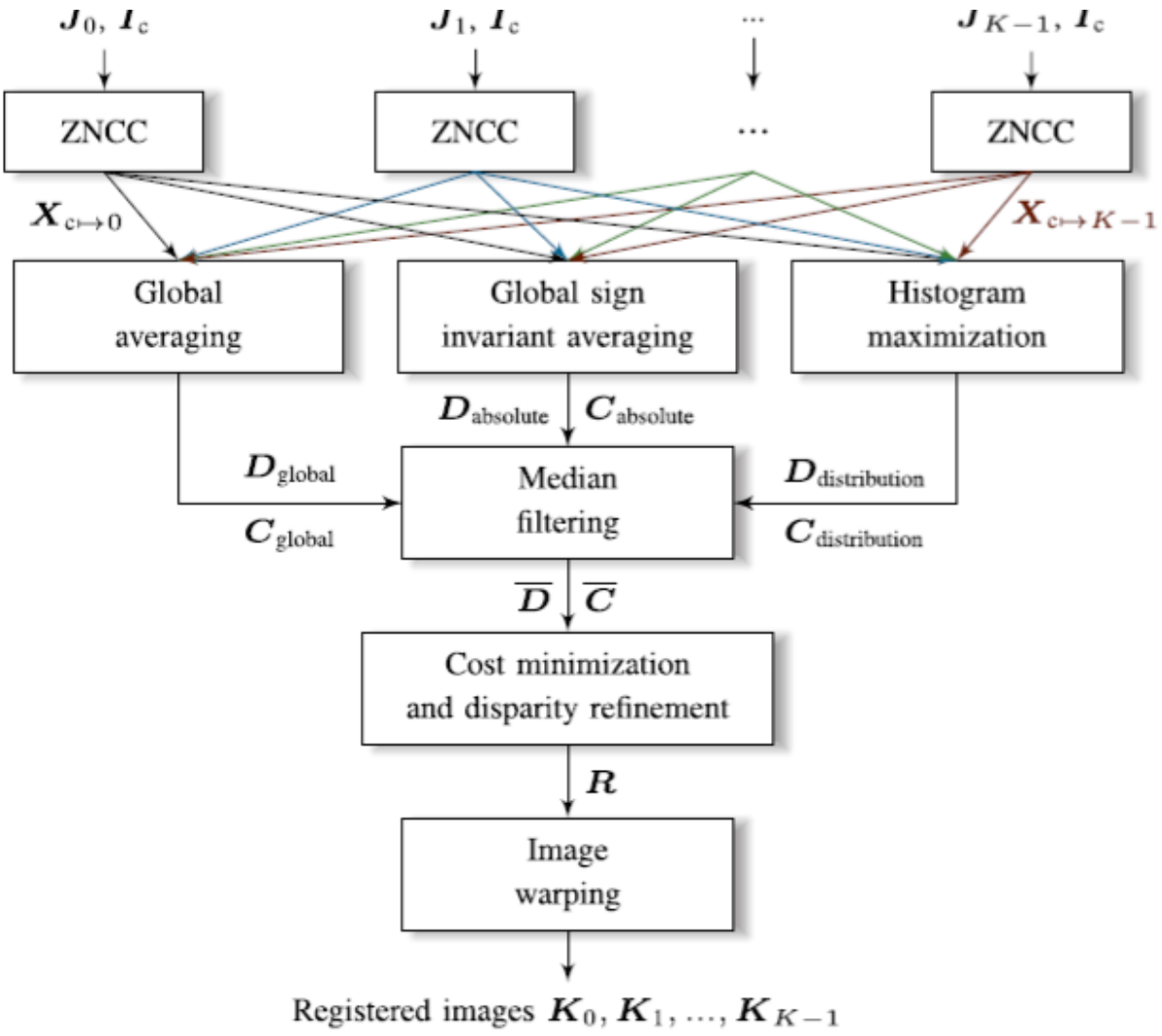
After recording the different images, the corresponding pixels are misplaced both in horizontal and vertical direction at the same time. In a perfect setup, the displacement should be purely horizontal, vertical, or diagonal, which is not applicable in practice.The aim of the proposed calibration is to align the views back on the virtual CAMSI array grid and to resolve the arbitrary displacement to a one-dimensional disparity.All cameras in the array must be calibrated with respect to the central camera before a registered multi-spectral image can be calculated.

The aim of the proposed calibration is to align the views back on the virtual CAMSI array grid and to resolve the arbitrary displacement to a one-dimensional disparity. By recording a calibration pattern in front of all cameras,it is possible to estimate the homography of the planes which includes the pattern.The displacement for the peripheral and the center camera. The transformation between all peripheral cameras and the center is estimated, while only the peripheral positions are warped onto the center. The decision regarding the calibration pattern was made in favor of using a checkerboard as this is a well-investigated approach, which is proven to work robustly and accurately.Firstly, the calibration pattern must be detected in every recording.By providing prior knowledge about the pattern, e.g., number of boxes and geometry and by utilizing a pattern that is visible in all spectral components, the estimation of the features works reliable even for the different color and bandpass filters mounted in front

of the cameras.By warping the peripherals view independently, all spectral components are rectified to the central camera position, but only for the depth in which the calibrated chessboard pattern was located.



3.3 REGISTRATION



After calibration and rectification, the images still yield a one-dimensional disparity. In the following, an appropriate disparity estimation method is presented, which is suitable for multi-spectral imaging in contrast to existing state-of-the-art algorithms. Firstly, the relationship between depth and disparity is summarized.Secondly, a novel fast one dimensional disparity estimation for multi-spectral imaging is introduced, which is calculated globally for all cameras.In the end, a pixel transform is derived, which combines both calibration and registration information to avoid multiple concatenated image resampling that would introduce blurring.

$$d = \frac{B \cdot f}{z \cdot p}$$

with baseline B, focal length f , sensor pixel size p and depth z. This relation holds if the optic axes of the cameras are parallel, which is assumed for the CAMSI setup.Furthermore, the relationship between disparity and depth is inversely proportional. Hence, a high depth indicates a low disparity and vice versa.The smaller the hardware implementation can be conducted,the smaller is the disparity due to the reduced baseline distance between the cameras. Consequently, it is advisable to shrink the hardware design as much as possible to achieve a faster and more robust disparity search.

Disparity information can be transferred to a depth map using the triangulation relationship .A novel global cross-correlation registration is proposed.That estimates a central disparity map by taking many pairwisely estimated cost functions into account, while overcoming the cross-spectral extinction problem. For CAMSI, the disparity estimation is performed for each peripheral position with respect to the center view. Hence, for the proposed array size of K = 9,eight estimations have to be conducted . A one dimensional representation is chosen, such that the calibrated views have only a pure horizontal, vertical, or diagonal relationship. Therefore, it is proposed to search for candidates only horizontally, while all other cases are handled by rotating the images pairs by +90, +45 and −45 degrees, respectively.

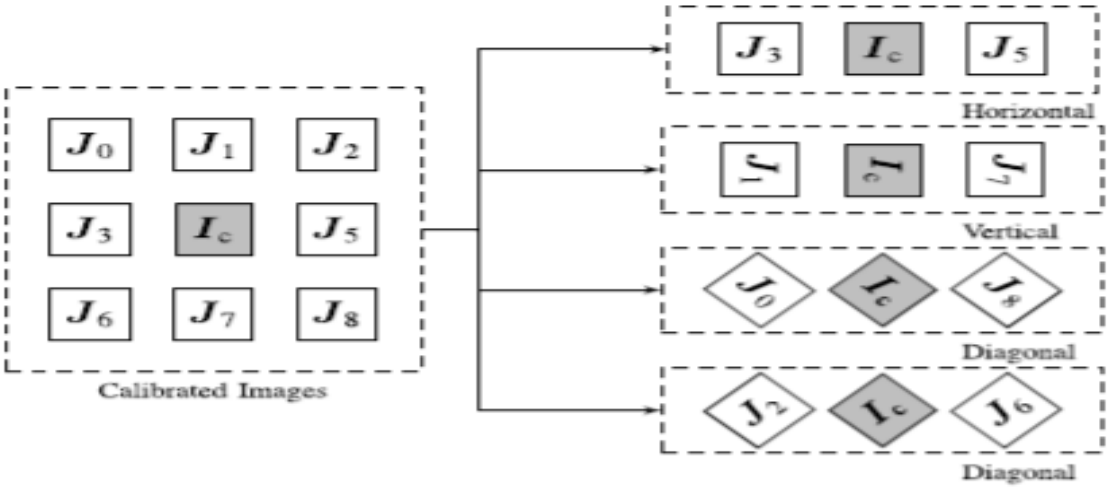


Fig. 8. The two-dimensional relationship of the different camera views is simplified by rotating the recordings in accordance with their position.

The registration problem is divided into different subtasks.The first step is to calculate the matching costs using Zero Mean Normalized Cross-Correlation (ZNCC) , which compares the structure of two zero-mean signals. In contrast to other metrics, e.g., SAD, SSD, or Census , the ZNCC is substantially more robust for the CAMSI setup as the image content differs significantly for multi-spectral images. For two vectors a and b, ZNCC is defined as

$$\mathrm{ZNCC}(\boldsymbol{a}, \boldsymbol{b}) = -\frac{\langle \boldsymbol{a} - \overline{a}, \boldsymbol{b} - \overline{b} \rangle}{||\boldsymbol{a} - \overline{a}||_2 \cdot ||\boldsymbol{b} - \overline{b}||_2}$$

For ZNCC, costs can take values between −1 and +1, while −1 describes the maximum positive correlation, zero depicts no correlation and +1 denotes the maximum negative correlation between the signals a and b.For an image of size M × N and D depth levels that shall be investigated, costs are obtained by evaluating the ZNCC between the center and peripheral view k.

$$(M_{c\mapsto k})_{x,y,d} = \mathrm{ZNCC}(I_c(r), J_k(r+d))$$

After calculating the cost matrix Xc→k for every peripheral camera view, a cost fusion is conducted to achieve a robust estimate, which makes use of all camera views.

for every pixel coordinate (x, y) and all investigated depth levels d ∈ D, one obtains costs for every matrix entry (Mc→k)x,y,d of the cost matrix Mc→k . In the following, let (Z)x,y,d be the operator for accessing an element at index (x, y, d) in the exemplary matrix Z, while the notation c → k states that the center recording is mapped to the camera position Pk . In the above equation, r denotes the region of interest around the current position (x, y) using an experimentally determined support window of size W = 7. Varying W controls the sharpness and noise level of the disparity estimation.Thus, small window sizes lead to sharper disparity maps, but include noisy mispredictions and vice versa. For pixels at the image borders, missing entries are padded by repeating.After calculating the cost matrix Xc→k for every peripheral camera view, a cost fusion is conducted to achieve a robust estimate, which makes use of all camera views.According to Fig. 9, this step consists of three different methods that are combined. Firstly, the matrices are averaged to obtain the global cost matrix high-quality aggregation.Given a smooth cost matrix, a total cost minimization results by selecting the disparity that stores the lowest costs per pixel.Secondly, to increase the robustness of the registration even further, the correlation properties of ZNCC are taken into account. As recorded objects can flip brightness over the spectral components, it is advisable to investigate not only the positive but also the negative correlation. Hence, the global sign invariant cost matrix results . In contrast to the disparity map results in a total cost maximization as costs close to 0 indicate no correlation and costs close to +1 shows a high positive or negative correlation.Thirdly, the disparity distribution itself is taken into account.Therefore, the distribution matrix distribution is calculated, which stores the disparities that are estimated for the various camera views.

C distribution stores a histogram for every pixel position (x, y) that shows the absolute frequency of the disparity estimates per disparity level d for all maps Consequently, the disparity map is calculated by choosing the most likely disparity. This third criterion can be interpreted as a non-linear majority decision for every disparity entry and all camera recordings.Then, a cost adaptive median filter is applied, which adjusts the filter size of the median filter according to the trustworthiness of the disparity estimation.

The filter strength F is estimated for every position (x, y) as

$$(\boldsymbol{F})_{x,y} = \begin{cases} 0, & (\overline{\boldsymbol{C}})_{x,y} > F_{\text{th}} \\ \left\lfloor \dfrac{F_{\text{min}} - F_{\text{max}}}{F_{\text{th}}} \min\limits_{(d)} \{(\overline{\boldsymbol{C}})_{x,y}\} + F_{\text{max}} \right\rfloor, & \text{else} \end{cases}$$

With minimal filter window size Fmin = 3, maximal window size Fmax = 15, and cost threshold Fth = 0.5

3.4 RECONSTRUCTION

The registered images Kk contain losses at various positions due to occlusions and mispredictions. The center image is fully preserved and can serve as reference for any distortion in every image. Moreover, it is very likely that losses are located at different positions for the various images as occlusions are dependent on the camera position. Consequently, a promising approach is to reconstruct the missing information by exploiting the spectral similarity as multiple references are available for every lost pixel. Hence, the reconstruction is interpreted as combined linear regression problem to estimate the distorted areas in the images Kk .Therefore, every image Kk is partitioned into square-shaped blocks For every block, the reconstruction area L is defined, which surrounds B and contains the known pixels A and the inner and outer losses Bi and Bo.

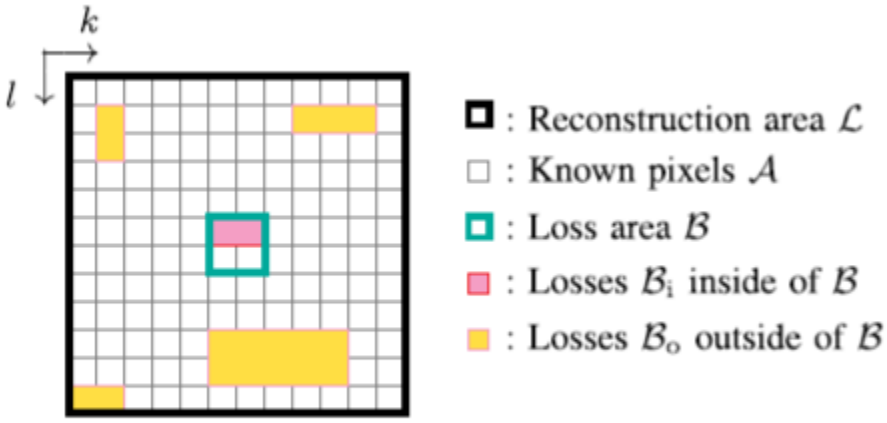


Fig. 13. Partitioning of the block-based reconstruction method for CAMSI.

Each reconstruction of a block B is treated as an independent problem and an arbitrary processing order can be chosen. We use the optimized processing order for each image Kk . The algorithm ensures that distortions are closed from the outer margin to the inside, so that the available support area is maximized. Moreover, the execution can be significantly fastened by processing unconnected losses in parallel, so that a higher computational efficiency is achieved.

After dividing the current distorted image Kk into a set of blocks B, the block-based reconstruction is performed.Therefore, the current distorted block B is written as vector s and the according reference blocks from the images Ki with i ∈ {0, . . . , K − 1} \ k are depicted as ri . The goal is to reconstruct the distorted vector by taking all undistorted references into account. For model generation, it is necessary to exclude the unknown samples from the distorted and the reference vectors. By allowing only undistorted entries of s, vector −s is obtained. The adjusted references ˜ri result after discarding samples, where s contains unknown entries.As reference views which contain distortions themselves, are unsuitable to reconstruct s, the set U is defined that only allows completely conserved reference views.We calculate a linear regression between every reference block and the distorted block to approximate with slope and offset scalars au and bu. These are determined by minimizing the squared model error of the linear regression model original samples remain unchanged.The block-based reconstruction is repeated until all distortions in every image are concealed. Thereby, already reconstructed pixels are used for the reconstruction of further pixels, but only in the same image. Thus, the reconstructions of different images are independent of each other and the reconstruction can be fully parallelized.

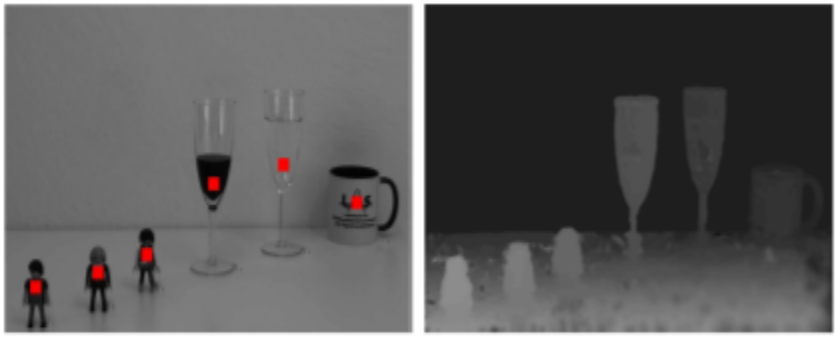
Chapter 4

Conclusions

4.1 Evaluation

To evaluate the performance of the proposed CAMSI system,several image sets and videos were recorded. Moreover,ground truth data was acquired for the images by sequentially mounting all utilized filters in front of the center camera. These references can be interpreted as Tunable Filter setups, e.g.,filter wheels, which provide a high spatial, but a restricted temporal resolution. Consequently, ground truth data can be provided for the images and not for the video data. The first image set NIR-RGB-UV Lab shows different objects,e.g., plants, liquids, and test patterns, in different depths that are placed on a table. For the recordings, the filter configuration NIR-RGB-UV from has been used.Furthermore, the second image set depicts different artificial plants, e.g., trees, orchids, which are captured under bad light conditions. For this, the filter set NIR-RGB-NIR from Tab. III was mounted. Both data sets consist of the images taken at the nine array positions as well as ground truth images recorded from the center position. The distance between camera and objects was between 2.3 and 3.0 meters. Moreover, a third data set using the configuration NIR-RGB-UV shows an outdoor video with buildings, forested areas and a cloudy sky. Indeed, only CAMSI images can be provided as the scene is dynamic so that ground truth is not anymore acquirable. The distances between camera and the recorded objects range between several meters and a few kilometers. In the following, a small part of the data set is shown. All recordings together with the CAMSI reconstructions are publicly available to invite researchers for participating in the proposed approach.At first, the registration performance of the novel CAMSI approach is compared to state-of-the-art cross-spectral disparity estimation techniques.

To measure the performance of disparity and depth estimation, an image set was recorded that contains six objects in varying depth levels while providing the ground truth camera distances for each object. The highlighted regions are used to calculate an average depth μz as well as depth deviation σz.This procedure is repeated by generating further disparity maps by applying two state-of-the-art cross-spectral stereo matching algorithms, namely Census + SGM and CCNG+SGM.



Object		Distance	Census [58] + SGM [64]	CCNG [47] + SGM [64]	CAMSI
1	μ_z	1305 mm	1234 mm	1257 mm	**1305 mm**
	$\pm\sigma_z$		$\pm$ 53 mm	$\pm$ 10 mm	**$\pm$ 0 mm**
2	μ_z	1355 mm	1279 mm	1318 mm	**1355 mm**
	$\pm\sigma_z$		$\pm$ 20 mm	$\pm$ 5 mm	**$\pm$ 2 mm**
3	μ_z	1420 mm	1357 mm	1392 mm	**1422 mm**
	$\pm\sigma_z$		$\pm$ 12 mm	$\pm$ 8 mm	**$\pm$ 0 mm**
4	μ_z	1495 mm	1499 mm	**1497 mm**	**1497 mm**
	$\pm\sigma_z$		$\pm$ 37 mm	$\pm$ 75 mm	**$\pm$ 1 mm**
5	μ_z	1585 mm	1463 mm	1563 mm	**1587 mm**
	$\pm\sigma_z$		$\pm$ 67 mm	$\pm$ 86 mm	**$\pm$ 22 mm**
6	μ_z	1615 mm	1553 mm	1641 mm	**1616 mm**
	$\pm\sigma_z$		$\pm$ 60 mm	$\pm$ 60 mm	**$\pm$ 0 mm**

It depicts the depth resolution accuracy for the two state-of-the-art methods and the proposed CAMSI algorithm. Apparently, CAMSI achieves a significantly higher accuracy and a low deviation. Due to the global cost aggregation, a robust disparity map is estimated that does not suffer from cross-spectral extinction.As all three approaches compare the structural similarity of image patches, no significant deviation regarding the overall performance of CAMSI is expected, which can be seen in TabA. Tab. B depicts the evaluation results for the simulated MSFA and the proposed CAMSI approach.

	PSNR	**SSIM**	**VMAF**	**BRISQUE**
Image set *NIR-RGB-UV Lab*				
Census [58]	32.95 dB	0.92	96.32	19.54 [13.37]
CCNG [47]	32.20 dB	0.91	96.25	19.93 [13.37]
CAMSI ZNCC (4)	**33.58 dB**	**0.93**	**96.44**	**19.48 [13.37]**
Image set *NIR-RGB-NIR Low Light*				
Census [58]	41.86 dB	0.95	97.16	29.95 [25.37]
CCNG [47]	39.27 dB	0.95	97.13	27.96 [25.37]
CAMSI ZNCC (4)	**42.04 dB**	**0.96**	**97.26**	**27.66 [25.37]**

PROCESSING TIME OF THE CAMSI IMPLEMENTATION WRITTEN IN MATLAB ON A DESKTOP COMPUTER AND A MOBILE NOTEBOOK

Algorithm	***Desktop***	***Notebook***
Registration	67.4 s	211.6 s
Reconstruction	4.8 s	18.3 s
Total	72.2 s	229.9 s

FInally, the computational complexity of the proposed CAMSI framework shall be discussed. As camera calibration has to be conducted only once, the processing time of CAMSI is dominated by registration and reconstruction. The cost metric is evaluated for every camera position Pk and all disparity candidates D, hence registration is more complex than solving regression tasks for the occluded and mispredicted pixels during reconstruction. Both registration and reconstruction mainly calculate mean, variance, and covariance values of image patches, so that an efficient implementation could be derived by using integral images. Additionally,more advanced techniques can be applied to lower the computation time even further or the algorithms can be implemented on GPU . For the sake of an intuitive implementation,CAMSI was written in MATLAB without the usage of integral images, or any further optimization. To demonstrate the computational complexity, the CAMSI framework was evaluated on the test system Desktop, which encloses an Intel i9-7940X CPU and 64 GB RAM.

Additionally, the test system Notebook was used that is equipped with a i7-6700HQ CPUand 16 GB RAM. The mean computation time averaged over 100 executions and the specified camera resolution. The algorithms have been parallelized to exploit the multi-core computer architecture.

4.2 Conclusion

In this contribution, a novel multi-spectral imaging system is presented. This becomes necessary as a review on state-of-the-art multi-spectral imaging techniques has revealed that none of the existing approaches is capable of capturing videos with a high spatial, temporal, and spectral resolution at the same time. Often, the dynamic application of filters is equally difficult as designing a handy setup, which is of compact size. A novel approach that can remedy the existing challenges is the proposed camera array for multi-spectral imaging,which provides additional depth information. However, this comes at the price of recording differing spectral images at various viewpoints, such that algorithmic post-processing is required after acquisition. By introducing geometrically constrained multi-camera sensor layout, the formulation of novel registration and reconstruction algorithms becomes possible to globally set up robust models for generating consistent multispectral videos. The performance of the novel camera array for multi-spectral imaging is analyzed by giving an extensive visual and objective evaluation. In comparison to recently published multispectral filter array imaging systems, the novel acquisition approach achieves an average gain of 2.5 dB PSNR. Moreover, the recorded data sets are provided online together with the reconstructed images. In the future, it will be evaluated whether the novel concept is expandable to hyper-spectral imaging by extending the geometrically constrained multi-camera sensor layout.

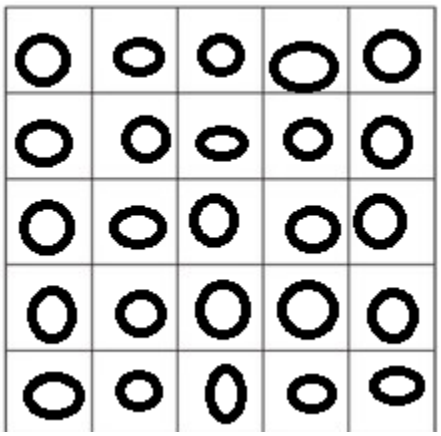
4.3 FUTURE SCOPE

In this paper the camera used was basler ace A1600-60GM.Many advancements has been done in this camera like the given below:

Basler Ace acA640-90gm (CS-Mount)	659 px x 494 px	VGA	ICX424	90 fps	Mono	GigE
Basler Ace acA640-120gm	659 px x 494 px	VGA	ICX618	120 fps	Mono	GigE
Basler Ace acA640-120um	659 px x 494 px	VGA	ICX618	120 fps	Mono	USB 3.0
Basler Ace acA640-120gm (CS-Mount)	659 px x 494 px	VGA	ICX618	120 fps	Mono	GigE
Basler Ace acA640-121gm	659 px x 494 px	VGA	ICX618 Replacement	134 fps	Mono	GigE
Basler Ace acA720-290gm	720 px x 540 px	VGA		291 fps	Mono	GigE
Basler Ace acA720-290gc	720 px x 540 px	VGA	IMX287	291 fps	Color	GigE

All this camera has better resolution and features than the one here.so if we use any one camera from this it will produce a better output.

The sensor layout can be easily extended to obtain hyper-spectral images by using 5 × 5, 7 × 7 or even larger array grid sizes.



REFERENCE

[1] Nils Genser,Jurgen Seiler,Andre kaup: Camera array for multispectral imaging.,IEEE Transactions onImage Processing,vol 29 2020.

[2] S. Mendis, S. E. Kemeny, and E. R. Fossum, “CMOS active pixel image sensor,” IEEE Trans. Electron Devices, vol. 41, no. 3, pp. 452–453,Mar. 1994.

[3] J. Brauers, N. Schulte, and T. Aach, “Multispectral filter-wheel cameras:Geometric distortion model and compensation algorithms,” IEEE Trans.Image Process., vol. 17, no. 12, pp. 2368–2380, Dec. 2009.

[4] D. Knipp, H. Stiebig, S. R. Bhalotra, E. Bunte, H. L. Kung,and D. A. B. Miller, “Silicon-based micro-Fourier spectrometer,” IEEE Trans. Electron Devices, vol. 52, no. 3, pp. 419–426, Mar. 2005.

[5] F. Kazemzadeh, S. A. Haider, C. Scharfenberger, A. Wong, and D. A. Clausi, “Multispectral stereoscopic imaging device: Simultaneous multiview imaging from the visible to the near-infrared,” IEEE Trans.Instrum. Meas., vol. 63, no. 7, pp. 1871–1873, Jul. 2014.

[6] Longoni, F. Zaraga, G. Langfelder, and L. Bombelli, “The transverse field detector (TFD): A novel color-sensitive CMOS device,” IEEE Electron Device Lett., vol. 29, no. 12, pp. 1306–1308, Dec. 2007