A Paper on Camera Array and Its uses

Neeraj Kaushik
Department of Electronics and Communication Engineering
Faculty of Engineering, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India

ABSTRACT: Camera array is defined as the arrangement of multiple cameras in a row so as to acquire better image quality and performance of camera. The reduced cost of cameras and the complex scenes make it possible and necessary to replace the monocular camera with camera array. In this paper, we first include a description of the current camera arrays and group them according to the array structure, then offer an overview of the image properties that benefit from the camera array including dynamic range, resolution, viewing by occlusions and estimating depth. Finally, a novel camera array-based airborne optical system is proposed to meet the requirements for wide field of view (FOV) security, surveillance and reconnaissance (ISR), high dynamic range and resolution, multi-view and multi-dimensional imaging. Some key technologies that need to be developed in future research, such as self-calibration and synthetic aperture imaging by camera arrays on mobile platforms, are highlighted for the sake of onboard implementation of this device in operation.

KEYWORDS: Camera arrays, pictures, videos, high resolution, spatiotemporal sampling, synthetic aperture.

INTRODUCTION

As a result of technical development, the size and complexity of network networks has grown dramatically over the past few decades. This effect is evident in the field of surveillance systems where the number of closed-circuit television (CCTV) cameras is constantly increasing, especially around critical safety areas such as stadiums, airports, road networks and railway stations. Using these devices, it is possible to respond rapidly or even anticipate situations where crowd health may be jeopardized [1]. However, to do so effectively, it is typically a human operator's task to recognize such situations and alert relevant authorities. Since human operators are restricted in processing several images simultaneously, automating this process is of interest and CCTV cameras have threats identified and reported automatically. We have built a test bed towards this end to benchmark autonomous patrolling strategies where the aim is to catch a smart evader.

One of the semiconductor industry's economic tenets is products which sell in large volumes are inexpensive, whereas products which sell in smaller volumes are more costly, almost regardless of the part's complexity. For computers, this relationship has altered people's thinking about building high-end systems; instead of constructing a custom high-end processor, using a large number of generic processors is more cost-effective [2].

We see similar patterns in digital imagery now. While the popularity of digital cameras increases, low-end imagers' efficiency continues to increase, while the high-end cameras' costs remain fairly constant. Researchers have also shown that multiple images of a static scene can be used to extend those cameras' output envelope. Examples involve image production with higher resolution or dynamic range. In other research, Schechner and Nayar used spatially different filters on a revolving camera to create high-resolution panoramas that also had a high dynamic range or high spectral resolution. Another use for multiple views is interpolation of view to create the impression of a smoothly moving virtual camera in a static or dynamic scene.

Most of these efforts employ a single high-quality, moving camera that views a static image. Multiple cameras are required to get similar results on complex scenes. In 1999, this inspired us to consider developing a compact array that included a large number of inexpensive video imagers.

The resulting multiple camera array comprises 100 video cameras, each attached to its own processing unit. The processing boards are capable of computing local images, as well as compressing MPEG2.

Early camera arrays

The earliest systems used a single translating camera to capture scenes from different viewpoints, and were limited to static scenes. In using a linear array of still cameras, Dayton Taylor applied the concept to a dynamic picture. He created the illusion of virtual camera movement through a frozen dynamic scene by activating the cameras simultaneously and jumping from one camera image to the next. Manex Entertainment used wider spaced cameras and introduced an adjustable trigger delay between cameras to capture photographs that suited a simulated high-speed camera lying around their scenes. Each of these systems used still cameras, and they were restricted to capturing a particular trajectory of a virtual camera through space and time defined by the camera setup [3].

The researchers switched to arrays of video cameras to capture a more general data collection. Video cameras need to be synchronized like still cameras but they also pose a new challenge: big data levels. The Visualized Reality TM project is the groundbreaking multiple video camera array concept. Our goal was to catch several views of an interpolation scene for video sharing. Their system's new version records video using VCRs, giving those nearly limitless recording times but poor quality. Their second edition uses 49 video cameras that capture main memory from a Laptop. This method is of higher quality (VGA resolution at 30 frames per second), but is limited to capture durations of nine seconds [4]. Growing third camera captures the video in color. For each three cameras they need one PC to manage the bandwidth of the video cameras.

Although the project Virtualized Reality TM uses fairly high-quality cameras, two other organizations have been working with vast numbers of cheap cameras. First community distributed Light Field Camera allows live dynamic light fields from an 8x8 array of webcams for commodities. The Self-Reconfigurable Camera Array of the second category uses 48 commodity Ethernet cameras with horizontal electronic translation and pan controls to maximize the interpolation performance. While the nature of these devices makes them much cheaper in terms of the cost per camera than Virtualized Reality TM, major sacrifices have been made in using these commodity cameras. Second, none of the arrays could be aligned, triggering reconstructions of artefacts in the vision [5]. Furthermore, neither device tackled the bandwidth problems of constructing a general purpose large camera array because they were looking at single applications. First group has opted to incorporate a finite-view scheme, meaning that each camera only transmits enough data to reproduce a limited number of light field views per frame time. Second group cameras use JPEG compression, but their choice of Ethernet and a single device to run the array limits them to 320x240 pixel resolution at 15-20 frame per second.

LITERATURE REVIEW

This paper presents a calibration system centered on a flat lenticular array producing a color-coded light-field whose observed color varies depending on the angle it is viewed from. We derive an approach for estimating the camera's focal length and relative pose from a single image of an object. We describe camera calibration output across different focal lengths and camera types, and show the benefits of focal length estimation when rendering a virtual object in a video with continuous zooming [1]. This paper explores aspects of modeling sensors and photogrammetric calibration, focusing emphasis on automated self-calibration techniques. After an initial review of the history and state of the art, selected topics of current interest for close-range photogrammetry are discussed within the calibration framework. These involve modeling of sensors, with a description of standard, extended and common calibration models and non-traditional camera systems.

Self-calibration is then explored in both targeted planar arrays and targetless scenes that are appropriate for SfM-based exterior orientation, after which aspects of calibration and measurement accuracy are addressed [2]. This paper presents an error metric with improved stitching sensitivity in regions with a defined structure. Use this metric, our system efficiently seeks optimum ordering of pair-wise warps with minimal parallax artifacts for robust stitching. In non-overlap areas, weighted extrapolation of warps maintains temporal consistency while at the same time preventing visual discontinuities during transitions between views. The residual global deformation caused by the warps is distributed over the entire panorama domain using restricted relaxation, thus residual as close to the original views of the data as possible [3]. This paper maintains and stabilizes filter effects while remaining agnostic to the filter's inner workings. It captures filter effects in the gradient domain, and then uses gradients of the input frame as a guide to enforce continuity in time and space. Our formulation of the least squares adds minimal overhead compared with the processing of naive results.

In addition, when filter costs are high, we implement a filter transfer strategy that decreases the number of filtering computations per frame by an order of magnitude, with only a small reduction in visual quality. We demonstrate our algorithm on multiple formats of camera array including stereo images, light fields and small baselines [4]. Present a high-speed video algorithm using a camera array with good perceptual consistency in realistic scenes that may be cluttered and have a complex context. We synchronize the cameras in such a way that they each record an image at a different offset time. The algorithm processes the interleaved jittery frames and generates a video that is stabilized. Our approach consists of: synthesizing views from a virtual camera to compensate for cameras perspective discrepancies, and video compositing to remove remaining objects, especially around disocclusions.

More specifically, we process the raw video's optical flux to measure the difference to the target virtual frame for each raw frame [5]. This paper introduces Light Field Messaging (LFM), a method of embedding, transmitting, and receiving hidden video information that is projected on a screen and captured by a camera on handheld. The system's goal is to reduce perceived visual artifacts of the message embedding, while at the same time optimizing the accuracy of the camera-side message recovery. LFM includes photographic steganography for the embedding of display and camera-captured messages. Unlike digital steganography, the criteria for embedding are substantially more difficult due to the combined effect of the radiometric emission function of the lens, the sensitivity function of the camera and the camera display. This paper proposes a new method for combining the spherical lens and planar sensors effectively, eliminating costly and bulky bundles of fibre. To simulate a wide-FOV multi-sensor scenario, we create a single sensor LF camera prototype, rotating the sensor relative to a fixed principal lens. Finally, we define a toolchain for processing, including a handy spherical LF parameterization, and display depth estimation and post-capture refocus for indoor and outdoor panoramas with 1515 u1600 ubiquitous 200 pixels (72 MPix) and 138 ° FOV [6]. This paper poses technological problems and a study of the different solutions. It underlines the relation and incorporation of different modules in various environments and scenarios for application.

Some problems can be solved together to increase performance and accuracy according to the most recent works. With the rapid growth of surveillance systems, camera network sizes and complexities are growing and the surveyed environments are becoming increasingly complicated and crowded. This paper examines how to deal with these emerging challenges [7]. Stereo videography is a effective tool for quantifying animal kinematics and behaviour, but it can be difficult to use when setting outdoor fields. We present here a workflow and related tools to perform calibration of cameras positioned in a field setting and to estimate the accuracy of the resulting stereoscopic reconstructions. The workflow is illustrated by example stereoscopic reconstructions of bat and bird flight. We provide software tools to plan experiments and process the resulting calibrations that could be used by other researchers to calibrate their own cameras. Our field protocol can be implemented in one afternoon, requiring only short video clips of light, portable devices with calibration [8].

METHODOLOGY

The cameras in the array should be positioned with each camera's optical axis in a certain pattern, with an equivalent baseline and angel. The camera array can be divided into the parallel arrangement and radial arrangement based on the geometric pattern.

Parallel Arrangement

Cameras in the series are arranged by parallel arrangement in such a way that cameras' optical axes are identical to one another. It includes linear array (1D) and area array (2D) for acquisitions of multi-view images, as shown in Fig. 1. Rich overlapping FOV and multiple parallaxes created by baselines between each optical axis are presented in a camera array in parallel arrangement, generating the corresponding depth maps. The parallel camera array can be categorized as the single projection center and multiple projection centres, depending on the specific baseline range between each camera.

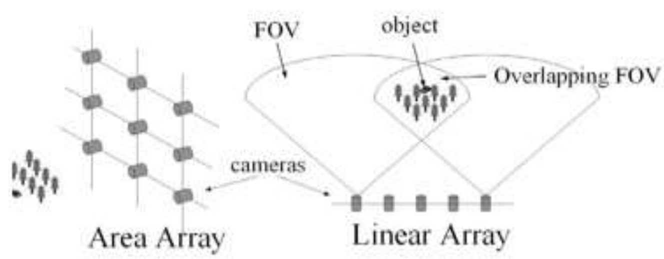

FIGURE 1: The parallel arrangement of camera array

Single projection center

The single projection center denotes the cameras are closely patched together to approximate being aligned to a specific perspective. The objective is, on the one hand, to acquire enhanced imagery with high resolution, high signal-noise ratio (SNR) and high dynamic range rendering in the large FOV overlap. On the other hand, this type is for imagery with synthetic aperture. Baker and Tanguay have built a network of up to 24 cameras with attractive FOV for high-performance photography, whose images are mosaiced yielding video with higher pixel resolution than the individual object. A prototype super-resolution camera array system consisting of 12 low-cost Web camera systems to achieve a S / N ratio of approximately 2 dB. Bennett created a larger scale implementation of the system using 96 cameras to capture the 7 megapixel video, and they also expanded the dynamic range by trading off the field of view and using multiple exposures through their cameras. Next they used 52 wide-angel cameras with complete FOV overlap with 30 staggered frames per second to achieve a 1560fps high-speed video.

Their network, however, used numerous workstations and their infrastructure, such as camera grid, interconnections and workstations, was huge, making them less suitable for on-site tasks. First, a work introduced a compact 3x3 camera array system of Pointgray Flea2 cameras to obtain dynamic fluid surfaces and used a high-speed imaging scheme similar to the concept that the exposure time at each camera was interleaved. The device was able to operate at 60fps, provided that each camera had a fixed frame rate of 30fps. For synthetic aperture imaging that has gained tremendous attention from researchers in the last two decades. Multiple projection centers Multiple projection centers suggest that cameras are spread over a wide baseline to capture the scene from several different points of view to provide multi-view details to stitch a large FOV (with small overlapping camera array FOV) or three-dimensional imagery (with large overlapping camera array FOV). Despite this, in the radial system the distinguishing properties of several projection centers can be apparent.

Radial Arrangement

Cameras in the array are arranged by radial arrangement to form a circle (2D) or hemisphere (3D) for multi-view acquisition of images [10]. Radial arrangement can be further divided into two types according to the directions of the cameras' optical axes: toed-in camera array and toed-out camera array, as shown in figure 2.

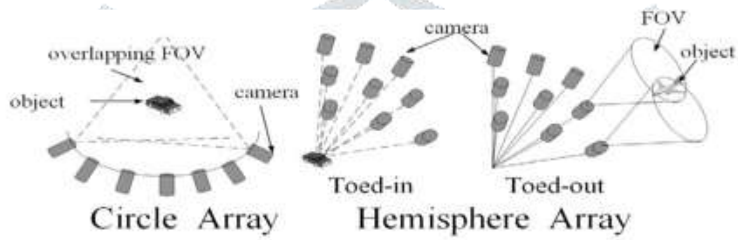

FIGURE 2: The radial arrangement of camera array

Toed-in camera array

Inside the toed-in camera array, the optical axis of each cam- era converges to a common level. In addition to the bullet time in cinematography, another use of this type is virtual reality (VR) where a large-scale camera array was used to capture 4D dynamic scenes and the virtual scene reconstruction was done by machine vision. They also addressed seven important features and criteria of the toed-in camera array for digitizing complex events: scalability, non-intrusiveness, naturalness, high quality, unhindered immersion, tele or post experience, active participation.

FIGURE 3: Toed-in camera array.

Toed-out camera array

The toed-out camera array method has been introduced in recent years, in which cameras' optical axes diverge from a specific stage. Consequently, it is often called a monocentric camera array. The camera array of dozens or tens of cameras grouped in a circle is primarily for omnidirectional imaging, for example. Gopro 360 camera series, or large FOV photography, e.g. camera array consisting of hundreds or thousands of cameras arranged in a hemisphere is primarily for large FOV, high resolution, and high dynamic range, as AWARE-2 will independently monitor each camera's exposure time to increase dynamic range. This device designed as part of the U.S. Defense Advanced Development Projects Agency AWARE program stands out as the notable milestone in the hemisphere shape, with a resolution of 120 by 50 FOV and 0.96 gig pixel. This device has numerous applications in large field microscopy, event recording, continuous tracking and spatial awareness.

CONCLUSION

While much work has been done in expanding the camera array arrangement for special tasks and exploring its potential properties, at the current stage the camera array mounted on the mobile platform like UAV is rarely involved. Some of the major barriers that impede this application are real-time camera array calibration when moving and synthetic camera aperture imaging. This paper presents a report on recent camera array research developments like device structure, array configuration, and proper output relations, which are main impact factors for their application. Then a new airborne optical system based on camera array was implemented and the corresponding main problems were also highlighted that need to be solved in the future, which will greatly facilitate the development of the next generation of airborne detection system.

REFERENCES

[1] K. M. O'Connor, L. R. Nathan, M. R. Liberati, M. W. Tingley, J. C. Vokoun, and T. A. G. Rittenhouse, "Camera trap arrays improve detection probability of wildlife: Investigating study design considerations using an empirical dataset," *PLoS One*, 2017, doi: 10.1371/journal.pone.0175684.

[2] T. Luhmann, C. Fraser, and H. G. Maas, "Sensor modelling and camera calibration for close-range photogrammetry," *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016, doi: 10.1016/j.isprsjprs.2015.10.006.

[3] F. Perazzi *et al.*, "Panoramic Video from Unstructured Camera Arrays," 2015, doi: 10.1111/cgf.12541.

[4] N. Bonneel *et al.*, "Consistent Video Filtering for Camera Arrays," *Comput. Graph. Forum*, vol. 36, no. 2, pp. 397–407, 2017, doi: 10.1111/cgf.13135.

[5] J. V. Dueholm, M. S. Kristoffersen, R. K. Satzoda, T. B. Moeslund, and M. M. Trivedi, "Trajectories and maneuvers of surrounding vehicles with panoramic camera arrays," *IEEE Trans. Intell. Veh.*, vol. 1, no. 2, pp. 203–214, 2016, doi: 10.1109/TIV.2016.2622921.

[6] D. G. Dansereau, G. Schuster, J. Ford, and G. Wetzstein, "A wide-field-of-view monocentric light field camera," 2017, doi: 10.1109/CVPR.2017.400.

[7] A. S. Glen, S. Cockburn, M. Nichols, J. Ekanayake, and B. Warburton, "Optimising Camera Traps for Monitoring Small Mammals," *PLoS One*, 2013, doi: 10.1371/journal.pone.0067940.

[8] D. H. Theriault *et al.*, "A protocol and calibration method for accurate multi-camera field videography," *J. Exp. Biol.*, 2014, doi: 10.1242/jeb.100529.