

Journal of Real-Time Image Processing: first issue of volume 11

Nasser Kehtarnavaz¹ · Matthias F. Carlsohn²

Published online: 16 December 2015
© Springer-Verlag Berlin Heidelberg 2015

The Journal of Real-Time Image Processing is entering its 11th year of publication with its stature in the area of image processing steadily improving as evident by the increase in its impact factor to 2.02 last year.

We begin this first issue of volume 11 in 2016 noting that there will be two volumes appearing in 2016, each volume consisting of 4 issues, some of which may appear as double special issues. In other words, the good news is that there is going to be a total of 1600 print pages of accepted papers that will appear in 2016. The additional volume 12 in 2016 is arranged for the purpose of practically removing the backlog of accepted papers that have already appeared as Springer online-first articles but are waiting to appear in print. Our hope is that this would address one of the major challenges we have faced in the last few years related to the page budget or limitation aspect of print issues.

As normally done in the last few years, the editorial board of JRTIP will be meeting at the upcoming SPIE Conference on Real-time Image and Video Processing in April 2016 in Brussels. The conference program is included in the back matter of this issue. We will report on the outcome of this meeting in an editorial appearing after the conference.

This first issue of volume 11 comprises a total of 19 papers with 18 papers as original research articles and one

paper as a survey article on stereo image processing. Also, a one-page erratum to a paper is included immediately after the appearance of the paper. Four themes are noted in these papers which involve: real-time stereo and multi-view image processing (2 papers), real-time image processing via dedicated architectures (4 papers), real-time implementations (6 papers), and real-time performance of various applications (7 papers). A brief summary of these papers is outlined below which is ordered by these themes.

The first paper by Tippetts et al. is a survey paper which provides a comprehensive review of stereo vision algorithms with the focus placed on real-time performance for resource-limited systems. Accuracy and runtime performance of various stereo vision algorithms developed in the past decade are compiled and presented in this survey paper. The algorithms are grouped into three categories: (1) those having real-time or near real-time performance on standard processors, (2) those having real-time performance on GPU, FPGA, DSP, and ASIC, and (3) those not having been performed in real-time.

The second paper by Pan et al. deals with a related subject by introducing a fast mode decision algorithm to reduce the computational complexity of multi-view depth video coding based on the mode correlations between depth video and its corresponding texture video, motion prediction, and coded block patterns. The experimental results reveal that the introduced algorithm can achieve about 67 to 69 % encoding time saving for even and odd views, respectively, while maintaining a comparable rate-distortion performance.

The third paper by Ghosh et al. presents an efficient VLSI architecture of a hierarchical block matching algorithm for motion estimation. Simulation results show that the designed architecture is more area efficient and faster than several existing full-search, three-step-search and

✉ Nasser Kehtarnavaz
kehtar@utdallas.edu

✉ Matthias F. Carlsohn
Matthias.Carlsohn@t-online.de

¹ University of Texas at Dallas, Richardson, USA

² Engineering and Consultancy for Computer Vision and Image Communication, Bremen, Germany

multi-resolution architectures. Noting that this architecture requires only two-port memory, which is common in consumer electronics systems, it can easily be integrated into an existing system at the expense of a small increase in the chip area. An erratum to this article appears at the end of this paper as a correction was provided by the authors after its online-first publication.

The fourth paper by Humayun et al. compares various shape-from-focus methods, which are normally time consuming and occupy a considerable amount of memory, within the context of a parallel computing environment. The speedups of various focus-measuring methods are analysed using different numbers of cores for the purpose of determining the optimal number of cores for these methods.

The fifth paper by Saidani et al. proposes a fast and efficient VLSI hardware architecture design of context formation for Embedded Block Coding with Optimal Truncation EBCOT module in JPEG 2000. A high-speed parallel bit plane coding (BPC) hardware architecture for the EBCOT module is designed and implemented. The experimental results indicate that the developed design outperforms well-known techniques with respect to the processing time, leading to 70 % reduction when compared to bit plane sequential processing.

The sixth paper by Hoffman et al. presents a high-throughput hardware implementation of the context-based adaptive variable length coding (CAVLC) encoder. The scanning solution presented determines all the required data for the encoding phase in a minimized and constant number of clock cycles. This modified scanning phase approach is shown to offer significant throughputs for CAVLC; for example, at 200 MHz the architecture is capable of encoding 1,080 p video files at 95 fps.

The seventh paper by Melo et al. addresses a real-time software-based solution for correcting the radial distortion of HD video on a personal computer equipped with a conventional graphics processing unit (GPU) and a video acquisition card. It involves acquiring the video data directly from an endoscopic camera control unit, and warping each frame using a heterogeneous parallel computing architecture. The paper shows that this heterogeneous approach together with the utilized efficient memory access patterns in the GPU improved the performance leading to frame rates of more than 250 fps.

The eighth paper by Kumar et al. describes parallel implementations of video object detection algorithms such as Gaussians mixture model (GMM) for background modelling, morphological operations for post-processing, and connected component labelling (CCL) for blob labelling. Both parallelization strategies and optimization techniques are deployed to exploit the computational capacity

of CUDA cores on GPUs. Experimental results indicate that the developed parallel GPU implementations achieve significant speedups of about 250 times for binary morphology, about 15 times for GMM and about twice for CCL when compared to sequential implementations running on the Intel Xeon processor.

The ninth paper by Szwoch et al. presents the outcome of background subtraction algorithms implemented on a supercomputer platform. The choice of the algorithm, a number of threads and a task scheduling method are utilized together to provide both accuracy and efficiency for real-time processing of high-resolution camera images. The experiments were performed on a supercomputer cluster using a single machine with 12 physical cores. The accuracy and performance of the implementations were evaluated for varying image resolutions and numbers of concurrent processing threads.

The tenth paper by Köstler et al. discusses an efficient multi-grid algorithm to solve the problem of inverse transformation from gradient space. A comprehensive performance analysis is conducted to derive a performance model for the multi-grid algorithm generating an improved implementation with an overall performance of more than 25 frames per second for 16.8 M images at full high dynamic range compression including data transfers between CPU and GPU.

The eleventh paper by Márquez-Neila et al. introduces a procedure for reducing the number of samples required for fitting a homography to a set of noisy correspondences via a random sampling method. This procedure uses a geometric constraint that detects invalid minimal sets. In the experiments reported, it is shown that this constraint not only reduces the number of random samples at a negligible amount of computation but also balances the processor workload over time preventing visual stalls that are observed on mobile devices.

The twelfth paper by Nieto et al. covers a real-time lane modelling method using an efficient design and implementation of particle filtering. The method is used to determine the position of a vehicle inside its own lane and the curvature of the road ahead as an advanced driver assistance system. The effectiveness of the method has been demonstrated by implementing a prototype and testing it on road sequences with different illumination conditions, pavement types, and traffic density at processing times below 2 ms per frame for laptop CPUs, and 12 ms for embedded CPUs.

The thirteenth paper by Jung et al. proposes a Korean–English bilingual videotext recognition method which is computationally efficient and achieves recognition accuracy comparable to existing methods. This method uses a split–merge strategy by merging split segments into characters while avoiding unnecessary computation using

geometric features. Both the efficiency and effectiveness of the proposed method are verified by experiments on a challenging database containing 51,290 text images (176,884 characters).

The fourteenth paper by Dubská et al. discusses how the Hough space is used as a 2D signal instead of just detecting and processing local maxima. It is shown that this approach is computationally efficient, robust, and accurate. The proposed approach is then applied to detect and localize matrix codes (e.g. QRcode, Aztec, DataMatrix) and chessboard-like calibration patterns in an efficient and accurate manner.

The fifteenth paper by Burian et al. deals with the design and the implementation of an image recognition system based on FPGA. It explores an n -tuple methodology and the organization of the neural networks data on an n -tuple memory. The system was tested on recognition of road signs. The test results reveal that the designed system can serve as part of more complex recognition systems.

The sixteenth paper by Liu et al. presents a real-time and robust approach to recognize two types of gestures consisting of seven motional gestures and six finger spelling gestures. This approach utilizes stereo images captured by a stereo webcam to achieve robust recognition under realistic lighting conditions and in various backgrounds. The results obtained indicate that high recognition rates under realistic conditions are obtained in real-time on PC platforms at the rate of 30 frames per second.

The seventeenth paper by Yamada et al. improves a previously developed system for embedding watermarks in video content by incorporating real-time transcoding. The prototype testing of the system demonstrated the feasibility of the developed video-on-demand service by streaming up to 20 individually watermarked videos. In addition, it was shown that the embedded watermarks were robust against encoding while the quality of the watermarked images was mostly maintained.

The eighteenth paper by Jeon et al. covers a robust fuzzy-bilateral filtering method and its application to video deinterlacing. The proposed bilateral filtering method utilizes range and domain filters based on a fuzzy metric. This approach is adaptively applied to both existing pixel activity and the associated positions between existing neighbouring pixels. The simulation results indicated that the proposed method can efficiently interpolate the interlaced field while enhancing details.

The last or nineteenth paper by Zivkovic discusses modifications of the iterated conditional mode (ICM) algorithm to achieve a so-called gentle descent in the optimization framework during the first iterations of the algorithm leading to a significant improvement in performance. It is shown that the modified ICM is computationally competitive to similar optimization frameworks for a set of vision problems such as stereo depth estimation, image segmentation, image denoising, and inpainting.