



Real-time image processing for augmented reality on mobile devices

Zhihan Lv¹ · Jaime Lloret² · Houbing Song³

Published online: 23 March 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

In the last few years, remarkable progress has been made in mobile consumer devices. Modern smartphones and tablet computers offer multi-core processors and graphics processing units, which have opened up new application possibilities such as augmented reality, virtual reality, and 3D reconstruction. Augmented Reality (AR) is a key technology that is going to facilitate a paradigm shift in the way users interact with data and has only just recently been recognized as a viable solution for solving many critical needs.

1 Introduction

The AR technology can be used to visualize data from hundreds of sensors simultaneously overlaying relevant and actionable information over environments through a headset. Mobile devices make the AR technology easily accessible in the field. With easier and more accessible use, a variety of different applications (besides video gaming) become more feasible. In summary, AR is an upcoming technology which is going to have a major impact on many applications and over a wide type of technologies [1]. Currently, conducting AR is time consuming on mobile devices. This special issue is meant to address the implementation or real-time aspects of AR. The components in an AR system include color consistency, feature tracking, bundle adjustment, camera (self-) calibration, stereo matching, and multi-view stereo. Moreover, the high-level applications based on augmented reality,

e.g. smart city, simulation, and industrial engineering, would require additional computational time.

This special issue aims to offer the latest results addressing real-time image processing techniques for augmented reality. The thrust of this special issue is on the real-time aspects of augmented reality.

2 Accepted papers

The paper entitled “High-speed real-time augmented reality tracking algorithm model of camera based on mixed feature points” [2] aims at the balance between real-time performance and stability, proposes a new method model of real-time camera motion tracking based on mixed features. This model comprehensively utilizes feature points and feature lines as scene features, and combines multiple methods to construct hybrid features and performs parameter estimation. Also mentioned in the paper are image feature optimization methods based on scene structure analysis, and use of iterative filtering of feature lines to obtain a stable feature line set, which can adaptively construct hybrid features based on the composition of scene features and geometric features, thus improving the tracking stability of the camera. Experiments done based on this approach show that the proposed target detection and tracking algorithm has good real-time performance and robustness, and improves the success rate of target detection and tracking.

The paper entitled “A Cylindrical Shape Descriptor for Registration of Unstructured Point Clouds from Real-Time 3D Sensors” [3] proposes an unstructured point cloud registration method to process the data set from the real-time

✉ Zhihan Lv
lvzhihan@qdu.edu.cn

Jaime Lloret
jlloret@dcom.upv.es

Houbing Song
h.song@ieee.org

¹ School of Data Science and Software Engineering, Qingdao University, Qingdao, China

² Instituto de Investigación para la Gestión Integrada de Zonas Costeras (IGIC), Universitat Politècnica de Valencia, Valencia, Spain

³ Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA

3D sensor of RGB-D or TOF camera. A VTP-ISC descriptor is developed, which can significantly reduce the matching error of the twist and fold positions on the grid. In addition, the proposed three-dimensional cylindrical shape descriptor shows the latest results of unstructured point cloud registration. This method is shown to achieve satisfactory results in point cloud registration, and exhibits significant improvements in the standard benchmark.

In the paper entitled “Detection of real-time augmented reality scene light sources and construction of photorealistic rendering framework” [4], the main network of the multi-channel light source is improved and a multi-channel image fusion for joint training is realized. Then, in order to avoid the influence of batch reduction on the model distribution, a regularization method is used to train small batches. Combined with the method of regional candidate network, the final detection accuracy and the accuracy of candidate frame regression are improved. Finally, a variety of image enhancement techniques are used for rendering. The results obtained through experiments not only achieves good lighting and shadow effects, but also meets the realistic rendering of frames in more complex scenes. The experimental results also show that the virtual light source automatically generated by the algorithm can approximate the lighting of the real scene. Under the illumination by one or more real light sources, the virtual object and the real object can produce approximately the same lighting effect in an augmented reality environment.

In order to solve the problem of human errors caused by design deviations in the current mobile augmented reality library document pushes, the paper entitled “Mobile Augmented Reality Technology for Design and Implementation of Library Document Push System” [5] proposes a new data visualization and interactive design process. Based on the analysis of user needs, the augmented reality library document push is compared with the traditional library document push, and the demand is transformed into an application. A combination of overview and details of visual elements are considered. The system test results show that the library document push system designed in this paper can effectively reduce the memory consumption of the system, increases the refresh frame rate of the view, and improves the user’s interaction fluency.

For effective elimination of tracking errors caused by moving targets, an augmented reality interaction method based on the improved Kinect sensor command is proposed in the paper entitled “Error Elimination Method in Moving Target Tracking in Real-time Augmented Reality” [6]. This paper designs an action tracking architecture. A previous algorithm to solve the pose together with the optical flow method to track the feature points are used to avoid the operation of extracting, matching and removing

the mismatch of the feature points of each frame, and thus improving the speed of the algorithm.

The paper entitled “Network Algorithm Real-time Depth Image 3D Human Recognition for Augmented Reality” [7] studies the application of augmented reality real-time depth image technology in 3D human motion recognition technology. The accuracy and real-time performance of sensor-based 3D human reconstruction are affected by visual characteristics and illumination changes. It is challenging to extract features while tracking. This leads to faults in the three-dimensional reconstruction of the human body. Based on the relationship between the depth image and the distance and reflectivity, a model for correcting the distance error and reflectivity error of the depth image is established to improve the accuracy of the depth image, and finally the overall accuracy of the three-dimensional reconstruction of the human body.

The paper entitled “Using Augmented Reality and Deep Learning to enhance Taxila Museum Experience” [8] envisages to enrich the user’s museum experience through relevant multimedia information and establishes better connections with well-preserved cultural relics. A smartphone application based on augmented reality (AR) is proposed, which uses deep learning to recognize artifacts in real-time and retrieve supporting multimedia information for visitors. Use of Convolutional Neural Network (CNN) to correctly identify artifacts provides users with accurate content. Through user-centered questionnaire surveys, it is compared with traditional human guidance or free user visits. A comprehensive statistical evaluation and discussion are provided.

The paper entitled “Gradient Information Distillation Network for Real-Time Single Image Super-Resolution” [9] proposes a gradient information distillation network. On one hand, the advantages of fast and light-weight are maintained through information distillation. On the other hand, gradient information is used to improve the Super-Resolution performance. The network mainly includes two branches, the gradient information distillation branch (GIDB) and the image information distillation branch (IIDB). A residual feature transfer mechanism (RFT) is introduced to combine the features of the two branches. With the help of GIDB and RFT, the network mentioned in the paper can retain rich geometric structural information, making the edge details of the reconstructed image more clear. Experimental results show that this method is better than existing methods while restricting model parameters, computational complexity and run time accordingly. It provides the framework for its real-time and mobile applications.

In order to realize the three-dimensional reconstruction of dynamic objects, the paper entitled “Multi-view data capture for dynamic object reconstruction using handheld augmented reality mobiles” [10] proposes a system that obtains almost synchronized frame streams from multiple

mobile handheld devices. The system uses a decentralized trigger strategy and a data relay architecture to collect frames and movement gestures in real-time. The architecture can be deployed by edge computing or in the cloud. The paper demonstrates the effectiveness of the system and uses it for 3D skeleton and volume reconstruction. The trigger strategy achieves the same performance as the synchronization method based on Network Time Protocol (NTP), but provides higher flexibility. A challenging new data set, named 4DM, is introduced in the paper which involves 6 handheld augmented reality phones that record people's activities performing outdoor sports.

In the paper entitled “Human Motion Tracking and Positioning for Augmented Reality” [11], an effective human motion feature extraction and descriptor algorithm is discussed based on the characteristics of smart terminals' weak processing capabilities and limited hardware resources. More specifically, the paper proposes a feature point detection and positioning method suitable for smart terminals, which can be used to solve the structural mismatch problem. In addition, this paper also proposes a recursive tracking algorithm for human motion oriented augmented reality. The combination of ORB feature descriptor and KLT algorithm eliminates the phenomenon of virtual object jitter. The experimental results show that the recursive tracking scheme is superior to the detection tracking scheme in terms of time and accuracy.

The paper entitled “Image Real-time Augmented Reality Technology Based on Spatial Color and Depth Consistency” [12] uses the color image as a reference, computes the confidence of the original depth map, and performs stereo matching according to the feature points. Depth mapping is mainly enhanced by the color, edge and segmentation results of the color image. A deep learning computing system based on the augmented reality technology is designed. The system uses binocular cameras to collect target images in real-time and correct the calibrated images to obtain better parallax images. A semi-global block matching (SGBM) algorithm and depth calculation are used to realize the tracking and registration of virtual objects. Experiments in different environments show that the system has good real-time performance, illumination invariance and depth consistency.

The paper entitled “Fast Incremental Structure from Motion Based on Parallel Bundle Adjustment” [13] starts from the motion framework, and presents a novel structure, which not only reduces the computational costs, but also realizes a parallel beam adjustment on the GPU device, achieving large-scale optimization. In this framework, local beam adjustment is added to the framework of the incremental motion structure, which is then adjusted by parallel beams. Extensive experiments are performed on several challenging datasets and compared with the

most advanced methods. Experimental results show that this method has the best performance in terms of accuracy and efficiency.

To improve the flexibility of the 3D reconstruction system without loss in accuracy, the paper entitled “MARS: Parallelism based Metrically Accurate 3D Reconstruction System in Real-Time” [14] presents a solution of unconstrained Metrically Accurate 3D Reconstruction System (MARS) for 3D sensing based on a consumer camera. By using a simple initialization from a depth map, the system achieves incremental 3D reconstruction with a stable metric scale. Experiments are conducted using both real-world data and public datasets. Competitive results are obtained using the proposed system compared with several existing methods.

The paper entitled “Study on Computer Vision Target Tracking Algorithm Based on Sparse Representation” [15] uses the basic tracking algorithm and BOMP algorithm for image reconstruction. The BOMP algorithm is used as the core method of a sparse representation model. Then, a target observation model is established and a sparse display is introduced into the particle filter framework. Then, the sparse representation coefficient is updated to enable the norm to reach the optimal solution and ensure the accuracy of tracking the target. Finally, through simulation experiments, the success rate of target coverage is obtained. The results show that the BOMP algorithm maintains high tracking accuracy and strong stability when objects change their appearances.

The paper entitled “Accurately Realtime Visual SLAM Combining Building Models and GPS for Mobile Robot” [16] presents a novel 7 DOF (i.e. orientation, translation, and scale) visual simultaneous localization and mapping (vSLAM) system for mobile robots in outdoor environments. At the front end of the vSLAM system, a fast initialization method is designed for different vSLAM backbones. In the background of vSLAM, a non-linear optimization mechanism is proposed to achieve more robust optimization through the combination of multi-modal data. By combining the pose estimated by visual information with the position received from GPS, drift is further alleviated. Experimental results show that this method has good initialization performance, real-time operation, real-time scale, high accuracy and robustness, and is thus very suitable for external augmented reality applications in outdoor environments.

References

1. Lv, Z., Lloret, J., Song, H.: Internet of Things and augmented reality in the age of 5G. *Comput. Commun.* **164**(1), 158–161 (2020)
2. Sun, W., Mo, C.C.: High-speed real-time augmented reality tracking algorithm model of camera based on mixed feature points. *J. Real-Time Image Process.* **16**, 1–11 (2021)

3. He, Y., Chen, S., Yu, H., et al.: A cylindrical shape descriptor for registration of unstructured point clouds from real-time 3D sensors. *J. Real-Time Image Process.* **1**, 1–11 (2021)
4. Ni, T., Chen, Y., Liu, S., Wu, J.: Detection of real-time augmented reality scene light sources and construction of photorealistic rendering framework. *J. Real-Time Image Process.* **30**, 1–11 (2021)
5. Lu, J.: Mobile augmented reality technology for design and implementation of library document push system. *J. Real-Time Image Process.* **17**, 1–11 (2021)
6. Shi, Y., Zhao, Z.: Error elimination method in moving target tracking in real-time augmented reality. *J. Real-Time Image Process.* **19**, 1–11 (2021)
7. Huang, R., Sun, M.: Network algorithm real-time depth image 3D human recognition for augmented reality. *J. Real-Time Image Process.* **12**, 1–13 (2021)
8. Khan, M., Israr, A.S., Almogren, A.S., Din, I.U., Almogren, A., Rodrigues, J.J.: Using augmented reality and deep learning to enhance Taxila Museum experience. *J. Real-Time Image Process.* **12**, 1–12 (2021)
9. Meng, B., Wang, L., He, Z., Jeon, G., Dou, Q., Yang, X.: Gradient information distillation network for real-time single-image super-resolution. *J. Real-Time Image Process.* **5**, 1–12 (2021)
10. Bortolon M, Chippendale P, Messelodi S, Poiesi F.: Multi-view data capture using edge-synchronised mobiles. arXiv preprint. <https://arxiv.org/abs/2005.03286>. (2021)
11. Yue, S.: Human motion tracking and positioning for augmented reality. *J. Real-Time Image Process.* **5**, 1–12 (2021)
12. Zhai, L., Chen, D.: Image real-time augmented reality technology based on spatial color and depth consistency. *J. Real-Time Image Process.* **26**, 1–9 (2021)
13. Cao, M., Zheng, L., Jia, W., Liu, X.: Fast incremental structure from motion based on parallel bundle adjustment. *J. Real-Time Image Process.* **13**, 1–14 (2021)
14. Zhang, S., Wang, T., Li, G., Dong, J., Yu, H.: MARS: parallelism-based metrically accurate 3D reconstruction system in real-time. *J. Real-Time Image Process.* (2021). <https://doi.org/10.1007/s11554-020-01031-5>
15. Ma, W., Xu, F.: Study on computer vision target tracking algorithm based on sparse representation. *J. Real-Time Image Process.* (2021). <https://doi.org/10.1007/s11554-020-00999-4>
16. Liu, R., Zhang, J., Chen, S., Yang, T., Arth, C.: Accurate real-time visual SLAM combining building models and GPS for mobile robot. *J. Real-Time Image Process.* (2021). <https://doi.org/10.1007/s11554-020-00989-6>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.