

High-speed target tracking system based on multi-interconnection heterogeneous processor and multi-descriptor algorithm

Jiaqing WANG^{1,3}, Yongxing YANG¹, Liyuan LIU^{1,3} & Nanjian WU^{1,2,3*}

¹State Key Laboratory for Superlattices and Microstructures, Institute of Semiconductors,
Chinese Academy of Sciences, Beijing 100083, China;

²Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences,
Beijing 100083, China;

³University of Chinese Academy of Sciences, Beijing 100083, China

Received 14 August 2017/Revised 29 March 2018/Accepted 3 April 2018/Published online 24 October 2018

Citation Wang J Q, Yang Y X, Liu L Y, et al. High-speed target tracking system based on multi-interconnection heterogeneous processor and multi-descriptor algorithm. *Sci China Inf Sci*, 2019, 62(6): 069401, <https://doi.org/10.1007/s11432-017-9406-0>

Dear editor,

Target tracking has been a subject of intense scholarly research in past decades. Many parallel processors [1–4] and efficient tracking algorithms based on feature extraction [3–5] have been proposed to meet the requirements of high-speed tracking. However, the tracking algorithms are all based on a single descriptor so that the description for the target is not comprehensive and the target is prone to be missed in complex scenery. The tracking algorithms involve histogram statistics, which would consume most of the algorithm running time, and the previous parallel processors cannot support a fast calculation of a statistical histogram.

We propose a high-speed target tracking system based on multi-interconnection heterogeneous processor and multi-descriptor algorithm. The proposed processor contains a processing element (PE) array and a row processor (RP) array. Each PE (RP) is connected to the 1st, 2nd, 4th, 8th, 16th, 32th and 64th nearest neighbour PEs (RPs). The new PE array and RP array can exponentially speed up the execution of histogram statistics and accumulation operation, respectively. The pro-

posed tracking algorithm exploits multi descriptors, which includes local binary patterns (LBP), histograms of oriented gradients (HOG) and gray-level local binary patterns (GLLB), to describe the target from different aspects. It can improve the robustness of the tracking system. Results indicate that the system can track moving object with 950 fps in a complex scenery.

Multi-interconnection heterogeneous processor. The proposed processor architecture is shown in Figure 1(a). It mainly contains three parts: an $M \times M$ single instruction multiple data (SIMD) PE array processor, an $M \times 1$ SIMD RP array processor and a microprocessor unit (MPU). The PE array processor performs pixel-parallel low-level image processing, such as filtering and local feature extraction. The RP array processor performs row-parallel middle-level and high-level processing to acquire global image information and to carry out the object recognition. The MPU is responsible for the overall chip management. Compared to the traditional PE array [2–4], each PE in our design has multi-interconnections with more neighbour PEs. It can access the data not only from the west and the south, but also from far west and

* Corresponding author (email: nanjian@red.semi.ac.cn)

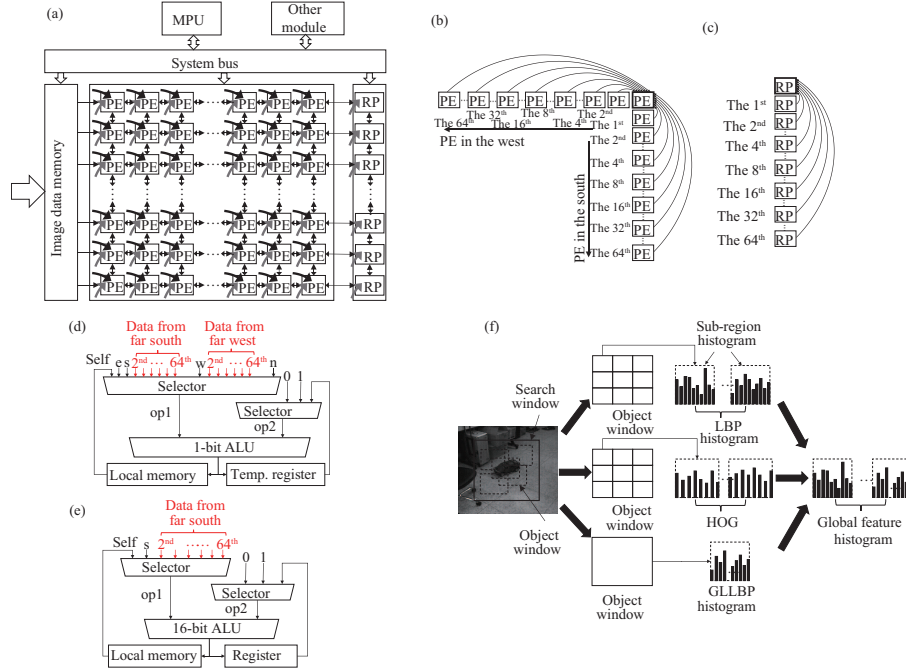


Figure 1 (Color online) Proposed processor architecture and tracking algorithm. (a) The whole processor architecture; (b) connection between PEs; (c) connection between RPs; (d) PE circuit; (e) RP circuit; (f) multi-descriptor tracking algorithm.

far south. The details on the connections of each PE are shown in Figure 1(b). Each PE can access data from the 1st, 2nd, 4th, 8th, 16th, 32th and 64th PE in the west or south. Each RP also has multi-interconnection with more neighbour RPs. It can access the data from far south. As shown in Figure 1(c), each RP can access data from the 1st, 2nd, 4th, 8th, 16th, 32th and 64th RP from the south. If the transmitted data is out of PE array or RP array, it is set to zero. Figures 1(d) and 1(e) show the detailed PE circuit and RP circuit.

Compared with the existing architecture, the multi-interconnection architecture can exponentially speed up the computation of histogram statistics. Take 64×64 imaging patch as an example. First, each PE accesses the west nearest PE's data and calculates a 2-PEs histogram. Second, each PE visits the west 2nd nearest PE and calculates 4-PEs histogram. If we keep on doing the similar operations six times, we can finish a 64-PEs histogram statistics. Third, each PE visits the histogram from the south and does the similar operations as for the west direction. Thus the PE array finishes effective 64×64 -PEs histogram statistics rapidly. As a result, it only takes 2×6 steps to perform a 64×64 image feature histogram statistics by this PE array, while it takes 2×63 steps by the previous PE array [2–4]. Generally, for an $N \times N$ image patch, this PE array needs $2 \times \log_2 N$ steps to finish histogram statistics, com-

pared to $2 \times (N - 1)$ steps using the previous PE array. Based on the same analysis, this processor architecture allows the RP array to get data from any PE column within $\log_2 N$ steps to perform row-parallel middle-level processing, smaller than $N - 1$ steps using the previous processors. And RP array can execute accumulation operation within $\log_2 N$ steps, compared to $N - 1$ steps using the previous RP array.

Multi-descriptor tracking algorithm. The tracking algorithms [3–5] are all based on a single descriptor. Obviously, these are not robust in various application scenarios because the description of the target is not comprehensive. We propose a tracking algorithm based on multi-descriptors, including GLLBP [3], LBP [6], and HOG [7]. LBP is an efficient descriptor for texture description. The GLLBP descriptor combines the LBP value and the grey value (GL). It describes not only local texture but also distribution of luminance. As shown in Figure 1(f), we actually use one-sub-region histogram of GLLBP and nine-sub-region histogram of LBP to describe the target. The one-sub-region histogram of GLLBP describes the texture in different gray values and the sub-region histogram of LBP describes the texture on different positions. HOG contains the information of gradient orientation that consists of a large amount of silhouette information. These three histograms can describe complementary aspects of the target, and

each descriptor has good performance in tracking process. If the target and background are different in silhouette, in texture, or in gray even though similar in texture, the combination of the three histograms can easily distinguish them. The operations involved are all simple, which could have high computation efficiency based on the proposed processor.

The proposed tracking algorithm is shown in Figure 1(f). First of all, an image is captured, and a search window is set artificially. The size of search window can either be the same to, or be smaller than the image's size. Then, object windows are picked out as many as possible in the search window. Each object window is divided into a grid of sub-regions, followed with the LBP histogram, GLLBP histogram and HOG of each sub-region simply (with the weight of each sub-region set to 1) concatenated to form a global feature histogram of the object window. The sub-region number of different features can be different. Finally, the global feature histogram as a vector is input to the trained adaboost classifier and gets a score. The object window getting the highest score is considered containing the target. The algorithm is implemented on the proposed processor architecture in a parallel way with accelerated execution of histogram statistics, which can achieve a high speed even though it involves three descriptors.

High-speed tracking system implementation. A high-speed tracking system consists of a camera, an actuator and a field programmable gate array (FPGA). The resolution of the camera is 2048×1088 . The proposed processor is implemented on an FPGA. It mainly contains a 128×128 PE array, a 128×1 RP array and an MPU. Each PE and RP are equipped with a 64 bits RAM and a 256 words RAM, respectively. A PE array is used for low-level processing and fast histogram statistics, and an RP array is used for middle-level processing and accumulation operation of adaboost classifier.

The proposed algorithm is executed on the multi-interconnection heterogeneous parallel processor. Initially, the target is selected artificially in the first frame, or the object window scans the whole 2048×1088 image to find the target. A search window of 128×128 pixels is set according to the target location of last frame. Then, the target is located in the search window, and the absolute position in the whole image can be calculated. Finally, the actuator adjusts the camera to the right direction according to the absolute position of the target and makes the target locate in the center of the whole image. If the object win-

dow is set to 96×96 pixels in the proposed algorithm, the maximum pixel difference between the adjacent frames is restricted to below 16.

Experimental results. We carry out experiment with a high-speed tracking system to evaluate the performance of the tracking system. Results show that the proposed algorithm is more robust than other algorithms. The proposed processor executes the algorithm more efficiently in high speed. For detailed results of the experiment, please refer to Appendix A.

Conclusion. We propose a high-speed tracking system based on multi-interconnection heterogeneous processor architecture and multi-descriptor tracking algorithm. This architecture can speed up the execution of histogram statistics so that the algorithm can use multi descriptors, the number of search can achieve up to 961 times, and the final tracking speed can get to 950 fps. The proposed tracking algorithm exploits three descriptors to describe the target, making the target representation more comprehensive and more robust in complex scenery.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61434004, 61234003, 61504141).

Supporting information Appendix A. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

- 1 Komuro T, Ishii I, Ishikawa M, et al. A digital vision chip specialized for high-speed target tracking. *IEEE Trans Electron Dev*, 2003, 50: 191–199
- 2 Zhang W, Fu Q, Wu N J. A programmable vision chip based on multiple levels of parallel processors. *IEEE J Solid-State Circ*, 2011, 46: 2132–2147
- 3 Yang Y X, Yang J, Liu L Y, et al. High-speed target tracking system based on a hierarchical parallel vision processor and gray-level LBP algorithm. *IEEE Trans Syst Man Cybern Syst*, 2017, 47: 950–964
- 4 Yang J, Liu L Y, Shi C, et al. Heterogeneous vision chip and LBP-based algorithm for high-speed tracking. *Electron Lett*, 2014, 50: 438–439
- 5 Liu L Y, Wu N J, Yang J, et al. High-speed visual tracking with mixed rotation invariant description. *Electron Lett*, 2016, 52: 511–513
- 6 Ojala T, Pietikainen M, Maenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intel*, 2002, 24: 971–987
- 7 Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, San Diego, 2005. 886–893