

Architecture description and prototype demonstration of optoelectronic parallel-matching architecture

Keiichiro Kagawa, Kouichi Nitta, Yusuke Ogura, Jun Tanida,
and Yoshiki Ichioka

**Department of Material and Life Science, Graduate School of Engineering,
Osaka University

Abstract. We propose an optoelectronic parallel-matching architecture (PMA) that provides powerful processing capability for distributed algorithms comparing with traditional parallel computing architectures. The PMA is composed of a parallel-matching (PM) module and multiple processing elements (PE's). The PM module is implemented by a large-fan-out free-space optical interconnection and parallel-matching smart-pixel array (PM-SPA). In the proposed architecture, each PE can monitor the other PE's by utilizing several kinds of global processing by the PM module. The PE's can execute concurrent data matching among the others as well as inter-processor communication. Based on the state-of-the-art optoelectronic devices and a diffractive optical element, a prototype of the PM module is constructed. The prototype is assumed to be used in a multiple processor system composed of 4×4 processing elements, which are completely connected via 1-bit optical communication channels. On the prototype demonstrator, the fundamental operations of the PM module such as parallel-matching operations and inter-processor communication were verified at 15MHz.

1 Introduction

Parallel distributed processing is an effective method to accelerate the performance of computing system. In the parallel distributed processing, a task is divided into a number of processes executable concurrently. The processes are distributed and executed over multiple processing elements (PE's), so that the total processing time can be reduced.

A heuristic optimization described by a distributed algorithm is a good application of a parallel computing system. In the algorithm, the solution space is divided into multiple pieces of segments, in which the candidates of the solution are sought concurrently by multiple PE's. In the framework of the traditional parallel computing architecture, global processing to calculate multiple data from all the PE's can be a processing bottleneck. Because communication between the PE's and processing are implemented separately, the heavy traffic occurs on the

** kagawa@mls.eng.osaka-u.ac.jp

network path to or from the PE that executes the global processing. The bottleneck causes throughput reduction of the whole parallel computing system. This bottleneck can not be eliminated by simply increasing the communication capacity of network. Therefore, the traditional parallel computing architectures are not always suitable for the distributed algorithms.

In this paper, we propose an optoelectronic parallel-matching architecture (PMA) which is an effective parallel computing architecture suitable for the distributed algorithms. The PMA is based on an optoelectronic heterogeneous architecture formerly presented by Tanida *et al.*,[1] which is composed of electronic parallel processors for local processing and an optical network processor for interconnection and global processing between the electronic processors. The optical network processor is assumed to be embodied by the optical interconnection and the smart-pixes[2] for wide communication bandwidth and dense connectivity between the PE's. In the architecture, both electronic and optical processors work in complementary manner. An electronic processor shows high performance in the local processing, whereas an optical processor is good at the global processing. The system based on the PMA also has ability to execute the global processing without degrading the throughput of network. Detection of the PE's satisfying a given condition and summation of absolute differences over the multiple PE's are typical examples of the global processing. The optical network processor of the PMA is called a parallel-matching (PM) module, which consists of a large-fan-out free-space optical interconnection and a parallel-matching smart-pixel array (PM-SPA). The proposed architecture can reduce the execution time for the fundamental global data processing: global data matching, detection of the maximum (minimum) data, and ranking of the data, compared with the other traditional architectures with photonic networks.

2 Parallel Matching Architecture

We assume a multiple-instruction multiple-data stream (MIMD) parallel computing system consisting of N PE's embodied by the smart-pixel technology. The PE's are connected each other via a photonic network. A heuristic optimization algorithm based on the distributed algorithm is a good application of the parallel computing systems, which can be applied to the problems that do not always have a rigorous solving method. A general procedure of the distributed optimization algorithm is composed of distribution of the data, parallel processing, and integration of the calculated data. First, the candidates of solutions are distributed to the PE's. Second, each PE locally calculates the fitness function of candidate. Finally, good candidates are selected among the candidates based on the values of the fitness function. Note that this operation is achieved by global processing over the multiple PE's.

Figure 1 shows the system compositions for the distributed algorithms by the traditional MIMD parallel computing system and the parallel matching architecture. The traditional architecture has a hierarchy composed of a master PE and multiple slave PE's as shown in Fig. 1(a). The rolls of the master PE are

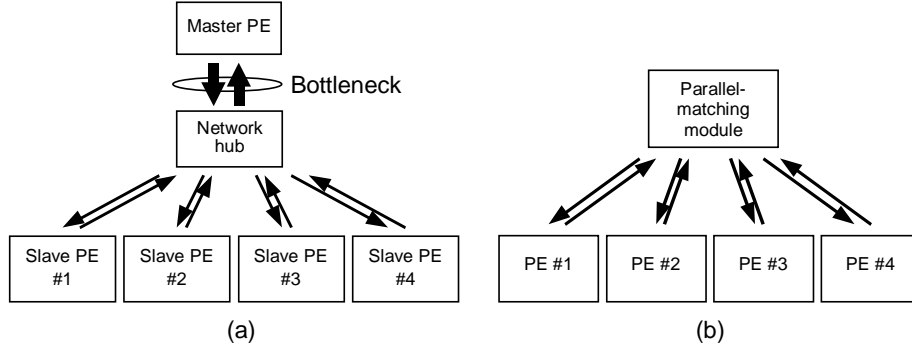


Fig. 1. Configurations of parallel computing architectures: (a) a traditional master-slave architecture and (b) the parallel-matching architecture (PMA)

data distribution, integration, and global processing. The master PE distributes the data to the slave PE's and integrates the resulting data from them through the network. After data integration, the master PE executes global processing locally. Because the amount of the network traffic in the data distribution and integration is very large, these procedures can be processing bottlenecks. This bottleneck can not be eliminated by simply increasing the communication capacity of network, for the total amount of the fanned-in data from N slave PE's to the master PE is N times as large as the bandwidth of the communication path between the network communication module and the PE's.

On the other hand, the PMA has a different composition as shown in Fig. 1(b). The PMA is composed of the PM module and the multiple PE's. In the PMA, the fitness of each candidate is compared with the candidates on the other PE's by using the global processing mechanism of the parallel-matching (PM) module. The PM module offers both networking and global processing, so that the master PE for data distribution and integration is not required. The PE's in the system have the same priority because the global processing is executed inside the PM module; that is the system has no hierarchy. As a result, there is no bottleneck in the proposed architecture in global processing.

The PM module consists of large-fan-out free-space optical interconnection and a parallel-matching smart-pixel-array (PM-SPA). The PM module is regarded as a kind of the network hub in which a specific mechanism for the global processing is built-in. The global processing in the PMA is data comparison among the data sent from the PE's. The PM module monitors the output data from all PE's, and concurrently compares the datum from each PE with the data from the other PE's. When a PE requires the compared result, it is sent back to the PE through the network communication channel. As mentioned above, data distribution and integration increase the network traffics and the processing overheads at a PE. However, because the global processing is ex-

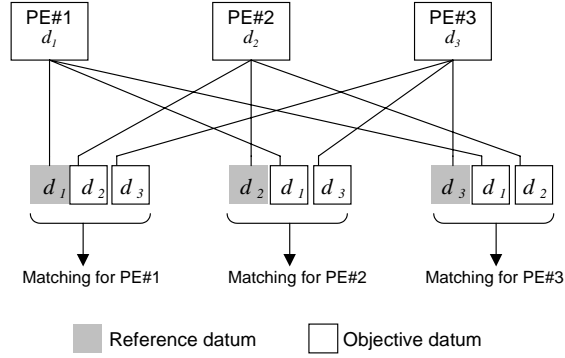


Fig. 2. Reference datum and objective data in the parallel matching. d_1, d_2 , and d_3 denote the output data of PE#1-#3, respectively

executed inside the PM module without occupying the network bandwidth, the throughput of the total system does not become decreased.

We define the datum from each PE as the reference datum and the one from the other PE's as the objective data as shown in Fig. 2. The reference datum and the objective datum to be compared are called a matching pair. The PM module tests the reference datum and each of the objective data for the following conditions: 1) the reference datum is equivalent to the objective datum, 2) the reference datum is less than the objective datum, and 3) the reference datum is more than the objective datum. The result of the global comparison is expressed by a set of logical values. When the condition is satisfied, the returned value is 1 (true), otherwise 0 (false). These operations are called parallel-matching operations, which are denoted by **pEQU**, **pMORETHAN**, and **pLESSTHAN**, respectively. (The prefix **p** means 'parallel.') We also define the fourth parallel-matching operation: summation of the absolute differences denoted by **pDIFF**. This operation provides the summation of the absolute difference between the reference datum and the objective datum. Utilizing the **pDIFF** operation, each PE can obtain the quantitative value of the difference.

Figure 3 shows a schematic diagram of the parallel matching with 5 PE's. In the figure, PE's A, B, and C obtain 4-bit binary values representing the results of parallel matching: **pEQU**, **pMORETHAN**, and **pLESSTHAN**, respectively. PE-D obtains the result of the **pDIFF** operation. The numbers in the boxes of PE's are the output data from the PE's. After the output data are fanned out and exchanged, they are concurrently compared by the parallel-matching operations in the PM module. Then, one of the parallel-matching results or the objective datum is selected by the multiplexer on the request from the PE's. In general, for m -bit data format, up to $(m + 1)$ PE's can be compared at the same time. Finally, the selected result is sent back to each PE. In Fig. 3, example values of the parallel-matching results are shown. The operation mode of PE-E

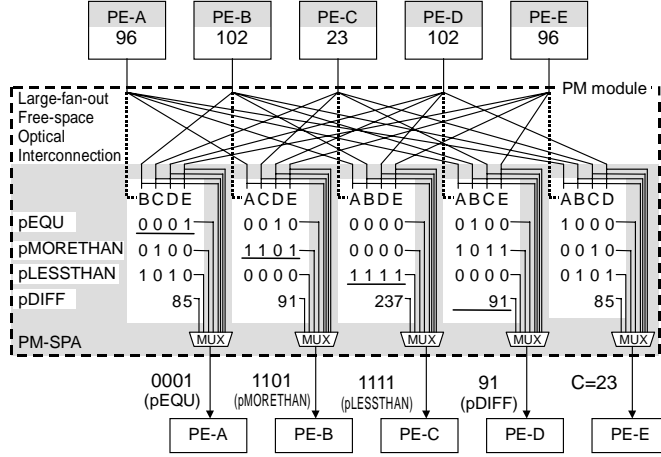


Fig. 3. Fundamental operations of the parallel-matching architecture. MUX means a multiplexer

is different from the others. That is the communication mode in which the data from PE-C is sent to PE-E transparently.

3 Experimental prototype system

We construct a prototype system of the PM module to demonstrate its fundamental operations. In designing the prototype, we assume the parallel computing system shown in Fig. 4. The parallel computing system consists of 4×4 PE's, which are completely connected via the PM module. The PE's are located on a two-dimensional grid, and each of them is connected to the PM module with bit-serial optical fiber channels. Each PE is embodied by smart-pixels coupled with an optical fiber. The data from the PE's are sent to the PM module by the optical fibers. As mentioned below, a complete-connection network is implemented by optical data fanning. With the optically fanned-out signals, the parallel-matching operations and the processing for inter-PE communication are executed by the PM-SPA. The resulting data are emitted from the PM-SPA, and returned to the PE's through the optical fibers.

Figure 5 shows the schematic diagram of the optoelectronic complete-connection. As shown in Fig. 5(a), the optical signals from 4×4 PE's in the bit-serial format are assumed to be aligned on a two-dimensional grid as an input image toward the PM module. Because the whole image of the light signals is required for one PE, 4×4 replica images shown in Fig. 5(b) are prepared for 4×4 PE's. In the prototype, an 8×8 -VCSEL array (GigalaseTM; Micro Optical Devices; emitting wavelength, 850nm; pixel pitch, $250\mu\text{m}$) is used as a light emitter array. In the prototype, the function of the PM-SPA is emulated by a CPLD

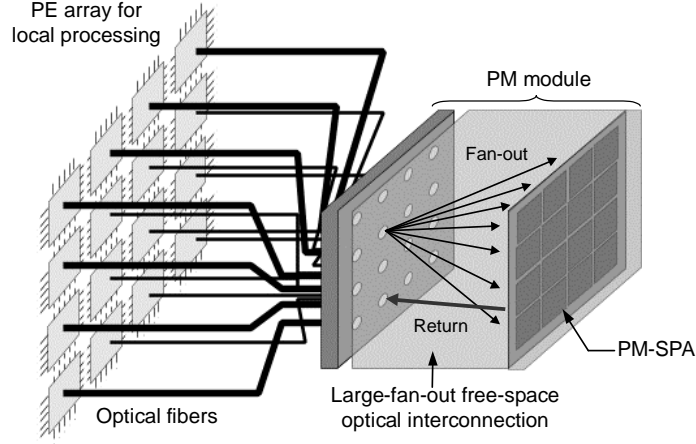


Fig. 4. Target prototype system of the PMA

(Model FLASH374i, Cypress) coupled with a 4×4 -complementary-metal-oxide-semiconductor photodetector (CMOS-PD) array (Model N73CGD) supplied by United States-Japan Optoelectronic Project (JOP). As shown in Fig. 5(c), one of the replicas is detected by a CMOS-PD array, then transferred to the CPLD, and the fundamental operations of the PMA are executed.

For the large-fan-out optical interconnection, a conventional $4f$ optical correlator was adopted. We constructed a Fourier transform lens system whose focal length is 160.0mm for wavelength 850nm. In designing the lens system, CodeVTM of Optical Research Associates was used.

As an optical fan-out element that generates complete-connection pattern shown in Fig. 5(b), we designed a phase-only computer-generated hologram (CGH) filter with two-level phase modulation based on the Gerchberg-Saxton algorithm.[3] Figure 6(a) shows the ideal mapping on the output plane of the interconnection optics. The output pattern contains 16 replicas of the VCSEL image arranged on a grid, in which each quadrant contains 2×2 replicas of the VCSEL image. Each replica corresponds to the optical signals for a single PE. Because the equipments used in fabrication of the CGH filter do not have enough fabrication accuracy to eliminate the 0th light spot, the copied images are located not to be overlapped with the 0th image in the design. The pitch and the margin of adjacent replicas of the VCSEL image are 2.5mm and 1.5mm, respectively. Figure 6(c) shows the filter pattern with two-level phase modulation. The CGH filter was fabricated by the electron beam (EB) lithography. Figure 6(b) shows the reconstructed interconnection pattern of the fabricated CGH filter for 4×4 VCSEL's when the filter was incorporated in the $4f$ optical correlator.

Finally, we operated the prototype system without the CGH filter to verify the fundamental parallel-matching operations and inter-PE communication. The

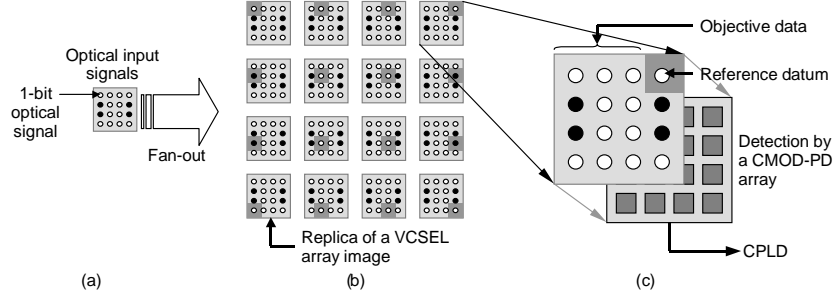


Fig. 5. Schematic diagram of optoelectronic complete-connection: (a) Output data displayed on the VCSEL array, (b) replica images of the VCSEL array for the complete-connection network, and (c) a replica image of the VCSEL image for one PE

data transfer was in the bit serial format, and the word length of the data was set to 4. From the experimental results, we have verified that the fundamental operations of the prototype were executed exactly at 15MHz. The operational speed was limited by the one of the CMOS-PD array. The bit rate of communication per PE and the total bit rate of the prototype were 15Mbps (bit per second) and 240Mbps, respectively. The frequencies of the parallel-matching operation for each PE and the whole system were 0.68M operations/sec and 11M operations/sec, respectively.

4 Conclusions

We have proposed an optoelectronic parallel-matching architecture (PMA) as an effective parallel computing architecture. The fundamental operations of the PMA, **pEQU**, **pMORETHAN**, **pLESSTHAN**, and **pDIFF**, have been defined. This architecture is specialized for the global data processing and has capability to accelerate execution of distributed algorithms, because the PMA has a specific mechanism for parallel-matching operations over multiple processing elements. The prototype system of the PMA was constructed to demonstrate the fundamental global operations of the PMA based on the state-of-the-art optoelectronic devices and a phase-only CGH filter. In the prototype, the PM-SPA, which was the core module of the PM module, was emulated by the CPLD and the CMOS-PD array. The prototype was assumed to be used with 4×4 PE's that are completely connected via the PM module with 1-bit optical channels. For optical interconnection of the prototype, a Fourier transform lens system was designed. As a fan-out element, the phase-only CGH filter with two-level phase modulation was designed based on the Gerchberg-Saxton algorithm, and was fabricated by the EB lithography. We confirmed that the prototype performed the fundamental parallel-matching operations and the inter-PE communication at 15MHz. For the whole system, the bit rate of inter-PE communication and the

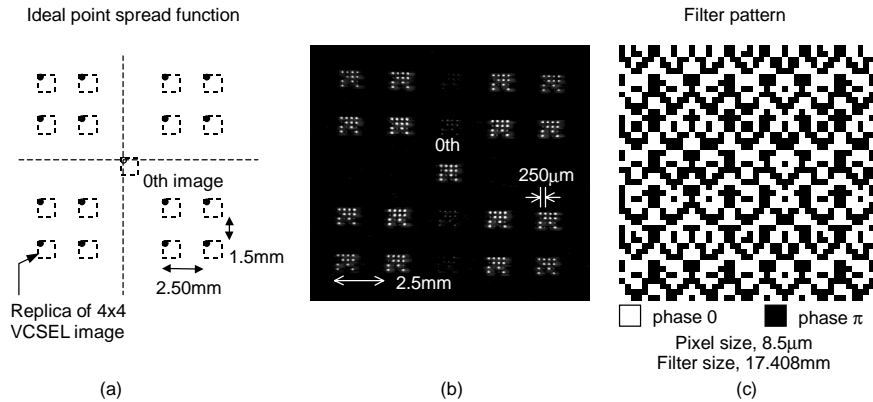


Fig. 6. (a) Designed optical interconnection pattern for complete-connect network composed of 4×4 PE's, (b) a part of the obtained CGH filter with two-level phase modulation, and (c) experimental result of the optical interconnection by the CGH filter

frequency of the parallel-matching operation were 240 Mbps and 11M operations per second, respectively. The operational speed of the prototype was limited by the CMOS-PD array. The performance can be improved by using high-speed photodetectors with high sensitivity such as MSM photodetectors coupled with transimpedance photo-amplifiers.

Acknowledgment

This research was supported by the JOP user funding under the Real World Computing Partnership (RWCP). The authors would like to appreciate the activities of the JOP. This work was also supported by Development of Basic Tera Optical Information Technologies, Osaka Prefecture Joint-Research Project for Regional Intensive, Japan Science and Technology Corporation.

References

1. P. Berthomé and A. Ferreira, *Optical interconnections and parallel processing: trends at the interface* (Kluwer Academic Publishers, London, 1998).
2. T. Kurokawa, S. Matso, T. Nakahara, K. Tateno, Y. Ohiso, A. Wakatsuki, and H. Tsuda, "Design approaches for VCSEL's and VCSEL-based smart pixels toward parallel optoelectronic processing systems," *Appl. Opt.* **37**, 194–204 (1996).
3. R. W. Gerchberg and W. O. Saxton, "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," *OPTIK* **35**, 237 – 246 (1972).