Functional-Memory-Based Systems Enabling Recognition and Learning

Tetsushi Koide, Hans Jürgen Mattausch, Ali Ahmadi, Takashi Morimoto, and Kosuke Yamaoka
Research Center for Nanodevices Systems, Hiroshima University, Higashi-Hiroshima, 739-8527, Japan

Abstract— The effective implementation of pattern recognition and learning, which are basic functions for building artificial systems with capabilities similar to the human brain [1], is of great technical and practical importance. To realize the intelligent information processing, a very high system performance, which does not suffer from the memory bottolneck, can be required. In this paper, an efficient solution for satisfying the above requirement by a unification of memory and processing unit. Practical application results to picture segmentation and pattern matching for the unification method as well as to system applications like real-time tracking of moving objects and character recognition/learning are reported.

I. Unification of Processing and Memory Part

The unification of processing and memory part is a promising development direction, which can combine high performance of an integrated system with high integration density and low power consumption. In the following sections, two examples for the unification of memory and processing unit, which have been proposed in our group.

A. Picture segmentation with a cell network

The extraction of the different objects contained in an input picture is called picture segmentation, which is an indispensable basic operation for the enabling intelligent processing such as object recognition or object tracking. For this purpose we have developed a reliable, hardware-oriented region-growing algorithm [2] as well as two LSI architectures with real-time processing capability [3, 4].

The cell-network approach, which exploits the unification of memory and processing part, enables fully parallel processing with a regular array of pixel-related processing elements (PEs) and local memory. In this way, the segmentation of QVGA size pictures with a frame rate of 300fps and a data exchange of 36,000Gbps between memory and processing part becomes possible. Figure 1 shows a cell-network design with 41x33 pixels in 350nm CMOS technology. From the measured performance of this test chip we estimated that QVGA picture segmentation in 300 sec is realistic at the 90nm CMOS technology node.

Above segmentation speed is much higher than the standardized video frame rate of 30fps, so that a block-based processing with a reduced number of PEs, as shown in Fig. 2, is sufficient. For supporting this approach a large wordlength of the memory is realized by dividing the memory into many banks, so that the necessary data for all PEs can be accessed in parallel in each processing step. 320 PEs and an access bandwidth of 80Gbps are sufficient for real-time segmentation of QVGA size pictures. This so-called image-scan architecture, for which we have constructed an FPGA-based evaluation system, is suitable for FPGA as well as ASIC realization[5].

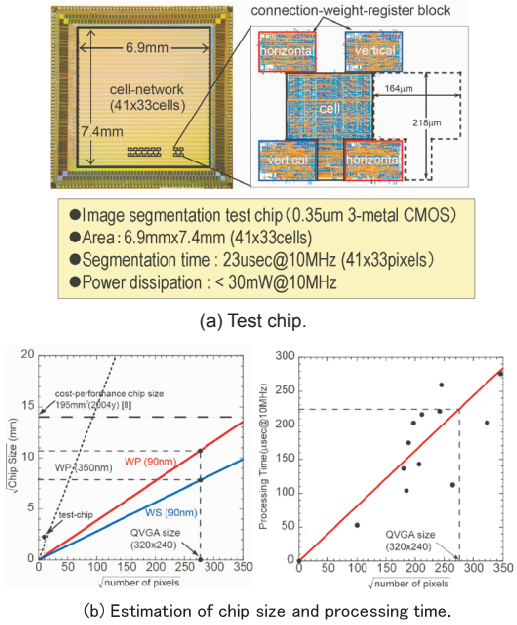


Fig. 1. Picture-segmentation network with 41x33 pixels in 350nm CMOS technology.

B. Pattern-matching operation with an associative memory

An associative memory has the capability of determining the nearest match between input-data words and a stored basis of reference-data words according to a distance measure. Especially for real-time recognition it will be necessary to implement fast matching up to large absolute minimum distances.

We have developed associative memory integrated-circuits on the basis of mixed digital-analog circuitry (see Fig. 3), which unifies the memory and the processing part in the form of an associative memory [6, 7, 8, 9, 10, 11]. The comparators for the reference patterns are integrated in digital form into the memory. The distance-measure results are digitally determined and transformed into analog signals for each reference pattern in parallel.

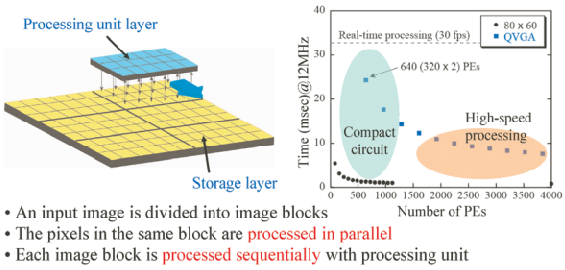


Fig. 2. Conceptual view of the image-scan segmentation architecture.

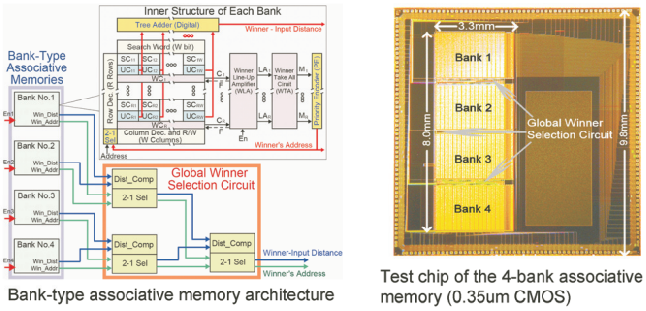


Fig. 3. Architecture diagram of the mixed digital-analog associative memory.

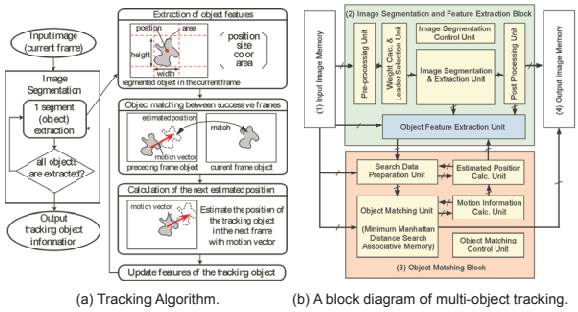


Fig. 4. Tracking algorithm and architecture of multiple objects based on picture segmentation and pattern matching.

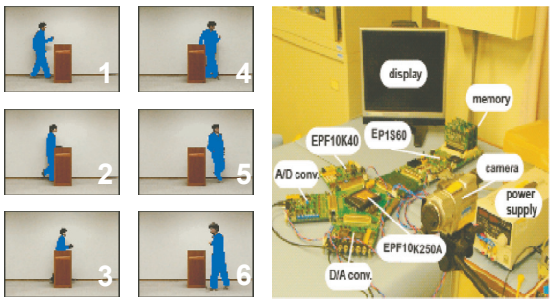


Fig. 5. Prototype system for real-time object tracking based on an FPGA implementation.

II. Application to Information Processing Systems

In the following 2 application examples are discussed, which utilize several memory/processing-unified sub-systems for e.g. segmentation, pattern matching, feature extraction and pattern generation as well as recognition and learning

A. Real-time multi-object tracking with segmentation and pattern matching

The tracking of moving objects is an important technology essential for the realization of e.g. high quality traffic observation systems or intelligent robots. We have proposed an algorithm for moving-object tracking, which is based on picture segmentation, pattern matching, and an efficient feature extraction method for the objects of the segmented pictures [12, 13, 14, 15]. As explained in Fig. 4(a), the segmented objects of each input frame are submitted to a feature-extraction unit for determining data such as position, horizontal/vertical dimensions, size and color. A matching unit is then used to determine the most similar object in the previous frame and in this way the tracking of the movement or non-movement of all objects is made possible.

The resulting VLSI architecture consists of three main circuits, namely (1) a segmentation part which applies either the cell-network or the image-scan architecture, (2) a circuit, which uses the 2-dimensional frame-segmentation results for determining height/width, area, position etc. of segmented objects in parallel, and (3) a circuit for the matching operation with the objects of the previous frame (Fig. 4(b)). A real-time demonstration and evaluation system of the proposed object-tracking architecture for 80x60-pixel images has been built on the basis of an FPGA implementation of the proposed architecture and is shown in Fig. 5(a). Based on the evaluation results, real-time object tracking with up to 1000 objects is estimated possible with QVGA motion pictures (30fps), even with an FPGA implementation. Figure 5(b) shows an example of object tracking result with developed FPGA-based prototype system.

B. Character recognition and learning based on associative memory

We investigate a learning OCR system which is based on using a short/long-term memory for ranking the reference patterns and updating the reference memory with new input samples, constantly [16, 17, 18]. The goal is to obtain an OCR hardware design which is capable of recognition of new input characters written in different fonts, in a speed level of 10 ms per word. Moreover, we plan to use the mixed analog-digital fully-parallel associative memory for nearest Hamming/Manhattan-distance search as the main classifier of the system to enhance the pattern search speed.

Figure 6 shows an outlook of the whole prototype system. The main steps considered for the OCR process are: data reading, binarizing, noise removal, image labeling, segmentation, classification, ranking, and optimization. In the data reading step, the data of each text line are scanned continuously as a

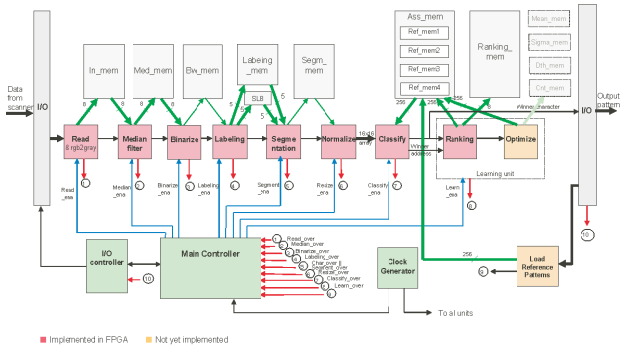


Fig. 6. Block diagram of the system architecture.

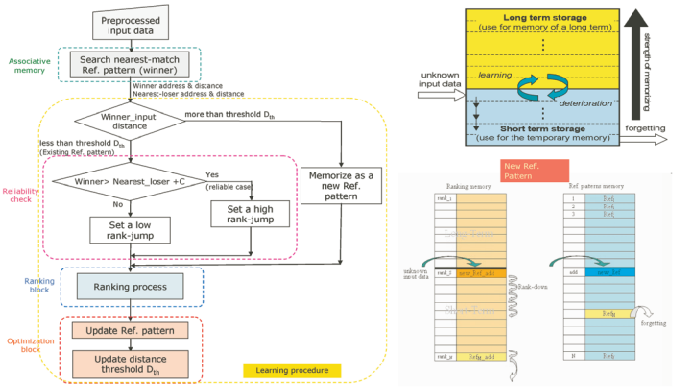


Fig. 7. Flowchart of learning process including updating steps and reliability evaluation.

sequence of thin frames by moving a reading device (scanner sensor) on the text. In order to remove noise from the frame, a median filter with neighborhood of 3×3 is applied. Next, by employing a sequential labeling algorithm, different segments of the image frame are labeled and segments larger than a threshold level are recognized as a single character. The last and main step of the process is character classification which is carried out by a nearest-distance search algorithm applying the associative memory.

The core concept of learning in our model is based on a short/long term memory and is very similar to memorizing procedure in the human brain. For this, the memorized reference patterns are classified into two areas. One is a short-term storage area where new information is temporarily memorized, and the other is a long-term storage where a reference pattern can be memorized for a longer time without receiving the direct influence of incoming input patterns. The transition of reference patterns between short-term and long-term storages is carried out by means of a ranking algorithm (Fig. 7). Besides, an optimization method is applied to update the reference vectors magnitude and optimize their distribution as well as the threshold values used for ranking.

The model was implemented in the FPGA platform of the Altera Stratix family (EP1S80) shown in Fig. 8. As for reading device, we used the line-scan sensor ELIS-1024 of Panavision Co. with a resolution level of 1024 pixels per line. Using Stratix DSP development kit EP1S80 and a clock frequency of 15 MHz, a total number of 5,447 logic cells and 2.1 Mbits of SRAM memory are used for placement and routing of the whole system.

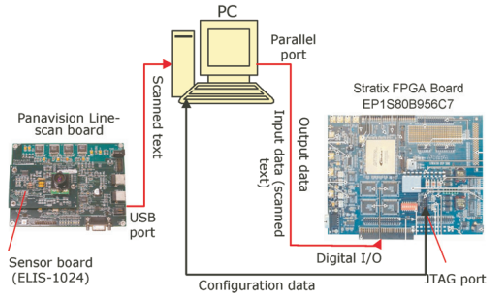
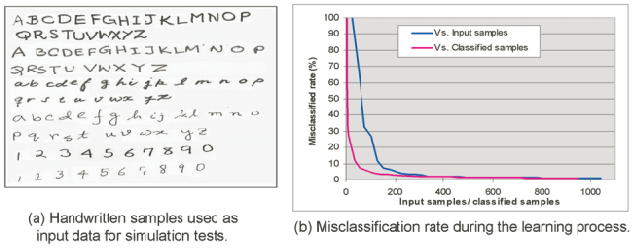


Fig. 8. A schematic of the connection between FPGA board and I/O device



(a) Handwritten samples used as input data for simulation tests.

(b) Misclassification rate during the learning process.

Fig. 9. Changes in the misclassification rate during the learning process.

The learning performance of the system was evaluated with a number of 35 datasets of English handwritten characters written by four different writers were used for the experiments (Fig. 9(a)). The classification results during and after the learning process are reported in Table I. It is worth-noting that the system has started learning without any initial reference patterns or a predefined dataset. The average misclassification rate is around 5% which is reasonable for this type of application.

Figure 9(b) depicts the misclassification rate during the learning process. In the first approach, the new reference pattern is initially given the highest unoccupied rank in the long-term memory as long as unoccupied ranks are available, whereas in the second approach it gets the top rank in the short-term memory from the beginning.

III. Conclusion

In this paper two application examples were discussed, which utilize several memory/processing-unified sub-systems for example segmentation, pattern matching, feature extraction and pattern generation as well as recognition and learning.

First, we have developed a real-time multi-object tracking architecture based on image segmentation and object matching

TABLE I
CLASSIFICATION RESULTS FOR FOUR DATASETS FROM DIFFERENT WRITERS DURING THE LEARNING PERIOD (TAKING NO INITIAL REF. PATTERNS).

Writer	**Dataset A** (52 samples)		**Dataset B** (52 samples)		**Dataset C** (52 samples)		**Dataset D** (52 samples)	
	Mis classify	New Ref. added	Mis classify	New Ref. added	Mis classify	New Ref. added	Mis classify	New Ref. added
1	9	42	12	25	11	19	19	14
2	11	36	9	26	16	19	10	31
3	10	36	6	20	10	21	12	12
4	11	39	14	20	13	22	5	20
Total %	19.7	73.6	19.7	43.8	24	38.9	22.1	37

and a prototype system for FPGA/ASIC implementation. With region-growing image segmentation, we detect all objects in the image no matter whether they are moving or not. The structure of the image-scan processing element array is exploited for simultaneous extraction of simple object features. Using image segmentation and object-feature-extraction results of successive frames, we exploit object matching in the object-feature space for tracking of the objects. The proposed architecture is realized with modern FPGA hardware and is verified to enable real-time tracking of up to 230 objects for QVGA-size video picture at 20MHz clock frequency.

Next, we have also proposed a learning model based on a short/long-term memory and an optimization algorithm for constantly adjusting the reference patterns. The major blocks of the system were implemented in an FPGA platform and tested with real data samples of handwritten and printed English characters. The simulation results showed an acceptable performance of classification and learning. In order to enhance the search speed in the classification block, we are planning to use a fully parallel associative memory implemented in an LSI architecture, as the main classifier of the system via an ASIC design.

Consequently the unification approach can lead to high performance, high density solutions and is demonstrated in the picture segmentation and associative-memory applications. In particular the efficient pattern-matching performance with an associative memory leads to elegant solutions for the realization of intelligent functions like object tracking, recognition and learning in integrated systems.

ACKNOWLEDGEMENTS

The VLSI chips were fabricated by VDEC, University of Tokyo, in collaboration with Hitachi Ltd. and Dai Nippon Printing Corp, and Rohm Corporation and Toppan Printing Corporation.

REFERENCES

[1] L. C. Jain, U. Halici, I. Hayashi, S.B. Lee, and S. Tsutsui, "Intelligent Biometric Techniques in Fingerprint and Face Recognition," CRC Press, 1990.

[2] T. Morimoto, Y. Harada, T. Koide, and H. J. Mattausch, "Efficient video-picture segmentation algorithm for cell-network-based digital CMOS implementation," IEICE Trans. on Information & Syst., vol. E87-D, no. 2, pp. 500-503, Feb., 2004.

[3] T. Morimoto, Y. Harada, T. Koide, and H. J. Mattausch, "Pixel-parallel digital CMOS implementation of image segmentation by region growing," IEE Proc. Cir., Devices & Syst., vol. 152, no. 6, pp.579-589, Dec., 2005

[4] H. Adachi, T. Morimoto, K. Yamaoka, T. Koide, and H. J. Mattausch, "Image-scan architecture for efficient FPGA/ASIC implementation of video-segmentation by region growing," Proc. of the International SoC Design Conference, pp.301-304, Seoul, Korea, Oct. , 2005.

[5] K. Yamaoka, T. Morimoto, H. Adachi, K. Awane, T. Koide, and H. J. Mattausch, "Multi-object tracking VLSI architecture using image-scan based region growing and feature matching," Proc. of the IEEE Int'l Symp. on Circ. and Syst. (ISCAS), pp. 5575-5578, 2006.

[6] H.J. Mattausch, T. Gyohten, Y. Soda and T. Koide, "An Architecture for Compact Associative Memories with Deca-ns Nearest-Match Capability up to Large Distances," IEEE International Solid-State Circuits Conference Digest of Tech. Papers (ISSCC'2001), pp. 170-171, 2001.

[7] H.J. Mattausch, T. Gyohten, Y. Soda and T. Koide, "Compact Associative-Memory Architecture with Fully-Parallel Search Capability for the Minimum Hamming Distance", IEEE Journal of Solid-State Circuits, 37, pp. 218-227, 2002.

[8] Y. Yano, T. Koide, and H.J. Mattausch, "Associative Memory with Fully Parallel Nearest-Manhattan-Distance Search for Low-Power Real-Time Single-Chip Applications," Proc. of the Asia and South Pacific Design Automation Conference (ASP-DAC'2004), pp. 543-544, 2004.

[9] M. A. Abedin, K. Kamimura, A. Ahmadi, T. Koide, H. J. Mattausch, "Minimum Euclidean Distance Associative Memory Architecture with Fully-Parallel Search Capability," Workshop on Synthesis And System Integration of Mixed Information Technologies (SASIMI'2006), pp. 350-354, 2006.

[10] Md. A. Abedin, Y. Tanaka, A. Ahmadi, T. Koide and H.J. Mattausch, "Nearest Euclidean-Distance-Search Associative Memory Architecture with Fully Parallel Mixed Digital-Analog Match Circuitry," Extended Abstracts of the 2006 International Conference on Solid State Devices and Materials (SSDM'2006), pp. 282–283, 2006.

[11] Md. A. Abedin, Y. Tanaka, A. Ahmadi, T. Koide and H.J. Mattausch, "Fully Parallel Associative Memory Architecture with Mixed Digital-Analog Match Circuit for Nearest Euclidean Distance Search," Proc. of the IEEE Asia Pacific Conference on Circuits and Systems (APCCAS'2006), pp. 1311–1314, 2006.

[12] T. Morimoto, O. Kiriyama, Y. Harada, H. Adachi, T. Koide, and H.J. Mattausch, "Object Tracking in Video Pictures based on Image Segmentation and Pattern Matching," Proc. IEEE Int. Symposium on Circuits and Systems (ISCAS'05), Kobe 2005, 3215-3218 (2005.5)

[13] K. Yamaoka, T. Morimoto, H. Adachi, T. Koide, and H.J. Mattausch, "Image Segmentation and Pattern Matching Based FPGA/ASIC Implementation of Real-Time Object Tracking," Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC'2006), pp. 176–181, 2006.

[14] K. Yamaoka, T. Morimoto, H. Adachi, K. Awane, T. Koide, and H.J. Mattausch, "Multi-Object Tracking VLSI Architecture using Image-Scan based Region Growing and Feature Matching," Proc. IEEE Int. Symposium on Circuits and Systems (ISCAS'06), pp. 5575–5578, 2006.

[15] T. Morimoto, H. Adachi, K. Yamaoka, K. Awane, T. Koide, and H.J. Mattausch, "An FPGA-Based Region-Growing Video Segmentation System with Boundary-Scan-Only LSI Architecture," 2006 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS2006), pp. 946–949, 2006.

[16] A. Ahmadi, M. A. Ritonga, Md. A. Abedin, H.J. Mattausch, and T. Koide, "A Learning OCR System Using Short/Long-term Memory Approach and Hardware Implementation in FPGA," Proc. of the 2006 IEEE Congress on Evolutionary Computation (WCCI'2006), pp. 2702–2708, 2006.

[17] A. Ahmadi, Y. Shirakawa, Md. A. Abedin, K. Kamimura, H.J. Mattausch, and T. Koide, "An LSI Hardware Design for Online Character Recognition Using Associative Memory," Proc. of MWSCAS'2005 (IEEE Int'l Mid-west Symposium on Circuits and Systems), pp. 464–467, Cincinnati, USA, Aug. 2005.

[18] A. Ahmadi, Md. A. Abedin, H.J. Mattausch, T. Koide, Y. Shirakawa, and M. M. A. Ritonga, "Developing a Reliable Learning Model for Cognitive Classification Tasks Using an Associative Memory," IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP'2007), Hawaii, USA, April 2007. (accepted).