An Object Detection/Recognition System using a 3-Dimensional Integration with Local and Global Wireless Interconnections

Hiroshi Ando, Seiji Kameda, Nobuo Sasaki, Daisuke Arizono, Kentaro Kimoto†, Norimitsu Fuchigami, Kouta Kaya, Mamoru Sasaki, Takamaro Kikkawa† and Atsushi Iwata

Graduate School of Advanced Sciences of Matter, Hiroshima University
†Research Center for Nanodevices and Systems, Hiroshima University
Phone: +81-824-22-7358, E-mail: ando@dsl.hiroshima-u.ac.jp

1. Introduction

In order to realize hyper brain system which can recognize various objects in real-time/real-world, numbers of chips with massively parallel processing and wideband interconnection capabilities are needed. To assemble these multi chips with low-power and Gbps bandwidth interconnection, new integration techniques which replace the conventional System in Package techniques are required.

To solve the problem, we have proposed the 3-dimensional custom stack system (3DCSS) using two kinds of wireless interconnections: inductive coupling local wireless interconnect (LWI) and antenna coupling global wireless interconnect (GWI) [1]. In the system LWI is used to transmit/receive 2D image data between neighboring chips in parallel, and GWI is used to transmit/receive global system clocks and serial data such as control signals or database between all stacked chips. In the asynchronous LWI scheme without any clocking, the high bit rate of 1Gbps/ch and low power dissipation of 0.95mW/ch has been achieved by a 0.18μm CMOS technology [2]. The generation of ultra short Gaussian monocycle pulse which is the fundamental element for implementing GWI has been also demonstrated in the same technology [3].

To implement the multi-object recognition system, the processing algorithm and system/chip architecture which are suitable to the 3-dimensional integration technique have to be developed. Although many kinds of algorithm have been reported in a field of human face recognition [4], the most of these were developed aiming at software realization and did not apply to LSI implementation, because of complex large-scale calculation and huge memory capacity.

In this research, we have developed the architecture adopting "Eigenfaces" method based on PCA (Principal Component Analysis) which is one of the well-known face detection/recognition [5]. By combining the Eigenfaces method with 3DCSS, we have proposed architecture of the multi-object recognition system [6]. We have also implemented the prototype system developing two types of chips with a 0.18μm CMOS technology. The chip design utilizing the advantages of LWI and GWI has been also described.

2. Object detection/recognition algorithm

The Eigenfaces method should be suitable for object recognition hardware architecture because of several advantages in both of recognition performance and hardware implementation. This method has the equivalent or higher and robust recognition performance comparing with other recognition algorithms [7]. The various kinds of object can be detected and recognized by only preparing each individual database of them without changing processing. In hardware implementation, we can implement it with massively parallel circuit architecture and conventional digital circuit techniques without nonlinear processing, and design a chip which is commonly applied to detection and recognition without increasing in circuit area.

We explain a fundamental of the Eigenfaces algorithm briefly. An i-th face image consists of M pixels is represented as a row vector $\mathbf{\Gamma}_i$. A preprocessed face $\mathbf{\Phi}_i$ is defined by $\mathbf{\Phi}_i = \mathbf{\Gamma}_i - \mathbf{\Psi}$, where $\mathbf{\Psi}$ represents the average face of N images in DB(database), that is $\mathbf{\Psi} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{\Gamma}_n$. The "eigenfaces" can be calculated as the eigenvectors $\boldsymbol{a}_k$ (in an ascending order, k=1, 2, $\cdots$, m, $m \ll M$) of the covariance matrix C of DB, $C = \frac{1}{M}\sum_{n=1}^{M}\mathbf{\Phi}_n\mathbf{\Phi}_n^t$, where $\mathbf{\Phi}^t$ is a transposed matrix. A face image is transformed into so-called "eigen-space" ω_k by a simple operation: $\omega_k = \boldsymbol{a}_k^t\mathbf{\Phi}_i$. The eigen-space ω_k forms a vector $\mathbf{\Omega} = [\omega_1\ \omega_2\ \ldots\ \omega_m]$ that describes the contribution of each eigenface for face image.

Face detection is performed by generally used thresholding methods. A reconstructed image $\mathbf{\Phi}_r$, defined by $\mathbf{\Phi}_r = \sum_{k=1}^{m}\omega_k\boldsymbol{a}_k$, is used as an input of evaluation function for thresholding. For example, Euclidean distance $\varepsilon = \|\mathbf{\Phi}_{in} - \mathbf{\Phi}_r\|$ is often used as evaluation function, where $\mathbf{\Phi}_{in}$ is a preprocessed unknown input image. If the value ε is lower than a threshold, an unknown input image is classified as a human face. Face recognition is also achieved with the same calculations except for evaluation function. If the face space vector $\mathbf{\Omega}_{DB_i}$ of i-th face image in DB leads to the minimum distance $\varepsilon_{min} = \|\mathbf{\Omega} - \mathbf{\Omega}_{DB_i}\|$, we can know that the input face is the same as i-th face.

3. Hardware implementation

A schematic of the proposed multi-object recognition system is shown in Fig. 1. This system consists of three kinds of chips, that is Visual Processing chip (VP3D) [8], Detection/Recognition chip (DR3D) and Reference Memory chip (RM3D). Each chip has 21×2ch LWIs which can transmit to and receive data from neighboring chips simultaneously and 2ch GWIs for clock and binary digital data receiving. The RM3D has 2ch GWIs and transmitter circuits for clock and data.

Now we explain the proposed methods of detecting and recognizing by this system. At first, original image data is stored in $RM3D_1$ and transmitted to neighboring VP3D in 21-pixel parallel PWM (pulse width modulation) signals (LWI-1). The transmission rate is about 160Mbps when the maximum bit width and time resolution of PWM signal is 8bit and 4ns (250MHz clock distribution by GWI-1), respectively. Second, massively parallel image pre-processing is implemented by several VP3Ds and resulted image data is transmitted to $RM3D_2$ with LWI-2 as same as LWI-1. Finally, the DR3D receives pro-

cessed image data and object database through LWI-3 (5.3Gbps=21bit/4ns), or after storing other database to $RM3D_2$ from $RM3D_N$ by GWI-2 (250Mbps), and detects and recognizes objects.

This system has the ability of 40GOPS (Giga Operation Per Second) at 250MHz operation. Therefore, we expect to derive 160GOPS performance at the maximum LWI operation (1Gbps/1ch at present).

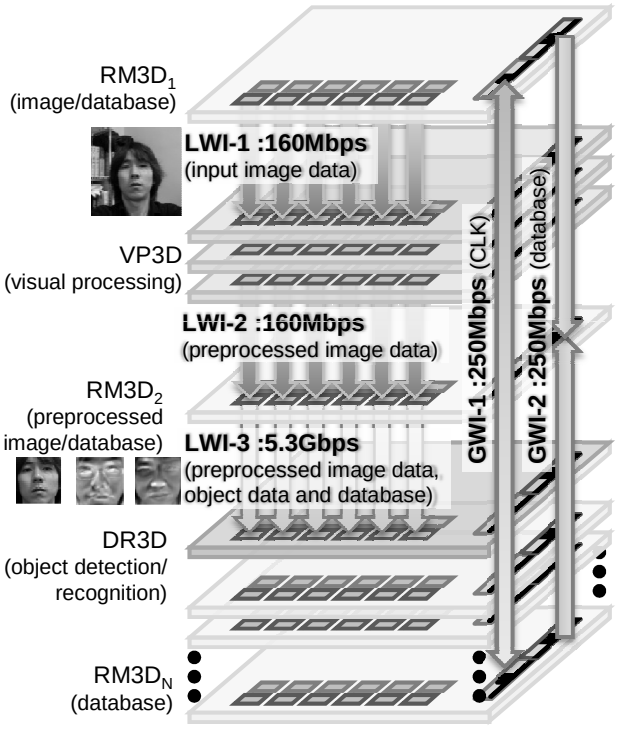


Figure 1: Multi-object recognition system with 3DCSS.

3.1 Reference memory chip - RM3D

Figure 2 (a) shows a block diagram of the proposed RM3D storing reference data of both of VP3D and DR3D. The capacity of SRAM is 56kbits for image data $\mathbf{\Gamma}$, 196kbits and 123kbits for database $\mathbf{\Psi}$ and $\boldsymbol{a}$. In communicating with VP3D, the binary digital image data is modulated to PWM signal by DPC (Digital-to-PWM Converter) and transmitted to VP3D in 21-pixel parallel with LWI. The visual processed data is received and stored after demodulation by PDC (PWM-to-Digital Converter). The 21bit digital bus data for one pixel (8bit $\mathbf{\Gamma}$, 8bit $\mathbf{\Psi}$ and 5bit $\boldsymbol{a}$) is transmitted to DR3D in pixel serial with LWI. The clock signal generated by VCO (Voltage Controlled Oscillator) and binary data stored in memory are transmitted to all of stacked chips. If we need huge memory capacity for database, we should only stack several RM3Ds because of wireless wideband communications by GWI.

3.2 Detection/recognition chip - DR3D

A block diagram of the proposed DR3D which enables to implement the object detection/recognition algorithm mentioned in Sec. 2 is shown in Fig. 2 (b). The DR3D can achieve the two operation modes of object detection and recognition in common circuits and 32-pixel parallel operation by utilizing the advantages of Eigenfaces algorithm.

At first, 21bit bus data are received by LWI and stored to each 32×32 shift register, where pixel size of object is 32×32, and converted to 32-pixel parallel data by shift-register. Second, in reconstructed image generator, $\mathbf{\Phi}_i = \mathbf{\Gamma}_i - \mathbf{\Psi}$ is calculated by subtracter, $\omega_k = \boldsymbol{a}_k^t \mathbf{\Phi}_i$ is calculated by multiplier and we obtain $\mathbf{\Phi}_r = \sum_{k=1}^{m} \omega_k \boldsymbol{a}_k$ by accumulator in 32-pixel parallel. Finally, Manhattan distance $\varepsilon_i = \|\mathbf{\Phi}_i - \mathbf{\Phi}_r\|$ is calculated with subtracter and compared in Winner-take-all circuits, detection or recognition process is finished.

Thus, the proposed multi-object recognition system could be implemented by making the most of LWI's and GWI's advantages that the Gbps multi channel communications enable to execute parallel processing and long-line wireless communications make it possible to stack several memory chips.

3.3 Fabrication and integration

Test chips of RM3D and DR3D fabricated in a 0.18μm CMOS technology are shown in Fig. 3. The chip size was 5×5mm^2, and the supply voltage and operation frequency were 1.8V and 250MHz, respectively. The detection time was 580μs and the one-object to one-database recognition time was 4.2μs at 84×84 image and 32×32 object size. The 20.6ms detection time and 12.7ms recognition time (30fps) should be achieved if we estimate the system ability at QVGA image which includes about 30 objects and 100 database objects.

The custom flexible printed circuit (FPC) shown in Fig. 4 was developed for testing each chip. Note that the most of area around a chip is needless because this FPC was used for preliminary measurements. We confirmed basic operation such as memory read/write, control signal generation. Now prototype 3DCSS is under development by stacking the measured chip.

4. Conclusion

The multi-object recognition system architecture was developed by utilizing the recognition algorithm based on Eigenfaces method and the 3-D integration scheme (3DCSS) with two types of wireless interconnections of LWI and GWI. The prototype system was designed with 3 types of chips for object detection/recognition, reference data storage and image pre-processing. Processing performance of 40GOPS at 250MHz was obtained by the chips with a 0.18μm CMOS technology. Object detection and recognition system performance of 580μs detection time and 4.2μs one-object to one-database recognition time was obtained.

References

[1] A. Iwata, et al., "A 3D-integration scheme utilizing wireless interconnections for implementing hyper brains", Digest of ISSCC2005, pp. 262-263, Feb 6-10, 2005.

[2] M. Sasaki, et al., "A 0.95mW/1.0Gbps spiral-inductor based wireless chip-interconnect with asynchronous communication scheme", Digest of Sympo. on VLSI Circuits, June 2005.

[3] N. Sasaki, et al., "A single-chip Gaussian monocycle pulse transmitter using 0.18μm CMOS technology for intra/interchip UWB communication", Digest of Sympo. on VLSI Circuits, 2006.

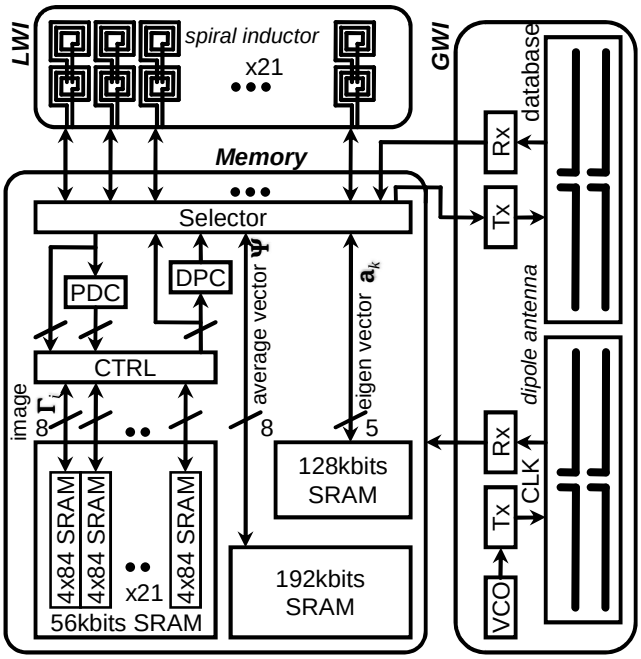
[4] W. Zhao, et al., "Face Recognition: A Literature Survey", *ACM CSUR archive*, vol. 35, pp. 399-458, 2003.

[5] M. A. Turk, et al., "Eigenfaces for Recognition", *CVPR'91*, pp. 586–591, 1991.

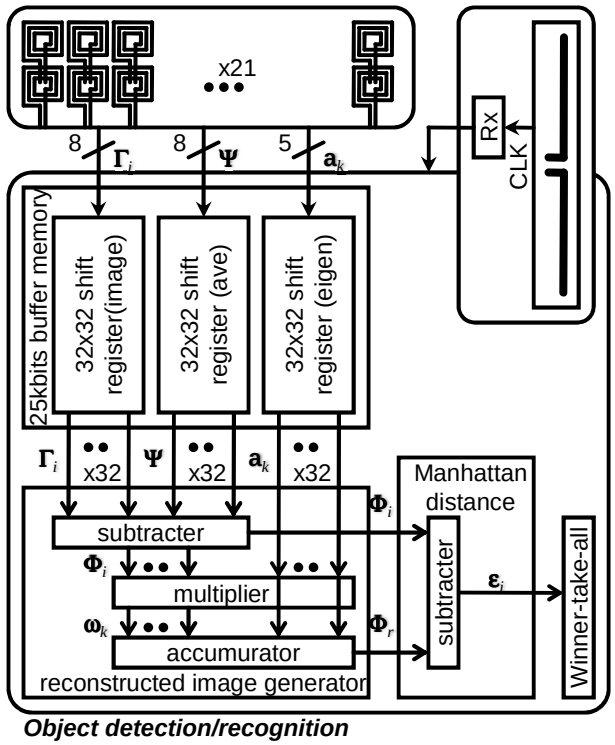
[6] H. Ando, et al., "A prototype software system for multi-object recognition and its fpga implementation", *Proc. Third Hiroshima International Workshop on NTIP*, pages 89–90, 2004.

[7] P. J. Phillips, et al., "The FERET evaluation methodology for face-recognition algorithms", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1090–1104, 2000.

[8] S. Kameda, et al., "A Brain-type vision system using a 3-Dimensional integration with local and global wireless interconnections",*Proc. Fourth Hiroshima International Workshop on NTIP*, pages 38–39, 2005.

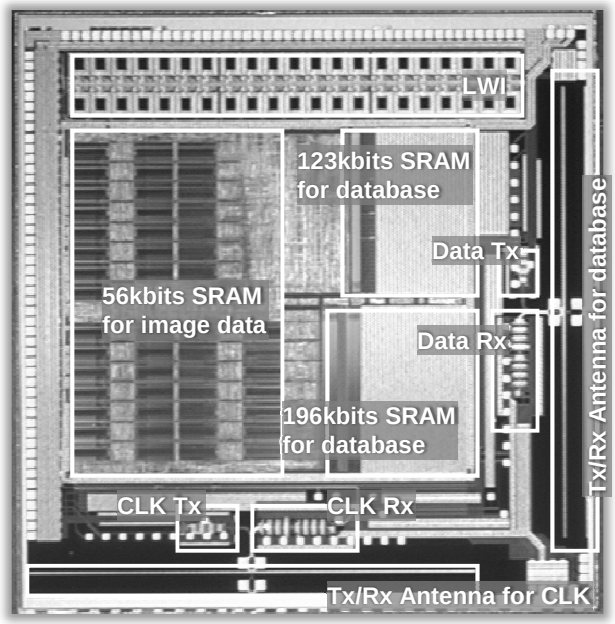


(a) RM3D

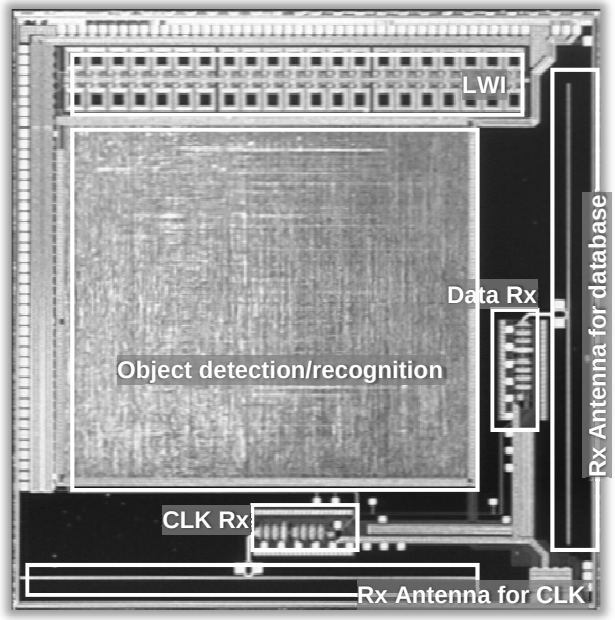


(b) DR3D

Figure 2: Block diagrams of RM3D and DR3D.



(a)



(b)

Figure 3: Microphotographs of RM3D (a) and DR3D (b).

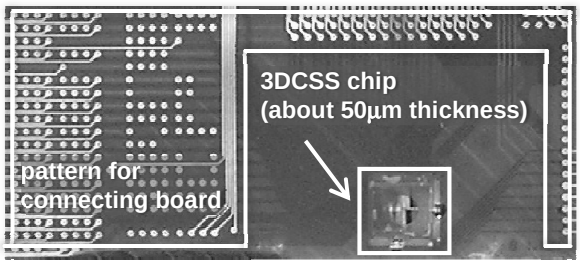


Figure 4: FPC board.