Tera-Scale Performance Machine Learning SoC (MLSoC) With Dual Stream Processor Architecture for Multimedia Content Analysis

Tse-Wei Chen, Chi-Sun Tang, Sung-Fang Tsai, Chen-Han Tsai, Shao-Yi Chien, *Member, IEEE*, and Liang-Gee Chen, *Fellow, IEEE*

Abstract—A new machine learning SoC (MLSoC) for multimedia content analysis is implemented with 16-mm^2 area in 90-nm CMOS technology. Different from traditional VLSI architectures, it focuses on the coacceleration of computer vision and machine learning algorithms, and two stream processors with massively parallel processing elements are integrated to achieve tera-scale performance. In the dual stream processor (DSP) architecture, the data are transferred between processors and the high-bandwidth dual memory (HBDM) through the local media bus without consuming the AMBA AHB bandwidth. The image stream processor (ISP) of the MLSoC can handle common window-based operations for image processing, and the feature stream processor (FSP) can deal with machine learning algorithms with different dimensions. The power efficiency of the proposed MLSoC is 1.7 TOPS/W, and the area efficiency is 81.3 GOPS/mm^2.

Index Terms—Digital circuit, hardware architecture, machine learning, multimedia content analysis, system-on-a-Chip (SoC).

I. INTRODUCTION

THE significance of mining and recognition in the next era of Tera [1] has received more and more attention in recent years. As a subfield in artificial intelligence, machine learning [2] provides a series of algorithms for mining and recognition, such as supervised learning algorithms, unsupervised learning algorithms, and reinforcement learning algorithms. These algorithms are widely employed in different applications in multimedia content analysis, including face detection [3], color image segmentation [4], and content-based image retrieval [5]. The repetitive operations for high-dimensional vector processing result in laborious computations for machine learning algorithms, so it is difficult to meet the real-time requirement by using traditional processors. Many hardware architectures and design methodologies for machine learning algorithms, such as Gaussian mixture model-based classification [6] and K-means clustering [7], [8], are proposed to accelerate the computational speed, but the hardware integration of different kinds of algorithms is still an open question. At the same time,

The authors are with the Graduate Institute of Electronics Engineering and Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan, R.O.C. (e-mail: sychien@cc.ee.ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

the development of high-resolution CMOS image sensors [9] has introduced the requirements for high-performance video processing, which is crucial to satisfy the real-time requirement in mobile systems. Different kinds of large-kernel image processing operations, such as the median filter, the sharpening filter, the bilateral filter [10], and the Gaussian filter, are essential to video noise reduction and video quality enhancement. Therefore, it is necessary to develop suitable architectures and platforms for the coacceleration of machine learning and image processing tasks.

Due to the rapid growth of consumer electronics and advances of semiconductor technology, mobile devices, such as digital cameras, portable computers, and cellular phones, are equipped with various kinds of high-performance processors. To analyze the content of image and video data, many kinds of VLSI architectures are proposed. Abbo *et al.* propose a massively parallel processor for video scene analysis [11], Kim *et al.* propose a processor with a visual attention engine to recognize objects [12], and Cheng *et al.* propose an SoC which combines a processor with a CMOS sensor [13]. Nevertheless, these processors do not have suitable architectures to accelerate machine learning algorithms for multimedia content analysis. Different from the previously reported works [11]–[13], a high-performance machine learning SoC (MLSoC) for multimedia content analysis is proposed. It focuses on the coacceleration of computer vision and machine learning algorithms, and the image stream processor (ISP) and the feature stream processor (FSP) are integrated into the dual stream processor (DSP). Both processors, the ISP and the FSP, are established on the 256-bit local media bus (LMB), which is directly connected to the high-bandwidth dual memory (HBDM). The HBDM offers the DSP the instant access of video data and feature data, and the hardware architecture can achieve a maximum throughput of 62.5 Gpixel/cycle for image processing operations and 16 vector/cycle for machine learning algorithms.

This paper is organized as follows. The proposed SoC architecture is first described in Section II. Then, the architectures of the ISP and the FSP are introduced in Sections III and IV, respectively. Next, the VLSI implementation of the proposed work is shown in Section V. Finally, a short summary is given in Section VI.

II. MLSOC ARCHITECTURE

Fig. 1 shows the system diagram of the MLSoC, which contains a complete platform for multimedia content analysis. There are two AMBA AHBs [14] of different bandwidth, and
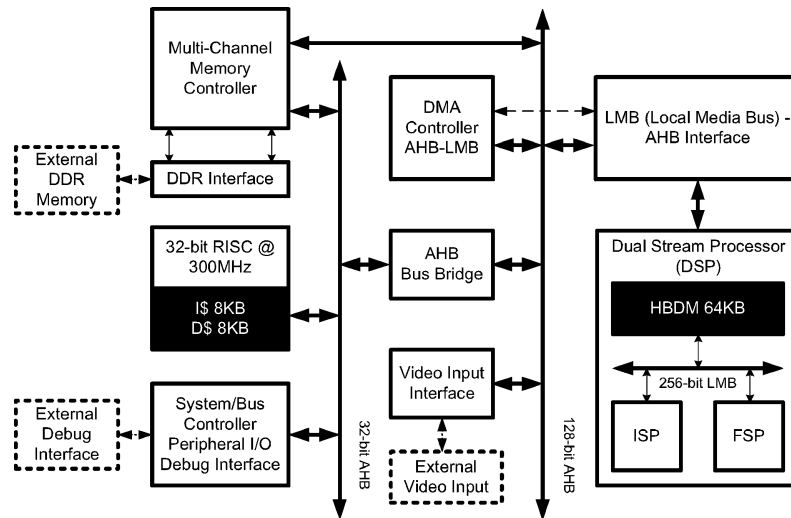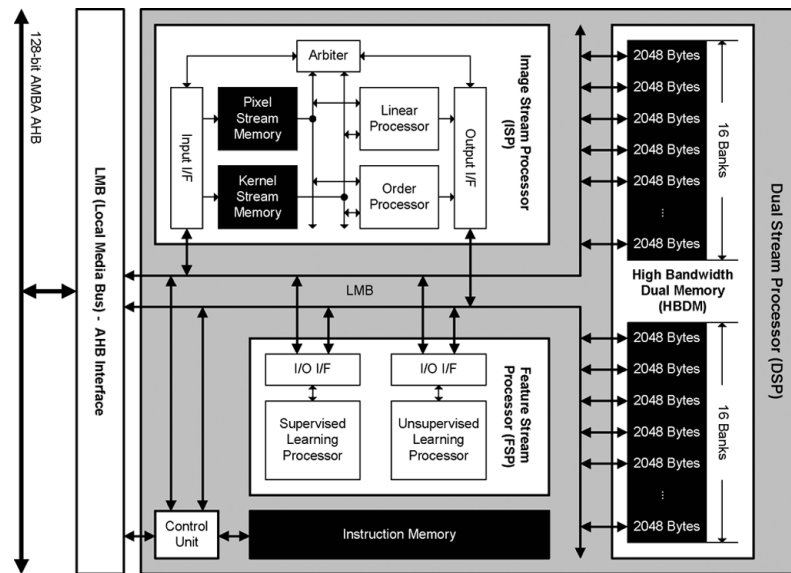
Fig. 1. Architecture of the proposed MLSoC.

Fig. 2. Architecture of the DSP.

the data in the two buses can be exchanged through the AHB bus bridge. Each silicon intellectual property (SIP) on two AHBs can access the external DDR memory through the multi-channel memory controller. To fully utilize the bus bandwidth, the RISC is connected to the 32-bit AHB, whereas the DSP is connected to the 128-bit AHB. The system and application tasks are executed on the RISC, and the DSP receives instructions from the RISC through the LMB-AHB interface, which is connected to the DMA controller to efficiently access video or feature data from the external DDR memory.

The architecture of the DSP is shown in Fig. 2. After receiving the instructions, the control unit in the DSP analyzes the instructions to manipulate the operations with two stream processors. In the DSP architecture, the data are accessed through the LMB, which is connected to the 32-bank HBDM. The half size of the HBDM is sufficient to store an image of 160×120 pixels, which can be obtained from the video input interface by down-sampling or slicing the image data. The purpose of the HBDM is to store the image data for the ISP and the feature data for the FSP, and the data in one memory of the HBDM can be fully copied to the other in 2048 cycles, where one cycle is defined as the inverse of the clock frequency of the system in this paper. The rapid data access accelerates the computation for multimedia content analysis and reduces the power consumption of data transmission between the RISC and the DSP because most of the operations are performed in the architecture of the DSP. Besides, the maximum input bandwidth of a subprocessor inside the DSP reaches 2048 bit/cycle.

III. IMAGE STREAM PROCESSOR

Large kernel operations for pixels are significant for image processing tasks in multimedia content analysis. For example, the Gabor transform [15] can be used for image texture extraction, and the Gaussian filter can be applied to the detection of scale-space extrema for feature points [16]. The image stream processor (ISP) employs massively parallel processing

elements to process image or video data, and the input bandwidth of image pixels is transformed from 16 pixel/cycle to 256 pixel/cycle inside the ISP by the input interface to accomplish window-based operations. In other words, a maximum of 16×16 pixels in a window can be processed in the same cycle based on the ISP architecture. Since the window-based operations are processed in the raster-scan order, the "Pixel Stream Memory," which is connected to the input interface in the ISP, can instantly offer the subprocessors the pixels of the current window by receiving the 16 new pixels from the HBDM and discarding 16 old pixels in the previous window. Therefore, the ISP can help achieve tera-scale performance even if the bandwidth of the LMB is only 256 bit. The ISP consists essentially of an arbiter, two subprocessors, and a set of shared memory. The arbiter controls the behavior of each processor according to the ISP instructions, and the processor not in use is automatically set to be inactive to save the power consumption. The set of shared memory, which includes "Pixel Stream Memory" and "Kernel Stream Memory," is used to store the pixel data and the kernel data for image processing tasks. Two subprocessors are the linear processor and the order processor, both of which are able to handle 256 pixel streams simultaneously, and the bit length of each pixel stream is 8 bit. These two subprocessors have the same output bandwidth, 1 pixel/cycle, and they deal with parallel data-in, scalar data-out image processing tasks.

In order to fetch the input pixel stream from the HBDM, a new memory architecture is developed for window-based operations in image processing tasks. The memory utilization scheme is illustrated in Fig. 3, where two examples are shown. The input image, whose width is denoted by W, is first partitioned into different slices, each of which includes $W \times 16$ pixels. These slices are stored to one of the HBDM and sequentially arranged. When the ISP demands pixels in a 16×16-pixel window, the HBDM sends 16 pixels simultaneously to the ISP in one cycle, and the memory bandwidth of the HBDM can be fully utilized. For example, in Fig. 3, the pixels in *Window 1* occupy *Slice 2* and *Slice 3*, and the pixels in *Window 2* occupy *Slice N*. No matter what the window position is, the pixels required by the ISP never occupy the same bank in the HBDM. Based on the memory architecture of the HBDM, the traditional line buffer memory [17], which is used for window-based operations, can be saved to reduce hardware costs.

A. Linear Processor

Fig. 4(a) shows the architecture of the linear processor, which handles linear operations for image processing. There are four levels of configuration network in this processor, and the amount of data is reduced level by level. The processing elements in the first level can deal with multiplications of a maximum of 256 parallel pixels and their corresponding coefficients from the window in the image, and the results are sent to the next level through the configuration network, which contains a set of context registers to manipulate the stream of pixels in each level. In the second level, the ALU trees handle subtraction and addition operations based on the results of multiplications in first level. The data are then collected in the next level through
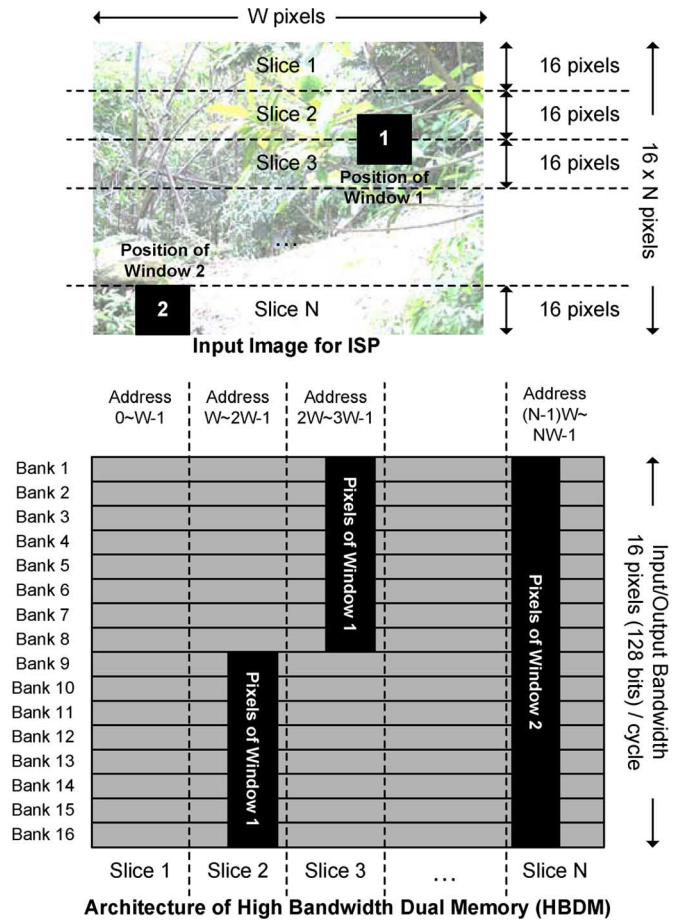

Fig. 3. Memory utilization scheme of the HBDM for the ISP.

the configuration network. In the third level, dedicated accelerators for face detection, pixel variance calculation, and correlation coefficients, are integrated into this processor to enhance the functionalities for video analysis. Based on the statistical analysis, these dedicated functionalities are frequently used in the image processing tasks for multimedia content analysis, and the high-throughput divider is used to compute one division operation per cycle. Then, the ALU in the fourth level performs simple instructions, such as additions and subtractions for the final output data. General image linear operations, including the Laplacian filter, the low-pass filter, the Gaussian filter, the Gabor transform, and the 16×16-pixel convolution, can be executed in one cycle with 40 cycles of latency.

In this architecture, the pixels can be repetitively processed by the processing elements in different levels of the configuration network. Different from traditional processors, the equivalent input bandwidth of the ISP is computed according to the total bandwidth of the input bandwidth to each of the pipeline stages. The bandwidth is distributed to the processing elements by the configuration network in four levels as shown in Fig. 4(a), and the performance of the linear processor can achieve 0.67 TOPS (Input Bandwidth = 0.31 TB/s + 0.98 TB/s + 0.04 TB/s = 1.33 TB/s) while the clock frequency is 300 MHz. This performance also results from the special design of the processor architecture, where the processing elements are cascaded in each
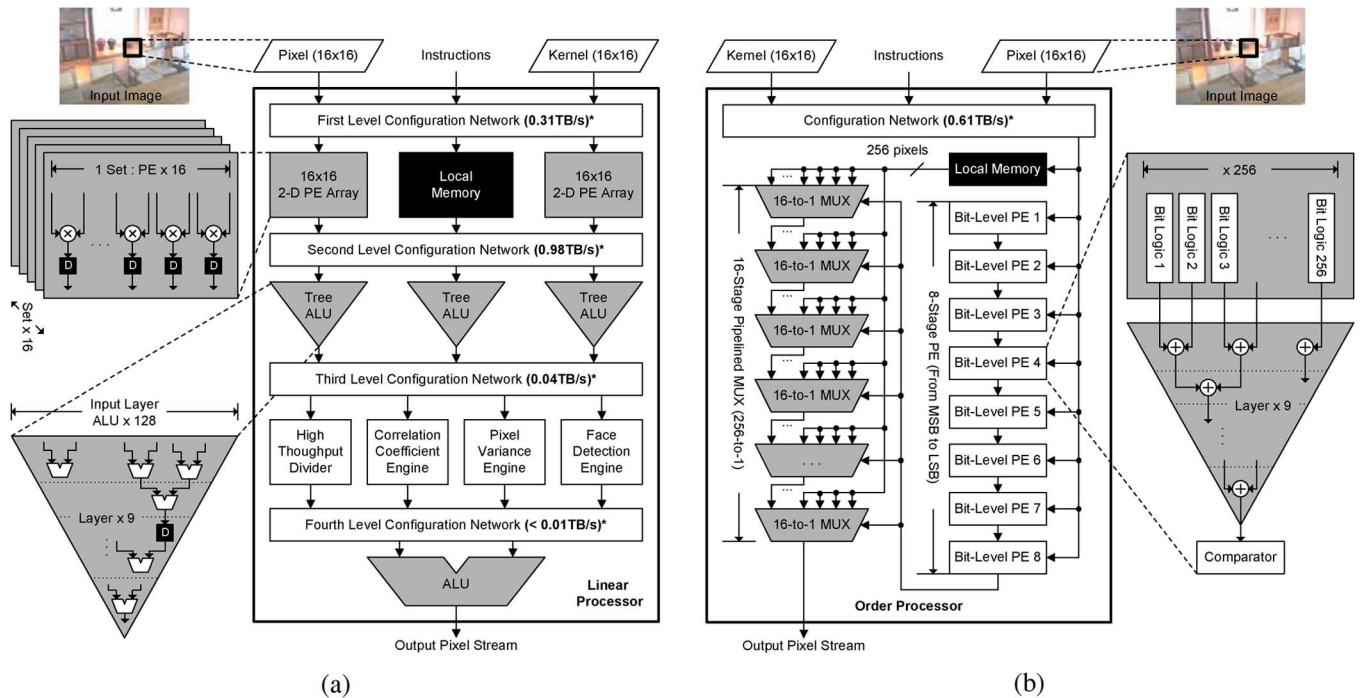
Fig. 4. (a) Architecture of the linear processor in the ISP. (*The bandwidth that is consumed in the processing elements in each level is listed.) (b) Architecture of the order processor in the ISP. (*The bandwidth that is consumed in the processing elements is listed.)

level. Note that in this paper, the 8-bit operations are considered for the performance evaluation.

B. Order Processor

Fig. 4(b) shows the architecture of the order processor, which deals with sorting operations for image processing. The sorting procedure [18], [19] is implemented by a set of reconfigurable hardware, which contains eight-stage processing elements for bit-wise operations and a set of multiplexers to compute the rank order of a set of parallel pixels. Each stage of processing elements contains 256 parallel bit-logic modules, an adder tree with nine layers, and a comparator to compare the results of summation between stages. The set of multiplexers is 16-stage pipelined to shorten the critical path for the real-time processing requirement. Common nonlinear image operations, such as the morphological filter, the 16×16-pixel median filter, the arbitrary kernel median filter, can all be executed in one cycle with 40 cycles of latency.

Similar to the linear processor, the configuration network of the order processor is responsible for the bandwidth allocation to eight "Bit-Level Processing Elements." Based on the configuration network, a total of 0.61 TB/s bandwidth can be supplied to each processing element at 300 MHz. The processing elements are also cascaded in each level to simultaneously process the input pixel streams, so the equivalent input bandwidth to the processing elements in the sub-processor is much higher than the total bandwidth of input pixels streams and kernel streams (Input Bandwidth $= 76.8$ GB/s $+ 76.8$ GB/s $= 0.15$ TB/s).

These two sub-processors, the linear processor and the order processor, enable the ISP to handle most of common operations used in the image processing tasks of computer vision algorithms, such as pre-processing and filtering. The performance

TABLE I
PERFORMANCE OF SOME ISP SINGLE OPERATIONS

Operation Description	Input Data Rate* (Gpixel/s)	Throughput* (fps)
3×3-Pixel Edge Detector	2.2	12173
5×5-Pixel Sharpening Filter	6.0	11860
7×7-Pixel Smoothing Filter	11.9	11558
9×9-Pixel Median Filter	19.6	11268
11×11-Pixel Dilation	29.4	10989
13×13-Pixel Erosion	41.1	10720
15×15-Pixel Laplacian Filter	54.9	10462
16×16-Pixel Gaussian Filter	62.5	10336
16×16-Pixel Variance	62.5	10336
16×16-Pixel Gabor Transform	62.5	10336

*The clock frequency of the ISP is 300MHz, and the frame size is 160×120 pixels.

of some ISP single operations for the DSP is listed in Table I. Although the proposed architecture is designed to handle the operations with a maximum of 16×16 pixels in a window, it can be used for other window sizes which are smaller than 16×16 pixels as well. As long as the AHB resources are available, the performance of window-based operations from 3×3 pixels to 16×16 pixels listed in Table I can be higher than 100 fps for an HDTV image (1920×1080 pixels) when the clock frequency is 300 MHz.

IV. FEATURE STREAM PROCESSOR (FSP)

The FSP, whose architecture is shown in Fig. 5, is intended to handle feature vectors extracted for multimedia content analysis, and it contains two subprocessors for machine learning algorithms: the supervised learning processor and the unsupervised learning processor. Although these two processors support different algorithms, the common property is that they both
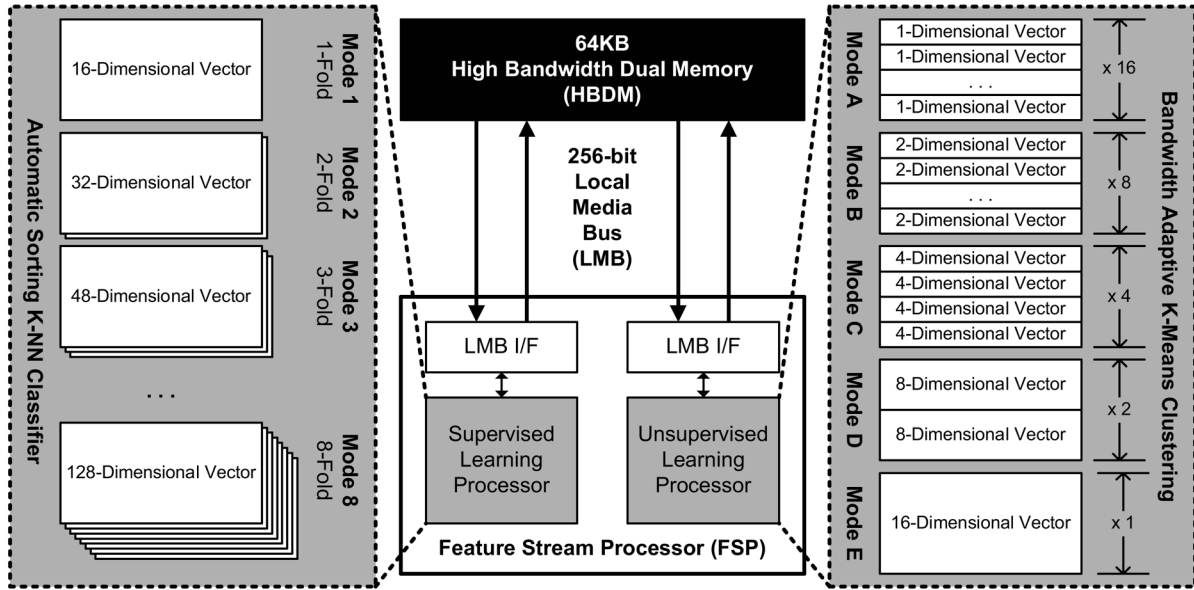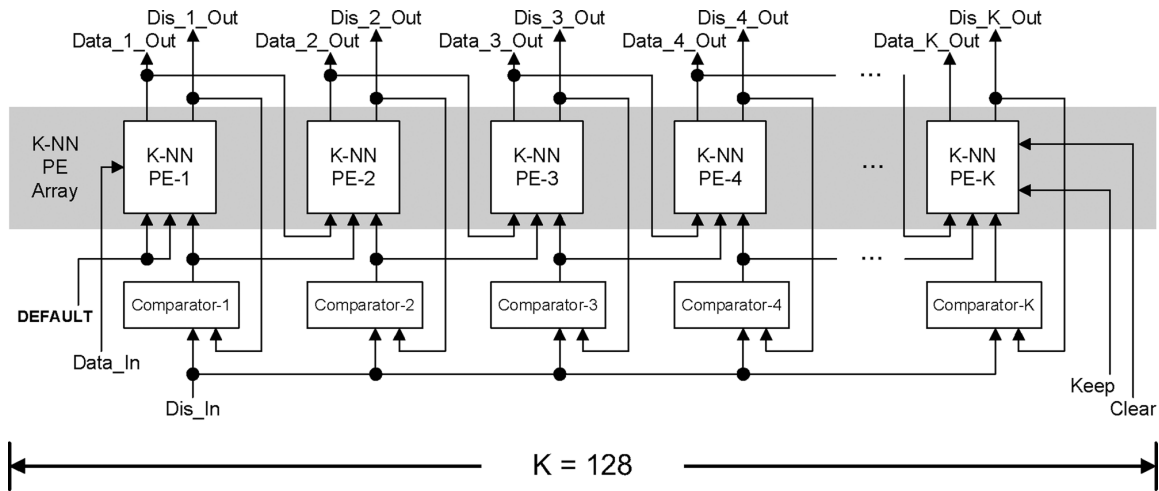
Fig. 5. Architecture of the FSP.

Fig. 6. Architecture of the automatic sorting mechanism for K-nearest neighbor in the FSP.

need a large amount of data bandwidth, and the supported algorithms are suitable for stream processing because of the regular data access pattern. Therefore, the two subprocessors are directly connected to the LMB to deal with the feature data in the HBDM.

Both of the sub-processors are able to compute either the Manhattan distance or the Euclidean distance by using the same hardware resources, and the distance measurement is selected according to the instructions. The supervised learning processor can handle the K-nearest neighbor algorithm, which is frequently employed in information retrieval applications, and it contains a set of parallel processing elements that is able to handle feature vectors with a maximum of 128 dimensions. It also employs the automatic sorting mechanism for distance ranking, and the results can be immediately dumped out without extra sorting stages. The architecture of the automatic sorting mechanism is shown in Fig. 6, which contains a total of 128 parallel processing elements to compute the ranking of dis-

tances. In addition, to fully utilize the bandwidth of the LMB, the unsupervised learning processor employs the bandwidth adaptive mechanism [20], which allocates different hardware resources for K-means clustering algorithm according to the number of feature vector dimensions. A simplified version of the architecture of the bandwidth adaptive mechanism is illustrated in Fig. 7, where four-parallel 1-D vectors are simultaneously processed in the three layers of different sets of processing elements in *Mode 1*, and two-parallel 2-D vectors and one-parallel 4-D vectors can be processed in *Mode 2* and *Mode 3*, respectively. Much more complicated than Fig. 7, the unsupervised learning processor focuses on the bandwidth adaptive mechanism for 16-D vectors, and the HBDM enhances the efficiency of the data access for the iteration process of K-means clustering. The performance of some FSP single operations for the DSP is listed in Table II, and the throughput of vectors can be adjusted to the bandwidth to obtain the optimal efficiency.
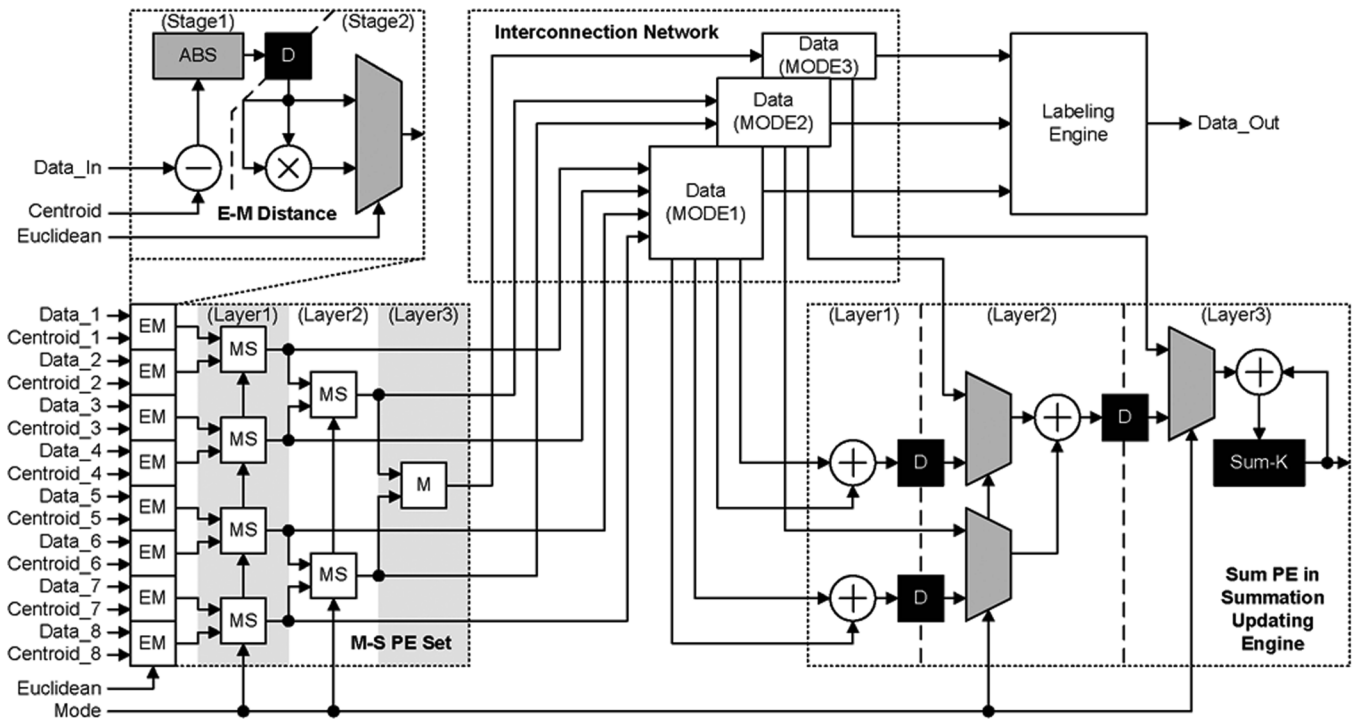
Fig. 7.   Architecture of the bandwidth adaptive mechanism for K-means in the FSP.

TABLE II
PERFORMANCE OF SOME FSP SINGLE OPERATIONS

Operation Description	Dimension	Throughput (vector/cycle)
K-Means Clustering	1	16
K-Means Clustering	2	8
K-Means Clustering	4	4
K-Means Clustering	8	2
K-Means Clustering	16	1
K-NN Ranking	16	1
K-NN Ranking	32	0.5
K-NN Ranking	64	0.25
K-NN Ranking	128	0.125

TABLE III
MLSoC SPECIFICATIONS

Technology	TSMC 90nm 1P9M CMOS
Die Size	16mm^2 (4mm × 4mm)
Power Supply	Core 1.2V, I/O 2.5V
Maximum Clock Frequency	300MHz
Logic Gate Count	Dual Stream Processor: 1.55M
(2-Input NAND Gate)	RISC and Others: 42.54K
On-Chip SRAM	HBDM: 64KB
	Others: 15KB
Peak Performance*	1.3TOPS
Power Dissipation*	768mW

*The peak performance and the power dissipation are measured when all sub-processors in the DSP (the linear processor, the order processor, the supervised learning processor, and the unsupervised learning processor) are active.

While applying the bandwidth adaptive mechanism with cluster number $K = 16$, a maximum of 16 cluster centroids are processed in parallel. Therefore, the input vector streams are repetitively sent to the "E-M Distance Calculator Set" [20] (Input Bandwidth = 0.38 TB/s) and the "M-S PE Set" [20]
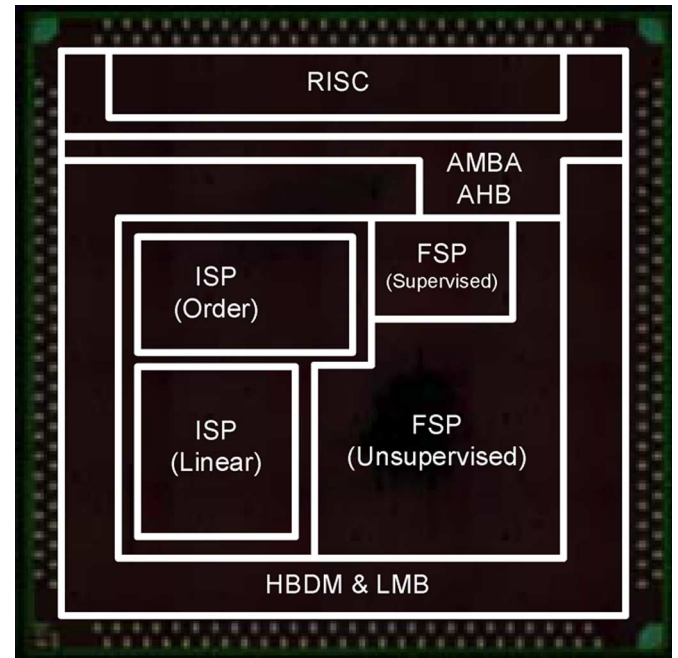

Fig. 8.   Chip micrograph.

(Input Bandwidth = 0.33 TB/s) to process two 256-D vector streams in parallel. Therefore, the performance of this processor can achieve 0.35 TOPS (Input Bandwidth = 0.38 TB/s + 0.33 TB/s) when the clock frequency is 300 MHz.

V. VLSI IMPLEMENTATION

The proposed MLSoC is implemented on a 16-mm^2 die using TSMC 90-nm 1P9M process. The maximum operating
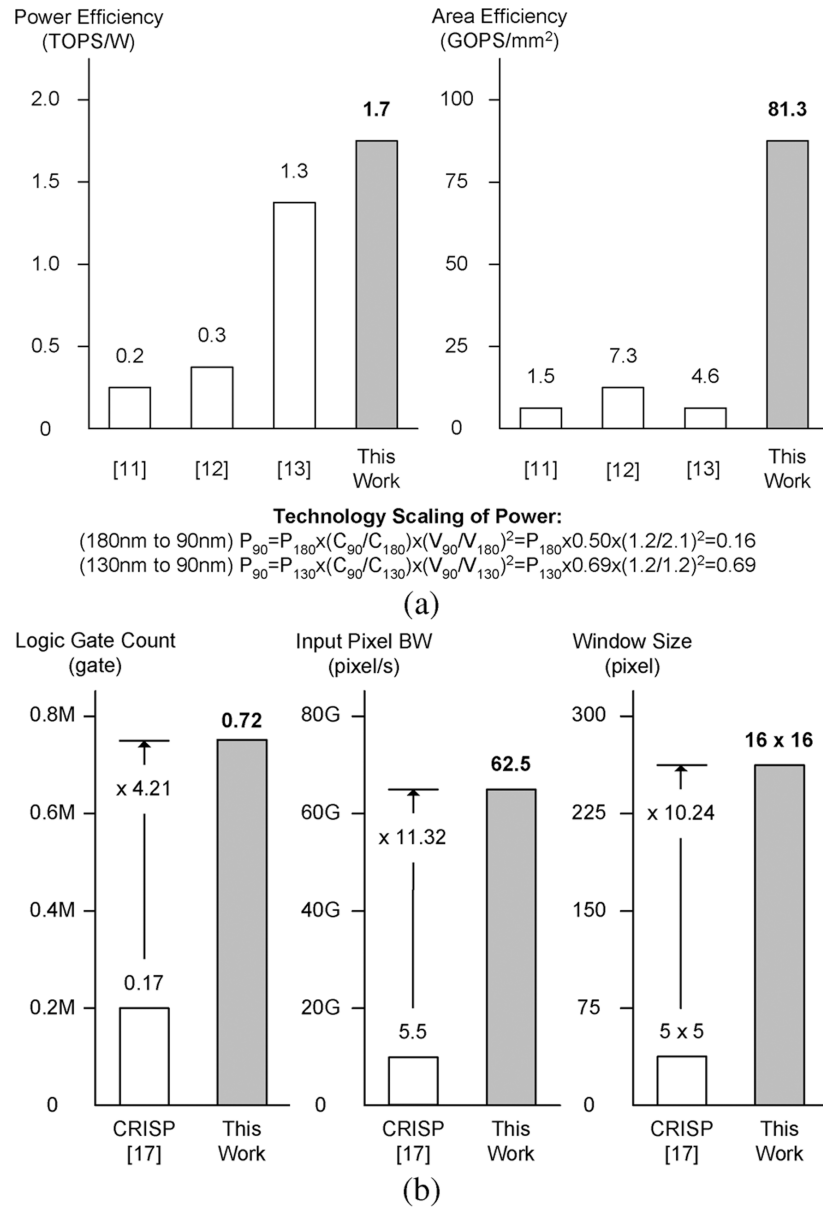
Fig. 9.   (a) Comparison of this work and the previously reported works [11]–[13] based on the power efficiency and the area efficiency. (b) Comparison of the ISP in the MLSoC and the CRISP [17].

frequency of the proposed MLSoC is 300 MHz, and the total on-chip memory is 79 KB. Table III shows the chip specifications, and the chip micrograph is shown in Fig. 8. The peak performance of the MLSoC achieves 1.3 TOPS (Input Bandwidth $= 1.33$ TB/s $+ 0.61$ TB/s $+ 0.38$ TB/s $+ 0.33$ TB/s), and the input bandwidth inside these processors can achieve more than 2.6 TB/s while one operation uses 2-Byte pixels for image data or 2-Byte vectors for feature data. The power efficiency and the area efficiency of this work are compared with the previously reported works [11]–[13] in Fig. 9(a). The proposed work achieves 1.7 TOPS/W in the power efficiency and 81.3 GOPS/mm^2 in the area efficiency, both of which are the highest among the four works.

The ISP in the MLSoC is compared with the related work, the coarse-grained reconfigurable image stream processor (CRISP) [17], and the result is shown in Fig. 9(b), which shows that the

proposed ISP uses 4.21 times of logic gate count to handle more than 11 times of input pixel bandwidth and to support 10.24 times of the maximum window size.

Fig. 10 shows the flowchart of an algorithm example using the proposed MLSoC. The image stream (160×120 pixels) is first loaded to the HBDM through the LMB-AHB interface. The ISP executes the instructions for "median filter" and sends the processed image stream to the HBDM. The FSP regards the data stored in the HBDM as the feature stream and executes K-Means clustering by using the bandwidth adaptive mechanism. When input vectors are 1-D, 16 times the speed acceleration can be achieved. The clustered feature vectors are stored to the HBDM again and the process of image segmentation is completed. Then the data can be transferred to the external memory by the DMA controller. The total time to compute the algorithm example in the DSP, where 32 iterations are performed in K-means, is
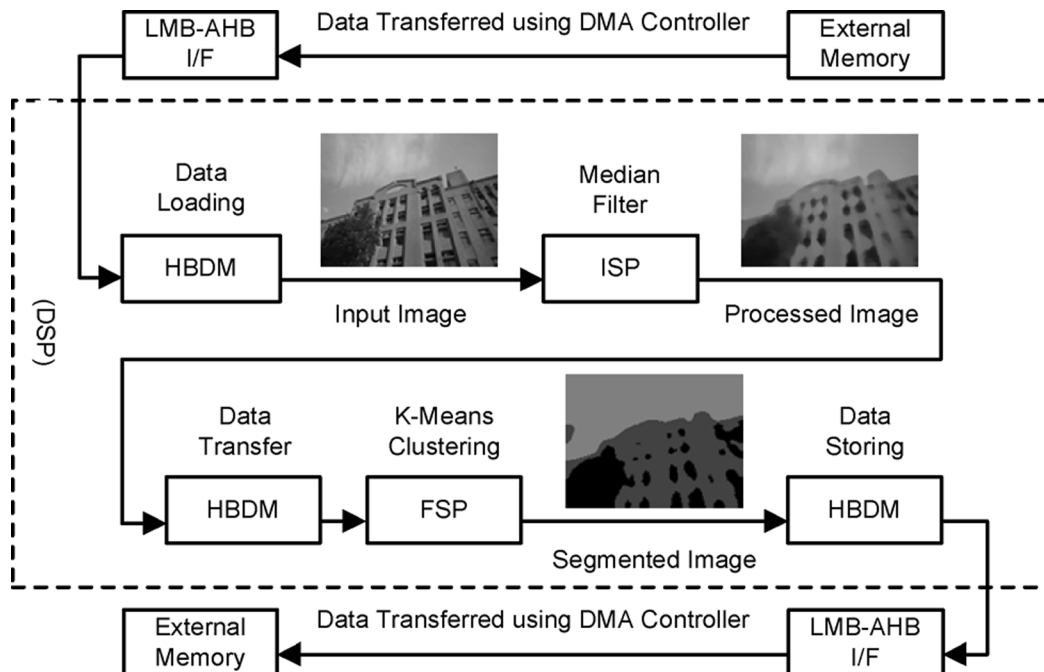
Fig. 10. Algorithm example of the proposed MLSoC.

less than 0.5 ms. Other complicated examples, such as image retrieval and object recognition, can also be accomplished in the proposed MLSoC with the aid of the RISC and the external DDR memory.

VI. CONCLUSION

A 1.7 TOPS/W 16-mm^2 MLSoC is implemented in TSMC 90-nm CMOS technology. The new SoC architecture meets the flexibility and performance requirements of multimedia content analysis for consumer electronics, and the tera-scale performance of the DSP enables the coacceleration of computer vision and machine learning algorithms. Moreover, the proposed MLSoC achieves higher power efficiency and area efficiency than other related works.

REFERENCES

[1] P. Dubey, "Recognition, mining and synthesis moves computers to the era of Tera," *Technology@Intel Mag.*, pp. 1–10, Feb. 2005.

[2] T. M. Mitchell, *Machine Learning*. Englewood Cliffs, NJ: McGraw-Hill, 1997.

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.*, Dec. 2001, vol. 1, pp. 511–518.

[4] Y. Deng, B. S. Manjunath, and H. Shin, "Color image segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.*, Jun. 1999, vol. 2, pp. 446–451.

[5] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[6] M. Shi and A. Bermak, "An efficient digital VLSI implementation of Gaussian mixture models-based classifier," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 14, no. 9, pp. 962–974, Sep. 2006.

[7] T.-W. Chen, C.-H. Sun, J.-Y. Bai, H.-R. Chen, and S.-Y. Chien, "Architectural analyses of K-Means silicon intellectual property for image segmentation," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2008, pp. 2578–2581.

[8] Y. Ma and T. Shibata, "A binary-tree hierarchical multiple-chip architecture for real-time large-scale learning processor systems," *Jpn. J. Appl. Phys.*, vol. 49, no. 4, pp. 04DE08–, 2010.

[9] S. Yoshihara, Y. Nitta, M. Kikuchi, K. Koseki, Y. Ito, Y. Inada, S. Kuramochi, H. Wakabayashi, M. Okano, H. Kuriyama, J. Inutsuka, A. Tajima, T. Nakajima, Y. Kudoh, F. Koga, Y. Kasagi, S. Watanabe, and T. Nomoto, "A 1/1.8-inch 6.4 MPixel 60 frames/s CMOS image sensor with seamless mode change," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2998–3006, Dec. 2006.

[10] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. Int. Conf. Comput. Vis.*, 1998, pp. 839–846.

[11] A. Abbo, R. Kleihorst, V. Choudhary, L. Sevat, P. Wielage, S. Mouy, and M. Heijligers, "XETAL-II: A 107 GOPS, 600 mW massively-parallel processor for video scene analysis," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2007, pp. 270–271.

[12] K. Kim, S. Lee, J.-Y. Kim, M. Kim, D. Kim, J.-H. Woo, and H.-J. Yoo, "A 125 GOPS 583 mW network-on-chip based parallel processor with bio-inspired visual attention engine," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2008, pp. 308–309.

[13] C.-C. Cheng, C.-H. Lin, C.-T. Li, S. Chang, C.-J. Hsu, and L.-G. Chen, "iVisual: An intelligent visual sensor SoC with 2790fps CMOS image sensor and 205GOPS/W vision processor," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, Feb. 2008, pp. 306–307.

[14] "AMBA Specification (Rev 2.0)," ARM Ltd., Cambridge, U.K., 1999 [Online]. Available: http://infocenter.arm.com/help/topic/com.arm.doc.ihi0011a/

[15] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 10, pp. 959–971, Oct. 1996.

[16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, 2004.

[17] J. C. Chen and S.-Y. Chien, "CRISP: Coarse-grained reconfigurable image stream processor for digital still cameras and camcorders," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1223–1236, Sep. 2008.

[18] I. Hatirnaz, F. K. Gürkaynak, and Y. Leblebici, "Realization of a programmable rank-order filter architecture using capacitive threshold logic gates," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jul. 1999, vol. 1, pp. 435–438.

[19] S.-C. Hsia and W.-C. Hsu, "A parallel median filter with pipelined scheduling for real-time 1D and 2D signal processing," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E83-A, no. 7, pp. 1396–1404, 2000.

[20] T.-W. Chen and S.-Y. Chien, "Bandwidth adaptive hardware architecture of K-Means clustering for video analysis," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 18, no. 6, pp. 957–966, Jun. 2010.

Shao-Yi Chien (S'99–M'04) received the B.S. and Ph.D. degrees from National Taiwan University (NTU), Taipei, Taiwan, in 1999 and 2003, respectively.

During 2003 to 2004, he was a Member of Research Staff with Quanta Research Institute, Tao Yuan County, Taiwan. In 2004, he joined the Graduate Institute of Electronics Engineering and Department of Electrical Engineering, NTU, as an Assistant Professor. Since 2008, he has been an Associate Professor. He serves as an associate editor for *Circuits, Systems and Signal Processing* and has served as a guest editor for the *Journal of Signal Processing Systems* in 2008. He also serves on the technical program committees of several conferences, such as ISCAS, A-SSCC, and VLSI-DAT. His research interests include video segmentation algorithm, intelligent video coding technology, perceptual coding technology, image processing for digital still cameras and display devices, computer graphics, and the associated VLSI and processor architectures. He has published more than 120 papers in these areas.

Dr. Chien serves as an associate editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.

Tse-Wei Chen received the B.S. degree from National Taiwan University (NTU), Taipei, Taiwan, in 2006, and the Ph.D. degree from Media IC and System Lab, Graduate Institute of Electronics Engineering (GIEE), NTU, in 2010.

In 2010, he became a Foreign Joint Researcher with Nagoya University, Nagoya, Japan. His research interests include computer vision, pattern recognition, machine learning, and associated VLSI architectures.

Chi-Sun Tang received the B.S. and M.S. degrees in electrical and electronics engineering from National Taiwan University (NTU), Taipei, Taiwan, in 2004 and 2007, respectively, where he is currently working toward the Ph.D. degree at the Graduate Institute of Electronics Engineering, NTU.

His research interests include SoC architecture and optimization, new techniques in intelligent image processors, and embedded video communication applications.

Sung-Fang Tsai was born in Hsinchu, Taiwan, in 1983. He received the B.S. degree in electrical engineering and M.S. degree in electronic engineering from National Taiwan University, Taipei, Taiwan, in 2005 and 2007, respectively, where he is working toward the Ph.D. degree in electronics engineering.

His major research interests include low-power stream processors for computer vision and low-power video coding architectures.

Chen-Han Tsai received the B.S. degree in electrical engineering from National Taiwan University (NTU), Taipei, Taiwan, in 2002, where he is currently working toward the Ph.D. degree at the Graduate Institute of Electronics Engineering, NTU.

His major research interests include face detection and recognition, motion estimation, H.264/AVC video coding, digital TV systems, multimedia SoC, and related VLSI architectures.

Liang-Gee Chen (S'84–M'86–SM'94–F'01) received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 1979, 1981, and 1986, respectively.

In 1988, he joined the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan. During 1993–1994, he was a Visiting Consultant with the DSP Research Department, AT&T Bell Labs, Murray Hill, NJ. In 1997, he was a Visiting Scholar with the Department of Electrical Engineering, University of Washington, Seattle. During 2004–2006, he was the executive Vice President and General Director of the Electronics Research ad Service Organization (ERSO), Industrial Technology Research Institute (ITRI), Taiwan. Since 2007, he has served as a Co-Director General of National SoC Program. Currently, he is the Deputy Dean of the Office of Research and Development and a Distinguished Professor of the Department of Electrical Engineering, NTU. He has authored or coauthored over 400 publications and 30 patents. He has been an associate editor of the *Journal of Circuits, Systems, and Signal Processing* in 1999–2008 and a guest editor for the *Journal of Video Signal Processing Systems*. He has been an associate editor for the *Journal of Information Science and Engineering* since 2002. Since 2007, he has served as an associate editor of *Research Letters in Signal Processing* and for the *EURASIP Journal on Advances in Signal Processing*. His research interests are DSP architecture design, video processor design, and video coding systems.

Dr. Chen is a member of Phi Tau Phi. He served as an associate editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY in 1996–2008, the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS in 1999–2001, and the IEEE TRANSACTIONS CIRCUITS AND SYSTEMS II in 2000–2001. During 2001–2004, he was also the Associate Editor of the Proceedings of the IEEE. He was the General Chairman of the 7th VLSI Design/CAD Symposium in 1995 and of the 1999 IEEE Workshop on Signal Processing Systems: Design and Implementation. He is the Past-Chair of the Taipei Chapter of the IEEE Circuits and Systems (CAS) Society and is a member of the IEEE CAS Technical Committee of VLSI Systems and Applications, the Technical Committee of Visual Signal Processing and Communications, and the IEEE Signal Processing Technical Committee of Design and Implementation of SP Systems. He was the Chair-Elect of the IEEE CAS Technical Committee on Multimedia Systems and Applications. During 2001–2002, he served as a Distinguished Lecturer of the IEEE CAS Society. He has been the program committee member of IEEE ISSCC in 2004–2007. He was the TPC chair of the 2009 IEEE ICASSP and will be for ISCAS 2012. He was the recipient of the Best Paper Award from the R.O.C. Computer Society in 1990 and 1994, the Long-Term (Acer) Paper Awards from 1990 to 2005, the Best Paper Award of the 1992 Asia-Pacific Conference on Circuits and Systems in the VLSI design track, and the Annual Paper Award of Chinese Engineer Society in 1993. In 1996, 2000, and 2002, he was the recipient of the Outstanding Research Award from the National Science Council, and, in 2000, the Dragon Excellence Award from Acer. He guided students who won the DAC/ISSCC Student Design Contest four times since 2004.