

An Image Scanning Method with Selective Activation of Tree Structure (Special Issue on New Concept Device and Novel Architecture LSIs)

著者	Akita Junichi, Asada Kunihiro
雑誌名	IEICE transactions on electronics
巻	E80-C
号	7
ページ	956-961
発行年	1997-07-25
URL	http://hdl.handle.net/2297/6983

An Image Scanning Method with Selective Activation of Tree Structure

Junichi AKITA^{†a)} and Kunihiro ASADA[†], *Members*

SUMMARY We propose a new scanning method for image signals using a tree structure of automata. The tree is scanned selectively along the signal path for realizing both lower power consumption and a kind of image compression by skipping non-active elements. We designed the node automata along with photo-detectors of 32×32 in a $7.2 \text{ mm} \times 7.2 \text{ mm}$ chip using a $1.5 \mu\text{m}$ CMOS technology. We demonstrate applications of the tree structure using its feature of selective activation; a moving picture compression using inter-frame difference, an adaptive resolution scan like human eyesight and a motion compensation as examples.

key words: image scan, tree structure, selective activation, automaton, image encoding

1. Introduction

With the increasing needs for image information, such as multi-media, channel capacity between image sensors and image processing units in the conventional system is becoming one of critical factors that prevents complex and high speed image sensing [1].

From a view point of VLSI technology, highly developed VLSI technologies enable us to fabricate both the photo-detectors and signal processing circuits in one chip, which is called "computational sensor [2], [3]," for overcoming the bottleneck of the channel capacity. In most of previous studies in the computational sensors, the main theme was aiming of so-called early vision processing, and few studies on image compression has been reported.

In this study, at first we propose a new signal scan method with a kind of image compression capability by skipping scan of non-active elements using a tree structure of automata, where scan is carried out only for active elements selected by the transitions of node automata. It is shown that data compression about 1/7 is possible by the present method against the conventional raster scan method for samples of binary images. It is more effective especially in pictures with large ratio of non-active elements, such as inter-frame difference of moving pictures. Here, note that the "active" element is defined depending on situations. It is anyway defined as an important element which carries information. Next,

we show circuits of tree-structured node automata and photo detectors to be integrated on one chip using a $1.5 \mu\text{m}$ CMOS technology.

Finally we also show applications of the present method utilizing its feature of hierarchical selective activation, such as a moving picture compression with inter-frame difference, an adaptive resolution scanning like human eyesight and a direct motion compensation for moving picture compression.

In this study, we treat image data as binary images for simplicity by digitizing their intensity before processing, though we will comment an extension method to the gray scale images in conclusion.

2. Tree Structure for Image Scan

2.1 Signal Scanning Strategy

In the conventional raster scan method, that is used in CCD devices, as shown in Fig. 1 (a), all the photo detectors in a sensor are always scanned, even if the number of active elements is expected small enough. This is especially the case of the difference between the current and the previous frames of moving pictures. It implies that there are redundant cycles in the conventional scan method, resulting in useless consumption of power and time.

Figure 1 (b) shows a new method of signal scan using a tree structure of automata, which can reduce the redundant cycles in the image scan. Here the photo detectors, whose outputs are binary, are placed in the lowest level of the tree structure. The automaton in each node has a value as logical-OR of its lower levels' automata, and returns the value to the higher level when

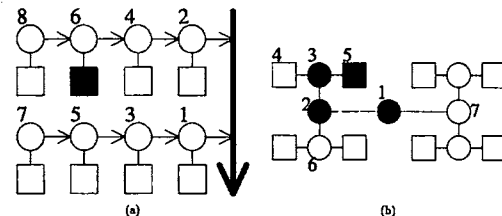


Fig. 1 Two signal scan methods. (a) the conventional raster scan, (b) the present scan method using tree structure of automata. The circles are scan circuits and squares are photo detectors. Numbers in the figure indicate the sequence of scan.

Manuscript received November 20, 1996.

Manuscript revised January 27, 1997.

[†]The authors are with the School of Engineering, the University of Tokyo (VLSI Design and Education Center, University of Tokyo), Tokyo, 113 Japan.

a) E-mail: akita@silicon.t.u-tokyo.ac.jp

it is scanned by the higher level. The value of 0 implies that the values of all the lower nodes are 0, so that it is not needed to scan the lower level further. When the value is 1, the automaton starts scanning the lower level nodes in order.

All the node automata are essentially identical, the state-transition of which proceeds using a gated clock signal (*transition* signal) from the upper node to the lower, and a *completion* signal from the lower node to the upper. The value returned by the lower nodes propagates to the upper nodes, finally up to the outside of the sensor. For example in the case shown in Fig. 1 (b), the initial *transition* signal is sent to node 1 from the outside. Node 1 returns 1 then sends the *transition* signal to node 2. Similarly, node 2 returns 1 then sends the *transition* signal to node 3. Node 3 returns 1, followed by 0 and 1, respectively, corresponding to the photo detectors; 4 and 5. When returning the final value, node 3 enables the *completion* signal as well. Responding to the *completion* signal, node 2 sends the *transition* signals to node 6, which returns 0 with enabling the *completion* signal. Responding to the *completion* signal, node 1 sends the *transition* signal to node 7, which again returns 0 with enabling the *completion* signal. Responding to the *completion* signal, node 1 enables the *completion* signal to the outside, indicating the end of data scan. Thus data sequence of 1110100 is generated, while the conventional raster scan results in 00000010.

2.2 Mean Code Length

Here we analyze the characteristics of the data sequence generated by the tree structure. Defining that b is the number of branches to lower nodes and N is the number of levels in the tree structure, the number of photo detectors, n , is equal to b^{N-1} . Note that the top node is at level N and the photo detectors is at level 1, then the level is counted from lower to upper. Assuming no correlation in photo detectors' data and the active (black) probability of photo detectors is p_1 , a node in $(i+1)$ -th level becomes active when there are at least one node whose value is active in i -th level. Thus the activation probability of the node in $(i+1)$ -th level, p_{i+1} is formulated as

$$p_{i+1} = 1 - (1 - p_i)^b.$$

So p_i is described as $p_i = 1 - (1 - p_1)^{b^{i-1}}$. Using the active probabilities of the node automata, the mean code length, \bar{L} , is derived as follows, with keep in mind the fact that the code length is exactly equal to the total number of scans at all levels.

The first scan to the top node is invoked from the outside, so that \bar{L} contains this bit. As for a node automaton at the i -th level, it conditionally scans its lowers when it is active, so that the expected scan number is $p_i b$. Since there are b^{N-i} nodes at the i -th level, \bar{L} can

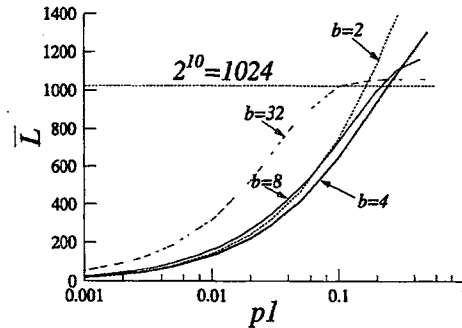


Fig. 2 The relation between \bar{L} and p_1 . $b = 4$ gives the shortest code length.

be described as

$$\begin{aligned} \bar{L} &= 1 + \sum_{i=2}^N b^{N-i} p_i b \\ &= 1 + \sum_{i=2}^N b^{N-i+1} \{1 - (1 - p_1)^{b^{i-1}}\}. \end{aligned}$$

The relation between \bar{L} and p_1 is illustrated in Fig. 2 for the 2^{10} photo detectors, which indicates that $b = 4$ gives the smallest code length. It can also be shown smallest in all cases of the number of photo detectors. In this study we consider the tree structure with $b = 4$, i.e. so called "quad tree," which we call 1:4 tree hereafter. The code encoded with this quad tree structure is called 1:4 tree code.

Though several applications of the quad tree for image signals are already reported from viewpoints of data structures suitable for image processing by software [4], we will show that it is also adequate to apply 1:4 tree for on-chip hardware implementation of image signals compression of computational sensors.

It is notable that the decode algorithm of 1:4 tree code is simply implemented by scanning a part of the code stream, which is easily implemented in hardware circuit with a small shift register.

2.3 Examples of Scanning

Figure 2 indicates that the possibility of active photo detectors larger than 0.24 results a code length larger than the conventional raster scan. Though p_1 is usually expected to be larger than 0.24 in case of still images, the \bar{L} is the code length of random picture. In other words, the spatial frequency is high. Practically it is experienced that the spectrum in spatial frequency is dominant in low frequency.

We carried out a simulation of 1:4 tree scan for moving pictures to study the code length \bar{L} for practical cases. We first encoded each frame in examples of moving pictures as a still image and compared it with the raster scan. These pictures have 256×256 pixels with 256 scales of intensity in each frame. We generated binary

Table 1 The 1:4 tree code length of example pictures. (\bar{p}_1 is the mean of active probability of elements and the number in parentheses is a ratio of the 1:4 tree code length to the raster scan.)

Name	\bar{p}_1	1:4 Tree Code Length[bit]		
		mean	min.	max.
MissA	8.92%	10674 (16.3%)	9853 (15.0%)	12029 (18.4%)
Neck	24.5%	25343 (38.7%)	24633 (37.6%)	25906 (39.4%)
Rail	32.3%	34091 (52.0%)	27465 (41.9%)	37389 (57.1%)
Rail2	16.4%	20276 (30.9%)	19625 (29.9%)	20777 (31.7%)

images by digitizing the intensity. Code lengths of 1:4 tree scan are shown in Table 1 with the ratio to the code length of raster scan. The results show that the 1:4 tree code length is smaller than that of the raster scan, even in case of active probability of photo detectors larger than 0.24, and the 1:4 tree code length is about 1/6 to 1/2 of that of the raster scan.

3. Implementation of 1:4 Tree

3.1 Circuit of Node Automaton

We designed a circuit for the node automaton in 1:4 tree structure. The state transition diagram of the node automaton is shown in Table 2. Here the state W represents *waiting*, and the state SA, ..., SD are the states of *scanning* lower nodes A, ..., D, respectively. C is the *transition* signal to the automaton, and CA, ..., CD are the *transition* signals to its lower nodes A, ..., D, respectively. EA, ..., ED are the *completion* signals of the scan from the lower nodes A, ..., D, respectively. VD is the logical-OR of VA, ..., VD, and V and E is the output value and *completion* signal of automaton, respectively. These signals are summarized in Fig. 3.

The designed circuit of the node automaton is shown in Fig. 4. In the tree structure, the clock signal is gated so as to propagate through the scanning path without activating other nodes. We call this scheme *selective activation*, which is effective for reducing power consumption. This node automaton contains 142 transistors. The number of nodes, N_A , in the tree structure of N level is $\sum_{i=1}^{N-1} 4^{i-1} \cong 4^{N-1}/3$, while the number of photo detectors is 4^{N-1} . So the number of transistors per photo detector is $142/3 \cong 47$, which is small enough to keep a practical fill-factor compared with previous studies [5], where the each photo detector contains about 40 transistors and one large capacitor.

We carried out spice simulation of tree structures composed of the circuits to estimate the power consumption for various images. Results of scanning 4x4 and 8x8 pixels are summarized in Table 3. In cases of

Table 2 Transition diagram of node automaton. (S and S' are the current and next state, respectively.)

Inputs				Outputs							
C	V	E	A	S	S'	CA	CB	CC	CD	V	E
0	-	-	-	W	W	0	0	0	0	VO	0
↑	0	-	-	W	W	0	0	0	0	VO	1
↑	1	-	-	W	SA	↑	0	0	0	VA	0
↑	-	0	-	SA	SA	↑	0	0	0	VA	0
↑	-	1	-	SA	SB	0	↑	0	0	VB	0
↑	-	-	0	SB	SB	0	↑	0	0	VB	0
↑	-	-	1	SB	SC	0	0	↑	0	VC	0
↑	-	-	-	SC	SC	0	0	↑	0	VC	0
↑	-	-	-	SC	SD	0	0	0	↑	VD	0
↑	-	-	-	SD	SD	0	0	0	↑	VD	0
↑	-	-	-	SD	W	0	0	0	0	VO	1

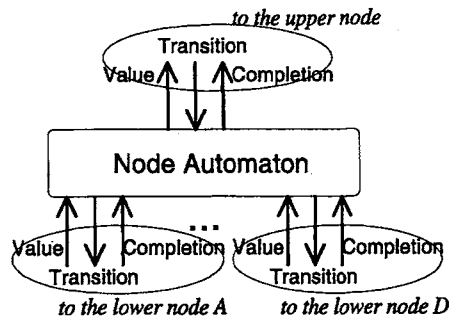


Fig. 3 Signals of node automaton. Each node has the signals of the *transition*, the *completion*, and the value signal for the upper node and the four lower nodes.

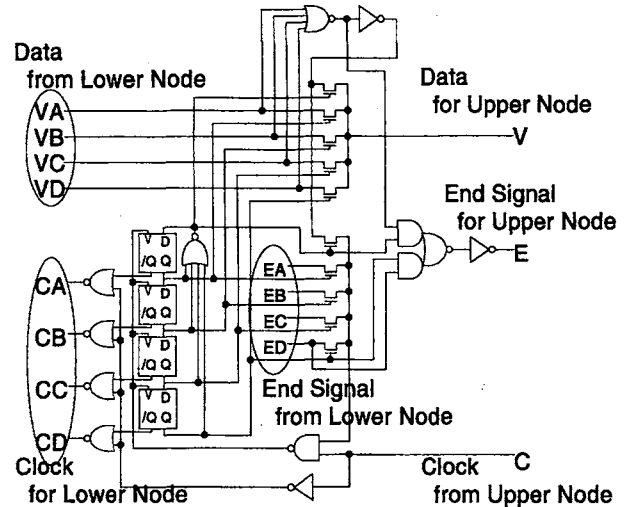







Fig. 4 Circuit of node automaton. for $b = 4$. This circuit contains 142 transistors.

8x8 pixels, the 1:4 tree has one top level node, 2x2 middle level nodes, 4x4 lower level nodes, and 8x8 photo detectors (these were given as voltage sources in this simulation) in order. Table 3 shows the mean energy consumption in the whole trees per one bit. The total

Table 3 The power consumption U of 1:4 tree structure in the scan of 4×4 or 8×8 pixels.

Image	# of pixels	Code length	U [pJ/bit]
	4×4	17bit	0.75
	4×4	21bit	0.76
	8×8	37bit	0.88
	8×8	69bit	0.89
	8×8	45bit	0.88

energy consumption scanning the whole image is the product of the energy per bit and the code length. This result shows that the energy consumption per bit keeps almost constant independent to the image pattern. The constant energy is not sensitive to the number of photo detectors, since the length of the signal path activated from the top to a photo detector is proportional not to the number of photo detectors but to the levels of the tree.

3.2 Layout of 1:4 Tree

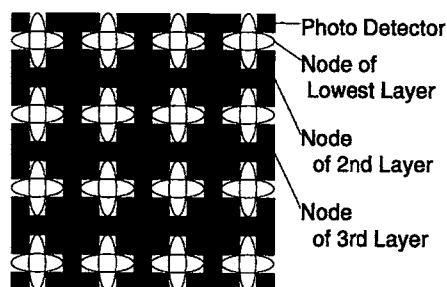
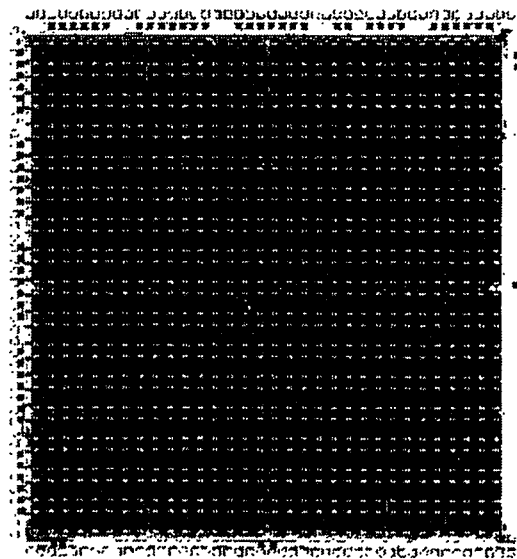
Since VLSI circuits have to be laid out on a surface of silicon chip, it is desirable that the node automata in tree structure is laid out so as to keep the area of photo detectors large enough. The photo detectors should be placed in the equal interval in focal plain.

A possible layout of photo detectors and node automata is shown in Fig. 5, that can place photo detectors uniformly and node automata with full-filling the surface of chip. In this layout the upper node automata can occupy the larger area and thus it can contain the larger size transistors. This is reasonable since the upper node circuits have to drive longer signal wires in the two dimensional tree layout. This is also suitable to minimize delays [6].

We designed this two dimensional tree structure including node automata and photo detectors for $1.5 \mu\text{m}$ CMOS with double metal layers, as shown in Fig. 6. Each photo detector has functions not only converting photo signal to binary signal, but also generating the inter-frame difference by having latches in each photo detector. This chip is now under fabrication, and measurement results will be reported in near future.

4. Applications of 1:4 Tree

In this section we discuss applications of 1:4 tree for

**Fig. 5** Possible layout of photo detector and node automaton of 1:4 tree structure.**Fig. 6** Chip layout of the 1:4 tree structure with 32×32 photo detectors by $1.5 \mu\text{m}$ CMOS. Chip size is $7.2 \text{mm} \times 7.2 \text{mm}$. The total number of transistors is 94,035.

some of image processing and its implementation.

4.1 Interframe Difference

In moving pictures successive frames usually do not have so much differences. One of the simplest algorithms for moving picture compression is to utilize the inter-frame difference by scanning only differences.

We carried out the 1:4 tree scan for moving pictures in Sect. 2 and compared it with the raster scan. Results of 1:4 tree scan are shown in Table 4 with the ratio to the code length of raster scan. The results show that the ratio of active elements in the inter-frame differences is only a few percents, and the 1:4 tree code length is smaller than in the case of still images. Compression ratio is about $1/14$ to $1/8$ compared with the raster scan.

This function can be implemented just by adding an exclusive OR gate and a latch to generate the inter-frame differences in photo detectors without modifying the node automata.

Table 4 Code length of inter-frame differences in example pictures. (\bar{p}_1^m is the mean of active probability in the inter-frame differences. The number in parentheses is a ratio of the 1:4 tree code length to the raster scan.)

Name	\bar{p}_1^m	1:4 Tree Code Length		
		mean	min.	max.
MissA	1.1%	4692 (7.2%)	1 (0.002%)	8733 (10.3%)
Neck	1.6%	6068 (9.3%)	1 (0.002%)	8285 (12.6%)
Rail	3.3%	8342 (12.7%)	1 (0.002%)	11729 (17.9%)
Rail2	2.1%	6819.6 (10.6%)	1 (0.002%)	7841 (11.7%)

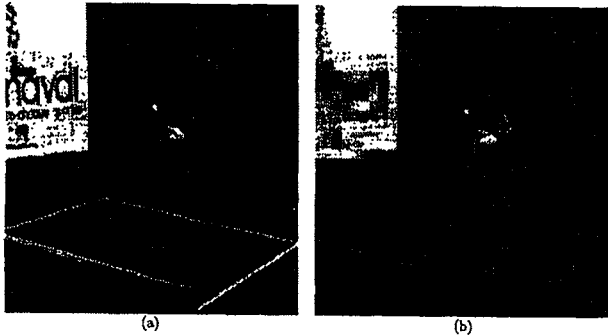


Fig. 7 The original sample image (a) which gives the code of 59,061 bits and the image with making attention only to its sub-area (b) which gives the code of 8,177 bits.

4.2 Adaptive Spatial Resolution Scan

Human eyesight has the function focusing to an area that we pay attention to, and more information is taken out from the area compared with the other area. In implementing this function in the sensor, we need to introduce adaptive mechanism not only in the signal intensity resolution [7], but also in the spatial resolution. For example, in case of seeing the picture in Fig. 7 (a), we usually focus to the area of Ping-Pong ball and the effective information in human eyesight may be illustrated as shown in Fig. 7 (b).

The scan in 1:4 tree structure proceeds from the top level to the bottom photo detectors in order, and the lower level gives the higher resolution, while the higher level gives the summarized information of the lower level. We can dynamically select two scan methods; either the normal scan for the focused area, or a scan intentionally stopped on the way for the other area. Thus we can obtain an image with adaptive resolution, where resolution is high only in the focused area.

For example, the focused image in Fig. 7 (b) gives just 8,177 bits, which is about 1/8 against the normal scan in Fig. 7 (a). In Fig. 7 (b), only the area of the Ping-Pong ball and the racket is normally scanned, while

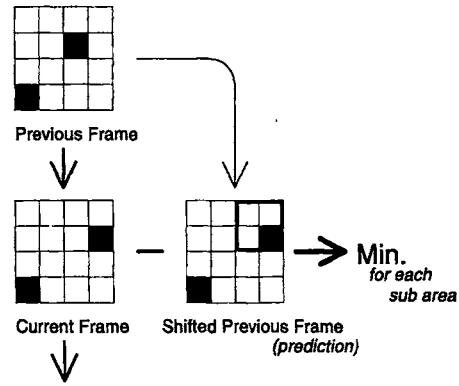


Fig. 8 Simple algorithms of motion compensation to select the direction which gives the minimum difference between the previous frame shifted to one direction and current frame for each subarea.

scan of the other areas are intentionally stopped on the way at the second lowest level. (Note that the pictures in Fig. 7 are illustrated as gray scale images for readers' convenience. These images are processed after digitizing to binary images in this study.)

This function is implemented by adding in each automaton an AND gate judging whether to continue the scan of lower nodes or to finish, controlled by a signal given from the outside.

4.3 Motion Compensation

The motion compensation is one of the most important and complex functions in practical moving picture compression such as MPEG2 [8]. The computational complexity of motion compensation is one of the key factors which prevents the single-chip, real-time MPEG2 encoding.

In the original 1:4 tree structure, each node returns a value of logical-OR of the values of lower nodes. We modify the value so that it represents the analog sum of the lower nodes' value. Thus the node value stands for the partial sum of the photo detectors' values. Using the difference of the current frame and a shifted previous frame as the pixel values, instead of photo detectors' outputs, as shown in Fig. 8, the node value gives the shifted inter-frame difference in subarea belonging to the nodes. The motion compensation is realized by finding the optimum shift direction which gives the minimum node value in magnitude, among shifted directions of, for example, N, NE, E, SE, S, SW, W and NW. Here N, S, E and W are the north, the south, the east and the west, respectively. Assuming very high frame rate, the motion distance of object is expect to be less than one pixel [9], and in this case the motion should be compensated only within one-neighbour pixels.

This method gives faster calculation compared with the conventional method, and the adaptive resolution scans can be applied in each shift direction. The reso-

lution is corresponding to the size of motion compensation area. This function is implemented by modifying the node automata so as to give the value of the sum of the lower nodes' values, and by adding the shift function of the photo detectors' outputs to its neighbors, along with the inter-frame difference method as described in Sect. 4.1.

The implementation of the sum function in each node results in a high amount of hardware of using a digital adder. It will be suitable to use analog adder here so as to keep the increase of hardware in reasonable amount.

5. Summary and Conclusions

We have proposed a novel signal scan method in image sensors using a tree structure of automata. The code length is expected to be much smaller than the conventional raster scan especially when applying to the difference of successive frames in moving pictures.

We have designed a circuit of the tree automata, where the clock signal is gated to the lower node selectively so as to minimize the power consumption. We have also designed a two dimensional layout of the 1:4 tree structure for a full-custom chip, in which both the node automata and 32×32 photo detectors are integrated in $7.2 \text{ mm} \times 7.2 \text{ mm}$ chip using $1.5 \mu\text{m}$ CMOS technology, and the chip is now under fabrication[†]. If using a $0.1 \mu\text{m}$ technology in future, it is estimated 480×480 photo detectors are integrated on a single chip.

We have proposed applications of 1:4 tree structure, such as a moving picture compression, an adaptive resolution scan like human eyesight, and a motion compensation for moving picture, which are implemented with a small extension in each node automata and photo detectors.

Though we treat all the images as binary in this study, it will be possible to extend the method to gray scale images by introducing the concept of the conditional replenishment method [10], where a flag signal associated with a photo detector is enabled when its image intensity has changed by a predefined amount of magnitude. For extending tree scan method to gray scale images, we can replace the binary image data simply by the flag data, while the gray scale image data is read out by an additional analog signal path.

References

- [1] W. Lawler and L. Harrison, "Performance of high-frame-rate, back-illuminated CCD imagers," *Proc. SPIE*, vol.2172, pp.90-98, 1994.
- [2] C. Koch, "Implementing early vision algorithms in analog

[†]The VLSI chip in this study has been fabricated in the chip fabrication program of VLSI Design and Education Center (VDEC), the University of Tokyo with the collaboration by Nippon Motorola and Dai Nippon Printing Corporation.

- hardware," *Proc. SPIE*, vol.1473, pp.2-16, 1991.
- [3] P. Seiz, D. Leipold, J. Kramer, and J.M. Raynor, "Smart and optical image sensors fabricated with industrial CMOS/CCD semiconductor process," *Proc. SPIE*, vol.1900, pp.30-39, 1993.
- [4] H. Samet, "Region representation: Quad trees from binary arrays," *Computer Graphics and Image Processing*, vol.13, pp.88-93, 1980.
- [5] A. Gruss, et al., "A VLSI smart sensor for fast range imaging," *Proc. IEEE Int. Conf. on Intelligent Robots and Systems*, 1992.
- [6] N.H.E. Weste and K. Eshraghian, "Principles of CMOS VLSI design: A systems perspective," Addison-Wesley, 1988.
- [7] B. Fowler, et al., "A CMOS area image sensor with pixel-level A/D conversion," *ISSCC '94*, pp.226-227, 1994.
- [8] ISO/IEC 13818-1, 13818-2, 13818-3 International Standard, 1994.
- [9] I. Ishii, et al., "Target tracking algorithm for 1 ms visual feedback system using massively parallel processing," *Proc. IEEE Int. Conf. Robotics and Automation*, pp.2309-2314, 1996.
- [10] F.W. Mounts, "A video encoding system with conditional picture—element replenishment," *BSTJ*, pp.2545-2554, 1969.



Junichi Akita was born in Aichi, Japan, on Aug. 22, 1970. He received B.S. and M.S. degrees in electronics engineering from the University of Tokyo, Tokyo, Japan in 1993 and 1995, respectively. He is currently studying in the doctor course in the University of Tokyo. His interest is in the mathematical modeling of integrated circuit systems and the smart sensors for low-power and intelligent applications.



Kunihiko Asada was born in Fukui, Japan, on June 16, 1952. He received the B.S., M.S., and Ph.D. in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1975, 1977, and 1980, respectively. He joined the faculty of the University of Tokyo as a research associate in 1980, and became a lecturer, an associate professor and a professor in the department of electronic engineering in 1981, 1985 and 1995, respectively. From

1985 to 1986 he stayed in Edinburgh University as a visiting scholar supported by the British Council. He moved to the VLSI Design and Education Center (VDEC) of the University of Tokyo, when it was newly established in 1996. He is currently a professor of VDEC, being also engaged in education in the department of electronic engineering. His interest is in design and evaluation of integrated systems and their component devices. He is a member of Institute of Electrical and Electronics Engineers (IEEE), Institute of Electrical Engineers of Japan (IEEJ). He served as the Editor of *IEICEJ Transactions on Electronics* from 1990 to 1992.