Low-Power Video Segmentation by Pipeline Processing of Tiled Images

T. Morimoto, H. Adachi, O. Kiriyama, Z. Zhu, T. Koide, and H. J. Mattausch

Research Center for Nanodevices and Systems, Hiroshima University, 1-4-2 Kagamiyama, Higashi-Hiroshima, 739-8527, Japan
Phone: +81-82-424-6265 Fax: +81-82-424-2499 e-mail: morimoto@rcns.hiroshima-u.ac.jp

1. Introduction

For object-based video processing such as recognition or tracking, the object-extraction process called *video segmentation* is an indispensable first step. One important field of object-based video processing are battery-based applications (ex. robot vision, intelligent transport systems, or mobile phones with camera). To realize these applications, 3 requirements have to be attained simultaneously, namely *real-time processing* (< 33msec/frame), *compact implementation* (single chip), and *low-power dissipation* (mW order). However, visual data has generally high complexity and contains a multitude of information, so that it's difficult to achieve these requirements with general purpose hardware like FPGAs, microprocessors or digital signal processors. Therefore, special purpose hardware is strongly required. Recently some hardware of video segmentation, based on the difference of frames, were proposed [1, 2, 3]. These approaches are effective for video images of a fixed camera. However it is difficult for mobile applications which use a moving camera, because all background objects are also moving in this case.

Previously, we proposed a digital video segmentation architecture [4, 5] as well as concepts for a compact implementation technique named tiled *subdivided-image approach (SIA)* [6] and a low-power processing technique named *boundary-active-only (BAO)* scheme. Since our region-growing segmentation algorithm is based on connection-weights with neighboring pixel, it works well not only for video signals from fixed cameras but also for video signals from moving cameras. This paper reviews the SIA segmentation algorithm, circuit details for BAO-implementation. BAO-performance evaluation with a full-custom designed CMOS test-chip in 0.35μm technology, including a 41×33 pixel-processing array, is also presented.

2. Video Segmentation Architecture with Tiled Images

In the proposed subdivided-image approach (SIA) [6], as shown in Fig. 1, an input video image is divided into a plurality of small-size tiles with an overlapped region of 1 row and 1 column. Each tile is subjected to segmentation using a small-size processing cell-network, and the processing of all tiles is successively performed in order. Finally the results for all tiles are put together to complete the segmentation of the whole image with the information of associated prelabeled regions.

Figure 2 shows the flowchart of the SIA algorithm and there are 3 main procedures called initialization, segmentation, and label restore. In the initialization procedure, the connection-weights W_{ij} in a tile image are calculated from the luminance (resp. RGB-data) differences $|I_i - I_j|$ for gray-scale images (resp. color images). Inclusion of a pixel i in a given segment is decided by examination of the W_{ij} with neighboring pixels j, which are already included in the grown region. Then leader pixels (self-excitable pixels), which are the seeds of the subsequent region-growing process, are determined from calculated connection-weights. Next, if there are already labeled pixels in overlapped prelabeled region, then segmentation and label restore procedures are executed and otherwise only the segmentation procedure is performed. In segmentation procedure, one of the leader pixels is self-excited and a new region is grown from this leader pixel. The priority of the overlap leader pixels is higher than that of other leader pixels, therefore these pixels are self-excited at first. In each growing step of the region, excitable pixels are determined with a threshold condition for the sum of connection-weights with excited neighbors, and the pixels fulfilling the excitation condition are automatically ex-

cited. The growing steps are repeated as long as excitable pixels exist. When no further excitable pixels are left, the growing process of the respective segment finishes and the excited pixels, constituting the new segment, are labeled and inhibited. The segmentation process is completed when all initially determined leader pixels are inhibited.

If the prelabeled cells which have different prelabel numbers should become the same region, a label conflict occurs. To avoid this situation, label numbers of prelabeled overlap regions are always observed in the label restore procedure. This procedure carries out following 2 main processes: (1) If cells with labels identical to excited cell exist, these cells are forced to be excited. The excitation condition on the prelabeled region must go on, because inclusion operation decided by the weight from excited neighbors. (2) If a growing region connects to the prelabeled region which has a different label (label conflict), conflicting label numbers are stored in a table. At the end of the region growing, the smallest number of conflict labels is allocated to the grown region and already labeled tiles are relabeled if necessary.

Figure 3 shows the construction of the cell-network, which is the core circuit of our video segmentation architecture. The label restore procedure is applied only to the cells on 1 row and 1 column in the overlapped region. The cell-network consists of cells, which are processing elements and correspond to the pixels, as well as connection-weight registers, which store the connection-weights. Each cell calculates the sum of the connection-weights with excited neighbors and determines its own new state (self-excited, excited, inhibited, labeled) according to a threshold condition. Due to this parallel processing of all cells, the power-dissipation increases in proportion to the number of cells (pixels). To avoid this increase, we propose the boundary-active-only (BAO) concept.

3. BAO Concept and Circuit Implementation

The BAO concept, which exploits the characteristics of the region-growing algorithm, is explained with Fig. 4. Due to the stepwise growth of each region, it is sufficient to activate only the cells which have an excitation possibility in the current growth step. Such cells must belong to the boundary of the currently grown region. More specifically, cells with an excitation possibility should not satisfy any of the following three conditions: (1) It is already excited ($x_{ij}=1$). (2) It has already a segment number ($l_{ij}=1$). (3) It is not excited and has no segment number, but there are no neighboring cells excited during the previous clock cycle t. In particular, condition (3) means that only a part of the complete boundary of the grown segment has an excitation possibility in the normal case (Fig. 4).

We implemented a BAO controller in each network cell, which realizes the BAO concept for reduced power dissipation by examining the above 3 conditions and controls the cell's stand-by mode by a clock-gating signal *cell_CLK$_{ij}$* (Fig. 5). Since the cell-network has long global clock lines with large capacitances, we additionally restrict clock distribution to potentially active network cells by using a clock controller. The controller distributes the clock signal in the next clock cycle only to rows including cells, which have been excited in the previous clock cycle, and their neighbor rows.

4. Test-Chip Design and Performance Measurements

We designed and fabricated a video segmentation test-chip which implements a cell-network with the described BAO architecture in a 0.35μm 2-Poly 3-Metal CMOS technology. Figure 5 shows the die photo of the fabricated chip including a

cell-network for 41×33 (1,353) pixels on an area of 51.1mm^2. The integration density achieved in the full-custom design is 26.5pixel/mm^2. Measured power dissipation for a worst-case input image (only one homogeneous region) is 94.0mW at 10MHz (0.069mW/pixel) in the segmentation phase. The worst-case power dissipation of a previously designed 10×10 cell-network without BAO, which has a twelve-times smaller cell number, is 30.9mW at 10MHz (0.309mW/pixel). Therefore, about 78% power-reduction per pixel have been achieved with the BAO concept. Average power dissipation, estimated with a 7 segment input image, is 45.8mW. Estimated segmentation time and Si-area consumption with BAO-architecture for QVGA-size images are <250μsec at 10MHz and <120mm^2 in a 90nm CMOS technology, respectively. The characteristic data of the test-chip are summarized in Table I.

5. Conclusions

We designed and fabricated a cell-network with 41×33 cells in 0.35μm CMOS technology for low-power video segmentation and experimentally confirmed the effectiveness of the proposed BAO architecture. Compared with our previously pro-posed segmentation architecture without BAO, about 78% power reduction per cell is achieved at 10MHz clock frequency.

Applying additionally the SIA approach, which effects only to the cells in prelabeled regions of the 1st row and column, VGA-size video-segmentation is expected to become possible with this 41×33 cell-network (16×15 tiles). The segmentation performance for VGA size input images is estimated as 7.49msec segmentation time at 10MHz clock frequency and < 94.0mW power dissipation.

References

[1] N. Ranganathan, et al., CVGIP: Image understanding, vol. 53 (2), pp. 189-197 (1991).
[2] S. Y. Chien, et al., Proc. of AP-ASIC2002, pp. 233-236 (2002).
[3] R. Yang, et al., Journal of Graphic Tools, vol. 7 (4), pp.91-100 (2002).
[4] T. Morimoto, et al., Ext. Abst. of SSDM2002, pp.242-243 (2002).
[5] T. Morimoto, et al., Ext. Abst. of SSDM2003, pp.146-147 (2003).
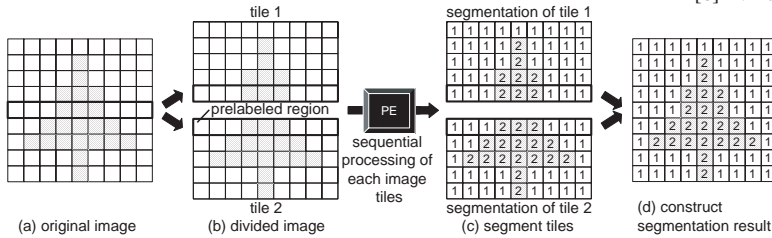[6] H. Adachi, et al., Proc. of SASIMI2004, pp.95-102 (2004).

Fig. 1: Processing example of the SIA approach. Prelabeled regions at the boundary of the tile enable correct segmentation of regions extending over several tiles.
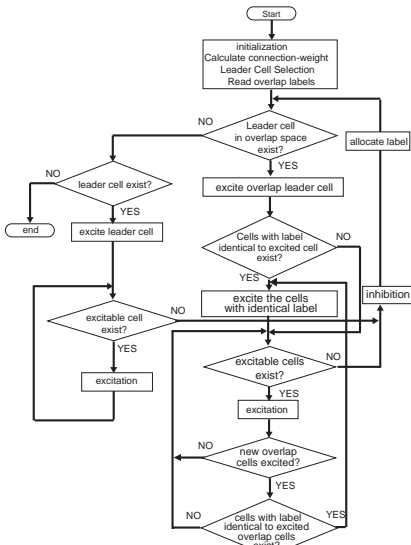
Fig. 4: Conceptual diagram of the proposed boundary-active-only (BAO) scheme. (a) shows the excited and newly excited regions at clock cycle t and $t+1$, respectively. (b) Only cells, which are neighbors of excited cells at $t+1$, are activated in clock cycle $t+2$.

Fig. 2: General flowchart of the SIA algorithm.
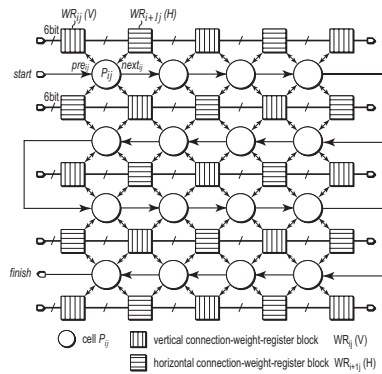
Fig. 3: Block diagram of the cell-network, imple-mented by alternately laying cells and connection-weight-register blocks.
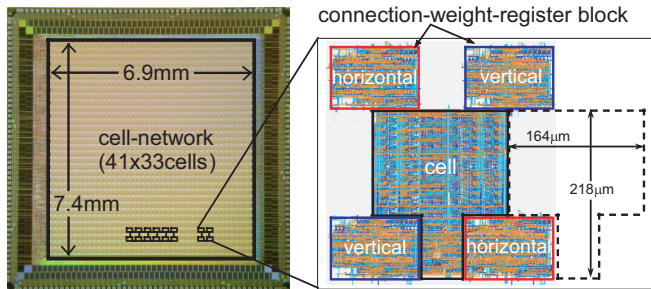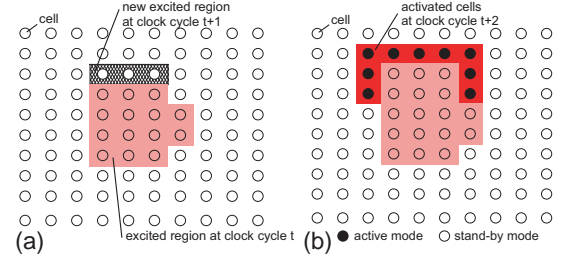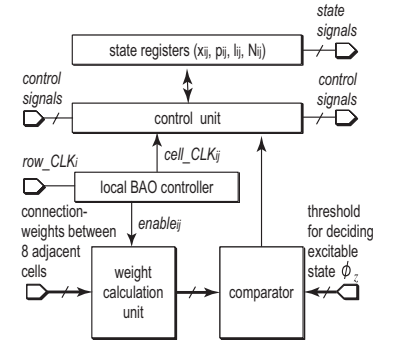
Fig. 5: Block diagram of the cell P_{ij} with BAO controller.

Fig. 6: Die photo of the network with BAO including 41×33 cells, designed in a 0.35μm 3-metal CMOS technology. The layout of cell and connection-weight-register blocks is magnified on the right side.

Table I: Characteristic data of the designed test-chip.

Technology	0.35μm, 2-Poly 3-Metal CMOS
Cell Architecture	Weight-Parallel (high-speed)[4]
Design Area	6.9mm×7.4mm (41×33 cells)
Supply Voltage	3.3V
Max Clock Frequency	10MHz
Segmentation Time (41×33 pixels)	34μsec@10MHz (Worst Case)
Worst Case Power Dissipation (Measured, 41×33 pixels)	94.0mW@10MHz (Segmentation) 192mW@10MHz (Initialize)
Pixel Density	26.5pixel/mm^2

Low-Power Video Segmentation by Pipeline Processing of Tiled Images

T. Morimoto, H. Adachi, O. Kiriyama, Z. Zhu, T. Koide, and H. J. Mattausch
Research Center for Nanodevices and Systems, Hiroshima University

NTIP
Hiroshima University

Introduction and an Short Overview

Video Segmentation
Extracting meaningful regions from natural input video pictures (30frame/sec) for higher level image processing applications.

Features of the Proposed Video Segmentation
- **High segmentation quality** as good as the conventional architecture[†]
- **78% power reduction** per pixel compared with our previously proposed architecture[†] without Boundary-Active-Only concept
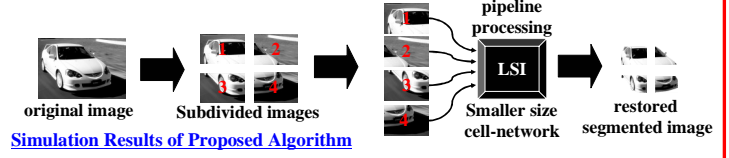- **VGA size (640×480pixels) single chip video segmentation**
  - Segmentation time < 7.49msec@10MHz,
  - Power dissipation at 10MHz < 94mW@10MHz

Boundary-Active-Only Scheme for Low Power
[†] T. Morimoto et al., 1st COE Workshop, 2003.
- Low-power technique for region-growing segmentation algorithm without sacrificing real-time processing
- Only boundary cells of the currently grown region have to be activated

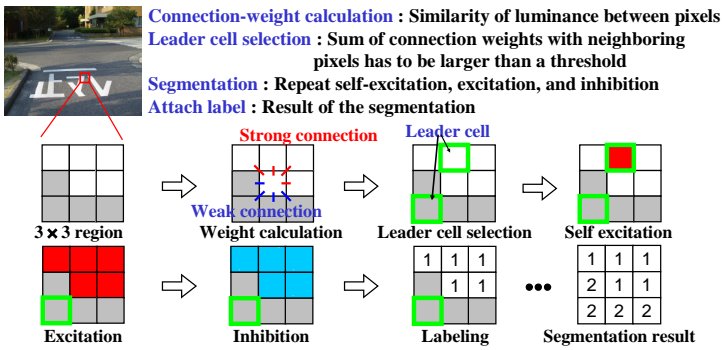### Subdivided-Image (SIA) Approach for Compact and Low Power

original image → Subdivided images → pipeline processing LSI → Smaller size cell-network → restored segmented image
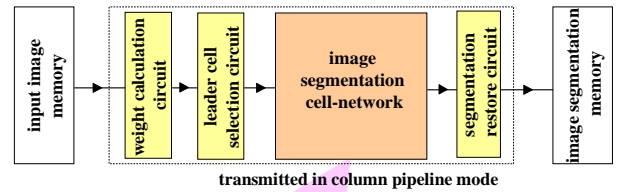
Simulation Results of Proposed Algorithm

Input image   Segmentation result

Simulation software
MATLAB (image processing toolbox)
Input image size
640×480 pixel (VGA size)
Pixel number per tile
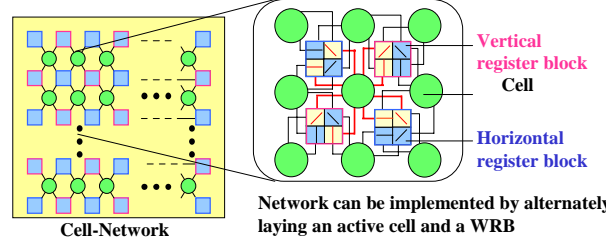41×33 pixel (1353 pixel)
Tile number
240 tiles (16×15)

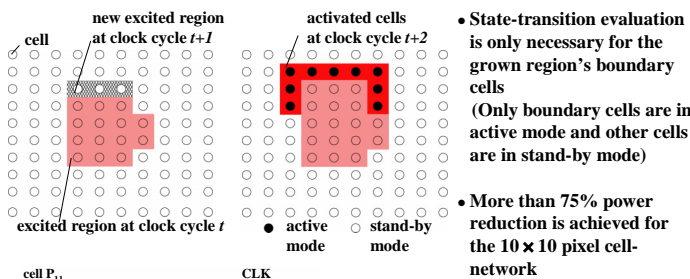## Video Segmentation Algorithm (Region Growing Approach)

Connection-weight calculation : Similarity of luminance between pixels
Leader cell selection : Sum of connection weights with neighboring pixels has to be larger than a threshold
Segmentation : Repeat self-excitation, excitation, and inhibition
Attach label : Result of the segmentation

Strong connection   Leader cell   Weak connection

3×3 region → Weight calculation → Leader cell selection → Self excitation
Excitation → Inhibition → Labeling → Segmentation result

Boundary-Active-Only Scheme (BAO)

new excited region at clock cycle t+1
activated cells at clock cycle t+2
cell
excited region at clock cycle t
● active mode   ○ stand-by mode

- State-transition evaluation is only necessary for the grown region's boundary cells (Only boundary cells are in active mode and other cells are in stand-by mode)
- More than 75% power reduction is achieved for the 10×10 pixel cell-network

cell P₁₁   CLK
CLK₁=0
ZOR₁=0
CLK₂=1
ZOR₂=0
CLK₃=1
ZOR₃=1
CLK₄=1
ZOR₄=0
CLK₅=0
ZOR₅=0
Clock Controller
cell-network

State transition in row 3 (rows 2, 3 and 4 have the possibility that excitable cells exist during the next clock cycle)
Clock controller distributes the clock signal only to rows 2, 3, 4

Clock control for each row :
- Only region-growing boundary rows are activated by gated clock signals
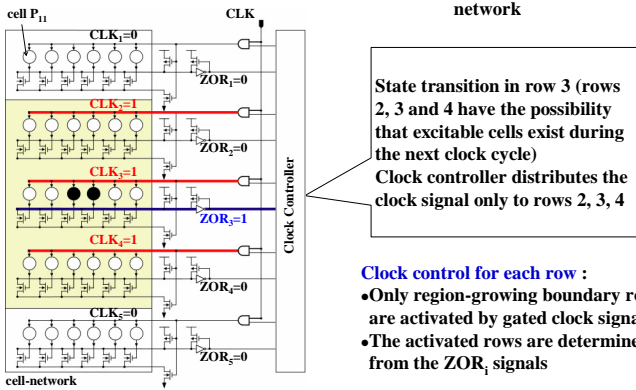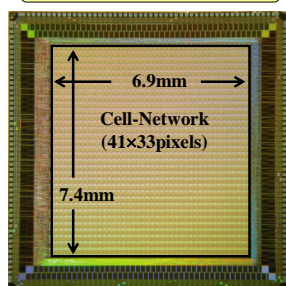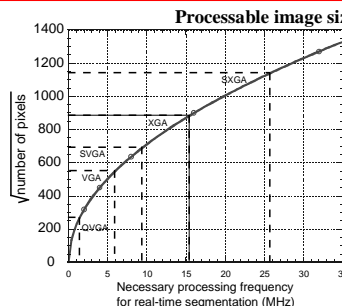- The activated rows are determined from the ZOR_i signals

Table: Comparison result between previously proposed architecture[†] and newly proposed BAO architecture [†] T. Morimoto et al., 1st COE Workshop, 2003.

	without BAO[†]	with BAO	ratio
Number of Transistors	1694/cell	1738/cell	+3% / cell
Processing Time (41×33pixels)	34μsec	34μsec	0%
Power Dissipation	0.309mW/cell	0.069mW/cell	-78% / cell

Realized Architecture

input image memory → weight calculation circuit → leader cell selection circuit → image segmentation cell-network → segmentation restore circuit → image segmentation memory

transmitted in column pipeline mode

Cell : Consists of registers and adders/subtractors
Changes its state $x_k \in \{1,0\}$ depending on $\sum W_{ik} \times x_k$ of 8 neighbors

Weight-register block (WRB) : Two register-block types (horizontal/vertical)
Store the 4 connection-weights between the adjacent active cells
Output weights $W_{ik} \times x_k$ to adjacent active cells

Vertical register block
Cell
Horizontal register block
Cell-Network

Network can be implemented by alternately laying an active cell and a WRB
Area minimization with effective sharing of WRBs among neighboring cells
High speed execution by pixel-based fully parallel processing

Subdivided-Image Approach (SIA)

Image Segmentation for the SIA algorithm

The labels at overlap region are used in other tile's segmentation
Label conflict Labels 2, 4, 5 are allocated to same region
active tile
conflicting labels
□ : Overlap region

Tile-Segmentation Flow

(a) Overlap leader cell self-excitation
(b) Overlap leader cell with identical label excited
(c) New overlap cells excited
(d) Cell with label identical to overlap cell also excite
(e) End of excitation, region is segmented
(f) Allocate label to excited region

Flowchart of tile-segmentation

Implementation Result

6.9mm
Cell-Network (41×33pixels)
7.4mm

Technology	0.35μm 2-Poly 3-Metal CMOS
Measured Max Frequency	10MHz
Power Dissipation (3.3V, segmentation)	45.8mW@10MHz (average)
	94.0mW@10MHz (worst)
Segmentation Time	34μsec@10MHz (worst)
Pixel Integration Density	26.5pixel/mm²

Processable image sizes in real-time and required clock frequency

√Number of pixels vs. Necessary processing frequency for real-time segmentation (MHz)

Estimated image size processable in real-time (**10msec/frame**) as a function of the clock frequency (Estimated with 41×33 pixel cell-network)

VGA size image (640×480 pixels) at 6MHz
SVGA size image (800×600 pixels) at 10MHz
XGA size image (1024×768 pixels) at 16MHz