Low-Power Video Segmentation by Pipeline Processing of Tiled Images


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 (e.g., 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 pixels, 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 0.35um technology, including a 41x33 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 \( W_i \) 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 \( W_i \) 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 excited. 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 prelabeled 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 \( x_i = 1 \). (2) It has already a segment number \( l_i = 1 \). (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. (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 signals 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 0.35um 2-Poly 3-Metal CMOS technology. Figure 5 shows the die photo of the fabricated chip including a
cell-network for 41x33 (1,353) pixels on an area of 51.1mm².
The integration density achieved in the full-custom design is
26.5pixel/mm². 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 10x10 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
VGA-size images are <250µsec at 10MHz and <120mm² 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 41x33 cells
in a 90nm CMOS technology for low-power video segmenta-
tion and experimentally confirmed the effectiveness of the 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 41x33 cell-network (16x15 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.

Acknowledgments
The test-chips in this study have been fabricated in the chip fabrication
program of VDEC, the University of Tokyo in the collaboration with Rohm
Corporation and Toppan Printing Corporation. Part of this work was sup-

References

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

<table>
<thead>
<tr>
<th>Technology</th>
<th>0.35µm, 2-Poly 3-Metal CMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Architecture</td>
<td>Weight-Parallel (high-speed)</td>
</tr>
<tr>
<td>Design Area</td>
<td>6.9mm²×7.4mm² (41x33 cells)</td>
</tr>
<tr>
<td>Supply Voltage</td>
<td>3.3V</td>
</tr>
<tr>
<td>Max Clock Frequency</td>
<td>10MHz</td>
</tr>
<tr>
<td>Segmentation Time</td>
<td>34µsec@10MHz (Worst Case)</td>
</tr>
<tr>
<td>Worst Case Power Dissipation</td>
<td></td>
</tr>
<tr>
<td>(Measured, 41x33 pixels)</td>
<td>94.0mW@10MHz (Segmentation)</td>
</tr>
<tr>
<td></td>
<td>192mW@10MHz (Initialize)</td>
</tr>
<tr>
<td>Pixel Density</td>
<td>26.5pixel/mm²</td>
</tr>
</tbody>
</table>
Low-Power Video Segmentation by Pipeline Processing of Tiled Images

Research Center for Nanodevices and Systems, Hiroshima University

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
- 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

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

Boundary-Active-Only Scheme (BAO)
- New excited region at clock cycle t+1
- Activated cells at clock cycle t+2
- Cell state transition: Only the activated cells which can excite its neighbors have to be excited
- Cell active mode: Excited from clock signal
- Cell stand-by mode: No excitation

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

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

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

<table>
<thead>
<tr>
<th>Without BAO</th>
<th>With BAO</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Transistors</td>
<td>1694/cell</td>
<td>1378/cell</td>
</tr>
<tr>
<td>Processing Time</td>
<td>34usec</td>
<td>34usec</td>
</tr>
<tr>
<td>Power Dissipation</td>
<td>0.390mW/cell</td>
<td>0.069mW/cell</td>
</tr>
</tbody>
</table>

Implementation Result
- Technology: 0.35μm 3-Poly 3-Metal CMOS
- Measured Max Frequency: 10MHz
- Power Dissipation
  - (3.3V, segmentation): 45.8mW@10MHz (average), 94.6mW@10MHz (worst)
  - Segmentation Time: 34usec@10MHz (worst)
  - Pixel Integration Density: 26.5pixel/mm²

Estimated image size processable in real time (≤ 10msec/frame) as a function of the clock frequency

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

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

Subdivided-Image Approach (SIA)
- Flowchart of tile-segmentation

Realized Architecture
- Cell: Consists of registers and adders/subtractors
  - Changes its state xk∈{1,0} depending on \( \sum W_k x_k \) 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 \( W_k x_k \) to adjacent active cells

Network can be implemented by alternately laying an active cell and a WRB