Image Reconstruction Scheme for Watersheds based Video Segmentation

HyunSil Shin	HyunWoo Park	HwangKyu Yang	TaeSoo Yun
Graduate School of Design & IT	Graduate School of Design & IT	Division of Internet Engineering	Division of Digital Contents
Dongseo University	Dongseo University	Dongseo University	Dongseo University
Busan, S. Korea	Busan, S. Korea	Busan, S. Korea	Busan, S. Korea

Abstract - *In this paper, we proposed a new solution to the reconstruction of image in order to solve the over-segmentation problem of watershed transformation which is widely used as an image segmentation method. The process for reconstruction of image consists of consolidation of edges and removal of local minima. The former(consolidation of edges) enhances the edge between the object in image background and the other(removal of local minima) erodes partially an image in order to eliminate the small local minima. Comparing to the existing solution, the proposed method could efficiently solve the over-segmentation problem within a very short time. And, using the modified similarity measure, we could acquire the result which is independently merged even though the criterion area is different. So, our result can be applied real-time video based application such as video conference, surveillance and MPEG-4 based systems.*

Keywords: Image Reconstruction, Watershed, Video Segmentation.

1 Introduction

Video segmentation is a process of partition an image sequence into moving objects and tracking the evolution of the moving objects along the time axis [1]. It requires image segmentation as essential step, whose goal is to partition an image into homogeneous regions and locate the contours of the regions as accurately as possible. Recently, as the standards for MPEG-4 [2], MPEG-7 [3] are established, the importance of image segmentation is more increasing. Recently, a lot of method using watershed algorithm has been proposed for the video segmentation and object tracking. Watersheds based methods result in a good performance in both computational load and segmentation accuracy. But traditional watershed transform has been suffered from over-segmentation due to small local minima included in gradient image which is input to the watershed transform. The effectiveness of the image segmentation methods is limited by the quality of the gradient image used in the transformation [4]. The traditional gradient algorithms based on step edge model cannot eliminate the local minima completely.

The first method to solve this problem is to merging adjacent regions by predefined criteria after initial partition through the watershed algorithm [5]. The other approaches are using markers [6]. These traditional methods focused on merging segmented regions using prior knowledge after initial partition. But our approach is a new attempt to enhance initial performance and reduce the computational burden by eliminating local minima resulted in over-segmentation at initial step.

This paper introduce new algorithm of reconstructing image for making the gradient image that is to be the input of watershed transform and extend it to apply object tracking. And we also introduce a similarity measure to merge segmented regions. The outline of the proposed overall process is displayed in Fig. 1. The outline of this paper is as follows. Section 2 describes the initial segmentation using watersheds transformation based our proposed reconstruction algorithm. Section 3 describes region merging using similrarity measure based on motion information. In section 4, the experimental results are presented to validate our proposed algorithm. Finally, section 5 describes conclusions and discusses topics for future research.

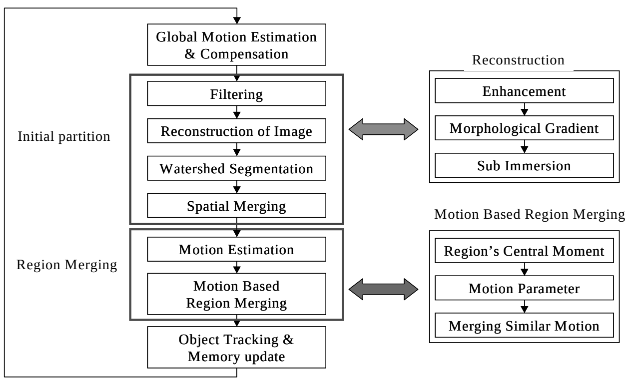


Fig. 1. The block diagram of overall process

2 Initial Partitioning

The initial partitioning consists of three steps. From the input image, we apply preprocessing to the input image to distinguish boundary between the objects more clearly. And then we obtain the gradient image to be used in watershed transform using morphological operator. Since the gradient image still includes minute local minima, we apply sub-immersion step to the gradient image to eliminate small local minima and locate the starting point of watershed transform. And finally, we spatially merged the separated regions using edge information of adjacent regions.

2.1 Image Filtering

This stage, histogram equalization is performed to eliminate the noise spread out the whole image. It is performed via median filtering technique. It is, however, difficult to remove completely the noise exist in the boundary between objects. Therefore, we use mathematical morphological operator, i.e., opening and closing, to remove noises exist boundary between.

2.2 Image Reconstruction

The goal of image reconstruction is to consolidate the boundary between object in each frame and effectively remove the local minima result in over-segmentation. This process is the most important point of our research results. The image reconstruction consists of boundary enhancement stage which distinguish edge between objects and background, and local immersion stage.

Utilizing the automatic thresholding technique [7], we can enhance the image according to the following equation.

$$f'(x,y) = \begin{cases} C_{th}, & if \;\; f(x,y) < C_{th} \\ f(x,y) \times \alpha, & otherwise \end{cases} \tag{1}$$

where, $f(x,y)$ is original input image, $f'(x,y)$ is a enhanced image, α is constant, and C_{th} is threshold value obtained by automatic thresholding technique.

And then gradient image is generated using morphological operator for watershed transformation [8]. The gradient image generated through the morphological operator may still small local minima result in over-segmentation in watershed transform. To eliminate these minute minima, we performed the sub-immersion of the gradient image to the appropriate water level. It is important that we should determine the critical point not only to clarify the boundary between objects, but also eliminate small local minima. We propose new sub-immersion algorithm that immerse the image until the mean of the image intensity converges to the constant value using the algorithm1 as follows.

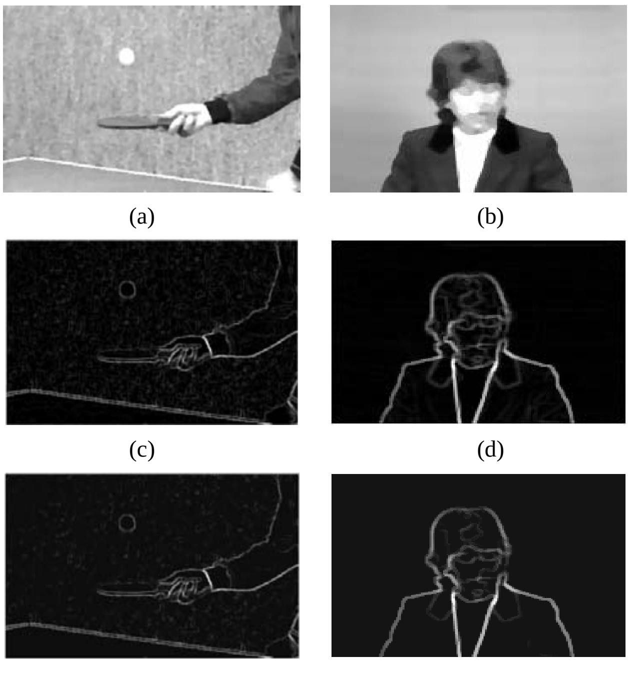
Algorithm 1 : The Algorithm of image reconstruction

Input : The gradient image
Output : The reconstructed image

```
// initialize the parameter and environment.
int mean, bak_mean=0;
do {
    mean = ( sum_{y=0}^{height-1} sum_{x=0}^{width-1} f(x,y) ) / ( height × width )
    if( f(x,y) < mean x β )
        f(x,y) = mean x β
    if(mean == bak_mean) break;
    else  bak_mean = mean;
} while(true)
```

$$mean = \frac{\sum_{y=0}^{height-1} \sum_{x=0}^{width-1} f(x,y)}{height \times width}$$

The results of reconstructed image are displayed in Fig. 2. The (a),(b) of Fig.2. is the result of edge enhancement and (c), (d) is reverse image of gradient image, (e),(f) is the reverse image of sub-immersed result.



(a) (b) (c) (d) (e) (f)

Fig. 2. The reconstructed

2.3 Spatial Merging

The over-segmented regions are still remained even though the watershed transformation is performed. These regions have to be merged according to a certain criteria. In this paper, adjacent regions are merged by the information of watershed line between two catchment basin B_A, B_B [1]. We define $Edge(R_i, R_j)$ as the number of pixel consists of adjacent region R_i, R_j, and $Weak(R_i, R_j)$ as the number of pixel of whose intensity are less then δ. For the adjacent region R_i, R_j, if the condition (2) is satisfied, the two regions are merged. The process of merging is displayed in Fig. 3.

$$\frac{Weak(R_i, R_j)}{Edge(R_i, R_j)} > \frac{1}{3} \qquad (2)$$

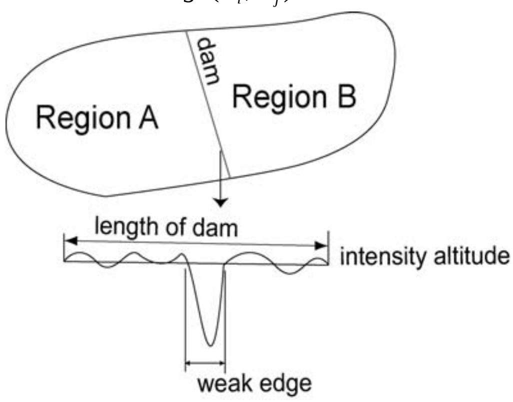


Fig. 3. The merging process of region based on edge information

3 Region Merging Based Motion Information

In this section, we describe the process of merging based on motion information of each region obtained by the initial partition. We also introduce the similarity measure to merge adjacent regions.

3.1 Motion Estimation

Each region separated in initial partition stage represents the spatial relations based on pixel intensity. To track the moving object, the regions have the same intensity and motion information have to be merged. Therefore, the motion of each region has to be estimated and merged based on similarity measure. In this paper, we estimate the motion information using block matching method which can be used widely. The block matching method has the advantage of being easily implemented in hardware. To estimate the motion information of the corresponding area of next frame for the each separated region in the current fame, we use affine motion model constituted 6-parameters [9].

Generally, motion information can be estimated using a traditional steepest descent method [10]. However, it has been found that the method requires much calculation time. Therefore, this paper used Least Square Method(LSM) to reduce time the amount of calculation. The size of one block is 16x16 rectangle and the motion vector u, v is calculated using Mean Square Error(MSE) [10].

In addition, since the full search method is not effective in contrast to time, we use adaptive motion estimation applying 3-step search and full search in hierarchical framework. That is, we performed apply block matching with 3-step in level-1 and level-2 of high resolution level, and full-search in level-3 of low resolution level.

Fig. 4. displays the result of estimation of motion vector using adaptive motion estimation in this research.

(a) motion for each region (b) segmented region

Fig. 4. The estimated motion via adaptive block matching

3.2 Region Merging

In this stage, we merge regions using similarity measure based on block motion estimated in previous stage. To calculate the similarity measure, mean-square motion compensation error [9] is used. To reduce the time of searching of entire image and to manage the space more efficiently, the separated regions are managed Region Adjacency Graph(RAG).

For all segmented region in initial stage, motion compensation error E_i is calculated using the following eqn. 3 for all segmented region.

$$E_i = \frac{1}{\| R_i \|} \sum_{r \in R_i} (f(r,t) - I[f(r + d(p_i), t-1)])^2$$
$$for\ j = i \text{ or } j : R_j\ adjacent \text{ to } R_i \qquad (3)$$

where, $\|R_i\|$ is the number of pixel included in region R_i, I is bilinear interpolator.

The mean-square motion compensation error for all adjacent region based on adjacent matrix is calculated according to the eqn. 4. As you can see from the eqn. 8, E_{ii} represents motion compensation error in case they are compensated using the motion of E_i itself, and E_{ij} denotes motion compensation error when it is compensated using the motion of adjacent region E_j.

The eqn. 4 is similar to the eqn. of [1]. But, contrary to the approach [1], our proposed similarity measurement could show the result which is independently merged for criterion area.

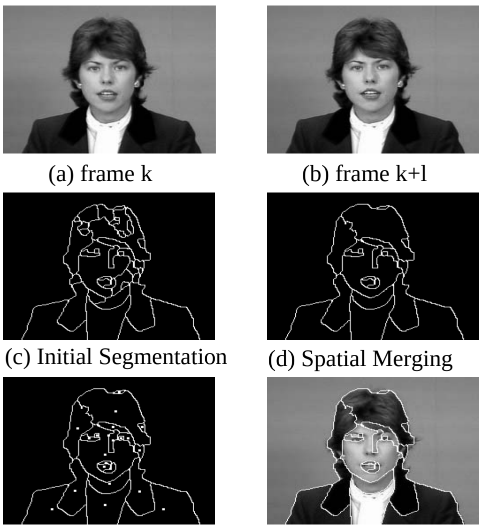
For the all adjacent region, when the value $\Delta_{i,j}$ is less than the threshold value T2, the adjacent region is merged.

$$\Delta_{i,j} = \frac{E_{i,j} + E_{j,i} - E_i - E_j}{N_i + N_j} \quad (4)$$

$$where,\ E_{i,j} = \frac{1}{\| R_i \|} \sum_{r \in R_i} (f(r,t) - I[f(r + d(p_j), t-1)])^2$$

Lastly, the value of E_i, Δ_{ij}, N_i for merged region is updated, and performed this step until Δ_{ij} is larger than T_2.

The result of merging based on similarity measure, initial segmentation, spatial merging is displayed in Fig. 5.



(a) frame k (b) frame k+l

(c) Initial Segmentation (d) Spatial Merging

(e) Momentum (f) Final Segmentation

Fig. 5. The result of region merging based on similarity measure

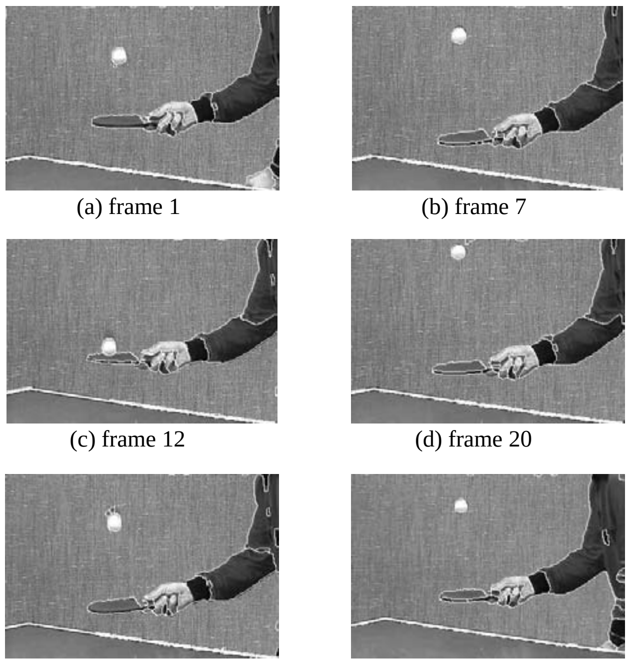
4 Experiment Results

To validate the proposed scheme, we implemented the algorithm and tested several test set frequently used in state of arts in research. We compared the performance with other techniques, such as methods using sobel operator, morphological operator and multi-scale morphological operator, in terms of number of regions and computational complexity. Because we do not obtain the source code for other techniques, we follow faithfully the steps proposed in that literature as much as possible. Table 1 shows the result of comparision.

Table 1: The comparision in the number of region

	1	4	8	12	16
Sobel	4059	4057	4018	3922	4009
Morphological	2032	1941	1996	1977	1989
Multi-scale morphological	1112	1111	1139	1102	1110
our method	320	276	292	291	291

Fig. 6 shows the result of final segmentation result through the spatio-temporal merging proposed in previous section. As shown in Fig. 6. We can get fine result because our method eliminate the local minima before the watershed transform.



(a) frame 1 (b) frame 7

(c) frame 12 (d) frame 20

(e) frame 24 (f) frame 33

Fig. 6. The final segmentation results

Table 2: The results of PSNR for Gaussian noise

	10%	20%	30%	40%	50%
Clare 10^{th}	31.819	31.398	31.622	20.589	29.759
Akio 15^{th}	43.038	41.674	40.107	37.217	35.081
Mother&Daughter 30^{th}	37.231	36.513	38.043	34.941	32.860
Table Tennis 20^{th}	30.280	29.178	25.496	25.860	26.541
Hall monitor 20^{th}	33.092	31.161	29.811	29.344	27.922

We also performed experiment on noise image to know how much tolerant is our method to noise. Table 2 shows the result of PSNR for gaussian noise.

5 Conclusions and Future Works

In this paper, we proposed a new image reconstruction method to solve to over-segmentation problem, and also proposed similarity measure to merge the region with same motion based on motion information estimated from initially segmented regions. And we established the fact that our result can be effectively applied object tracking in video sequence.

The advantage of the proposed method is that it can be applied real-time video based application because it is very fast compared to the traditional approach through simplifying the process making gradient image. And our method provides reliable result even if the image was cluttered and had textured background.

However, the proposed method has a problem of a small loss in the weak edge while the regions are merging in spatial domain. Future work should concentrate on incorporating color information and temporal information such as motion and change detection into the initial partition.

Acknowledgement

This research was supported by the Program for the Training of Graduate Students in Regional Innovation which was conducted by the Ministry of Commerce Industry and Energy of the Korean Government.

6 References

[1] Demin Wang, "Unsupervised Video Segmentation Based on Watersheds and Temporal Tracking," IEEE Trans. on circuit and systems for video Technology. vol. 8, No. 5, pp. 539~546, Sep. 1998.

[2] ISO/IEC JTC1/SC29/WG11, "Overview of the MPEG-4 standard," MPEG98/N2323, Dublin, Ireland, July, 1998.

[3] "MPEG-7 Requirements Document," MPEG Requirements Group, 1998.

[4] L. Vincent and P. Soille, "Watersheds in digital spaces: An effieient algorithm based on immersion simulations," IEEE Trans. Pattern Anal. Machine Intell., vol. 13, pp. 583-598, 1991.

[5] A. Tremeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," IEEE Trans. Image Processing, vol. 9, pp. 735-744, Apr. 2000.

[6] Liu, J., Yang, Y. H., 1994, Multiresolution color Image Segmentation, IEEE Trans. Pattern Anal. Machine Intell. 16 (7), 689-700.

[7] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Trans. SMC., vol. 9, pp. 62-66, 1979.

[8] M.Kim, J.G.Choi, D.Kim, H.Lee, M.H.Lee,C.Ahn, and Y.Ho, " A VOP generation tool:Automatic segmentation of moving objects in image sequences based on spatio-temporal information," IEEE Trans. on Circuits. syst. Video Technol., vol. 9, no. 8, Dec., 1999.

[9] L. Wu, J. Benoise-Pineau, Ph. Delagnes, and D. Barba, "spatio-temporal segmentation of image seqences for object-oriented low bit-rate image coding," Signal Processing:Image Commum., vol. 8, pp. 513-543, 1996.

[10] A. M. Tekalp, Digital Image Processing, Englewood Cliffs, NJ:Prentice-Hall, 1995.

[11] Yaakov Tasig and Amir Averbuch, "Automatic Segmentation of Moving Objects in Video Sequences: A Region Labeling Approach,", IEEE Trans. on Circuits Syst. video Technol., vol. 12, No. 7, July, 2002.