

# A REAL-TIME CURVE EVOLUTION-BASED IMAGE FUSION ALGORITHM FOR MULTISENSORY IMAGE SEGMENTATION

Yuhua Ding<sup>1</sup>, George J. Vachtsevanos<sup>1</sup>, Anthony J. Yezzi Jr.<sup>1</sup>, Wayne Daley<sup>2</sup>, Bonnie S. Heck-Ferri<sup>1</sup>

<sup>1</sup>School of ECE, Georgia Institute of Technology, Atlanta, GA 30332, USA  
{yhding, george.vachtsevanos, ayezzi, bonnie.heck}@ece.gatech.edu,  
Tel. (404) 894-6252, Fax. (404) 894-7358, <http://icsl.marc.gatech.edu/>

<sup>2</sup>Food Processing Technology Division, Electro-Optics, Environment and Materials Laboratory,  
Georgia Tech Research Institute, Atlanta, GA 30332-0837, USA  
wayne.daley@gtri.gatech.edu, Tel. (404) 894 3693, <http://www.gtri.gatech.edu/>

## ABSTRACT

A partial differential equation (PDE)-based feature-level image fusion approach is proposed for multisensory image segmentation. The energy functional of the proposed fusion model is a weighted sum of several functionals, each constructed based on the characteristics of the sensor image. The weight selection decides the way that the model handles redundant, conflicting, or complementary information involved in the multisensory data. The method is implemented using level sets and is fast enough for real-time segmentation tasks. Finally the algorithm is applied to the segmentation of x-ray and visual images, and the results show that the fusion algorithm is efficient, accurate, and robust.

## 1 INTRODUCTION

It has been shown that the segmentation accuracy has a significant effect on feature distributions and classification performance in object recognition applications [3]. However, segmentation on monosensory images may not produce satisfactory performance due to the intrinsic ambiguity and incompleteness associated with the data. In monosensory image segmentation, each image generates false regions or edges. Several approaches ([2,6,10,11]) have been developed to increase the segmentation accuracy on monosensory images. However, the processing times of these methods are multiple times longer and thus, cannot be used in real-time applications. Images acquired by different sensors are generally partially redundant and partially complementary, which can be used to reduce the imprecision and to interpret the scene more accurately. Unfortunately, the multisensory image segmentation problem is not a straightforward extension of its monosensory counterpart on multiple images. The fusion algorithm must make smart decisions to eliminate the redundancy, to include the complementary information, and to resolve the confliction.

Among all levels at which the fusion can be performed: pixel, feature, and decision, feature-level fusion is well suited for real-time applications. On one hand, the amount of data to be processed is greatly reduced by focusing on the higher level image representation of segmentation features. On the other hand, the information loss during the fusion is not essential since most significant features have been preserved.

By combining different modalities in segmentation, we expect to achieve the following goals:

- Locate more accurate boundaries for the objects in the scene.
- Eliminate false edges/regions as many as possible.
- For the false edges/regions that cannot be totally eliminated, the significant difference between the segmentation results on individual modalities can be utilized to further separate false edges/regions from true edges/regions.

In the PDE-based curve evolution methods, the contour evolves according to the optimal flow derived from the energy functional. Segmentation errors occur when the images do not satisfy the assumptions based on which the energy functional is constructed. Because this type of method is local in nature, the curves tend to get trapped by unexpected features before they reach the true edges in the image, which causes under- or over-segmentation errors. Since the unexpected feature that entrapped the contours may show itself differently in images acquired by other sensors, it may function as a driving force to push the contour toward the true edge and thus reduce the segmentation errors. In the proposed model, the energy functional is a weighted sum of several functionals, each constructed based on the sensors' characteristics. Therefore, the method is applicable to the case where the objects in the scene exhibit totally different views between images. *A priori* information of the scene and sensors can be employed to further boost the processing speed so that the algorithm is more suitable for real-time applications.

The paper is organized as follows: First, the multisensory image segmentation methods are briefly reviewed. Then, a real-time multisensory image segmentation approach is presented. Finally, the approach is applied to segmentation of the visual- and x-ray images of bone-contaminated poultry meat, and the results are presented.

## 2 MULTISENSORY IMAGE FEATURES AND SEGMENTATION METHODS

Depending on the characteristics of the physical sensors and target objects, the definition for edge and region on multisensory images may vary. Bonnin [1] defines multispectral edge point as a pixel where there are important variations at least in one direction in its local neighborhood of a property, in at least one spectral image. This definition implies a logical OR between the different spectral images of edge points. But it does not solve the problem of the obvious displacement of an edge between two spectral images. This problem has to be taken into account and solved in the thinning edge and chaining edge steps. In addition, the false edges caused by artifacts and noises are not considered. By applying the logical OR operation, the false edges will be accumulated in the final edge map. Therefore, caution needs to be taken when fusing edge maps. The majority voting rule in [9] can eliminate false edges if they are only observed in a limited range of bands. However, only the most important edge information is preserved. Similarly, the region of homogeneity implies a logical AND between the homogeneity predicates in all spectral images. The homogeneity criteria are usually thresholdings of the homogeneity measures. The measures can vary largely with different sensors, and even for the same measure, the thresholds may be different. The choice of the homogeneity measure and the adjustment of its thresholds are usually guided by the knowledge of the sensors' physical characteristics.

One common scheme for multisensory image segmentation, as shown in Fig. 1, is to obtain features from individual sensor images, then combine these monosensory features using simple fusion rules or within the mathematical frameworks of Bayesian Theorem, Fuzzy Theory, and Dempster-Shafer (DS) Theory. Note that in order to utilize the rules in such frameworks, the fusion problem first has to be modeled in that framework. One application is to fuse SAR and optical images to achieve better performance in detecting urban areas [4]. The fusion of individually segmented images is modeled as a nonlinear optimization problem, where the objective function is the sum of two terms: the mismatch between the two fused segmentation maps and the discrepancies between corresponding fused and original segmentations. The handling of unmatched monosource features are controlled by the penalization parameters on the second term. One concern about this method is that *ad hoc* calibrations of several parameters are necessary. Another issue is that the nonlinear optimization problem is hard to solve and can be too slow to be used real-time.

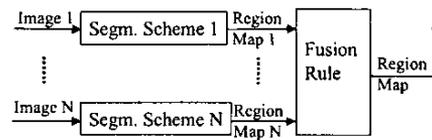


Fig. 1. Fusion of individual segmentation results.

One can also build a segmentation map directly on all images, as shown in Fig. 2. For example, in [5], multisensory data are classified pixel by pixel using DS theory to produce a labeled image (segmentation map). However, the segmentation map needs to be further refined to remove speckle errors. A natural PDE-based scheme is to use the Mumford-Shah (MS) functional by treating the multisource images as a multichannel image [7]. The limitation of this model is that it is not applicable to the case where at least one image cannot be approximated using a piece-wise constant image. Besides, it is well known that the curve evolution based on this model is extremely slow and therefore, cannot be used in real-time applications.

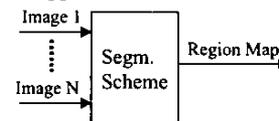


Fig. 2. Multichannel image segmentation.

## 3 A PDE-BASED MODEL FOR MULTISENSORY IMAGE SEGMENTATION

The multisensory image segmentation can be modeled as an optimization problem: locate the contours so that the following energy functional is minimized:

$$E(t) = \sum_{i=1}^N w_i(t) E_i(t), \text{ s.t. } \sum_{i=1}^N w_i(t) = 1 \quad (1)$$

where  $E_i$  is the energy functional,  $w_i$  is the controlling weight, on the  $i$ -th image, and  $N$  is the total number of images. First, each energy functional,  $E_i$ , is chosen according to the characteristics of the corresponding image. Thus, the model is applicable to essentially any combination of sensors. Second, the weights can vary as curve deforms. The weight selection is a demanding task since it determines how information is handled. Generally, the following rules can be applied:

- In case of redundancy, equal weights are used.
- In case of complementarity, the sources that contain the most useful information have highest weights.
- In case of conflicts, the sources with more reliable information have higher weights.

Therefore, the key issue is to identify the local information content at a specific time. Two common weight choices are constant weights and 0-1 weights.

- Constant weights: the weights do not change as contours evolve. The optimal force on the contour is a weighted sum of individual forces derived from each functional.
- 0-1 weights: all weights are set to 0 except for one set to 1 at any specific time. Images are segmented one after another, as shown in Fig. 3, each starting from the

previous segmentation. Since the initial estimation of on one image is from the segmentation results of its registered peers, the convergence is faster than segmenting each modality independently.

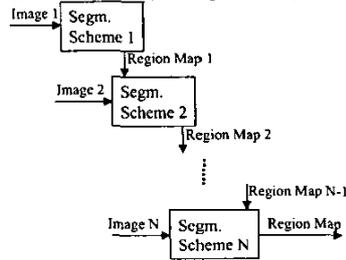


Fig. 3. Segmentation using 0-1 weights.

Simulations of the above two weight choices are performed on the gray-scale test images shown in Fig. 4, where  $I_0$  is the root image, and  $I_A, I_B$  are the two gray-scale test images generated by diffusing and shading  $I_0$ , respectively. In  $I_A$ , the edge information is totally lost during diffusion, while in  $I_B$ , although edge is well preserved, the region  $R$  cannot be separated from its background  $\bar{R}$  based only on the intensity information.

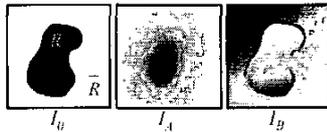


Fig. 4. Gray-scale test images.

The energy functionals in (1) for the test images are

$$E_1 = -\frac{1}{2}(u-v)^2 + \alpha \int ds, \quad (2)$$

$$\text{and } E_2 = \int \phi ds, \quad (3)$$

where  $u$  and  $v$  are mean intensities inside and outside the contour, respectively,  $\phi = \frac{1}{1 + \|\nabla I\|^2}$  is the edge feature

metric, and  $\alpha \in [0, 1]$  is a constant.  $E_1$  is used to maximize the difference of the mean intensities inside and outside of a smooth contour, while  $E_2$  is minimized when the contour falls on the edge. The optimal flows derived from  $E_1$  and  $E_2$  are the binary flow ([12])

$$\frac{d\bar{C}}{dt} = (u-v) \left( \frac{I-u}{A_u} + \frac{I-v}{A_v} \right) \bar{N} - \alpha \kappa \bar{N} \quad (4)$$

and the geodesic flow

$$\frac{d\bar{C}}{dt} = \phi \kappa \bar{N} - (\nabla \phi \cdot \bar{N}) \bar{N}, \quad (5)$$

where  $\bar{C}$  is the contour,  $A_u$  and  $A_v$  are areas inside and outside the contour, respectively,  $\bar{N}$  is the unit normal, and  $\kappa$  is the curvature.

The results of the proposed model using both the equal, constant weights and the 0-1 weights are shown in Fig. 5 (a) and (b), respectively. It can be seen that the convergence speed of (a) is faster than that of (b). In both cases, the final contour is superior to the monosensory results.

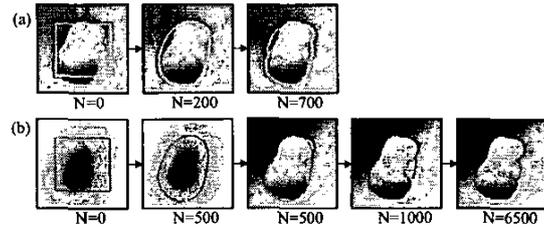


Fig. 5. Result on test images using (a) equal and constant weights, and (b) 0-1 weights.

#### 4 SEGMENTATION OF X-RAY AND VISUAL IMAGES

The proposed method is used to segment x-ray and visual images of bone-contaminated poultry breast. The bones to be detected are fan bones, the fan-shaped thin surface bones that show differently in x-ray and visual images. Monosensory segmentation on either mode is not satisfactory. Visual images contain shadow and edge features that may appear similar in color and shading to bones. In x-ray images, the local contrast is good, but the fan bone regions have similar intensity levels as the thicker meat does, and therefore, the segmentation is difficult.

An example of the registered images is shown in Fig. 6., where (a) and (b) are the clips of the visual and x-ray images, respectively. The fan bone is marked out in each image. The x-ray image and the red channel of the visual image are overlapped to form the synthesized image in (c). Significant mismatch between the two modalities is observed due to different sensor sensitivities: one common subject (such as a fan bone) may have distinct views in the images of different modalities.

Before fusion, the x-ray images are preprocessed to remove the slow intensity variation caused by uneven thickness of the meat using the feature extraction method in [8]. The visual feature image for the x-ray clip in Fig. 6 is shown in Fig. 7(a). Fig. 7 (b) is the segmentation result using (4) on the feature image. The segmentation algorithm failed to separate the fan bone region totally from the darker regions nearby.

The proposed fusion scheme is applied on the x-ray feature image and the red channel image. The binary flow in (2) is used throughout the segmentation procedure. The initial contour is obtained by thresholding the red channel of the visual image. The results using constant and 0-1 weights are shown in Fig. 8 and Fig. 9, respectively.

From the test, the result using constant weights is generally not as good as the result when 0-1 weights are used. A major reason is that the significant mismatch between the two images tends to cause over-segmentation when constant weights are applied. The scheme with 0-1 weights is not only effective in allocating fan bone boundaries, but also robust to the mismatch.

The algorithm is tested on 51 fan bone-contaminated image clips cut from the registered visual and x-ray

images taken on-line in a poultry plant. The segmentation results are assessed visually according to the closeness of the contour to the true boundary. Results show that the fusion schemes provide better performance than the single-sensor scheme does. It is also observed that the algorithm is capable of correcting monosensor segmentation errors caused by limitations of both the segmentation algorithm and sensor properties, as illustrated in Fig. 10.

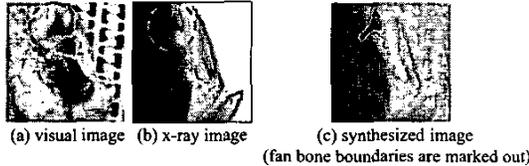


Fig. 6. An example of registered visual and x-ray images.

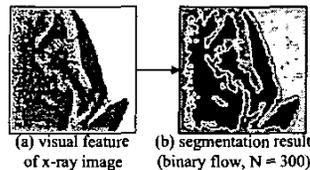


Fig. 7. The x-ray feature image and its segmentation.

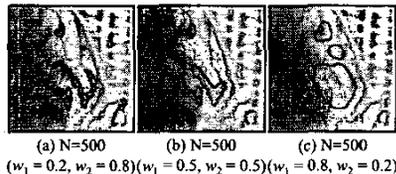


Fig. 8. The segmentation result using constant weights.

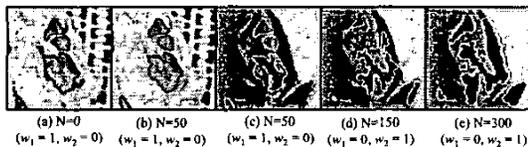


Fig. 9. The segmentation result using 0-1 weights.

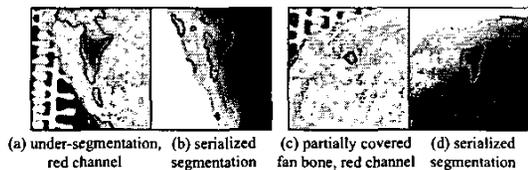


Fig. 10. Comparison of results obtained from visual-based segmentation and fusion using 0-1 weights.

The algorithm is implemented using level sets, and the processing time on 160x160 image clips is no more than 500ms on a 900MHz Pentium III PC. Therefore, the proposed scheme is suitable for real-time applications.

## 5 CONCLUSIONS

In this paper, a multisensory image segmentation scheme is presented. The key issues of this scheme are weight

selection and functional construction. The experimental results suggest that this algorithm is effective, accurate, and fast. Future research will be focused on dynamic weight selection and interpretation of segmentation results.

## ACKNOWLEDGEMENTS

The authors would like to express their appreciation to the State of Georgia and the Governor's Food Processing Advisory Council for their support in the conduct of this research. The authors also wish to thank Tyson Foods, Inc. and Cagle's Inc. for supplying chicken parts during the development and testing of the system.

## REFERENCE

1. P. Bonnin, B. Hoeltzener-Douarin, and E. Pissaloux, "A New Way of Image Data Fusion: The Multi-Spectral Cooperative Segmentation," *Proc Intl Conf Image Processing*, pp. 572-575, 1995.
2. B. Charroux, S. Philipp, and J-P. Cocquerez, "Image Analysis: Segmentation Operator Cooperation Led by the Interpretation," *Proc. Intl. Conf. Image Processing*, pp. 939-942, 1996.
3. Y. Ding, G. Vachtsevanos, A. Yezzi, Y. Zhang, and Y. Wardi, "A Recursive Segmentation and Classification Scheme for Improving Segmentation Accuracy and Detection Rate in Real-Time Machine Vision Applications," *Proc. 14th Intl. Conf. Digital Signal Processing (DSP2002)*, pp. 1009-1014, 2002.
4. L. Fatone, P. Maioni, and F. Zirilli, "Fusion of SAR/Optical Image to Detect Urban Areas," *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, pp. 217-221, 2001.
5. S. L. Hegarat-Masclé, I. Bloch, and D. Vidal-Madjar, "Application of Dempster-Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing," *IEEE Trans. Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 1018-1031, 1997.
6. F. Huet and S. Philipp, "Fusion of Images after Segmentation by Various Operators and Interpretation by a Multi-Scale Fuzzy Classification," *Proc. 14th Intl Conf. Pattern Recognition*, pp. 1843-1845, 1998.
7. G. Koepfler, C. Lopez, and L. Rudin, "Data Fusion by Segmentation. Application to Texture Discrimination," *Actes du 14me Colloque GRETSI*, pp.707-710, 1993.
8. E. Peli, "Feature Detection Algorithm Based on a Visual System Model," *Proc. IEEE*, vol. 90, no. 1, 78-93, 2002.
9. B. Solaiman, R.K. Koffi, M-C Mouchot, and A. Hillion, "An Information Fusion Method for Multispectral Image Classification Postprocessing," *IEEE Trans. Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 395-406, 1998.
10. C. Spinu, C. Garbay, and J.M. Chassery, "A Multi-Agent Approach to Edge Detection as a Distributed Optimization Problem," *Proc. 13th Intl. Conf. Pattern Recognition*, pp. 81-85, 1996.
11. D. Stewart, D. Blacknell, A. Blake, R. Cook, and C. Oliver, "Optimal Approach to SAR Image Segmentation and Classification," *IEE Proc. Radar, Sonar Navigation*, vol. 147, no. 3, pp. 134-142, 2000.
12. A. Jr. Yezzi, A. Tsai, and A. Willsky, "Medical Image Segmentation via Coupled Curve Evolution Equations with Global Constraints," *Mathematical Methods in Biomedical Image Analysis, 2000. Proc. IEEE Workshop on Biomedical Image Analysis*. pp. 12-19.