Real-world Pill Segmentation based on Superpixel Merge using Region Adjacency Graph

Sudhir Sornapudi[1], R. Joe Stanley[1], Jason Hagerty[1,2] and William V. Stoecker[2]
[1]*Missouri University of Science and Technology, Department of Electrical and Computer Engineering, Rolla, MO, U.S.A.*
[2]*Stoecker & Associates, Rolla, MO, U.S.A.*
{*ssbw5, stanleyj, jrh55c, wvs*}*@mst.edu*

Keywords: Segmentation, Clustering, Superpixels, Graph Theory, Region Adjacency Graph, Threshold Cut.

Abstract: Misidentified or unidentified prescription pills are an increasing challenge for all caregivers, both families and professionals. Errors in pill identification may lead to serious or fatal adverse events. To respond to this challenge, a fast and reliable automated pill identification technique is needed. The first and most critical step in pill identification is segmentation of the pill from the background. The goals of segmentation are to eliminate both false detection of background area and false omission of pill area. Introduction of either type of error can cause errors in color or shape analysis and can lead to pill misidentification. The real-world consumer images used in this research provide significant segmentation challenges due to varied backgrounds and lighting conditions. This paper proposes a color image segmentation algorithm by generating superpixels using the Simple Linear Iterative Clustering (SLIC) algorithm and merging the superpixels by thresholding the region adjacency graphs. Post-processing steps are given to result in accurate pill segmentation. The segmentation accuracy is evaluated by comparing the consumer-quality pill image segmentation masks to the high quality reference pill image masks.

1 INTRODUCTION

According to National Library of Medicine (NLM, 2016), unidentified and misidentified pills present a challenge to patients, family members and health professionals. Misidentified pills constitute a safety hazard. In the US, nine out of 10 people over age 65 take more than one prescription pill which may increase the chance of pill misidentification. This can lead to adverse drug events (ADE). This situation calls for automatic pill identification, enabling anyone to easily verify whether a pill with different shape, imprint or color is a generic equivalent to the drug he or she was already taking. In an era of increasing polypharmacy and widespread use of 7-day pill dispensers, rapid and accurate automatic pill identification has lifesaving potential.

During the last decade, the improvement in computational power and digital camera technology has facilitated advances in machine vision research, yielding significant progress in automation of medical and industrial computer vision systems. Automatic identification of prescription drugs is now an increasingly important biomedical research topic. Large prescription drug databases are now available to researchers. These databases include the National Library of Medicine (NLM) Pillbox database (Pillbox, 2016), DailyMedPlus (Dailymedplus, 2016), WebMD (WebMD, 2016), and Drugs.com (Drugs.com, 2016). These resources provide various features of a pill, where users can manually access information on pill size, color, shape, and imprint, to allow pill identification (Caban et al., 2012). However, identification by manual website access is error prone and time-consuming. There is a need for an automatic pill identification system that is

Figure 1: Consumer-quality pill images.

Sornapudi S., Joe Stanley R., Hagerty J. and V. Stoecker W.
Real-world Pill Segmentation based on Superpixel Merge using Region Adjacency Graph.
DOI: 10.5220/0006135801820187
In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, pages 182-187
ISBN: 978-989-758-225-7

fast, reliable and easy to use.

Segmentation is the first and most critical step in the pill identification process. Segmentation isolates the pill from background, enabling accurate analysis of the pill features. The images in Figure 1 are typical of pill images used in this project. These are examples of consumer-quality images provided by the NLM Pill Image Recognition Challenge 2016. Simple thresholding on the images in Figure 1 leads to significant segmentation errors, due to shadows and uneven lighting. These challenges are not present in the reference pill images shown in Figure 2.

The main objective of the NLM Pill Image Recognition Challenge was to use computer vision algorithms to rank lower-quality consumer images of prescription pills after training with high quality reference images as shown in Figure 2. These freely available high quality digital images and associated data (C3PI, 2016) were generated by NLM as part of the Computational Photography Project for Pill Identification. Although this challenge provided progress toward automatic pill identification, there is as yet (Fall 2016) no reliable and accurate automatic pill identification technology available.

The Consumer-quality images as shown in Figure 1 have issues such as low illumination, noisy background and pill shadows, all of which pose great challenges in pill segmentation. When pill images include noisy background, feature extraction algorithms can determine false features. Hence, there is a need to develop a segmentation algorithm to reduce these problems.

Figure 2: Reference pill images.

The proposed clustering segmentation algorithm includes three important steps (Xu and Wunsch, 2005). Initially, pre-processing is done to oversegment the pill images by obtaining superpixels based on the modified k-means clustering algorithm. Secondly, a region adjacency graph is obtained from the over-segmented pill image to merge the regions within a certain threshold. Finally, various post-processing steps are applied to obtain the desired mask.

The goal of this paper is to accurately segment consumer-quality pill images captured using commonly available digital cameras and smartphones. After successful segmentation of the pill, in future work, features like shape, imp- print and color will be extracted. These features help is comparing, correlating and ranking the consumer- quality images using the high-quality reference images.

2 METHODS

The main objective of the paper is to segment consumer-quality pill images which are affected by background noise and shadows. Once the pill is isolated, feature extraction is more reliable.

The proposed algorithm initially smoothes the image to reduce noise using a Gaussian smoothing filter. The simple linear iterative clustering (SLIC) algorithm (Achanta et al., 2012) algorithm is then applied to generate superpixels. The resultant image is con- verted into a region adjacency graph and thresholded to merge the superpixels. A final binary mask is obtained by thresholding color planes, applying an opening operation, filling holes, and applying a convex hull. A bounding box is applied to obtain only the segmented pill region.

2.1 SLIC Superpixels

The pre-segmentation of an image is a crucial step before applying region adjacency graphs. This step includes the generation of superpixels. Superpixels are a group of pixels which share similar characteristics with their neighboring pixels. They capture the image redundancy and subsequently reduce complexity in performing further image processing tasks. There are various approaches to generate superpixels (Boykov and Jolly, 2001)(Shi and Malik, 2000)(Comaniciu and Meer, 2002)(Felzenszwalb and Huttenlocher, 2004)(Achanta et al., 2012). This paper uses the SLIC algorithm to generate superpixels because it is faster, more memory efficient, and has better boundary adherence than its predecessors. A detailed step-by-step procedure of the SLIC algorithm is provided in Achanta et al (Achanta et al., 2012).

The pill image is initially pre-processed using a Gaussian smoothing filter with standard deviation 2. The SLIC algorithm, which generates superpixels based on k-means clustering (Kanungo et al., 2002),

is applied. The search space in the SLIC algorithm is limited to a specific region around a cluster centroid. This reduces the number of distance calculations which in turn reduces the complexity and run time. Also it considers a weighted distance approach by combining both color and spatial proximity. These features allow the algorithm to outperform existing state-of-the-art superpixel methods. The search is done for 10 iterations after initializing the cluster centroids. This generation of superpixels may be regarded as an over-segmentation process.

The output is a labelled image, as the algorithm assigns a unique label for each superpixel. An average color value of all pixels in a superpixel is calculated and assigned to the respective superpixel as shown in Figure 3.

Formally let, μ_R denote the mean of a set of colors $\mathbf{p}_0$, $\mathbf{p}_1$,..., $\mathbf{p}_N$ in region R, as given by equation (1):

$$\mu_R = \frac{1}{N}\sum_{i=0}^{N} p_i \tag{1}$$

where N is the total number of pixels in that region.

2.2 Region Adjacency Graph

A region adjacency graph (Tremeau and Colantoni, 2000) is created as a step towards merging of superpixels. The initial pre-segmentation, that is, initial generation of superpixels is crucial to create an associated adjacency graph. There is no loss of visual information in the pre-segmentation process. Pixels are only merged if they belong to same superpixel region.

The over-segmented image is now considered as a graph. The centroid of each superpixel in the image is a node in the graph. All nodes in the adjacent regions are joined to form an edge as shown in Figure 4.

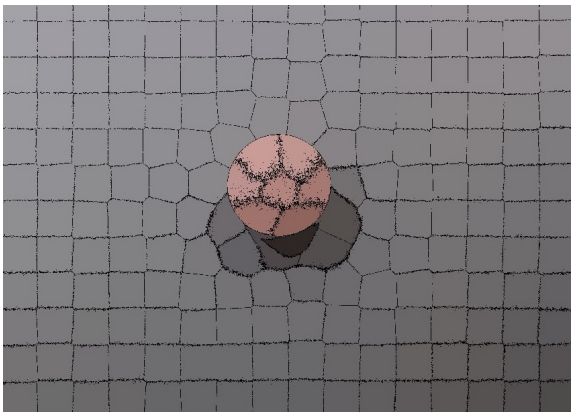

Figure 3: Pill segmented with superpixels with compactness factor = 12.

This collection of edges is called the region adjacency graph.

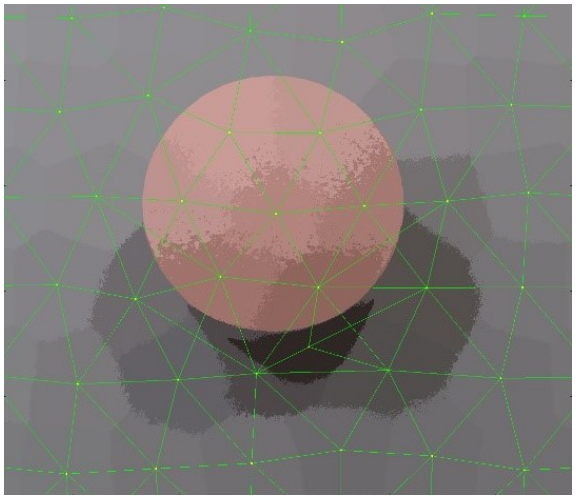

Figure 4: Labelled image (zoomed) with region adjacency graph, showing edges as lines.

The weight for the edge between two adjacent nodes (van der Walt et al., 2014) can be defined in various ways. The superpixels can be merged using these edge weights. As each superpixel is of uniform average color, the edge weights are defined by the difference of average color between the adjacent superpixel regions. The regions connected with a lower edge weight have similar color features and were merged using a threshold value of 29, empirically determined from a dataset of 30 random images from the provided consumer quality images. The adjacent superpixel regions are merged if the edge weight is lower than the pre-determined threshold value; if the edge weight is higher than the threshold value, the graph is cut as shown in Figure 5.

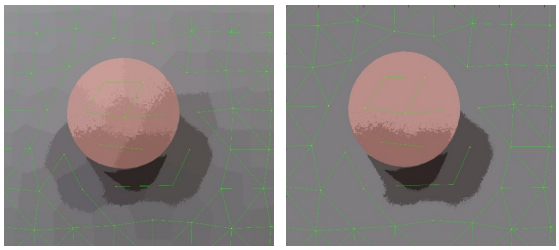

Figure 5: (a) Superpixels with graph cut (zoomed).
(b) Merged regions with graph cut (zoomed).

As a result, a fully connected region adjacency graph (RAG) is divided into disconnected regions with threshold-cuts as shown in Figure 6. The pixels of newly generated regions are assigned to the average color value of the merged regions. This reduces segmentation complexity substantially and results in

easier generation of the pill mask.

2.3 Post-processing

The image resulting from merging superpixel regions by RAG thresholding is still affected by the shadows of the pill. The outer shadow needs to be merged with the background and the inner shadow shall be merged with the object (pill).

When background color intensity is close to the Pill color intensity, segmentation errors occur upon merging. To overcome this problem, a histogram of the image resulting after RAG thresholding is plotted as shown in Figure 7. Since, the background occupies most of the area in the image, the majority of the pixels share the same intensity level as observed in Figure 7. The bin of the histogram with background pixels has the highest probability. All pixels sharing this most probable bin value are assigned to zero intensity. This eliminates the problem stated above.

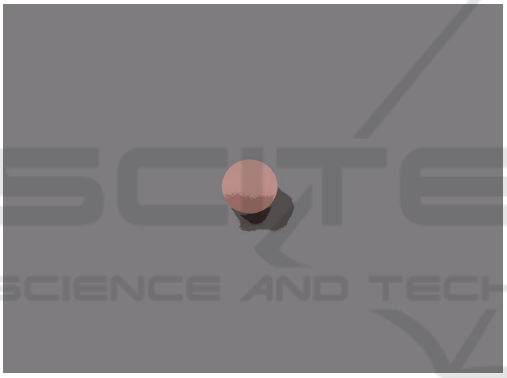

Figure 6: Superpixels merge using RAG.

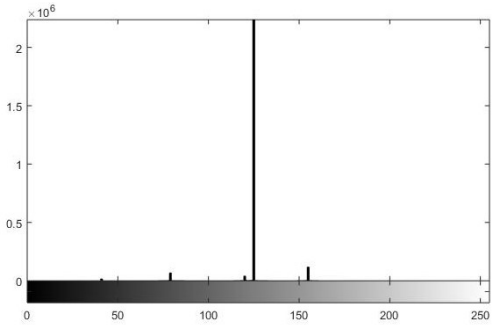

Figure 7: Histogram of image from figure 6.

On analyzing the color intensity values of various pill images, the red and blue planes contribute the majority of intensity changes from pill to its shadow. After reviewing 30 random consumer-quality images (previously used to determine the threshold for region connecting), threshold cutoff values of 105 and 83 were chosen for red and blue planes respectively. An

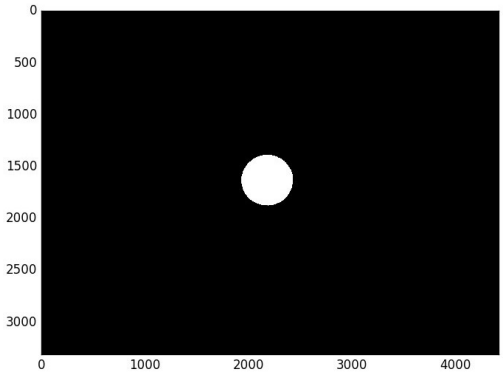

Figure 8: Binary Mask of the Pill.

OR operation is applied to masks from both planes to generate a single binary mask.

A morphological opening (erosion followed by dilation) is then done to remove blobs of radius less than 9 pixels with a circular structuring element. Any holes in the mask are filled with a flood fill operation. A final mask is generated by applying the convex hull operation on the filled mask as shown in Figure 8.

Figure 9: Boundary marked on the pill (zoomed).

A distinct boundary along the edges of the pill is shown in the overlay image for this mask, Figure 9. A bounding box is applied to this mask to obtain the pill region as shown in Figure 10.

Figure 10: Result of bounding-box.

3 EXPERIMENTAL RESULTS

The proposed pill segmentation algorithm showed favorable accuracy results for the 5000 consumer-

quality pill images provided by the NLM system. Since the algorithm uses a color segmentation approach, some of the pills with color similar to background color were completely merged with the background, resulting in a complete black mask. This is the primary limitation of the proposed algorithm.

The algorithm produced accurate segmentation results on the 2000 high-quality reference pill images as shown in Figure 12. These images are chosen as the benchmark for comparing the segmentation results of consumer-quality pill images.

The 5000 consumer-quality masked pill images were scored manually to analyze the accuracy of the segmentation with respect to segmentation of reference pill images. Results show accurate segmentation for 2243 pills, as shown in figure 11(left). For 1862 pills, some shadow is included along with the pill in the mask (Figure 11, center). The remaining pill images (17.9%) have false segmentation (Figure 11, right) due to the challenges mentioned above. In summary, the proposed algorithm produces acceptable segmentation accuracy for 82.1% of 5000 consumer-quality pills.

Figure 11: Bounding-box of segmented Consumer Pill Images.

Figure 12: Bounding-box of segmented Reference Pill Images.

The time taken to run the algorithm (written in python v2.7) on each pill image (of varying size with largest being 2400 x 1600) on Intel Core i5 2400 processor, 8 GBytes DDR3 RAM and 512 MBytes AMD RADEON HD 6350 graphics card is on average 683.95 seconds. In order to make the segmentation proceed faster, a scaling factor is introduced and applied to reduce and resize the input image. Also the number of superpixels, and the disk size for morphological operation are reduced as input image is scaled-down. But there is a trade-off with the quality of the mask generated as lower scaling factors are considered as shown in Figure 13. This is shown in Table

Figure 13: Segmentation results with scale factor 1(left), 0.4(center), 0.1(right).

1. The quality of generated binary masks on average is provided in Table 1 corresponding to 82.1% of consumer-quality images with acceptable segmentation accuracy. The quality of the binary mask produced from each of those images for varying scale factor ($i = 1.0, 0.9, 0.8, ..., 0.1$) are computed by equation 2.

$$Q_i = \left(1 - \frac{|p_i - p_{1.0}|}{p_{1.0}}\right) * 100 \qquad (2)$$

Where Q_i is the segmentation quality of the binary mask, p_i is number of pixels in the object region of binary mask and $p_1.0$ is the number of pixels in the object region of binary mask for a scale factor of 1.0. The speed factor, calculated as the ratio of the average run-time to process each image at a particular scale factor to that of the run-time to process the pill image with scale factor 1.0. To provide best segmentation results at a faster rate, a scaling factor of 0.4 is considered to be the optimum value upon reviewing all the image masks from the dataset.

Table 1: Effect of scaling factor on segmentation accuracy and speed factor for individual pills.

Scaling Factor	Speed Factor	Average Q value
1.0	1.00x	100%
0.9	1.11x	97.82%
0.8	1.66x	97.76%
0.7	1.95x	97.43%
0.6	2.93x	97.10%
0.5	4.12x	96.67%
0.4	**6.19x**	**98.46%**
0.3	10.30x	94.07%
0.2	19.08x	89.80%
0.1	40.30x	83.19%

4 CONCLUSIONS

The proposed method of merging superpixel regions using a region adjacency graph threshold-cut

approach to successfully segment consumer-quality pills with few limitations. Application of a resizing factor gave some promising results for algorithm speed, with a trade-off in quality of mask.

Although the process has eliminated the background noise and produced excellent results for most of the pills and capsules, the shadows caused by pill illumination is still a challenge for some pills. Pills with similar background color also pose a great challenge in boundary determination. Finding an adaptable solution that works for all 5000 pills is challenging. Further analysis needs to be done to get accurate segmentation for all the consumer-quality pills.

This project was originally developed as an entry to the Pill Image Recognition Challenge conducted by the National Library of Medicine. The 5000 consumer-quality image data- sets were accessed from the NLM database. Future work corresponds to extraction of various features that are crucial to match the given consumer-quality pill images to their reference images using rank scoring.

REFERENCES

Achanta, R., Shaji, A., and Smith, K. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *Pattern Analysis and Machine Intelligence*, 34(11):2274–2281.

Boykov, Y. Y. and Jolly, M. P. (2001). Interactive graph cuts for optimal boundary amp; region segmentation of objects in n-d images. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 105–112 vol.1.

C3PI (2016). Computational photography project for pill identification. https://lhncbc.nlm.nih.gov/project/c3pi-computational-photography-project-pill-identification. Last accessed on Aug 28, 2016.

Caban, J. J., Rosebrock, A., and Yoo, T. S. (2012). Automatic identification of prescription drugs using shape distribution models. In *2012 19th IEEE International Conference on Image Processing*, pages 1005–1008.

Comaniciu, D. and Meer, P. (2002). Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.

Dailymedplus (2016). Medicos consultants. http://www.dailymedplus.com. Last accessed on Aug 28, 2016.

Drugs.com (2016). Pill identifier. https://www.drugs.com/imprints.php. Last accessed on Aug 28, 2016.

Felzenszwalb, P. F. and Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., and Wu, A. Y. (2002). An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):881–892.

NLM (2016). National library of medicine : Pill image recognition challenge. http://pir.nlm.nih.gov/challenge/. Last accessed on May 31, 2016.

Pillbox (2016). Prototype pill identification system. http://pillbox.nlm.nih.gov/. Last accessed on Aug 28, 2016.

Shi, J. and Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905.

Tremeau, A. and Colantoni, P. (2000). Regions adjacency graph applied to color image segmentation. *IEEE Transactions on Image Processing*, 9(4):735–744.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., and the scikit-image contributors (2014). scikit-image: image processing in Python. *PeerJ*, 2:e453.

WebMD (2016). Pill identification. http://www.webmd.com/pill-identification/default.html. Last accessed on Aug 28, 2016.

Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678.