Abstract—Region segmentation and edge detection are standard image processing operations. Clustering can be used for region segmentation. However, often clustering results depend on the selection of various parameters, such as the number of clusters, or the clustering algorithm used. The framework presented here employs the result of edge detection on the original image, as well as on the clustering results of the same image, to automatically select (according to some agreement measure) the optimal number of clusters, and the corresponding (best) segmentation. The framework supports an extended pixel representation in which other information, such as texture, can be incorporated in addition to edge and region information. To illustrate this framework, the edge guided clustering algorithm presented here, uses the Canny edge detection approach to guide region identification through fuzzy k-means clustering. Experimental results on benchmark images for which manual segmentation is available as reference illustrate the effectiveness of this approach.

Index Terms—Image segmentation, fuzzy k-means clustering, canny edge.

I. INTRODUCTION

The goal of image segmentation is to partition an image into meaningful regions, i.e. regions corresponding to individual objects, or natural parts of objects. Segmentation could be used for tasks such as object recognition, occlusion boundary estimation within motion, image compression, image editing, or content-based image retrieval from large databases [1].

A. Background and Related Work

Several review papers point out that there are many segmentation approaches using various features extracted from the image, such as gray scale values, texture, and gradient [2], [3], [4], [5]. At an algorithmic level, the main approaches to segmentation are classified as non-contextual and contextual techniques [6]. The first category groups pixels together based on some similarity in features (e.g. similar gray scale value), whereas the second approach uses also relationships among features (such as closeness between pixels and boundary-based techniques).

The approach proposed here combines both techniques. It starts from the observation that boundary information and region information are, to some extent, complementary in nature. The boundary analysis relies on dissimilarity of neighboring pixels, whereas region detection techniques take advantage of certain similarities between pixels. Therefore, it is expected that, due to their complementarity, when combined these techniques will produce segmentation results which are better than those obtained by each of these techniques separately [3].

Considering the stage when the techniques are used, two main approaches are distinguished for segmentation methods integrating region and boundary information [3]:

- Embedded integration – when the edge information is used during the region finding step;
- Post-processing integration – when the two sets of information are developed in parallel and then are fused together.

Examples of post-processing integration include edge and region data [7], region growing guided by an evaluation function [8], constraint clustering approaches [9].

B. Current Approach

The framework presented here and illustrated in Fig. 1 is also a post-processing integration approach with the following features. Mainly, this approach uses the fuzzy k-means clustering for region detection, and edge detection to infer k, the (optimal) number of clusters. Optionally, it supports texture information and can isolate the foreground for better segmentation. To select k, the approach makes use of an evaluation component between edge detection from the original and clustered images. This component is implemented as a measure of goodness. Several such measures are considered here.

![Fig. 1. Segmentation using edge guided image clustering framework.](image)

The particular problem addressed in this research is automatic detection of various components within an image...
based on color information. Images are represented in the CIELUV color space [10], from which L, the luminosity space is used for edge detection. In this space the Euclidean distance is consistent with that perceived by the human eye.

C. Organization of this Paper

From this point on the paper is organized as follows: section II exemplifies a possible implementation of the framework illustrated in Fig. 1, namely the edge guided clustering algorithm (EGC), section III shows experimental results of the EGC algorithm, and section IV concludes this research and points out possible extensions.

II. THE EDGE GUIDED CLUSTERING ALGORITHM (EGC)

One of the main difficulties in using clustering in automatic image segmentation is the selection of number of clusters. In the current framework the number of clusters that maximizes the agreement between segmentation and the output of an edge detection operator is selected. For illustration purposes the work presented here uses the fuzzy k-means clustering algorithm [11], [12] and the Canny edge detection method [13]. The resulting EGC algorithm is described in Fig. 2. It is important to point out from the start that any other choices of clustering and/or edge detection approaches fit into this framework.

As shown in Fig. 2, the EGC algorithm consists of three major stages, as follows.

A. Stage I: Preprocessing

First, the RGB image data is mapped into CIELUV space. Optionally at this stage, the foreground is isolated from the background and the corresponding image is passed on to the subsequent stages of the algorithm. Foreground isolation works well for images where the background is connected to the margins of the image (e.g. nature images). However, this operation may not be needed for images with no background. Initial results (see section III) show great improvement on segmentation when the background is isolated. In particular on images from nature, where the background might consist of many shades of same color (such as blue sky or green grass), eliminating the background improves the chances of the clustering algorithm to find meaningful clusters that correspond to objects in the image.

The background segmentation approach used here selects as background that portion of the image with the smallest average minimum distance to the image margins. Fuzzy k-means algorithm, with various values of k∊K={1, 2, 3, ...}, is used to generate background candidates. More precisely, background segmentation is obtained as follows:

For every pixel \(x\) in the image the distance \(d_m\) to the image margin is defined by

\[d_m = \min\{d(x, m_i), i = 1, 2, 3, 4\}\]

where \(m_i\) \(i = 1, 2, 3, 4\) denote the four margins of the image.

For fixed \(k\) \(\in K\) denote by \(C_{j,k}\) the cluster \(j\) obtained from the corresponding fuzzy k-means algorithm with Euclidean distance between the pixels. For each cluster \(C_{k,j}\) denote by \(d_{k,j}\) the cluster average distance to the margin defined as:

\[d_{k,j} = \frac{\sum_{x \in C_{k,j}} dm(x)}{|C_{k,j}|}, \quad j = 1, \ldots, k\]

where \(|C_{k,j}|\) is the size (number of pixels) of \(C_{k,j}\).

The image background is selected as the cluster \(C_{k_0,j_0}\), where

\[j_0 = \arg \min \{d_{k,j} | j = 1, \ldots, k\}\]

\[k_0 = \arg \min \{d_{k,j_0} | k \in K\}\]

\[j_0 = j_{k_0}\]

In other words the background corresponds to the \(j_0\) cluster generated by the fuzzy k-means with \(k = k_0\). In the experiments carried here \(K = \{2, 3, 4\}\).

B. Stage II: Edge Detection and Segmentation

In this step the boundary and region information are extracted. For boundary, the Canny method [13] finds edges by looking for local maxima of the gradient in a gray scale image. The gradient is calculated using the derivative of a Gaussian filter. The method uses two thresholds, to detect strong and weak edges, and includes the weak edges in the output only if they are connected to strong edges. This
method is therefore less likely to be affected by noise, while at the same time more likely to generate meaningful true weak edges.

The fuzzy k-means algorithm [11] is used to compute the regional information, but other unsupervised clustering methods, or even other region detection techniques can be used at this stage. The fuzzy k-means algorithm has the advantage that clusters are more robust to variations of pixel values reflecting variations within the same color. Moreover, because each pixel belongs to more than one of the k clusters (with various degrees in [0, 1]), the membership values can be further used to generate different region areas.

Boundary information, the set of edges resulted from the Canny method denoted by CE L in Fig. 1, guides the fuzzy k-means using a measure to evaluate the agreement between CE L and CE C, the result of Canny edge detection on clusters. Measures to evaluate the agreement between different segmentations have been extensively studied. Many of these measures are based on the Rand index [14] and its various extensions [15], or measures of errors [15]. The current work uses measures of overlap between CE L and CE C, relative to CE Lk, CE Ck or their union, for various values of k. For a binary image I (the result of the edge detection), the overlap \( O(A, B) \) of subsets A and B of I is defined as

\[
O(A, B) = \sum_{x \in I} I_{X_{A \cap B}}(x)
\]

That is, \( O(A, B) \) counts the common pixels of A and B.

Next, for various values of k the following measures are defined:

\[
M^k_1 = \frac{O(CE_L, CE_C)}{\sum_{x \in I} I_{\frac{CE_L}{C}}(x)}
\]

\[
M^k_2 = \frac{O(CE_L, CE_C)}{\sum_{x \in I} I_{\frac{CE_C}{C}}(x)}
\]

\[
M^k_3 = \frac{O(CE_L, CE_C)}{\sum_{x \in I} I_{\frac{CE_L \cup CE_C}{C}}(x)}
\]

In the experiments carried out here \( K = \{2, \ldots, 20\} \). \( M^k_1 \) and \( M^k_2 \) measure the degree of inclusion between CE L and CE C, respectively, while \( M^k_3 \) quantifies the degree of their intersection relative to their union. Since \( M^k_1 \) and \( M^k_2 \) are not symmetric, a symmetric version, \( M^k_s \) is obtained by applying a symmetric binary operator to them. Here they are multiplied. That is,

\[
M^k_4 = M^k_s \times M^k_2
\]

C. Stage III: Segmentation Selection

Any measure \( M^k_s, i=1, \ldots, 4 \) can be used to determine the optimal value for k, the parameter of the fuzzy-k algorithm. More precisely,

\[
k^*_o = \arg \max \{ M^k_s | k = 2, \ldots, 20 \}
\]

That is, for each measure of agreement the optimal k is considered to be that for which the edge sets produced by Canny edge detector from the original image and from the clustered image agree the most.

D. Incorporating Texture into the EGC Framework

For some types of images (e.g. natural scene images) texture information can aid the segmentation process. The framework presented here can easily support the additional texture information, as shown in Fig. 1. While there are many texture operators, here the standard deviation calculated for each pixel within a vicinity of 11x11 and 3x3 pixels is used. Pixel representation is extended by adding texture information as another (4th) dimension. The EGC algorithm with texture is referred here as EGC-Tex.

III. EXPERIMENTAL RESULTS

Experiments with the framework described above were carried out on 40 images from The Berkeley Segmentation Dataset and Benchmark [16]. However, due to space limitation, the results included in this section illustrate a subset of representative (with various degrees of difficulty) images.

How good are these results? As it is well known, and often pointed in the literature, there are no unified evaluation methods for image segmentation algorithms [16] and [17]. This is due to the fact that segmentation evaluation is usually application dependent (e.g. context-based image retrieval, medical field, etc.). Yet, here, the objective is to obtain an approach which is as much as possible application independent. The segmentation results are evaluated relative to the manually segmented images obtained from [16] using the corresponding measures of edge agreement from the previous section.

In the results discussed next, the EGC segmentation algorithm with background/foreground segmentation step using \( M^k_1 \) measure is used by default, unless it is otherwise specified. Also it must be mentioned that after segmentation, each pixel is colored with the color obtained by averaging the colors of all pixels in its cluster. Note that here a cluster can consist of several spatially disjoint regions within the image (e.g. the yellow cluster in the Two flowers image consists of three regions visible in Fig. 5).

A. Canny Edge, Clustering Edge, and the Manual Segmentation

Fig. 3 shows in the top right plot the edges detected by the Canny algorithm, while the middle and bottom rows show the edges obtained by the k-mean clustering for k=3 with four EGC versions.

The first image in the middle row illustrates the EGC version with no texture information, while the second image uses the texture as the standard deviation in an 11x11 neighborhood. Similarly, the bottom row shows the no texture and texture EGC variations with “no background detection” version of the EGC (i.e. step two from the first stage of the EGC algorithm shown in Fig. 1 is skipped). The edges common to the clustering and Canny edge mapping are shown in a brighter white contour; the darker contour
corresponds to cluster edges not matched by the Canny edge mapping. As it can be seen in this figure, the Canny and k-means edges do not fully overlap, hence the necessity to compute the agreement between the two sets of edges, and further to maximize this agreement over all clusters within an image. Also, these figures illustrate how different EGC versions lead to different clustering and hence, to different final segmentations. As it will be illustrated later, the EGC with background selection (and sometimes with texture information) leads to better segmentation.

Fig. 4 shows the original seven images selected to illustrate the EGC segmentation algorithm, their corresponding manual segmentations [16], and the best segmentation found by the EGC algorithm.

B. EGC and Goodness Evaluation

Fig. 5 shows the segmentation results produced by the EGC in conjunction with the four measures defined in section II B. The authors’ more are less subjective evaluation declares M1 and M3 as better measures to guide the “best” segmentation (i.e. clustering). They rank the goodness of these measures as M1, M3, M4, and M2.

A less subjective analysis of the four evaluation methods uses similarity (as number of clusters) between the clustering obtained from EGC with Canny edge (CE) and the clustering obtained from EGC with Canny applied to manually segmented edge (referred here as ground truth, GT). This analysis is based on the data from Table I where, the optimal (according to the four measures) values for k in the fuzzy k-means clustering for CE and GT are listed. The table also shows, for each of the four measures, the mean and the standard deviation of the difference between the number of clusters k inferred from each category of images (GT and CE). The instances where the k values are the same for a given measure are shown in bold.

Intuitively, a smaller mean and standard deviation correspond to a better evaluation measure. For example, when all k values for GT and CE coincide, the mean and standard deviation of their difference will be zero. This observation, in conjunction with the data from Table I, leads to the following ranking of evaluation measures: M1, M3, M4, and M2. This is an interesting result which agrees with the human perception and interpretation of the segmentations from Fig. 5.

C. Variations of the EGC Algorithm

Fig. 6 contrasts four variants of the basic EGC algorithm: EGC with background segmentation step (second row), EGC with no background segmentation (third row), EGC-Tex with background segmentation (fourth row), and EGC-Tex with no background segmentation. The EGC algorithm with background segmentation consistently gives better results in all the images tested (recall that not all results are included here). This is partially due to the fact that variations within the background influence (and change) much similar (color-wise) clusters in foreground. So this step is almost always useful, however, the background/foreground segmentation can be tricky. That is, one must be sure that the true background is selected; otherwise all the subsequent steps are meaningless. The background selection method proposed in section II A works for all the images tested here.

The texture addition used here, i.e. the standard deviation within a neighborhood, does not improve the segmentation as much as the background/foreground selection step. However, other textures may boost more clearly the segmentation. For example, a pixel’s closeness to one or more edges might be more meaningful. An example of a (possibly) better segmentation due to texture information is the horse image (third image in the fourth row), where the shades on the horse are visible; in the other segmentations of this image those shades are not captured. This is deemed a “possibly better” segmentation, because the evaluation in image segmentation is always context-dependent: for some applications the detection of the horse is sufficient, whereas other applications may require a finer reconstruction of the original image.

IV. CONCLUSION

The edge guided clustering framework for segmentation proposed here leads to an automatic segmentation approach which uses boundary detection for segmentation selection. The approach is automated in the sense that it does not depend on a threshold selection as many segmentation algorithms do [5]. Rather, boundary information serves to automatically select the clustering that gives the best segmentation, according to some agreement measure between the boundary and the clustering information. The framework also supports texture information and various evaluations of goodness of segmentation relative to manual segmentation.

This approach is illustrated on several images through the edge guided clustering algorithm. The experimental results show that separating the background and continuing the analysis on the foreground only gives, in general, better segmentations.

Investigating different color spaces (e.g. HSV or RGB), various texture information, and different distances within the fuzzy k-means clustering (here the Euclidean distance was used) are considered for future work. Also, further use of pixel membership degrees obtained with fuzzy k-means clustering may lead to better segmentations of some images. For instance, only the pixels with a membership degree larger than 0.5 should be labeled as belonging to a given class; the remaining of the “weaker” pixels could be used for another round of processing that may use spatial information. Finally, comparing the EGC algorithm with other segmentation approaches and a more formal evaluation for assessing the segmentation quality are also important directions for future study.
TABLE I: COMPARISON BETWEEN THE NUMBER OF CLUSTERS OBTAINED THROUGH EGC GUIDED BY MANUAL SEGMENTATION (GT) AND CANNY OPERATOR (CE) THE SHADED COLUMN, CORRESPONDING TO MEASURE M1 CONTAINS THE BEST RESULTS, IN THE SENSE OF HIGHEST AGREEMENT BETWEEN THE CANNY EDGES FROM THE MANUALLY SEGMENTED IMAGE AND THE CLUSTERED IMAGE.

<table>
<thead>
<tr>
<th>Original Images</th>
<th>Number of clusters, k, determined by each evaluation measure guided by the manual segmentation (GT) and Canny operator (CE)</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M1xM2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GT</td>
<td>CE</td>
<td>GT</td>
<td>CE</td>
<td>GT</td>
</tr>
<tr>
<td>Starfish</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Acropolis</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Rhinoceros</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Horses</td>
<td>10</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Two red flowers</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Two Asian faces</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Building</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

Difference between GT and CE guided clustering

<table>
<thead>
<tr>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/7=2</td>
<td>16/7=2.28</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.83</td>
</tr>
</tbody>
</table>

Mean x Standard Deviation

<table>
<thead>
<tr>
<th>Mean x Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.66</td>
</tr>
</tbody>
</table>

Fig. 3. Illustration of canny edges (top right figure) and clustering edges using EGC with background selection (first figure, second line), EGC-Tex with background selection (second figure, second line), EGC with no background selection (first figure, bottom line), and EGC-Tex with no background selection. The edges common to the clustering and Canny edge are shown in a brighter white contour; the darker contour corresponds to cluster edges not matched by Canny edge.
Fig. 4. Original images, their corresponding manual segmentations, and the best EGC segmentation (EGC with background and measure M1).
<table>
<thead>
<tr>
<th>ORIGINAL IMAGE</th>
<th>M₁</th>
<th>M₂</th>
<th>M₃</th>
<th>M₄=M₁×M₂</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Original Image" /></td>
<td><img src="image2" alt="Segmentation M₁" /></td>
<td><img src="image3" alt="Segmentation M₂" /></td>
<td><img src="image4" alt="Segmentation M₃" /></td>
<td><img src="image5" alt="Segmentation M₄" /></td>
</tr>
<tr>
<td><img src="image6" alt="Original Image" /></td>
<td><img src="image7" alt="Segmentation M₁" /></td>
<td><img src="image8" alt="Segmentation M₂" /></td>
<td><img src="image9" alt="Segmentation M₃" /></td>
<td><img src="image10" alt="Segmentation M₄" /></td>
</tr>
<tr>
<td><img src="image11" alt="Original Image" /></td>
<td><img src="image12" alt="Segmentation M₁" /></td>
<td><img src="image13" alt="Segmentation M₂" /></td>
<td><img src="image14" alt="Segmentation M₃" /></td>
<td><img src="image15" alt="Segmentation M₄" /></td>
</tr>
<tr>
<td><img src="image16" alt="Original Image" /></td>
<td><img src="image17" alt="Segmentation M₁" /></td>
<td><img src="image18" alt="Segmentation M₂" /></td>
<td><img src="image19" alt="Segmentation M₃" /></td>
<td><img src="image20" alt="Segmentation M₄" /></td>
</tr>
<tr>
<td><img src="image21" alt="Original Image" /></td>
<td><img src="image22" alt="Segmentation M₁" /></td>
<td><img src="image23" alt="Segmentation M₂" /></td>
<td><img src="image24" alt="Segmentation M₃" /></td>
<td><img src="image25" alt="Segmentation M₄" /></td>
</tr>
<tr>
<td><img src="image26" alt="Original Image" /></td>
<td><img src="image27" alt="Segmentation M₁" /></td>
<td><img src="image28" alt="Segmentation M₂" /></td>
<td><img src="image29" alt="Segmentation M₃" /></td>
<td><img src="image30" alt="Segmentation M₄" /></td>
</tr>
</tbody>
</table>

Fig. 5. Comparison of best segmentations (using EGC with background detection) obtained with M₁, M₂, M₃, and M₄ measures. First column shows the original image.
Fig. 6. Segmentation results contrasting EGC with background detection (second row) vs. EGC with no background detection (third row) vs. EGC-Tex with background detection (fourth row) vs. EGC-Tex with no background detection (bottom row).

REFERENCES


Sofia Visa was born in Romania and received her BS in Mathematics and Computer Science from L. Blaga University, Sibiu, Romania in 1998, and her MS and PhD in Computer Science from the Electrical & Computer Engineering and Computer Science Department (ECECS), University of Cincinnati in 2002 and 2007, respectively. She is an Assistant Professor in the Mathematics and Computer Science Department at the College of Wooster, Wooster, OH. Before coming to Wooster, she was an Assistant Professor in the Computer Science Department at Saint Anselm College. Her research interests are in the area of machine learning including issues related to data classification, learning from imbalanced data sets, applications of learning algorithms in real world problems, and bioinformatics.

Mircea Ionescu was born in Romania and received his BS and MS in Electrical Engineering, Polytechnic University of Bucharest in 1996 and 1997, respectively. In 2006 he completed his PhD in Computer Science from the Electrical & Computer Engineering and Computer Science Department (ECECS), University of Cincinnati. He is a senior Computer Vision Scientist at Charles River Laboratories, Seattle, WA. Prior to this, he worked as Software Developer Engineer at several companies in Romania and Germany. He also had teaching appointments as Adjunct Faculty member in the Mathematics Department at University of Cincinnati and as part time Lecturer at the College of Wooster.

His areas of expertise are machine learning and computer vision. His main focus is on developing computer vision algorithms for biological images.