Sequential Possibilistic Local Information One-Means Clustering For Image Segmentation

Wenlong Wu, James M. Keller
Electrical Engineering and Computer Science Department
University of Missouri-Columbia
Columbia, MO 65211, USA
ww6p9@mail.missouri.edu, kellerj@missouri.edu

***Abstract*—Clustering has long been applied to the problem of image segmentation. Because of spatial connectivity constraints, several approaches have been proposed to incorporate local consistency into image segmentation by clustering. One popular method, the fuzzy local information c-means (FLICM) has been shown to produce good segmentation results. Like the fuzzy c-means (FCM) from which it is derived, FLICM requires that pixels "share" memberships across clusters, that is, the memberships of a pixel across all clusters need to sum to one. The possibilistic c-means (PCM) clustering was introduced to relax the membership sum-to-one constraint of the FCM, and has found a place in the clustering universe, particularly in those situations where the data contains outliers, is noisy, or highly overlapped. This paper extends the structure of FLICM to possibilistic versions for image segmentation. Two approaches are proposed. The first, called possibilistic local information c-means (PLICM) inserts the local information term of FLICM into the basic PCM model. PLICM, like PCM, can produce coincident cluster centers. Recently, a sequential application of PCM (with c = 1) has been developed to mitigate negative effects of the co-incident cluster formation. Three algorithms form the family of sequential possibilistic 1-means (SP1M). These algorithms are extended to account for local information in image segmentation (SPLI1M). After development of the approach, experiments are performed on images which show that the SPLI1M family has superior performance in image segmentation over FLICM, PLICM and other clustering algorithms that don't combine local spatial information of the image.**

Index Terms*—*clustering, possibilistic c-means, image segmentation, spatial constraints

I. Introduction

Image segmentation is an essential topic in image processing and a classic difficult problem in the computer vision field. The objective of image segmentation is to set apart different components in an image so that each such component is consistent and separated from neighboring components. Many image comprehension algorithms in computer vision, such as object detection and feature extraction, depend on the quality of image segmentation. Although numerous image segmentation methods have been proposed [1, 2, 3], a universal and practical approach to segment images has not been found. Segmentation remains a choke point in computer vision [4].

Clustering algorithms are powerful and useful techniques and have been popularly used in artificial intelligence. Cluster analysis research has a long history and has been combined with several research tools, including neural networks and probabilistic reasoning. Cluster analysis has a prominent role in data mining, pattern recognition and other search-related techniques, which helps researchers find inner structure of data. It has become a vital methodology in the discovery of data distribution and underlying patterns. A fundamental aspect of any clustering algorithm is its ability to construct meaningful partitions that organize objects from a dataset based on specific criteria. A clustering algorithm can also explore and discover as it uses its connectivity and density functions. It can define multiple-level granularity structures and hypothesize models for each cluster, then find the best fit of each model as it relates to another.

K-means clustering (here called the hard c-means) and mean shift clustering are viable and popular tools in image segmentation [5, 6]. In an effort to recognize the uncertainty in image segmentation, the fuzzy c-means (FCM) [7] was introduced for image segmentation [8], and it has robust characteristics for points with ambiguous characteristics. However, the FCM algorithm is sensitive to images with noise or outliers due to the membership sum-to-one constraint [9]. Furthermore, the FCM algorithm performs clustering in feature space, for example 3D color space, which sometimes leads to a failure in image segmentation due to the lack of using local spatial position information in the image.

Many clustering techniques consider a pixel of an image as an isolated element in image segmentation, and so these methods are vulnerable in a noisy environment. Hence, spatial clustering is incorporated to overcome this problem, because it exploits the spatial relationship between the pixels [10, 11]. The fuzzy local information c-means (FLICM) [12] was constructed to utilize local spatial information in the original image to improve the performance of image segmentation by FCM. FLICM does clustering in both feature space and in the original image spatial domain, which is an improvement for clustering algorithms in image segmentation. However, FLICM has to be run many times with different values of the cluster number 'c' to get the best number of clusters for images where the number of clusters is unknown.

The possibilistic c-means (PCM) [13] is a generalization of the FCM, which abandons the membership sum-to-one constraint in FCM. It forms a "typicality" partition of the data, and is more robust to outliers than either the crisp or fuzzy c-means. Due to the formulation of the PCM, a meaningful result can be obtained even when $c = 1$. The sequential possibilistic one-means (SP1M) [14, 15, 16] is a sequential version of PCM that iteratively hunts for one possibilistic cluster at a time. There is an important parameter, *eta*, for each cluster, in both PCM and SP1M that must be specified or estimated. SP1M with dynamic eta [16] was proposed to pick *eta* dynamically, which has been shown to be more robust than the original versions [14, 15]. PCM, and hence SP1M, are mode-seeking algorithms in clustering. Since there are no constraints between cluster centers, they have the property that coincident clusters can form, i.e., two cluster centers migrate to the same place. This might be a problem (they could miss some less dense clusters) or can be an advantage in that they can naturally help determine how many clusters a dataset may have. In this paper, we combine the measure of local spatial information of the image from FLICM with PCM and SP1M to derive two novel clustering algorithms in image segmentation: the possibilistic local information c-means (PLICM) and sequential possibilistic local information one-means (SPLI1M).

The rest of the paper is organized as follows. Sections II briefly reviews FCM, PCM, SP1M, and FLICM. Section III introduces PLICM and SPLI1M respectively. Section IV represents our experimental study, and Section V summarizes our conclusions and planned future work.

II. PRELIMINARY THEORY

II.A. Fuzzy C-Means

The fuzzy c-means (FCM) [7] algorithm is defined as the minimization of the objective function

$$J_{FCM}(U,V;X) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}||v_i - x_k||^2 \quad (1)$$

with the constraint

$$\sum_{i=1}^{c} u_{ik} = 1 \quad (2)$$

for all $k = 1, \ldots, n$. Here, the fuzzifier parameter $m \in (1, \infty)$. Optimization of the FCM model is performed by randomly initializing V and then alternately updating U and V using the necessary conditions for the extrema of J_{FCM},

$$u_{ik} = \left(\sum_{j=1}^{c}\left(\frac{||v_i - x_k||^2}{||v_j - x_k||^2}\right)^{\frac{1}{m-1}}\right)^{-1} \quad (3)$$

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} x_k}{\sum_{k=1}^{n} u_{ik}^{m}} \quad (4)$$

This alternating optimization algorithm is run until a suitable termination criterion holds, for example when successive estimates of V change less than a threshold ε,

$$\max_{i=1,\ldots,c} ||v_i - v_i'|| < \varepsilon \quad (5)$$

In this paper, we will always consider the Euclidean distance $||\cdot||$, the fuzziness index $m = 2$, and the threshold $\varepsilon = 0.01$. FCM can work well on many general datasets but does not perform well on outliers due to the constraint (2). Therefore, outliers can have a significant effect on the cluster centers found by FCM.

II.B. Possibilistic C-Means

In the hard c-means and even the FCM, outliers have relatively high membership due to the membership sum-to-one constraint. Actually, it is not uncommon for FCM to have a cluster center "stuck" at an outlier. One early attempt to ameliorate this problem was to include a noise cluster [17]. There, all points are considered to be equidistant from an amorphous noise cluster, and this distance is large compared with the distances of the "good" points to cluster prototypes. Therefore, noise points are attracted to the noise cluster. The membership values are still constrained as with FCM, and produce asymmetric cluster membership functions. Of course, there is also the question of what distance to use. The possibilistic c-means (PCM) [13] is more robust against outliers or noise because it abandons the membership sum-to-one constraint required in FCM to generate non-trivial necessary conditions for minimization. Outliers or noise have a large distance to all the existing clusters so that they naturally have low typicality to all clusters. Typicality is used in PCM instead of membership, because it measures how "typical" a particular point is to each cluster, that is, how close the point is to each cluster prototype. In PCM, the trivial solution $u_{ik} = 0$ is avoided by adding an additional penalty term in the objective function:

$$J_{PCM}(U,V;X) = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ik}^{m}||v_i - x_k||^2 + \sum_{i=1}^{c}\eta_i \sum_{k=1}^{n}(1 - u_{ik})^m \quad (6)$$

The task is to minimize (6) subject to

$$u_{ik} \in [0,1] \text{ for all } i \text{ and } k$$

$$0 < \sum_{k=1}^{n} u_{ik} < n \text{ for all } i$$

$$\max_{i} u_{ik} > 0 \text{ for all } k$$

The typicality update (necessary condition) in PCM is:

$$u_{ik} = \frac{1}{1 + \left(\frac{d_{ik}^2}{\eta_i}\right)^{\frac{1}{m-1}}} \quad (7)$$

The parameter d_{ik} denotes the distance between k^{th} data point and i^{th} cluster center. The parameter η_i determines the distance at which the typicality degree equals 0.5. As was suggested in [13], η_i can be determined as follows:

$$\eta_i = P\frac{\sum_{k=1}^{N} u_{ik}^{m} d_{ik}^2}{\sum_{k=1}^{N} u_{ik}^{m}} \quad (8)$$

Usually, P is set as one. The cluster center update is still the same as the cluster center update in (4).

PCM performs well against outliers or noise, but it may yield coincident cluster centers [18, 19]. Some research has been done with PCM to deal with the coincident cluster center issue, such as AM-PCM [20] and PFCM [21]. As stated above, this property of PCM can be a problem at times, or can actually be a positive attribute in case the number of actual clusters in the data is less than c [22].

II.C. Sequential Possibilistic One-Means

A major reason why PCM may find coincident clusters is due to initialization issues (see Fig. 2). While a heuristic rule is to always start with more clusters than you expect to find (and merge or identify coincident clusters), a more direct approach was created with the sequential possibilistic one-means (SP1M) family of algorithms [14, 15, 16]. The PCM algorithm has the characteristic that it can produce a meaningful typicality partition when $c = 1$ [23]. Hence, the SP1M algorithms overcome the coincident cluster drawback by generating one cluster at a time using P1M and stopping when all the "dense" regions are found.

The pseudocode for the latest version of SP1M with dynamic eta is shown in TABLE I where X is the dataset, c is the input for cluster number, ε is the threshold, K is defined to be the number of points whose maximum typicality is smaller than 0.5. Note that the (*) details of dynamic η computation in TABLE I is discussed in [16].

TABLE I: SP1M PSEUDO CODE

Algorithm: SP1M pseudo code	
Input: X, c, ε	
Output: U: Final membership partition	
V: Final cluster center group	
Initialize U, V as empty	
Do {	
Repeat <**loop to find a suitable cluster**>	
Pick $v \in X$ with probabilities	(9)
Repeat <**loop to execute P1M**>	
Compute η (i) dynamically	(*)
Compute u (v, X)	(7)
Compute v (u, X)	(4)
Until termination	(5)
Until termination	(10)
Append u to U	
Append v to V	
} While ($i++< c$ && $\#(P1M) < K$)	

In SP1M, the cluster centers are not initialized purely randomly. They are initialized from probabilities based on the typicalities of the previously found clusters. The initial cluster centers are picked from dataset X with probabilities

$$p(x_k) = \begin{cases} \frac{1}{n} & if\ i = 1 \\ 0 & if\ \max\limits_{j=1,\ldots,i} u_{jk} > 0.5 \\ \frac{1-\max\limits_{j=1,\ldots,i} u_{jk}}{n-\sum_{s=1}^{n}\max\limits_{j=1,\ldots,i} u_{js}} & \text{otherwise} \end{cases} \quad (9)$$

When P1M has found a new cluster center v, v may be very close to one of the cluster centers in V found so far. We only consider v a new cluster center if it has a distance of at least 2η from each cluster center in V, that is, if

$$\min_{w \in V} \left||v - w|\right| \geq 2\eta \quad (10)$$

If this condition does not hold, then v is discarded and P1M runs again to find a non-coincident cluster center. SP1M terminates when c non-coincident clusters have been found. In [16], a termination criterion was inserted to stop the program to avoid being trapped in an endless loop when no more new clusters can be found. Note that $\#(P1M) < K$ means on the k^{th} of P1M, if the times of abandoning coincident cluster is greater than K, we will stop the program, where K is the number of points whose maximum typicality is smaller than 0.5. The points whose maximum typicality is larger than 0.5 are likely to be those points strongly identified in an already found cluster. The value of K decreases as each P1M runs.

II.D. Fuzzy Local Information C-Means

In [12], Krinidis and Chatzis proposed the fuzzy local information c-means (FLICM) clustering algorithm that incorporates local spatial information of an image in a novel fuzzy way. FLICM can overcome the disadvantages of FCM, and at the same time, it enhances the clustering performance for image segmentation.

In FLICM, a fuzzy local factor G_{ik} is introduced:

$$G_{ik} = \sum_{\substack{j \in N_k \\ k \neq j}} \frac{1}{d_{kj}+1} (1 - u_{ij})^m ||x_j - v_i||^2 \quad (11)$$

where the k^{th} pixel is the center of the local window, i is the reference cluster and the j^{th} pixel belongs in the set of the neighbors falling into a window around the k^{th} pixel. d_{kj} is the spatial Euclidean distance between pixels k and j, u_{ij} is the j^{th} pixel membership degree in the i^{th} cluster, m is the weighting exponent on each fuzzy membership, and v_i is the cluster center of cluster i.

By adding G_{ik}, FLICM incorporates local spatial information into its objective function, defined as follows:

$$J_m = \sum_{k=1}^{N} \sum_{i=1}^{c} [u_{ik}^m \left||x_k - v_i|\right|^2 + G_{ik}] \quad (12)$$

The two necessary conditions for J_m to be at its local minimal extreme, with respect to u_{ik} and v_i in the FLICM paper are given as follows:

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\left||x_k - v_i|\right|^2 + G_{ik}}{\left||x_k - v_j|\right|^2 + G_{ij}}\right)^{1/m-1}} \quad (13)$$

$$v_i = \frac{\sum_{k=1}^{N} u_{ik}^m x_k}{\sum_{k=1}^{N} u_{ik}^m} \quad (14)$$

However, the above two equations are not necessary mathematical conditions for J_m to be a minimum because G_{ik} is treated as a constant value with respect to u_{ik}, while it is actually not [24]. However, the above two equations are used as update equations, and are shown to give good segmentation results in [12], and this is even the recommendation of [24], where this issue was discussed.

III. Sequential Possibilistic Local Information One-Means

III.A. Possibilistic Local Information C-Means

Like FCM, clusters in FLICM are coupled with each other, and the membership sum-to-one constraint may still lead to a failure in image segmentation if the local spatial position information of specific noise pixels is ambiguous. In this section, the membership sum-to-one constraint in FLICM is abandoned, and the resultant algorithm is called the possibilistic local information c-means (PLICM).

In PLICM, the fuzzy factor G_{ik} is included in the objective function:

$$J_m = \sum_{k=1}^{N} \sum_{i=1}^{c} [u_{ik}^m \left||x_k - v_i|\right|^2 + G_{ik}] + \sum_{i=1}^{c} \eta_i \sum_{k=1}^{N} (1 - u_{ik})^m \quad (15)$$

where the k^{th} pixel is the center of the local window, i is the reference cluster, u_{ik} is the degree of typicality of the k^{th} pixel in the i^{th}cluster, m is the weighting exponent on each possibilistic membership, and v_i is the cluster center of cluster i. Here, η_i determines the zone of influence (or bandwidth) of the i^{th} cluster.

Following the reasoning in FLICM, the typicality update equation is obtained as follows:

$$u_{ik} = \frac{1}{1 + \left(\frac{\left||x_k - v_i|\right|^2 + G_{ik}}{\eta_i}\right)^{1/m-1}} \quad (16)$$

The cluster center update equation is the same as equation (14). The PLICM pseudo code is shown in TABLE II.

TABLE II: PLICM PSEUDO CODE

Algorithm: PLICM pseudo code

Input: X, c, ε
Output: U: Final membership partition
V: Final cluster center group
Initial the fuzzy partition matrix U with FLICM
Repeat
  Compute membership values v_i using (14)
  Calculate the cluster center u_{ik} using (16)
Until convergence $v_i^{t+1} - v_i^t < \varepsilon$

Like PCM, PLICM still faces the coincident cluster center issue, which may leave out some critical information in an image. Thus, the initialization of PLICM is of vital importance. Initializing cluster centers with FLICM is a good choice for PLICM. This is similar to the recommendation of initializing PCM with FCM [13, 19].

We note that in [25, 26], a combined possibilistic fuzzy local information c-means was developed that produced both typicalities and a membership soft partition. The local information component was associated with the membership partition and not the typicalities. Typicalities are used mainly to flag outliers in synthetic aperture sonar images of the sea floor. Here, we investigate the utility of adding local information directly into the possibilistic clustering criteria function.

III.B. Sequential Possibilistic Local Information One-Means

In [13], PCM is computed by running the possibilistic one-means (P1M) c times sequentially. Similarly, PLICM can also be run sequentially. In both FLICM and PLICM, we update v_i by only using u_{ik}, independent of $u_{jk}, j = 1, \ldots, c, j \neq i$. In FLICM we update u_{ik} using all $v_j, j = 1, \ldots, c$, so the clusters are coupled to each other. Hence, PLICM can also be computed for $c = 1$. Anytime $c > 1$, even when $c \geq n$, PLICM can be executed by running c instances of PLI1M, independent of each other, sequentially or in parallel.

Hence, SP1M is combined with the local spatial information in the image into an algorithm called the sequential possibilistic local information one-means (SPLI1M). The basic idea is to run possibilistic local information one-means sequentially until all locally consistent clusters (segments) in the image are found.

The SPLI1M pseudocode is shown in TABLE III. This algorithm is developed based on the improved version SP1M in [16].

TABLE III: SPLI1M PSEUDO CODE

Algorithm: SPLI1M pseudo code

Input: X, c, ε
Output: U: Final membership partition
V: Final cluster center group
Initialize U, V as empty, $i = 1$
Do {
  Repeat **<loop to find a suitable cluster>**
    Pick $v \in X$ with probabilities (9)
    Repeat **<loop to execute PLI1M>**
      Compute η (i) dynamically (*)
      Compute u (v, X) (16)
      Compute v (u, X) (14)
    Until termination (5)
  Until termination (10)
  Append u to U
  Append v to V
} While ($i++ < c$ && $\#(P1M) < K$)

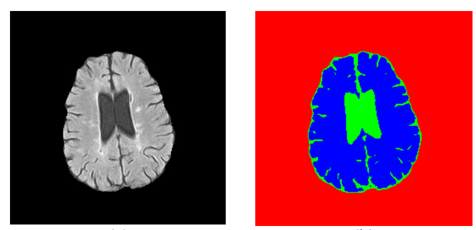
Note that the (*) details of dynamic η computation in TABLE III is discussed in [16].

IV. EXPERIMENTS

In this section, we present four experiments to compare FLICM, PLICM, SPLI1M and other clustering algorithms that don't combine local spatial information of the image. These four experiments use three images for comparison. One image is a medical image from the original FLICM paper [12], and the other two images are a balloon image and a fruit image. The first experiment illustrates FLICM and PLICM (using FLICM as initialization) as well as SPLI1M, showing them to be robust algorithms for image segmentation. We show that initialization of PLICM is of vital importance. The second experiment shows how SPLI1M outperforms SP1M and FLICM without an accurate cluster number as input. The third experiment demonstrates the soft partition results of SPLI1M and the superior performance of a possibilistic model over crisp and fuzzy approaches on an image with noise. The last experiment compares the proposed SPLI1M with two standard image segmentation clustering algorithms, k-means and mean shift.

IV.A. Basic segmentation examples

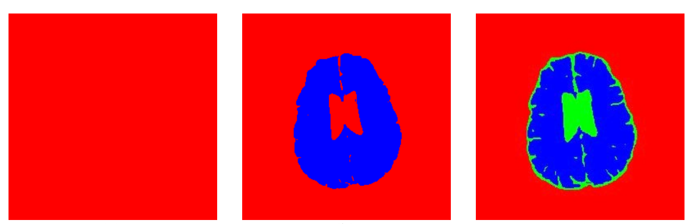
First, FLICM, PLICM and SPLI1M are run on a medical image from the original FLICM paper shown in Fig. 1(a). FLICM, PLICM (using FLICM as initialization) and SPLI1M all have a consistent good segmentation when hardened, i.e., with each pixel colored by the cluster with highest membership/typicality, as seen in Fig. 1(b).



(a) (b)

Fig. 1. (a) Original medical image and (b) Good image segmentation results (hardened clusters) of FLICM, PLICM (using FLICM as initialization) and SPLI1M. Here, for FLICM and PLICM, c was chosen as 3.

However, if PLICM uses random initialization, the segmentation results may not be stable; it depends on the initial cluster center choices. While all clustering algorithms need to deal with initialization, possibilistic versions are particularly sensitive to this choice.



(a) (b) (c)

Fig. 2. Three cases of hardened results of PLICM for Fig. 1(a). Here (a), all initial cluster centers were chosen in the background; (b) initial cluster centers picked in background and white matter; (c) initial cluster centers chosen in background, white and dark matter.

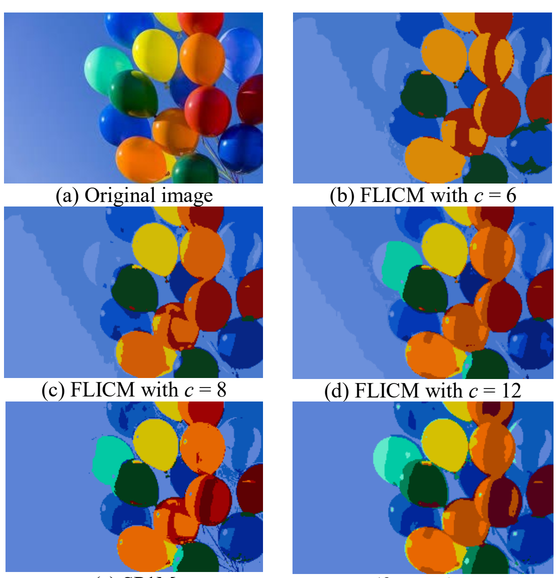
The hardened segments for three PLICM initializations of Fig. 1(a) are shown in Fig. 2. The first case is that all initial cluster centers fall in the background of the medical image, resulting in 3 coincident clusters, missing the brain completely (Fig. 2(a)). The second case resulted when some initial cluster centers fell on the background, and others on the white matter (Fig. 2(b)). The third case is that the background, white matter and dark matter all have initial cluster centers (Fig. 2(c)). Of course, we all like Fig. 2(c), but randomly choosing initialization opens the potential for vastly different outcomes.

Using FLICM for initialization, or sequentially running PLI1M (SPLI1M) are two good ways to avoid coincident cluster centers in image segmentation, depicted in Fig. 1(b). One advantage of running SPLI1M instead of FLICM is that SPLI1M does not need to specify the number of clusters (segments) in advance, producing a more robust clustering algorithm.

IV.B. Comparison of SPLI1M to FLICM and SP1M

When using FLICM in image segmentation, the number of cluster centers, or the number of colors in the image, must be specified. However, in the beginning, the number of clusters may not be easy to estimate for some images. In this experiment, a few values of cluster (color) numbers were specified for the balloon image in Fig. 3(a). SP1M (without combining local information) and SPLI1M were also tested on this image. The results in Fig. 3 are again the hardened partition values, with each pixel colored by the color of the cluster center with maximum membership or typicality.

As shown in Fig. 3 (b)-(d), different cluster (color) numbers were tried in FLICM. If the input number of clusters was smaller than the actual real number of colors, then obviously some clusters (colors) are missing in the hardened segmentation.



(a) Original image (b) FLICM with $c = 6$ (c) FLICM with $c = 8$ (d) FLICM with $c = 12$ (e) SP1M (f) SPLI1M

Fig. 3. (a) Original image, (b) FLICM where c = 6, (c) FLICM where c = 8, (d) FLICM where c = 12, (e) SP1M, automatically terminating at c = 9, (f) SPLI1M, automatically terminating at c = 11.

However, SP1M and SPLI1M do not need to specify the number of clusters (colors); furthermore, they will stop searching for new clusters (colors) when all possible dense clusters (colors) have been found. Note that in Fig. 3(f), some noisy pixels in the SP1M segmentation (Fig. 3(e)) were removed resulting in regions being smoother, due to the addition of local spatial information.

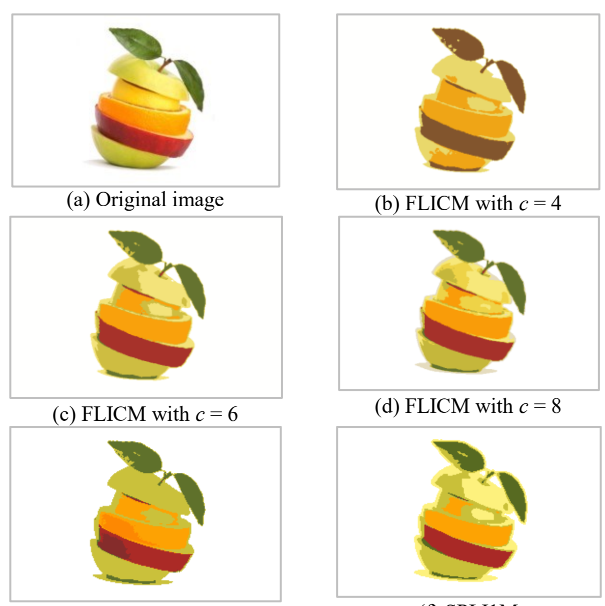
In this experiment, we also constructed the crisp partition matrix for the clusters found and compared it with the label partition matrix. The label partition matrix was generated by a human for the purpose of evaluation of the different clustering results. The actual number of clusters (colors) in the balloon image was chosen to be eight. We used the cluster comparison measure method in [27] to validate if the right cluster is detected. We run FLICM with different numbers of clusters ($c = 6, 8, 12$), and then SP1M and SPLI1M, hardening the final partitions. Then we compute the crisp Rand indices [27] of the final crisp partitions generated by each algorithm. If the Rand indices value is large, that means the partition matrix is more likely to match the label partition matrix. The Rand indices of FLICM with different number of clusters, SP1M and SPLI1M is shown in TABLE IV, showing that SPLI1M has an advantage in matching the crisp segmentation.

TABLE IV: The Rand Indices of FLICM, SP1M and SPLI1M

	Crisp Rand Indices
FLICM with c = 6	87.83
FLICM with c = 8	82.06
FLICM with c = 12	80.36
SP1M	87.64
SPLI1M	**89.94**

Another image (fruit image, Fig. 4(a)) was used to show the superior performance of SPLI1M over FLICM and SP1M. A few values of c were input to FLICM. The hardened segments are depicted in Fig. 4 (b)-(d). SP1M (without combining local spatial information) and SPLI1M were also run on the same image.

The hardened partitions from SPLI1M in Fig. 4(f) are better than those created by SP1M (Fig. 4(e)) and as good as or better than those from FLICM, regardless of the input choice for c (Fig. 4 (b)-(d)).



(a) Original image (b) FLICM with $c = 4$ (c) FLICM with $c = 6$ (d) FLICM with $c = 8$ (e) SP1M (f) SPLI1M

Fig. 4. (a) Original image, hardened partitions: (b) FLICM where c = 4, (c) FLICM where c = 6, (d) FLICM where c = 8, (e) SP1M, automatically terminating at c = 7, and (f) SPLI1M, automatically terminating at c = 6.

IV.C. Superior performance of SPLI1M on the image with noise

While hardened segmentation results are easy to visualize, one of the advantages of possibilistic and fuzzy approaches is that they produce soft partition matrices that can be used in subsequent processing. All the pixels have membership to each cluster (color). The soft partition results of SPLI1M on the fruit image used in the last experiment are shown in Fig. 5.

This soft segmentation provides a good indication of the strength of the cluster that can provide information of segmentation feature strength, and can allow for subsequent computer vision algorithms to take advantage of the typicalities/memberships during decision processes. For example, the goal of the soft segmentation in [25, 26] is to tailor and/or blend feature extraction and mine detection algorithms to various unclear seafloor backgrounds.

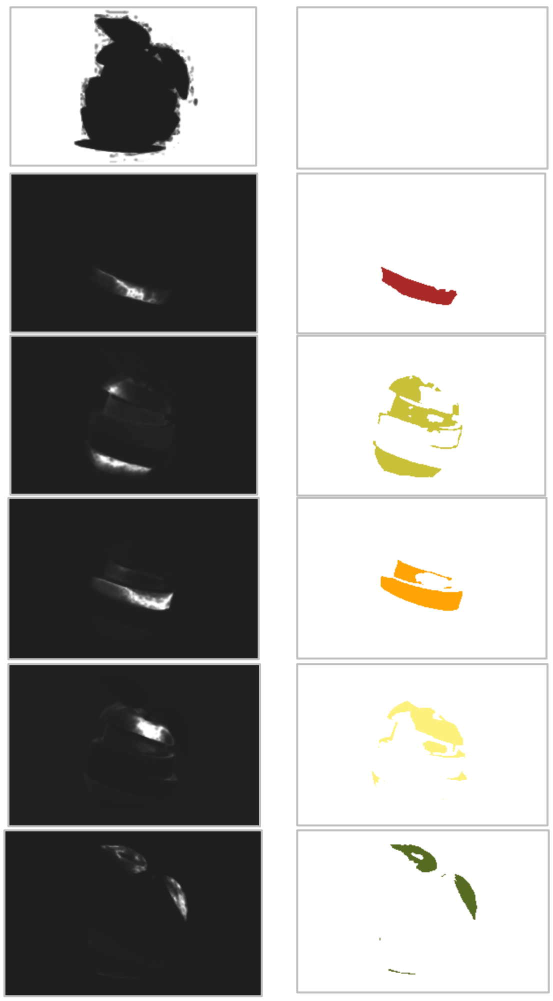
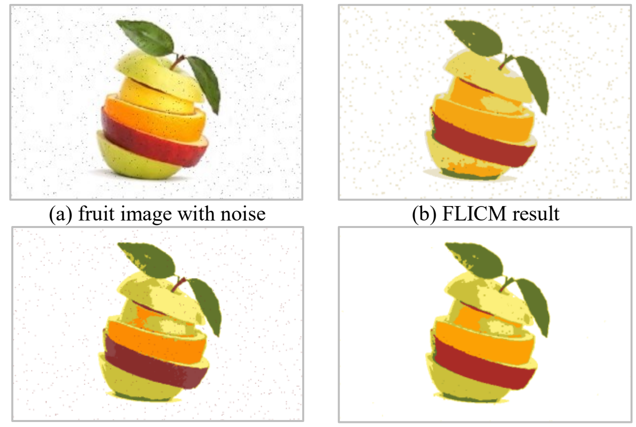


Fig. 5. The soft partition result of SPLI1M on the fruit image. Each image in the left column displays typicality of all pixels in each respective cluster, while the corresponding image in the second column shows the hardened segment.

One advantage of the possibilistic approach over the fuzzy one is that the possibilistic approach can detect outliers. We added some noise of a constant color to the fruit image. The noise is 5% uniform distributed in the whole range of fruit image. FLICM, SP1M (which doesn't combine local spatial information), and SPLI1M were run on this image with noise. The results of these three algorithms are shown in Fig. 6 (b)-(d).



(a) fruit image with noise (b) FLICM result

(c) SP1M result (d) SPLI1M result

Fig. 6. (a) Fruit image with noise, (b) Hardened FLICM result on part (a), c = 6, (c) Hardened SP1M result on part (a), automatically terminating at c = 7, (d) Hardened SPLI1M result on part (a), automatically terminating at c = 6.

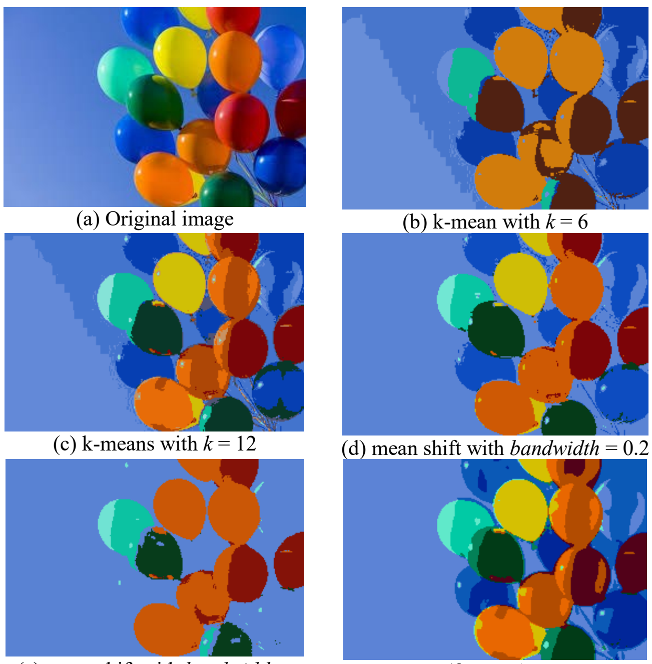
As we see, both FLICM and SP1M produced noise segments in the final segmentation result. SPLI1M shows superior performance over FLICM and SP1M on the image with noise. In fact, because of the insertion of local information in the clustering iterations, isolated noise points, or small groups of noise, will be absorbed into the surrounding image segments. This is why the noise has disappeared in the hardened segmentation of the fruit image in Fig. 6(d). The maximum typicality for each noise point is in the color cluster of the spatially surrounding region. The reason for this superior performance of SPLI1M over FLICM in an image with noise is that, even though FLICM contains the local information variable, it requires the number of expected clusters a priori. What happened in this example image is that one of the clusters got started in the noise color area and the sharing model of fuzzy memberships essentially kept it there, resulting in many small segments corresponding to that color. SPLI1M, on the other hand, hunts for one cluster at a time, allowing the noise points to be absorbed into the surrounding higher "volume" color regions.

IV.D. Comparison of SPLI1M to k-means and mean shift

The k-means and mean shift clustering algorithms are two effective clustering algorithms that are widely used in image segmentation. We now compare the proposed SPLI1M with k-means and mean shift. In the k-means clustering algorithm, the number of clusters must be specified in advance. In mean shift there is a parameter called *bandwidth* that needs to be adjusted. A few k values (k = 6, 12) in k-means and *bandwidth* values (*bandwidth* = 0.2, 0.25) in mean shift are tested on the balloon image.

As shown in Fig. 7 (b)-(c), the number of clusters values of 6 and 12 were used on the balloon image. If the input number of clusters k was smaller than the actual real number of clusters, then some clusters (colors) are missing in the final segmentation result. On the other hand, if the value of k was larger than the actual real number of clusters, then some redundant clusters will be found in the final segmentation result. As shown in Fig. 7 (d)-(e), *bandwidth* values 0.2 and 0.25 were used. As the *bandwidth* value grows, the number of clusters (colors) decreases. The *bandwidth* parameter has a similar influence on image segmentation as the parameter *eta* in the SPLI1M, which controls the cluster size or the color range for each segment. However, the k-means and mean shift are two crisp clustering algorithms. For example, they will have problems when clustering ambiguous color images.

SPLI1M, as a sequential clustering algorithm based on fuzzy set theory, is more robust against ambiguous or more continuously changing color images. Using SPLI1M with dynamic eta on the balloon image, the hardened segmentation is shown in Fig. 7 (f). It finds 11 clusters (colors). Each *eta* for its corresponding cluster is different. The *eta* value range is between 280 and 460. Picking the correct number of clusters as in Fig. 7(c) or the correct *bandwidth* as in Fig. 7(d) gives results close to those of SPLI1M. Of course, k-means or mean shift does clustering only on the 3D color feature space. It leaves out local spatial information in the image, producing anomalies in the final segmentation results.



(a) Original image
(b) k-mean with $k = 6$
(c) k-means with $k = 12$
(d) mean shift with *bandwidth* = 0.2
(e) mean shift with *bandwidth* = 0.25
(f) SPLI1M

Fig. 7. (a) Original image, (b) k-means with k = 6, (c) k-means with k = 12, (d) mean shift with bandwidth = 0.2, (e) mean shift with bandwidth = 0.25, and (f) Hardened partition of SPLI1M, automatically terminating at c = 11.

The crisp Rand indices [27] of k-means and mean shift algorithms were also computed to compare with the performance of SP1M and SPLI1M. The Rand indices of k-means with different value of *k*, mean shift with different values of *bandwidth*, SP1M and SPLI1M are shown in TABLE V. This also highlights the advantage of SPLI1M, and shows the sensitivity of k-means to the number of clusters and that of mean shift to the *bandwidth*.

TABLE V: The Rand Indices of k-mean, mean shift, SP1M and SPLI1M

	Crisp Rand Indices
k-means with $k = 6$	79.91
k-means with $k = 12$	83.98
mean shift with *bandwidth* = 0.2	89.66
mean shift with *bandwidth* = 0.25	84.08
SP1M	87.64
SPLI1M	**89.94**

V. Conclusions

In earlier work, we showed that sequential version of the possibilistic one-means (SP1M) performs better at finding *c* clusters than does FCM and PCM [14]. However, clustering in an image dataset is different from clustering in a standard dataset. A clustering algorithm running on the 3D color feature space leaves out the local spatial position information of the pixel, which sometimes may lead to poor segmentation results. FLICM extends FCM by incorporating local spatial information in the image and works better than FCM in image segmentation.

In this paper, PLICM extends FLICM by abandoning the membership sum-to-one constraint. However, initialization of PLICM has a significant impact on the final segmentation results. Using FLICM as initialization or sequentially running PLI1M (SPLI1M) are two good ways to avoid coincident cluster centers in image segmentation. One advantage of running SPLI1M instead of FLICM is that SPLI1M does not need to specify the number of clusters (colors) in advance, which is shown to produce a more robust segmentation, particularly in the presence of noise.

One issue in using any variant of PCM is that the value of the cluster-specific parameter, *eta*, must be specified or computed. We introduced SP1M with dynamic eta (SP1M-DE) [16] that determines *eta* dynamically at the beginning of each iteration of P1M. These algorithms have been extended to incorporate local spatial information for use in image segmentation. They perform very well when hardened partitions are used to create crisp image segments. In addition, the typicality partitions can be utilized to determine outliers, and in further computer vision tasks when image regions blend into each other. In the future, we will incorporate other methods of combining local spatial information of the image into our algorithms, such as clustering super-pixels of the image [28].

VI. References

[1] F. Precioso, M. Barlaud, T. Blu and M. Unser, "Robust real-time segmentation of images and videos using a smooth-spline snake-based algorithm", *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 910-924, 2005.

[2] L. Grady and C. Alvino, "The Piecewise Smooth Mumford–Shah Functional on an Arbitrary Graph", *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2547-2561, 2009.

[3] C. Li, R. Huang, Z. Ding, J. Gatenby, D. Metaxas and J. Gore, "A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities With Application to MRI", *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2007-2016, 2011.

[4] D. Liu, Y. Xiong, K. Pulli and L. Shapiro, "Estimating Image Segmentation Difficulty", *Machine Learning and Data Mining in Pattern Recognition*, pp. 484-495, 2011.

[5] M. Mignotte, "Segmentation by Fusion of Histogram-Based K-Means Clusters in Different Color Spaces", *IEEE Transactions on Image Processing*, vol. 17, no. 5, pp. 780-787, 2008.

[6] J. Jordan and E. Angelopoulou, "Mean-shift clustering for interactive multispectral image analysis", *2013 IEEE International Conference on Image Processing*, pp. 3790-3794, 2013.

[7] J. Bezdek, R. Ehrlich and W. Full, "FCM: The fuzzy c-means clustering algorithm", *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191-203, 1984.

[8] T. Pham, "Image segmentation using probabilistic fuzzy c-means clustering", *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, vol. 1, pp. 722-725, 2001.

[9] T. Runkler and J. Bezdek, "Alternating cluster estimation: a new tool for clustering and function approximation", *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 4, pp. 377-393, 1999.

[10] Y. Xia, D. Feng, T. Wang, R. Zhao and Y. Zhang, "Image segmentation by clustering of spatial patterns", *Pattern Recognition Letters*, vol. 28, no. 12, pp. 1548-1555, 2007.

[11] M. Krinidis and I. Pitas, "Color Texture Segmentation Based on the Modal Energy of Deformable Surfaces", *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1613-1622, 2009.

[12] S. Krinidis and V. Chatzis, "A Robust Fuzzy Local Information C-Means Clustering Algorithm", *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1328-1337, 2010.

[13] R. Krishnapuram and J. Keller, "A possibilistic approach to clustering", *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 2, pp. 98-110, 1993.

[14] T. Runkler and J. Keller, "Sequential possibilistic one-means clustering", *2017 IEEE International Conference on Fuzzy Systems,* pp. 1-6, 2017.

[15] T. Runkler, and J. Keller, "Sequential Possibilistic One–Means Clustering with Variable Eta," *Proceedings, 27th Workshop on Computational Intelligence, Dortmund, Germany, Dec.,* 2017 pp. 103-116, 2017.

[16] W. Wu, J. Keller and T. Runkler, "Sequential Possibilistic One-Means Clustering with Dynamic Eta", *IEEE International Conference on Fuzzy Systems*, pp. 1-6, 2018.

[17] R. Dave, "Characterization and detection of noise in clustering", *Pattern Recognition Letters*, vol. 12, no. 11, pp. 657-664, 1991.

[18] M. Barni, V. Cappellini and A. Mecocci, "Comments on 'A possibilistic approach to clustering'", *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 3, pp. 393-396, 1996.

[19] R. Krishnapuram and J. Keller, "The possibilistic C-means algorithm: insights and recommendations", *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 3, pp. 385-393, 1996.

[20] M. Yang and C. Lai, "A Robust Automatic Merging Possibilistic Clustering Method", *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 1, pp. 26-41, 2011.

[21] N. Pal, K. Pal, J. Keller and J. Bezdek, "A possibilistic fuzzy c-means clustering algorithm", *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 4, pp. 517-530, 2005.

[22] D. Anderson, J. Keller, O. Sjahputera, J. Bezdek and M. Popescu, "Comparing soft clusters and partitions", *2011 IEEE International Conference on Fuzzy Systems*, pp. 924-931, 2011.

[23] J. Keller and I. Sledge, "A Cluster By Any Other Name," *Proceedings, NAFIPS 2007 - Annual Meeting of the North American Fuzzy Information Processing Society*, pp. 427-432, 2007.

[24] T. Celik and H. Lee, "Comments on "A Robust Fuzzy Local Information C-Means Clustering Algorithm"", *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1258-1261, 2013.

[25] A. Zare, N. Young, D. Suen, T. Nabelek, A. Galusha and J. Keller, "Possibilistic fuzzy local information C-Means for sonar image segmentation", *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1-8, 2017.

[26] J. Peeples, D. Suen, A. Zare and J. Keller, "Possibilistic fuzzy local information C-means with automated feature selection for seafloor segmentation", *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIII*, vol. 10628, p.1062812, 2018.

[27] D. Anderson, J. Bezdek, M. Popescu and J. Keller, "Comparing Fuzzy, Probabilistic, and Possibilistic Partitions", *IEEE Transactions on Fuzzy Systems*, vol. 18, no. 5, pp. 906-918, 2010.

[28] S. Jia and C. Zhang, "Fast and robust image segmentation using an superpixel based FCM algorithm", *2014 IEEE International Conference on Image Processing*, pp. 947-951, 2014.