Possibilistic Fuzzy Local Information C-Means with Automated Feature Selection for Seafloor Segmentation

Joshua Peeples[a], Daniel Suen[a], Alina Zare[a], and James Keller[b]

[a]Electrical and Computer Engineering, University of Florida,Gainesville, FL, 32611
[b]Computer Science and Electrical Engineering, University of Missouri,Columbia, MO, 65211

ABSTRACT

The Possibilistic Fuzzy Local Information C-Means (PFLICM) method is presented as a technique to segment side-look synthetic aperture sonar (SAS) imagery into distinct regions of the sea-floor. In this work, we investigate and present the results of an automated feature selection approach for SAS image segmentation. The chosen features and resulting segmentation from the image will be assessed based on a select quantitative clustering validity criterion and the subset of the features that reach a desired threshold will be used for the segmentation process.

Keywords: Feature selection, SAS, segmentation, clustering, SONAR, cluster validity

1. INTRODUCTION

Synthetic aperture sonar (SAS) produces high resolution images of the sea environment.[1] This high quality data will assist in the detection of targets along the seafloor. In order to increase the accuracy and precision of the detection process, segmentation of the seafloor texture types will play a key role. Detecting objects of interest within SAS imagery with multiple seafloor texture types is a difficult problem, but if the SAS imagery is properly segmented into the various seafloor texture types, a detector for each region will perform better. Several segmentation algorithms have been applied to SAS imagery[2–6] and achieved promising results. One important component of segmentation algorithms is the features from the data. If the features are not optimal, the segmentation algorithm's performance will suffer. Feature selection is the process of choosing the best subset of features that are representative of the data and will lead to the best performance.[7] In this work, the Xie-Beni(XB) cluster validity measure[8] as well as a variation of the XB index were used to select the best number of features.

2. MOTIVATION

2.1 Previous Approaches

2.1.1 Segmentation Algorithms

K-means clustering was based on the approach developed by MacQueen.[9] K-means does a "crisp" clustering of the data by assigning each data point to only one cluster. In the SAS imagery, there are regions of transition that will need to be accounted for in the algorithm. To capture the regions of transition, fuzzy C-means clustering[10] can be used so that each data point can be have some membership in multiple clusters. Fuzzy C-means cluster does have issues with noisy images and image artifacts; as a result, a novel fuzzy factor term, G, was developed in the Fuzzy Local Information C-Means (FLICM)[11] to create a more robust segmentation algorithm. Within the SAS imagery, there may also be outliers that should not be clustered into the various seafloor textures. In order to identify these outliers, typicality can be used. The Possibilistic Fuzzy C-Means (PFCM) algorithm in[12] combines the fuzzy and possbilistic clustering approaches to produce membership and typicality values for the data. FLICM and PFCM both have desirable characteristics that will be useful for SAS image segmentation. The Possibilistic Fuzzy Local Information C-Means (PFLICM)[2] fuses both approaches and has been proven to segment SAS imagery well. PFLICM is summarized in section 3.1 of this work.

2.1.2 Feature Selection

Feature selection is an important area of study within Machine Learning. The advantages of using a subset of features include improved visualization and understanding of the data, efficient use of measurement and storage systems, decrease computational cost in training, and address the curse of dimensionality.[7]Feature selection methods can be divided into three main categories: wrapper, filter and embedded. The wrapper method selects the best subset of features by generating a model and evaluating each combination of features through some metric. Two wrapper methods are forward and backward feature selection.[13] The wrapper method is a greedy algorithm and can incur high computational cost, but this method should improve performance as features are added. The filter method evaluates the features before using a model. Some filter methods evaluate features based on relevance and other statistical measures.[13,14] The filter method is less computationally expensive but may remove features that will help improve the performance of the model or include features that can cause performance to deteriorate.[14] Embedded methods combine both approaches and perform feature selection during the training process.[7,15] Most feature selection processes use error as the evaluation metric but this is not a possible measure of performance for unsupervised applications. Some work has been done in unsupervised feature selection.[16–18] In this unsupervised feature selection framework, the cluster validity measure, XB, will be used to evaluate how well the features cluster the SAS imagery. The XB index and forward feature selection process are detailed in section 3.2 and 3.3 respectively.

3. METHODS

3.1 PFLICM

The Possibilistic Fuzzy Local Information C-Means (PFLICM) algorithm[2] was used to segment the SAS imagery. PFLICM integrates two previous clustering algorithms, the Fuzzy Local Information C-Means (FLICM)[11] and the Possibilisitic Fuzzy C-Means (PFCM).[12] PFLICM achieves four aims for SAS image segmentation: 1) account for gradients and regions of transition within the images, 2) identify outliers (i.e. targets, clutter), 3) include local spatial information, and 4) minimize computational cost. The objective function for PFLICM is shown in (1):

$$J = \sum_{c=1}^{C} \sum_{n=1}^{N} a u_{cn}^m \big(||\mathbf{x}_n - \mathbf{c}_c||_2^2 + G_{cn} \big) + b t_{cn} ||\mathbf{x}_n - \mathbf{c}_c||_2^2 + \sum_{c=1}^{C} \gamma_c \sum_{n=1}^{N} \big(1 - t_{cn} \big)^q \tag{1}$$

subject to the following constraints

$$u_{cn} \geq 0 \qquad \forall n = 1, ..., N \qquad \sum_{c=1}^{C} u_{cn} = 1 \tag{2}$$

where u_{cn} is the membership of the nth pixel in the cth cluster, x_n is a dx1 vector for the nth pixel, c_c is a $d \times 1$ vector of the cth cluster, t_{cn} is the typicality value of the nth pixel in the cth cluster, a, b and γ_c are weights on the membership and typicality terms respectively, and m and q control the "fuzziness" of the membership values for each cluster and identification of outliers in the data respectively. The G_{cn} term follows from[11] and incorporates local spatial information:

$$G_{cn} = \sum_{\substack{k \in \mathcal{N}_n \\ k \neq n}} \frac{1}{d_{nk} + 1} \big(1 - u_{ck} \big)^m ||x_k - c_c||_2^2 \tag{3}$$

where $\mathbf{x}_n$ is the center pixel of a local window, $\mathcal{N}_n$ is the neighborhood around the center pixel, and d_{nk} is the Euclidean distance between the center pixel and one of the neighboring pixels (x_k).

The cluster centers, membership and typicality values are found via alternating optimization. After these values are randomly initialized, the partial derivative with respect to the variables of interest is taken and set equal to 0 in order to produce three update equations. For the update equation for the membership values, a Lagrange multiplier term was added to enforce the sum-to-one constraint:

$$c_c = \sum_n \frac{\big(a u_{cn}^m + b t_{cn}^q \big) \mathbf{x}_n}{\big(a u_{cn}^m + b t_{cn}^q \big)} \tag{4}$$

$$u_{cn} = \frac{1}{\sum_{k=1}^{C} \left(\frac{(\mathbf{x}_n - c_c)(\mathbf{x}_n - c_c)^T + G_{cn}}{(\mathbf{x}_n - c_k)(\mathbf{x}_n - c_k)^T + G_{kn}} \right)} \tag{5}$$

$$t_{cn} = \frac{1}{1 + \left(\frac{b}{\gamma_c} ||x_k - c_c||_2^2 \right)^{\frac{1}{q-1}}} \tag{6}$$

γ_c is the mean of the separation of all the data points in the corresponding cluster ($||x_k - c_c||_2^2$). The fuzzy factor term, G, is treated as a constant in the membership update equation despite being a function of neighboring pixels. The actual derivative is not solvable[19] but using values from the past iterations produce good segmentation results. In order to reduce the run time of the algorithm, the image is first segmented using superpixels[20] and features are computed from the superpixels. After the segmentation is performed, PFLICM is implemented on the feature vector of each superpixel rather than pixel by pixel.

3.2 Cluster Validity Measure

Clustering is the process of grouping data into various group or clusters that share similar characteristics. Two important metrics for cluster validity is compactness and separation. Compactness refers to how similar the data points are within a cluster while separation measures how distinct the clusters are to one another [14]. In order to look at both metrics, Xie-Beni(XB) cluster validity measure[8] used in this work. The XB index is the ratio of the compactness to the separation of the clusters:

$$XB\ index = \frac{\sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^2 ||x_k - c_i||_2^2}{N \cdot \min_{i \neq k} ||c_i - c_k||_2^2} \tag{7}$$

The clusters that are well compact and separated should minimize the XB index. The XB index will be used as the evaluation metric to quantify how well segmented the SAS imagery based on each of the extracted features. The original XB index does not take into account the typicality of the data. The extended XB index (EXB) in[21] incorporates typicality by adding the membership and typicality values in the calculation of the index score. For this work, a variation of the XB index (VXB) is proposed

$$VXB\ index = \frac{\sum_{i=1}^{C} \sum_{j=1}^{N} (u_{ij} \cdot t_{ij})^2 ||x_k - c_i||_2^2}{N \cdot \min_{i \neq k} ||c_i - c_k||_2^2} \tag{8}$$

With the VXB, if the data point is an outlier, the membership and typicality values will be low while $||x_k - c_i||_2^2$ will be high. If the data point has high membership with the cluster, then the typicality value will be high while $||x_k - c_i||_2^2$ will be low. The VXB index is used to compare quantitatively PFLICM and PFCM because both algorithms incorporate possiblistic clustering.

3.3 Forward Feature Selection

The selection of the best features for the PFLICM algorithm is crucial to the segmentation process. A wrapper method, forward feature selection, was implemented in this work. After the features are extracted, one feature is selected. The selected feature is then used to cluster the data with PFLICM and the membership, typicality and cluster center values are estimated. The XB index is then used to score the segmentation produced. The next feature is then selected and the process is repeated until all features have been evaluated individually. The features are then ranked in descending order of the XB index and the best feature is removed from the original feature set and paired with the remaining features. The forward feature selection process will end once all features are exhausted from the original dataset. Figure 1 illustrates the forward feature selection process. The suite of features used for the selection process are Sobel,[22] histogram of oriented gradients,[23] local binary patterns,[24] mean, variance, shape,[25,26] Haralick texture features (contrast,correlation,energy, and homogeneity),[27] a Gabor filter,[28] a Gaussian filter,[29] lacunarity,[30] and a Laplacian of Gaussian filter.[6,29] An
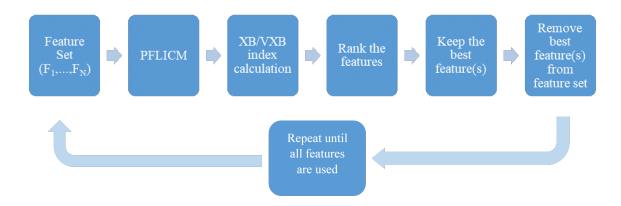
Figure 1. Forward Feature Selection Process

additional test for the forward feature selection process involved randomly selecting three features from a set of six features (histogram of oriented gradients, homogeneity, lacunarity, local binary patterns, shape, and Sobel) and comparing the randomly selected features with the three features chosen by the XB and VXB indices with PFLICM.

4. EXPERIMENTAL RESULTS

The forward feature selection process was implemented on the same image for three sets of features. A grid search was performed to find the optimal parameters that minimized the XB and VXB indices. Table 1 shows the initial value, final value and increments of the parameters used in the grid search. The weight of the typicality was set to be a tenth of the weight on the membership, i.e., $b = .1a$.

Table 1. Grid Search Experiment Values

Parameters	Initial Value	Final Value	Increment
a	2	12	2
b	.2	1.2	.2
m	1.2	3.0	.3
q	1.2	2.8	.2

The results of the grid search are found in Figure 2. For both the XB and VXB indices, the fuzzifier for the membership (m) and typicality (q) values were varied as the weights a and b were fixed to 8 and .8 respectively. As seen in Figure 2a-d, the values of m and q that minimize the indices values are 1.8 and 2.8 respectively. Figure 2a-d demonstrate how sensitive the scores are to each fuzzifier. As n becomes smaller, the index increases; additionally, if m is higher than 2, the index also becomes larger. Once the fuzzifier terms were found, the next step was to verify that the weights minimized the XB and VXB metrics. The weights, a and b, were varied with b set to be a tenth of a and the scores from this experiment are found in Figure 2e. The set of weights that minimized each score with m and q fixed were $a = 14$ and $b = 1.4$. The following parameter settings minimized both the XB and VXB: $a = 14$, $b = 1.4$, $m = 1.8$, $q = 2.8$, $\epsilon = $ 1e-6 (threshold for change in values) and 3 clusters.

(a) Log of XB

(b) Log of VXB

(c) Log of XB

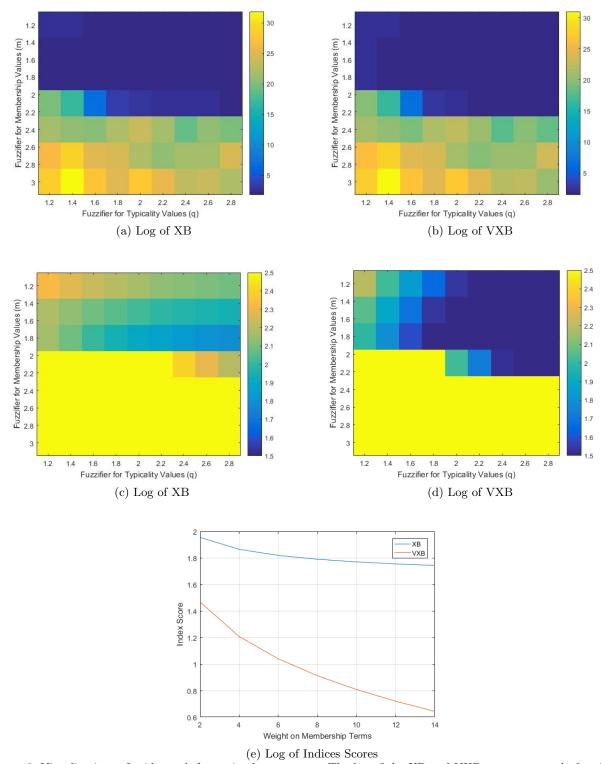(d) Log of VXB

(e) Log of Indices Scores

Figure 2. Visualizations of grid search for optimal parameters. The log of the XB and VXB scores were calculated while m and q were varied. The weights, a and b, were fixed to 8 and .8. (a-b) are scaled between 1.5 and 31 and (c-d) are scaled between 1.5 and 2.5. (e) is the plot of the log of the XB and VXB scores as the weights are varied and m and q are fixed.

For each clustering algorithm, the average XB and VXB index score over five runs with random initialization values was recorded and used to determine the selected features. Table 2 summarizes the index scores for each algorithm. The lowest for each index score is highlighted in bold and the second lowest score is underlined. The minimum score indicates the lowest value achieved for the optimal number of features selected by each algorithm. The same parameters for PFLICM were used in PFCM. The VXB was computed for only PFLICM and PFCM since FLICM and K-Means do not produce typicality values.

Table 2. Cluster Validity Measure Results (± 1 Standard Deviation)

Index	PFLICM	FLICM	PFCM	K-Means
Min XB	0.433±0.158	0.405±0.031	0.538±0.070	**0.284±3.543e-5**
Min VXB	**0.134±0.035**	-	0.153±0.017	-
XB (all features)	1.722±2.483e-16	**1.509±0.000**	10.994±0.000	2.089±1.172e-5
VXB (all features)	**0.565±0.000**	-	2.742±0.000	-

The progression of the XB and VXB scores as feature were added is recorded in Table 3 and 4 respectively. As seen in the tables, the scores should decrease as features are added until a minimum is reached (minimum score for each algorithm is highlighted in bold) and the score should begin to increase as the remaining features are added to the segmentation process.
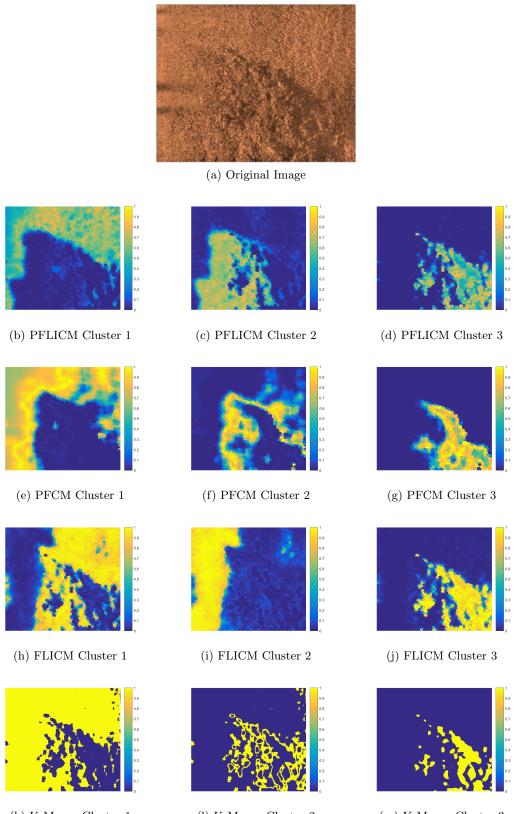
Table 3. XB Index Score Progression (± 1 Standard Deviation)

Number of Features	PFLICM	FLICM	PFCM	K-Means
1	0.678±0.161	0.613±0.138	**0.538±0.070**	**0.284±3.543e-5**
2	0.645±0.196	0.617±0.204	0.551±0.281	0.526±0.121
5	**0.433±0.158**	0.702±0.503	0.562±0.247	0.815±0.239
6	0.470±0.086	**0.405±0.031**	0.642±0.188	0.826±0.373
10	1.016±0.169	0.731±0.457	1.006±0.432	1.525±0.711
15	1.722±2.483e-16	1.509±0.000	10.994±0.000	2.089±1.172e-5
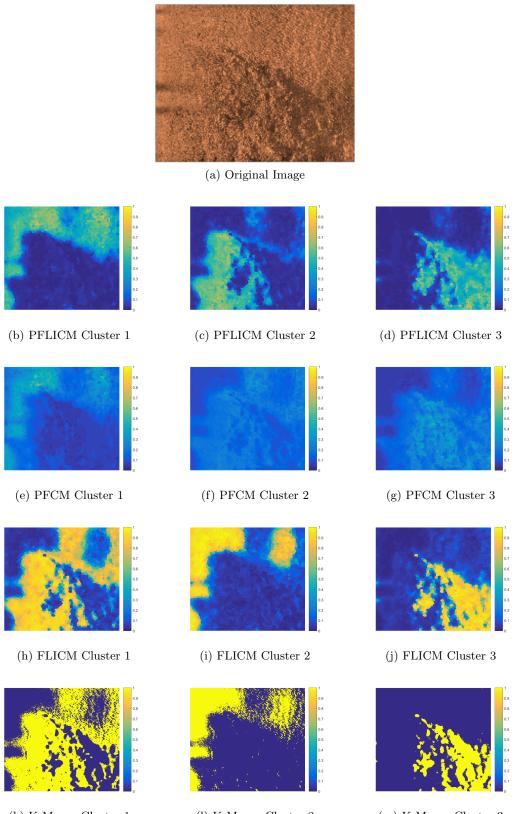
Table 4. VXB Index Score Progression (± 1 Standard Deviation)

Number of Features	PFLICM	PFCM
1	0.296±0.053	0.267±0.082
3	0.164±0.101	0.163±0.062
6	**0.134±0.035**	0.174±0.041
9	0.327±0.147	**0.153±0.017**
10	0.440±0.162	0.235±0.134
15	0.565±0.0000	2.742±0.000

The segmentation maps of the features selected using the XB index and all the features are shown in Figures 3 and 4 respectively. Figure 5 shows the progression of the segmentation results as features are added to PFLICM. The three randomly selected features vs the XB and VXB indices selection of three features are found in Figure 6. The selected features by the VXB index are shown for PFLICM and PFCM in Figure 7. The membership maps are shown for FLICM and K-Means and the product maps (membership*typicality) for PFLICM and PFCM.

(a) Original Image

(b) PFLICM Cluster 1

(c) PFLICM Cluster 2

(d) PFLICM Cluster 3

(e) PFCM Cluster 1

(f) PFCM Cluster 2

(g) PFCM Cluster 3

(h) FLICM Cluster 1

(i) FLICM Cluster 2

(j) FLICM Cluster 3

(k) K-Means Cluster 1

(l) K-Means Cluster 2

(m) K-Means Cluster 3

Figure 3. (a) SAS image with distinct regions: sand ripple, hard packed sand, and coral/rocky surface. The segmentation results for PFLICM (b-d), PFCM (e-g), FLICM (h-j), and K-Means(k-m) with the selected features using XB index.

(a) Original Image

(b) PFLICM Cluster 1   (c) PFLICM Cluster 2   (d) PFLICM Cluster 3

(e) PFCM Cluster 1   (f) PFCM Cluster 2   (g) PFCM Cluster 3

(h) FLICM Cluster 1   (i) FLICM Cluster 2   (j) FLICM Cluster 3

(k) K-Means Cluster 1   (l) K-Means Cluster 2   (m) K-Means Cluster 3

Figure 4. (a) SAS image with distinct regions: sand ripple, hard packed sand, and coral/rocky surface. The segmentation results for PFLICM (b-d), PFCM (e-g), FLICM (h-j), and K-Means(k-m) with all features.

(a) Original Image

(b) PFLICM Cluster 1

(c) PFLICM Cluster 2

(d) PFLICM Cluster 3

(e) PFLICM Cluster 1

(f) PFLICM Cluster 2

(g) PFLICM Cluster 3

(h) PFLICM Cluster 1

(i) PFLICM Cluster 2

(j) PFLICM Cluster 3

(k) PFLICM Cluster 1

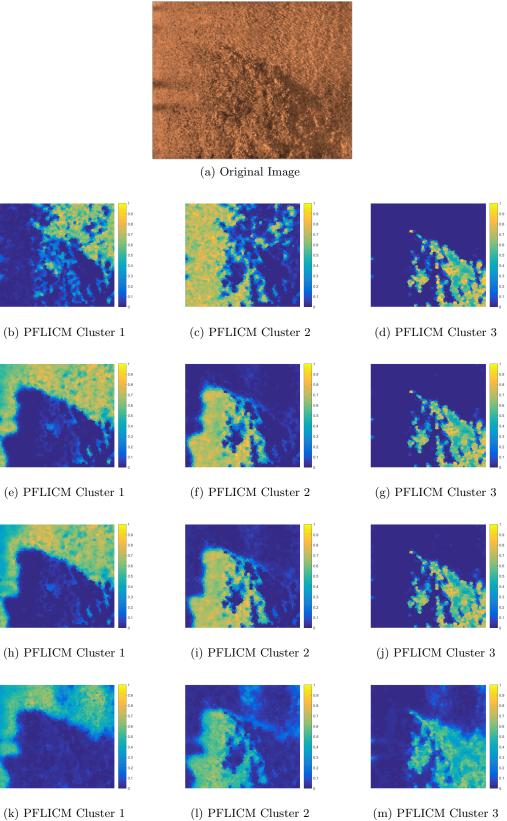
(l) PFLICM Cluster 2

(m) PFLICM Cluster 3

Figure 5. (a) SAS image with distinct regions: sand ripple, hard packed sand, and coral/rocky surface. The segmentation results for PFLICM with the first feature (b-d), three features (e-g), selected features (h-j), and all features (k-m).

(a) Original Image

(b) PFLICM Cluster 1

(c) PFLICM Cluster 2

(d) PFLICM Cluster 3

(e) PFLICM Cluster 1

(f) PFLICM Cluster 2

(g) PFLICM Cluster 3

(h) PFLICM Cluster 1

(i) PFLICM Cluster 2
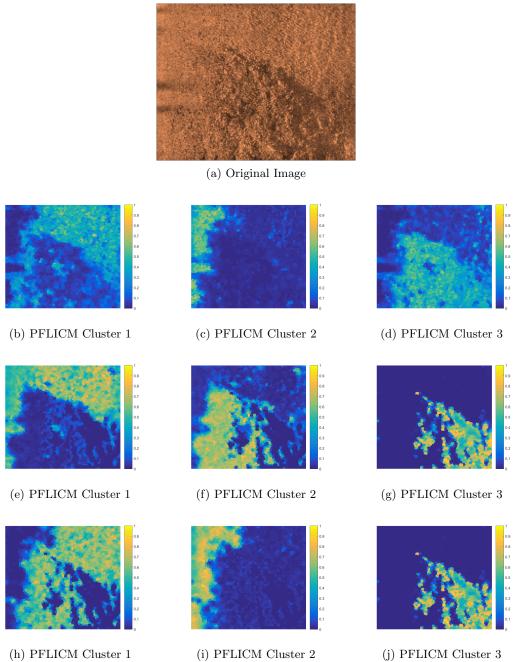

(j) PFLICM Cluster 3

Figure 6. (a) SAS image with distinct regions: sand ripple, hard packed sand, and coral/rocky surface. The segmentation results for randomly selected features (b-d), XB selected features (e-g) and VXB selected features (h-j).

(a) Original Image

(b) PFLICM Cluster 1

(c) PFLICM Cluster 2

(d) PFLICM Cluster 3

(e) PFCM Cluster 1
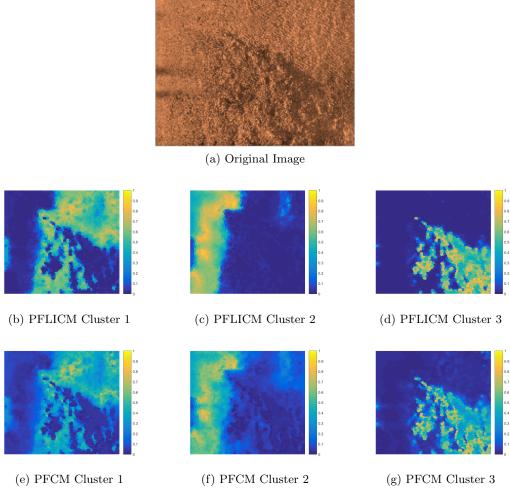

(f) PFCM Cluster 2

(g) PFCM Cluster 3

Figure 7. (a) SAS image with distinct regions: sand ripple, hard packed sand, and coral/rocky surface. The segmentation results for PFLICM (b-d) and PFCM (e-g) with selected features using VXB index.

5. DISCUSSION

Cluster validity indices are usually used to determine the optimal number of clusters. The XB and VXB indices can also provide information on the best features to use to cluster the data. As shown in Figures 2 and 3, the segmentation results for the features that minimized the XB score are visually better than using all of the features. In comparing the two metrics, VXB is useful for algorithms that perform fuzzy and possiblistic clustering. The VXB accounts for typicality and as seen in Table 2, the PFLICM VXB index was the lowest score in comparison to all algorithms. The segmentation results agree with this quantitative measure. If only the XB index was used, K-Means clustering was the best algorithm. For the SAS imagery, transition regions will need to be accounted for so "crisp" partitions will not be optimal for this application. For example, within this image, there are three seafloor textures. Between each seafloor texture, there is not a distinct boundary between where one seafloor type ends and the other begins. The K-Means algorithm clusters the sand ripple and hard packed sand regions together (Figure 3k) while the other fuzzy clustering algorithms do not. The initial hypothesis for why the XB index might favor K-Means is that the membership of each pixel is either 0 or 1 leading to fewer terms in the calculation of the index as opposed to fuzzy clustering algorithms in which pixels can have some membership in each cluster. Each algorithm was implemented with the feature selected by K-means (Sobel) and the results of this test (averaged over 5 runs) are found in Table 5. N is the number of valid pixels in the image and C is the total number of clusters. Despite there being fewer non-zero terms in the compactness calculation (numerator

of the XB index), each clustering algorithm has a compactness score on the same magnitude. The separation calculation (denominator of the XB index) is higher for K-Means which leads to a lower XB index.

Table 5. XB Index Calculation Details ($\pm$ 1 Standard Deviation)

Algorithm	Numerator	Denominator	XB Score	Number of Non-zero Terms in Numerator
PFLICM	2.009e4$\pm$0.000	1.827e4$\pm$0.000	1.100$\pm$0.000	N*C
PFCM	1.749e4$\pm$0.000	1.038e4$\pm$0.000	1.685$\pm$0.000	N*C
FLICM	**1.532e4$\pm$0.000**	1.894e4$\pm$0.000	0.809$\pm$0.000	N*C
K-Means	2.095e4$\pm$0.000	**7.492e4$\pm$0.000**	**0.280$\pm$0.000**	N

Figure 6 further validates the usefulness of the XB and VXB indices for feature selection. The randomly selected features (Figure 6a-c) chosen for segmentation were histogram of oriented gradients, local binary patterns and shape. The XB index selected lacunarity, Sobel, and homogeneity (Figure 6e-g) while the VXB index selected lacunarity, sobel and shape. The segmentation results for the selected features appear more distinct than those of the randomly chosen features.

6. CONCLUSION

PFLICM with forward feature selection was presented in this work. This work has two main contributions: the use of cluster validity as evaluation metric for feature selection and a variation of the XB index for fuzzy and possiblistic clustering algorithms. Unsupervised feature selection is a more difficult than supervised applications since error cannot be used to determine how well a model performs with a subset of features. Quantitative measures such as compactness and separation of the clusters obtained from the features validated the qualitative results (segmentation maps). Several extensions of this work are possible such as using a different feature selection process (filter or embedded methods) and different evaluation metrics (other cluster validity measures,[31] mutual information,[14] entropy).

ACKNOWLEDGMENTS

REFERENCES

[1] Hayes, M. P. and Gough, P. T., "Synthetic aperture sonar: a review of current status," *IEEE Journal of Oceanic Engineering* **34**(3), 207–224 (2009).

[2] Zare, A., Young, N., Suen, D., Nabelek, T., Galusha, A., and Keller, J., "Possibilistic fuzzy local information c-means for sonar image segmentation," *IEEE Symposium Series on Computational Intelligence (SSCI) Proceedings* (2017).

[3] Karoui, I., Fablet, R., Boucher, J.-M., and Augustin, J.-M., "Seabed segmentation using optimized statistics of sonar textures," *IEEE Transactions on geoscience and remote sensing* **47**(6), 1621–1631 (2009).

[4] Cobb, J. T. and Principe, J., "Autocorrelation features for synthetic aperture sonar image seabed segmentation," in [*Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*], 3341–3346, IEEE (2011).

[5] Cobb, J. T. and Principe, J., "Seabed segmentation in synthetic aperture sonar images," in [*Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVI*], **8017**, 80170M, International Society for Optics and Photonics (2011).

[6] Cobb, J. T. and Zare, A., "Multi-image texton selection for sonar image seabed co-segmentation," in [*Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XVIII*], **8709**, 87090H, International Society for Optics and Photonics (2013).

[7] Guyon, I. and Elisseeff, A., "An introduction to variable and feature selection," *Journal of machine learning research* **3**(Mar), 1157–1182 (2003).

[8] Xie, X. L. and Beni, G., "A validity measure for fuzzy clustering," *IEEE Transactions on pattern analysis and machine intelligence* **13**(8), 841–847 (1991).

[9] MacQueen, J. et al., "Some methods for classification and analysis of multivariate observations," in [*Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*], **1**(14), 281–297, Oakland, CA, USA (1967).

[10] Bezdek, J. C., "Objective function clustering," in [*Pattern recognition with fuzzy objective function algorithms*], 43–93, Springer (1981).

[11] Krinidis, S. and Chatzis, V., "A robust fuzzy local information c-means clustering algorithm," *IEEE transactions on image processing* **19**(5), 1328–1337 (2010).

[12] Pal, N. R., Pal, K., Keller, J. M., and Bezdek, J. C., "A possibilistic fuzzy c-means clustering algorithm," *IEEE transactions on fuzzy systems* **13**(4), 517–530 (2005).

[13] Kohavi, R. and John, G. H., "Wrappers for feature subset selection," *Artificial intelligence* **97**(1-2), 273–324 (1997).

[14] Peng, H., Long, F., and Ding, C., "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on pattern analysis and machine intelligence* **27**(8), 1226–1238 (2005).

[15] Chandrashekar, G. and Sahin, F., "A survey on feature selection methods," *Computers & Electrical Engineering* **40**(1), 16–28 (2014).

[16] Mitra, P., Murthy, C., and Pal, S. K., "Unsupervised feature selection using feature similarity," *IEEE transactions on pattern analysis and machine intelligence* **24**(3), 301–312 (2002).

[17] Cai, D., Zhang, C., and He, X., "Unsupervised feature selection for multi-cluster data," in [*Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*], 333–342, ACM (2010).

[18] Tang, C., Zhu, X., Chen, J., Wang, P., Liu, X., and Tian, J., "Robust graph regularized unsupervised feature selection," *Expert Systems with Applications* **96**, 64–76 (2018).

[19] Celik, T. and Lee, H. K., "Comments on a robust fuzzy local information c-means clustering algorithm," *IEEE Transactions on Image Processing* **22**(3), 1258–1261 (2013).

[20] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S., "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence* **34**(11), 2274–2282 (2012).

[21] Zhang, C., Zhou, Y., and Martin, T., "A validity index for fuzzy and possibilistic c-means algorithm," in [*Proc. IPMU*], **8**, 877–882 (2008).

[22] Frigui, H. and Gader, P., "Detection and discrimination of land mines in ground-penetrating radar based on edge histogram descriptors and a possibilistic k-nearest neighbor classifier," *IEEE Transactions on Fuzzy Systems* **17**(1), 185–199 (2009).

[23] Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection," in [*Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*], **1**, 886–893, IEEE (2005).

[24] Guo, Z., Zhang, L., and Zhang, D., "A completed modeling of local binary pattern operator for texture classification," *IEEE Transactions on Image Processing* **19**(6), 1657–1663 (2010).

[25] Abraham, D. A. and Lyons, A. P., "Novel physical interpretations of k-distributed reverberation," *IEEE Journal of Oceanic Engineering* **27**(4), 800–813 (2002).

[26] Cobb, J. T. and Stack, J. R., "In situ adaptive feature extraction for underwater target classification," in [*Applied Imagery Pattern Recognition Workshop, 2007. AIPR 2007. 36th IEEE*], 42–47, IEEE (2007).

[27] Haralick, R. M., Shanmugam, K., et al., "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics* (6), 610–621 (1973).

[28] Jain, A. K. and Farrokhnia, F., "Unsupervised texture segmentation using gabor filters," *Pattern recognition* **24**(12), 1167–1186 (1991).

[29] Malik, J., Belongie, S., Leung, T., and Shi, J., "Contour and texture analysis for image segmentation," *International journal of computer vision* **43**(1), 7–27 (2001).

[30] Williams, D. P., "Fast unsupervised seafloor characterization in sonar imagery using lacunarity," *IEEE transactions on Geoscience and Remote Sensing* **53**(11), 6022–6034 (2015).

[31] Liu, Y., Li, Z., Xiong, H., Gao, X., and Wu, J., "Understanding of internal clustering validation measures," in [*Data Mining (ICDM), 2010 IEEE 10th International Conference on*], 911–916, IEEE (2010).