35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Potential of Hybridization Methods to Reducing the dimensionality for multispectral Biological Images

Jihan Khoder, Rafic Younes, And Fethi Ben Ouezdou

Abstract— **we address the problem of unsupervised band reduction in multispectral imagery. We propose to use a new hybridization of dimensionality reduction method by combining two categories of bands selection method with projection method and apply it to multispectral data. The algorithm employs the concepts of fuzziness and belongingness (Fuzzy K-means) to provide a better and more adaptive clustering process. However, the Fuzzy hybridized algorithm is applicable to medical imagery. A cluster validity function associated with Bezdek's partition coefficient is employed for evaluation of the dimension reduction's performance for this multispectral data. Experiments conducted in this paper confirm the feasibility of the new hybridization for multispectral dimensionality reduction and shows the potential of the proposed approach.**

I. INTRODUCTION

Multispectral imaging has become an active research topic in recent years due to its wide-spread applications in areas such as resource management, agriculture, mineral exploration and environmental monitoring, with the number of channels in the hundreds instead of in the tens, multispectral imagery possesses much richer spectral information [1]. This paper addresses the problem of band reduction in multispectral imagery. One problem comes from the high dimensionality of multispectral data which is often an obstacle for data of dimension reduction and interpretation, since there are generally not enough samples to fill the hyperspace spanned by the large set of variables and thus to infer some structure within the data.

Mathematically, given n points $\boldsymbol{x_1}, \dots, \boldsymbol{x_n}$ in a high dimensional subspace of $\boldsymbol{\mathcal{R}^D}$ the goal of dimensionality reduction is to find a mapping: F: $\boldsymbol{\mathcal{R}^D} \longrightarrow \boldsymbol{\mathcal{R}^d}$, $\boldsymbol{y_i}$= F ($\boldsymbol{x_i}$) Where, i=1… n and d is the dimensionality of the embedding space(d<D). Several approaches exist for dimensionality reduction in multispectral data and can be split into two major groups. The first group relates to transformation-based approaches. Such methods aim at projecting the original data set onto adequate subspaces, chosen for their relevance to explain the data. This group includes not only linear methods are presented in [3,5,6,7],but also nonlinear techniques in [8, 9, 10, 11, 12, 13, 14, 15 and 16]. These methods, when applied to multispectral data, suffer from a lack of "explainability" from a physical point of view since, generally, the original spectral information is not preserved in the local/global geometrical structure on reduced data volumes. The second group includes feature-selection based approaches. Most of them require knowledge of the ground truth. Some methods use either entropy or mutual information [17, 22] and the spectral information divergence for Ward's linkage strategy using divergence (WaLuDi) [23], to derive appropriate band selection criteria. Only few of these approaches are partially unsupervised [18, 19] or totally unsupervised [20]. Affinity Propagation (AP) is a clustering algorithm that identifies a set of ''exemplars'' that represents the dataset [24], but the algorithm needs the initial similarity as a parameter. In [29], BandClust is a commonly used feature selection reduction method, it was found to provide encouraging results on real multispectral data in terms of overall classification accuracy .

We made a comparative study of non-parametric and unsupervised techniques for dimensionality reduction [27]. Numerous criterions where applied on artificial data while distinguishing similarity and stability evaluation criterions [25, 26]. We can ascertain the following points from this work. First, we classify in three categories projection reduction techniques as well as spectral Band selection techniques. Secondly, this study proposes a hybridization of BandClust/MDS which has permitted very encouraging results on the stability and similarity of reduced artificial data sets [25, 26]. This hybridization has the advantage of combining two relatively simple algorithmic-wise methods, and to conserve rare and static information as well as minimizing spectral redundancy [25].

In this paper, we propose an unsupervised approach to band reduction for multispectral images named Feature Extraction Using PCA with Fuzzy K-Means Clustering. It consists in splitting the initial range of spectral into disjoint clusters of bands, and minimizes the similarity of each class between the bands spectral in a reduced space.

In section II, we describe the proposed method. We define the data used and also discuss the type of analyses applied to test the capability of the proposed algorithm in section III. Describe performed experiments and obtained result obtained on multispectral data sets in section IV. We conclude and give some perspectives of this work in Section V.

II. PROPOSED METHOD

As original fuzzy-K-means Clustering algorithm often does not work well to reduce the high dimension, hence, to improve the efficiency, we proposed to hybridize PCA on clustering datasets, to obtain a reduced datasets containing possibly uncorrelated variables. The proposed model is

J. Khoder is with the LISV Laboratory, Versailles University, Velizy, France (phone: 33139254982; fax: 33139254985; e-mail: jihan.khoder@hotmail.com).

R. Younes and F. Ben Ouezdou are with is with the LISV Laboratory, Versailles University, Velizy, France (ryounes@ul.edu.lb, fethi.benouezdou@uvsq.fr)

illustrated in Figure 1. The algorithm is composed of the following steps:

1. Input datas X= $(X_1, X_2,..,X_i,..,X_n)$// X_i=$(x_{i1}, x_{i2},.., x_{im})$ of i spectral bands for i=1,..,n.
2. Apply the Fuzzy k–means Clustering Algorithm for the original dataset X= $(X_1, X_2, .., X_i,.., X_n)$ X_i=$(x_{i1}, x_{i2},.., x_{im})$ of i spectral bands for i=1,..,n. Thus, points on the edge of a cluster may be in the cluster to a lesser degree than points in the center of cluster. For each point x we have a coefficient giving the degree of being in the k^{th} cluster $u_k(x)$. Finally, this algorithm aims at minimizing an objective function , the sum of those coefficients is defined by

$$FKM(X, C)=\sum_{i=1}^{n}\sum_{j=1}^{k} u_{ij}^{m} \left\| x_i^{(j)} - c_j \right\|^2 \quad (1)$$

Where $\forall$ x $\sum_{k=1}^{num\ clusters} u_k(x)$ =1.

With fuzzy k-means, the centroids of a cluster are the mean of all points, weighted by their degree of belonging to the cluster:

$$\text{Center}_k = \frac{\sum_x u_k(x)^m x}{\sum_x u_k(x)^m} \quad (2)$$

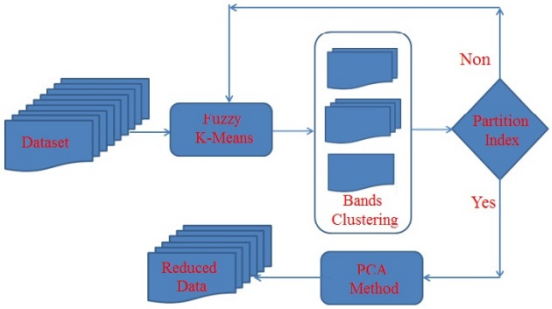

Figure.1. Feature Extraction by PCA with Fuzzy K-means Clustering.

The degree of belonging is related to the inverse of the distance to the cluster center: $u_k(x)= \frac{1}{d(\text{center}_k,x)}$ when thecoefficients are normalized and fuzzy –fied with a real parameter m > 1 so that their sum is 1. So

$$u_k(x)= \frac{1}{\sum_j \left(\frac{d(\text{center}_k,x)}{d(\text{center}_j,x)}\right)^{2/m-1}} \quad (3)$$

3. In step three, to validate and assess the quality of classifications, we applied the Partition Coefficient (F) [30,31] commonly used in the multispectral imaging. The index of validity studied is calculated from the membership degrees and centers of classes estimated by FCM defined by for each pair (k, m) argmax $\{F(K)\}_{K\in\{2,..,n-1\}}$, where

$$F(K) = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{K}(u_{ik})^2 \quad (4)$$

If the value of validity index is smaller than a threshold determined, then the algorithm FKM is reapplied to increase one class. On the other hand, if the value of coefficient partition (F) is higher than a threshold fixed, then we go to the final step in the reduction by projection method Principal component analysis.

4. Finally, we apply the PCA to each class bands, to obtain a representative spectral band of each class.

III. EXPERIMENTAL ACTIVITIES

This section, we provide the experimental using a medical image. These multispectral bio-images are obtained from a high throughput Liquid Crystal Tunable Filter which allows the capture of 33 spectral bands between 500 nm and 650 nm with a 4.5454 nm stepwise between each successive band.

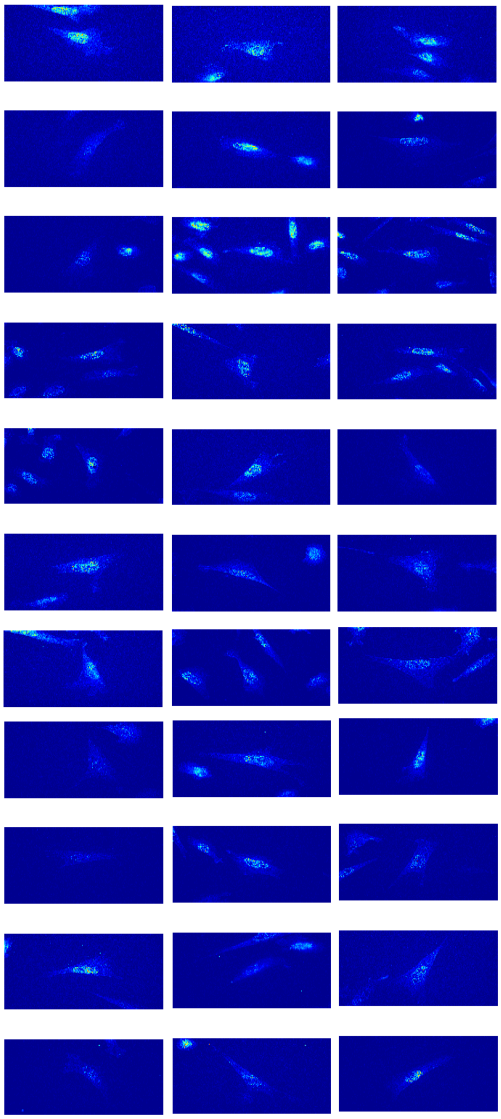

Figure.2. Carcinoma type of abnormally colon cells

Therefore, there are 33 sequential bands images for each type of malignant cancer grades in our work we applied on

cancer cells of Carcinoma type "*Ca*" for different frequencies or wavelength between 500 and 650 nm as illustrate in figure 2. This type of medical image is texture multispectral images [28] where we used medium objective magnification (X25).

*Ca**: abnormal tissue development corresponding to cancer.

X25*: mean zoom to 25 from optical microscopy.

IV. RESULTS DISCUSSION

In this section, these criteria are directly derived from classical statistical measures. Every value is individually considered according to the spatial and dimensions spectral. The multispectral image is represented as a three-dimensional matrix $\boldsymbol{I(x,y,\lambda)}$, with x is the position of the pixel in the line; y is the number of the line and $\boldsymbol{\lambda}$ the considered spectral band. $\boldsymbol{n_x, n_y, n_\lambda}$ Are respectively the number of pixels per line, the number of lines and the number of spectral bands.

In the previous sections, we evaluated the proposed hybridization algorithm on multispectral data sets with 33 spectral bands having 7 attributes as shown in figure3. The result has been shown in the table I, obtained for the tested unsupervised criteria presented above of the NCC, SC and MAE. In the following, we present a performance obtained from simulations by preserving the information's quantity structure of the original data.

A. *Structural content (SC):*

SC is the ratio of Power Spectral Density (PSD) of the two images (image reduced $\tilde{\mathbf{I}}(\mathbf{x},\mathbf{y},\boldsymbol{\lambda})$ on the reference image I(x, y, λ)) presented in [32] and is defined as

$$SC=\frac{\sigma_{\tilde{I}}^2+\mu_{\tilde{I}}^2}{\sigma_I^2+\mu_I^2} \quad (5)$$

And multispectral images: $SC = \frac{\sum_{x,y,\lambda}[\tilde{I}(x,y,\lambda)]^2}{\sum_{x,y,\lambda}[I(x,y,\lambda)]^2}$

B. *Normalized Cross-Correlation (NCC):*

The Normalized Cross-Correlation (NCC) is mentioned in [31] as those proposing to use fidelity.

$$NCC=\frac{\sum_{x,y,\lambda} I(x,y,\lambda)\tilde{I}(x,y,\lambda)}{\sum_{x,y,\lambda}[I(x,y,\lambda)]^2} \quad (6)$$

The closer the value of NCC is to 1, the better it is.

For the multispectral images, the proposed algorithm performed on multispectral data and the number of retained bands corresponding to the highest of similarity index value of structural content (93.57%). However, which attains its maximum value for the Normalized Cross Correlation (98.77%), as well as the results obtained using all the bands of the original datasets. This fact reveals the total performance on the quantities of data stored by this set of the hybrid method. On the other hand, the criterion of NCC and SC reflects the potential of a reduction method to preserve the geometrical structure of the information, when the ratio of information value is nearest to the unit.

C. *Mean Absolute Error (MAE) based on* $\boldsymbol{l_1}$:

In statistics, the mean absolute error is a quantity used to measure how close forecasts or predictions are to the eventual outcomes. MAE is given by:

$$MAE(\mathbf{I},\tilde{\mathbf{I}}) = \frac{1}{\mathbf{n_x n_y n_\lambda}}\sum_{\mathbf{x,y,\lambda}}\left|\mathbf{I(x,y,\lambda)} - \tilde{\mathbf{I}}\mathbf{(x,y,\lambda)}\right| \quad (7)$$

D. *Mean Squared Error(MSE) based on* $\boldsymbol{l_2^2}$:

The MSE of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated.

$$MSE(\boldsymbol{I},\tilde{\boldsymbol{I}}) = \frac{1}{n_x n_y n_\lambda}\sum_{x,y,\lambda}\left(\boldsymbol{I(x,y,\lambda)} - \tilde{\boldsymbol{I}}\boldsymbol{(x,y,\lambda)}\right)^2 \quad (8)$$

E. *Maximum Absolute Distortion (MAD)*

This criterion used by Motta $\boldsymbol{l_\infty}$[34], the error bound over the entire image reduced by the reference $\boldsymbol{I(x,y,\lambda)}$ and reduced$\tilde{\boldsymbol{I}}\boldsymbol{(x,y,\lambda)}$.

$$MAD=\boldsymbol{l_\infty}\left(\boldsymbol{I}-\tilde{\boldsymbol{I}}\right) = \boldsymbol{max_{x,y,\lambda}}\{\left|\boldsymbol{I(x,y,\lambda)} - \tilde{\boldsymbol{I}}\boldsymbol{(x,y,\lambda)}\right|\} \quad (9)$$

This criterion reflects the potential of a reduction method to preserve information, of which the distance is the most distant of the value of average information. The hybridized algorithm stopped with an optimal number of 7 selected bands and a corresponding (72.84%) of mean absolute error, the maximum value for mean quadratic error(78.32 %) is attained by Fuzzy k means with PCA, and the highest provides Max absolute (83.63%). Thus, these three criteria (MAE, MSE and MQE) tested on the hybridized algorithm reflects a good performance to preserve the geometrical structure on reduced data volumes.

In the interval above, represented by the classification order of 33 spectral bands. Classification: [2 1 3 5 5 6 6 4 6 6 4 6 6 4 6 4 6 6 4 7 7 7 7 7 7 7 5 7 5 7 7 7]

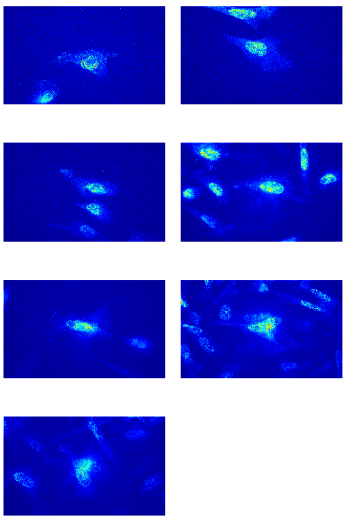

Figure.3. 7 bands reduced of Abnormally colon cells

TABLE I. THE VARIOUS REDUCTION METHOD OF HYBRIDIZATION USED, IN ORDER TO CATEGORIZE WITH NCC, SC, RATE MAE, RATE MQE , RATE MAD ON MULTISPECTRAL BIOLOGICAL IMAGES DATA.

MeanAbs	**72.84 %**
MeanQuad	**78.32 %**
MaxAbs	**82.44 %**
StrucContent	**93.57 %**
NormCC	**98.77%**

II. CONCLUSION AND PERSPECTIVES

The growth of high dimensional data creates a need of dimensionality reduction techniques to transform the data into a smaller, more manageable set which can be easily visualized. In this study, we have proposed a new approach for multispectral dimension reduction. This approach is unsupervised; i.e does not require either ground truth data or said number of retained bands. This hybridization method is very easily implementable and was found to provide encouraging results on real multispectral data. In general, the hybridized algorithm has several benefits: First, it does not require any ground truth knowledge; second, it automatically provides an estimate of the optimal number of bands for further classification purposes; third, it preserves the physical meaning of the multispectral data, which is useful in a band specification procedure for a given application. Finally, it is easy to implement. Future work can benefit by using the validity index of Bezdek's and thus, extend the applicability of dimensionality reduction techniques for hyperspectral data in comparison with the methods of classic projection.

References

[1] H. Grahn and P. Geladi, Eds., Techniques and Applications of Hyperspectral Image Analysis. Chichester, U.K.: Wiley, 2007.

[2] R. Bellman, Adaptive Control Processes. Princeton, NJ: Princeton Univ. Press, 1961

[3] Maaten L.J.P., Postma E.O. and Herik H.J. van den, 2007. Dimensionality reduction: A comparative review, Tech. rep. University of Maastricht.

[4] Davy Michael and Luz Saturnino, 2007. Dimensionality reduction for active learning with nearest neighbour classifier in text categorization problems, Sixth International Conference on Machine Learning and Applications, pp. 292-297.

[5] H. Hotelling. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24:417–441, 1933.

[6] J. Tenenbaum, V. de Silva, and J. Langford, J. A global geometric framework for nonlinear dimensionality reduction. Science, 290 (5500): 2319– 2323. (2000).

[7] Z. Zhang and H. Zha. Principal manifolds and nonlinear dimensionality reduction via local tangent space alignment. SIAM Journal of Scientific Computing, 26(1):313–338, 2004.

[8] C.K.I. Williams, Cand D. Barber. Bayesian classification with processes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20:12, 1342–1351. (1998).

[9] S. Lafon and A.B. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(9):1393–1403, 2006.

[10] C.K.I. Williams, Cand D. Barber. Bayesian classification with processes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20:12, 1342–1351. (1998).

[11] G. Hinton and S. Roweis. Stochastic neighbor embedding. Advances in Neural Information Processing Systems, vol. 15, MIT Press, Cambridge, MA, USA, 2003, pp. 833–840.

[12] X. He and P. Niyogi. Locality preserving projections. Advances in Neural Information Processing Systems, volume 16, page 37, Cambridge, MA, USA, 2004. The MIT Press.

[13] Yaozhang Pan , Shuzhi Sam Ge Abdullah Al Mamun. "Weighted locally linear embedding for dimension reduction". Pattern Recognition, Volume 42, Issue 5, May 2009, Pages 798-811.

[14] He Xiaofei, Deng Cai, Shuicheng Yan, Hong-Jiang Zhang. Neighborhood preserving embedding. Tenth IEEE International Conference on Computer Vision, Volume 2, Issue , 17-21 Oct. 2005

[15] D.L.Donoho and C. Grimes. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. PNAS, 100: 5591-5596. 2003.

[16] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput., 15: 1373-1396. 2003.

[17] L.van der Maaten, E. Postma, and H. van den Herik, "Dimensionality reduction:A comparative review," Tilburg Univ., Tilburg, The Netherlands,Tech. Rep. TiCC-TR 2009-005, 2009.

[18] Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982.

[19] B. Guo, R. Damper, S. Gunn, and J. Nelson, "A fast separabilitybased feature-selection method for high-dimensional remotely sensed image classification," Pattern Recognit., vol. 41, no. 5, pp. 1653–1662, May 2008.

[20] A. Martínez-Usó, F. Pla, J. M. Sotoca, and P. García-Sevilla, "Clusteringbased hyperspectral band selection using information measures," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 12, pp. 4158–4171, Dec. 2007.

[21] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," IEEE Geosci. Remote Sens. Lett., vol. 5, no. 4, pp. 564–568, Oct. 2008.

[22] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 2, pp. 153–158, Feb. 1997.

[23] B. Guo, S. Gunn, R. Damper, and J. Nelson, "Band selection for hyperspectral image classification using mutual information," IEEE Geosci. Remote Sens. Lett., vol. 3, no. 4, pp. 522–526, Oct. 2006.

[24] B.J. Frey and D. Dueck. Clustering by passing messages between data points. Science, 315:2007, 2007.

[25] J. Khoder, R. Younès, F.B. Ouezdou: " Stability of Dimensionality Reduction Methods Applied on Artificial Hyperspectral Images", Lecture Notes in Computer Science series, Volume 7594, August 2012.

[26] J. Khoder, R. Younès, F.B. Ouezdou: "Similarity of Dimensionality Reduction Methods Applied on Artificial Hyperspectral Images", The 2012 World Congress in Computer Science, Compute Engineering, and Applied Computing,pp.1208-1214 ,16-19 July, 2012, Las Vegas, Nevada, USA.

[27] J. Khoder, R. Younès: "Dimensionality Reduction on Hyperspectral Images: A Comparative Review Based on Artificial Datas", The 4th International Congress on Image and Signal, 15-17 October 2011, Shanghai, China.

[28] A. Chaddad, C.Tanougast, A. Dandache, A. Al Houseini, A. Bouridane. "Improving of Colon Cancer Cells Detection Based on Haralick□s Features on Segmented Histopathological Images." IEEE Conference on Computer Applications and Industrial Electronics, December 2011, pp.87 - 90.

[29] C. Cariou, K. Chehdi ans M. Le Loan. BandClust: An Unsupervised Band Reduction Method for Hyperspectral Remote Sensing. IEEE GeoScience and Remote Sensing Letters, Vol. 8, No. 3, May 2011.

[30] J.C.Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press New York, 1981.

[31] M. Eskicioglu and P. S. Fisher. A survey of quality measures for gray scale image compression. 9th Computing in Aerospace Conference, 93-4514. AIAA, oct. 1993.

[32] B. Lamiscarre and D. Léger. Nouveaux critères pour l'évaluation globale de la qualité des images ; applications aux images satellitaires restaurées ou comprimées. Rap. tech, CERT/DERO, sept. 1995