INFORMATION
TECHNOLOGY JOURNAL

A New Bands Selection Algorithm for Hyperspectral Image using Hyperspectral Derivative on Clifford Manifold

Yanshan Li

ATR National Key Lab. of Defense Technology, Shenzhen University, Shenzhen, 518060, China

Abstract: In order to select best bands from the hyperspectral image with high computation efficiency, this paper proposed a new bands selection algorithm of hyperspectral image using the hyperspectral derivative on Clifford manifold. It firstly analyze the hyperspectral image in the Clifford algebra which had high efficiency on computation and analysis. Firstly, it discusses the properties of the Clifford algebra and Clifford manifold. Secondly, it gives the definitions of the hyperspectral derivative and the curve change rate on high dimensional space in Clifford algebra for finding the key points. Based on the theories mentioned above, a new bands selection algorithm is given with detail steps. At last, the results of the experiments are shown. Compared with the existing algorithms, the proposed algorithm is highly efficient in respect of computation and time. The study provided a new method for bands selection algorithm of hyperspectral image with the theories of Clifford algorithm.

Key words: Hyperspectral image processing, Clifford algebra, bands selection, manifold, hyperspectral derivative

INTRODUCTION

With the development of the hyperspectral image processing technology, this technology has been applied to more and more fields, such as remote sensing (Landgrebe, 2000; Mazzoni *et al.*, 2007), medicine (Rajpoot and Rajpoot, 2004), military, etc. The main reason is that the hyperspectral image can provide high spectral information which is useful for analyzing the composition of material. However, it contains huge redundant data. The redundant data not only increases the difficulty of data processing but also reduces the efficiency of data processing. Because the traditional image processing methods are developed on the 2D or 3D images, the traditional image processing method is not feasible for the high dimensional image. It is necessary to reduce the band size of the hyperspectral image. In recent years, the technology of hyperspectral image processing has become a new research hotspot of the image processing (Mianji *et al.*, 2009; Mianji *et al.*, 2008; Windham *et al.*, 2005; Misman *et al.*, 2010). This study introduces the Clifford algebra theory for hyperspectral image processing to provide a new way.

Bands selection is a very important way to reduce the dimension of hyperspectral image besides the feature extraction. It figures out the bands subset of hyperspectral data. The selected bands yield the best performance based on some criteria which are usually linked to some measure of end-performance. Compared to the feature extraction, bands selection has advantage of saving the original information of the bands. There are many algorithms presented for feature selection. The Jeffreys-Matusita distance (JM-distance) (Swain and King, 1973; Bruzzonne *et al.*, 1995) is proposed based on divergence and it derives two criteria, the average separability criterion and the saturating transform of divergence for bands election. Keshava (2001) quantifies the distance between the spectra of two materials at corresponding spectral bands and then analyzes the separability of these two spectra using the Spectral Angle Mapper (SAM) metric. The spectral bands are selected in order to maximize the SAM metric, that is, the angle between the two spectra. Stein *et al.* (1999) apply Spectral Matched Filter (SMF) which is the likelihood ratio detection statistics for a known additive signal in a Gaussian background. They then evaluate the SNR losses and select spectral bands in order to optimize an objective function which is defined in terms of probability of detection. Kira and Rendell (1992) assign weights to features individually to indicate their relevance to a given task and select features with greater weights. Briefly, these algorithms including many other bands selection algorithms have good performance in specified tasks. The feature selection method based on Hybrid Genetic Algorithm (HGA) can select feature subsets which not only contain fewer features but also provide better detection performance for steg-analysis (Xia *et al.*, 2009). However, there are still some challenges in applying these techniques effectively, such as computational cost, timely cost and presence of local minima problems, etc.

This study try to develop a new bands selection algorithm with less computational cost and timely cost based on Clifford algebra.

CLIFFORD ALGEBRA

Basic of Clifford algebra: Clifford algebra (Li *et al.*, 2001; Sommer, 1998; Penrose and Rindler, 1984) provides an efficient framework to conduct the computations without coordinates. It can reduce the computation complexity and improve the efficiency of the computation. This algebraic approach contains all geometric operators and permits specification of constructions in a coordinate-free manner. It has widely been used in physics (Byrnes, 2006), computer vision (Schlemmer, 2004; Brackx and Schepper, 2005; Yanshan, 2008) and so on, due to two main advantages detailed as follows (Batard *et al.*, 2009):

- Clifford algebra is generally known that Clifford algebra provides a very efficient framework to conduct computations without coordinates. It can be used in the inner, wedge and geometric products as well as the related geometric interpretations of these operations. More precisely, the acquisition space of the image, i.e. the space where the image gets its values, is embedded into Clifford algebra. The local variations of the pixel values can be easily measured using geometric information and transformations
- Clifford algebra provides very convenient ways for computing in high dimensional space. Clifford algebra contains elements of different degrees (scalar, vectors, bivector, pseudoscalar, etc.). This allows combining information of different natures in a single multivector which process data as a global and concise way

Let $\mathbb{R}_n$ be n-dimensional Euclidean space, with the positive definite inner product $\langle .,. \rangle$. $\mathbb{R}_n$ has an orthogonal basis $e_1, e_2, \ldots, e_n$, such that $\langle e_i, e_j \rangle = \delta_{ij}$.

The Clifford algebra $\alpha(\mathbb{R}, n)$ is the free algebra generated by $\mathbb{R}_n$ modulo as the relation:

$$X^2 = \langle X, X \rangle \tag{1}$$

It can be considered as being the exterior algebra $\wedge \mathbb{R}^n$, where the extra information defining the inner product is included. From this point of view one has:

$$xy = x \wedge y - \langle x, y \rangle \tag{2}$$

for arbitrary vectors. Clearly, for $n>1$, $\alpha(\mathbb{R}, n)$ is not commutative.

For practical reasons, it can simple $\alpha(\mathbb{R}, n)$ to $\alpha(n)$ which then contains elements of the form a+ib.

where, a and b are from $\alpha(\mathbb{R}, n)$. Multiplication then assumes that i commute with all elements of the Clifford algebra. Based on $\alpha(n)$ we can define an anti-automorphism by:

$$\bar{X} = -X, \bar{i} = -i, \overline{ab} = \bar{b}\bar{a} \tag{3}$$

for arbitrary x in $\mathbb{R}_n$ and a and b in $\alpha(n)$. There is a natural identification of $\mathbb{C}$ as a subspace of the finite-dimensional algebra $\alpha(n)$ and the scale part in a Clifford number is clearly presented which will be denoted as $[a]_0$.

Clifford manifold: A vector manifold (Hestenes and Sobczyk, 1984) M is a set of vectors called points of M with certain properties to be described presently. A vector a(x) is said to be tangent to a point x in M if there is a curve $\{z(\tau); 0<\tau<\varepsilon\}$ in M extended from the point x(0) = x such that:

$$a(x) = a \cdot \partial x \equiv \left. \frac{dx(\tau)}{d\tau} \right|_{\tau=0} = \lim_{\tau \to 0} \frac{x(\tau) - x}{\tau} \tag{4}$$

If the chord of a curve are not null vectors, the curve can be parameterized by the magnitudes of its chords, $\sigma = \sigma(\tau) \equiv |x(\tau)-x|$, so:

$$a = \frac{d\sigma}{d\tau} \frac{dx}{d\sigma}, |a| = \frac{d\sigma}{d\tau}, \hat{a} = \frac{dx}{d\sigma} = \frac{dx}{|dx|}$$

The set A(x) of all vectors tangent to M at x is called the tangent space at x. At each interior point x of M, the tangent space A(x) is an m-dimensional vector space and M denotes an m-dimensional manifold.

The tangent space A(x) at an interior point generates a unique Clifford algebra which we call the tangent algebra of M at x and denote by G (x)≡G(A(x)). We assume that A(x) is nonsingular in the algebraic sense that it possesses a unit pseudoscalar I = I(x) which we call the (unit) pseudoscalar of M at x.

HYPERSPECTRAL DERIVATIVE IN $\alpha(\mathbb{R}, n)$

Representation for hyperspectral data set: The spectrum recorded is a combination of the actual spectrum of the real surface, modulated by the effects of the solar curve and the complexity of the ground. The spectrums of the points belonged to same class are usually different, since the effects on different points are different. Figure 1 shows the spectrums of the points selected from the water area. It is apparently that the curves belonged to a same
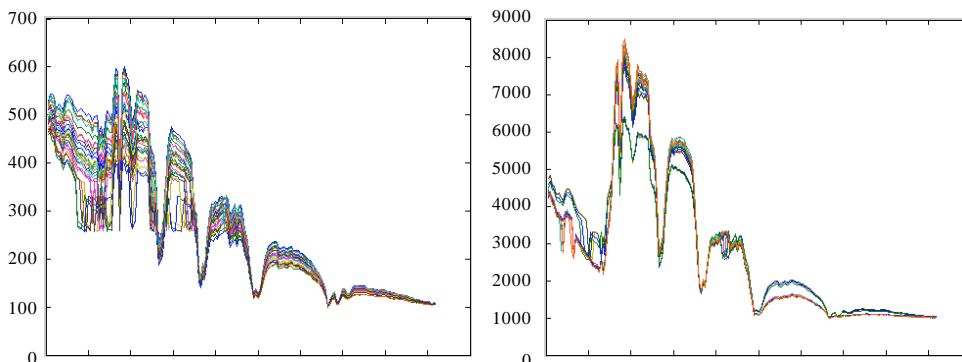
Fig. 1: The spectral curves of some materials. Every point of the hyperspectral image form a individual curve, one material always forms many curves

material are different and It is also shown that just only using one spectrum of one point or the average spectrums value of one class can hardly represent the whole spectral property of one class. In our work, we represent the data belonged to one class with a curve in high dimensional space, so the spectral properties of all samples can entirely be used in the feature selection.

Let us consider a hyperspectral image with C class labels and n spectral bands. We represent the data from one class with a curve in high dimensional space, so it produces C curves in the space whose dimension is the maximum count of the samples in classes.

Given a set of samples $X = \{x_1, x_2,\dots, x_n\} \in \mathbb{R}^{d \times n}$, x_i is the pixel of the hyperspectral image. We define a map function $\varphi := \mathbb{R}^n \to \mathbb{R}^m$, where m is the maximum count of the samples in classes. We assume that there are C curves on the manifold M, denoted by f_i, where $1 \le i \le C$. The curve f_i include the samples labeled by class i, denoted by $S = \{s_1, x_2,\dots, s_d\}$, where $s_i = B_i(x_{c1})e_1 + B_i(x_{c2})e_2 + \dots + B_i(x_{cm})e_m$ and $B_i(x_{cm})$ represents the ith band of the pixel x_{cm}. The tangent space is denoted by A(x).

Hyperspectral derivative in $\alpha(\mathbb{R}, n)$: The normal first order derivative can be estimated by:

$$\left.\frac{ds}{d\lambda}\right|_i \approx \frac{s(\lambda_i) - s(\lambda_j)}{\Delta\lambda} \qquad (5)$$

where, the s is the spectral function, λ_i is the wavelength of the ith band, $s(\lambda_i)$ is the reflectance value at wavelength λ_i, $\Delta\lambda = \lambda_j - \lambda_i$ and $\lambda_j > \lambda_i$.

From the hyperspectral curve in $\alpha(\mathbb{R}, n)$, we give the definition of the first order derivative as below:

$$\underline{S} = \left.\frac{dS}{d\lambda}\right|_i \approx \frac{D(S(\lambda_i), S(\lambda_j))}{\Delta\lambda} \qquad (6)$$

where, S is the high dimensional curve of some samples with same class label, D(.) is the distance metric function between the $S(\lambda_i)$ and $S(\lambda_j)$ and $\Delta\lambda = \lambda_j - \lambda_i$.

There are many distance metric existed in high dimensional space. In this study, we use the distance metric function based on the Clifford algebra as follows:

$$D(S(\lambda_i)), S(\lambda_j)) = |(S(\lambda_i) \wedge S(\lambda_j))/S(\lambda_j)| \qquad (7)$$

Using Eq. 6, one can compute derivatives of the hyperspectral image simply and identify the derivative features of one target classes. These features can then be included in the bands or feature selection operation to improve the efficiency. It is important to remember that no new information is created by using derivatives. The purpose is to identify the helpful spectral features that may be too subtle to be captured by other methods, so that the data dimension of the image can be kept, also was possible and still achieve accurate classification results.

BANDS SELECTION ALGORITHM USING HYPERSPECTRAL DERIVATIVE

Key points of hyperspectral curve: Suppose pixels belonging to one target class form a curve on the Clifford manifold, there must be many points which can identify the curve from other curves belonging other classes. These points are called key points in this study. the key points are noted by $X_{key} = \{x_{key1}, x_{key2},\dots, x_{keyk}\} \subset X$. In order to determine whether a band point is a key point, a criterion is defined by:

$$\psi = \frac{1}{\prod\limits_{j=1}^{n} CR\left(x_c, x_j\right)} \left(\frac{|x_c - x_{c-1}| + |x_{c+1} - x_c|}{\sum\limits_{i=1}^{n-1}|x_{i+1} - x_i|} \right) \tag{8}$$

where, $CR(x_i, x_j) = |\underline{S}|$ and $\forall x_i, x_j, x_c \in X$) belong to one target class.

Equation 8 can only give the geometric features of one class but it can not measure the separability of two classes. In order to select out the appropriate bands, the separability measure is necessary. In this study the separability measure of a point x_c among classes can be computed by:

$$\varphi = \sum\limits_{j=1}^{d}|x_c - x_j|) \tag{9}$$

Algorithm: Suppose that the training data set contains d classes and n bands, k optimal bands is required to be selected. $X_{key} = \{x_{key1}, x_{key2}, \ldots, x_{keyk}\} \subset X$ is the set of the key points and $X_s = \{x_{s1}, x_{s2}, \ldots, x_{sm}\} \subset X$ is the selected band points. So, The detailed steps of bands selection framework is as below.

Step 1: Assume d curves on the manifold from d classes training data.
Step 2: for i=1:d
    Compute the value of Eq. 8 and pick up m key points X_{keyi}.
    End
Step 3: Get the intersect X_{key} of $X_{key, i}$ by $X_{key} = \bigcap\limits_{i=1}^{d} X_{keyi}$. The amount of

the elements in X_{key} is k1. Get the union $\overline{X}_{key}$ of $X_{key, i}$ by $\overline{X}_{key} = \bigcup\limits_{i=1}^{d} X_{keyi}$.

The amount of the elements in $\overline{X}_{key}$ is k2.
Step 4: If m≤k1 than
    Compute the separability measure of a point x_c in X_{key} by Eq. 9
    and pick up m elements with largest values.
    Else if m≤k2 than
    Compute the separability measure of a point x_c in $\overline{X}_{key}$ by
    Eq. 9 and pick up m elements with largest values.
    End if

EXPERIMENTAL RESULTS AND ANALYSIS

Dataset description: The hyperspectral dataset used in our experiments is a section of a scene taken over northwest Indiana's Indian Pines by the AVIRIS sensor in 1992. From the 220 spectral channels acquired by the AVIRIS sensor, 20 channels were discarded because affected by atmospheric problems. From the 16 different land-cover classes available in the original ground truth, seven were discarded; since only few training samples were available for them (this makes the experimental analysis more significant from the statistical viewpoint). The remaining nine land-cover classes were used to generate a set of 4757 training samples (used for learning the classifiers) and a set of 4588 test samples (exploited

for assessing their accuracies). The experiments were run on a computer with Intel core 2 Dou CPU E7400.

Experimental results and analysis: Several experiments with many different bands selection algorithms are completed on the same hyperspectral image. Figure 2 shows the experimental result of different bands selection algorithms including the proposed framework, Sequential Forward Selection (SFS) and ReliefF. It shows that all the average recognition correct rates using the Bayesian classification on the selected bands. The recognition correct rate are raised when the number of the selected bands is rising. The rate of our method is higher than the SFS and ReliefF, especially when the number of the selected bands is small. Figure 3 shows the time cost of
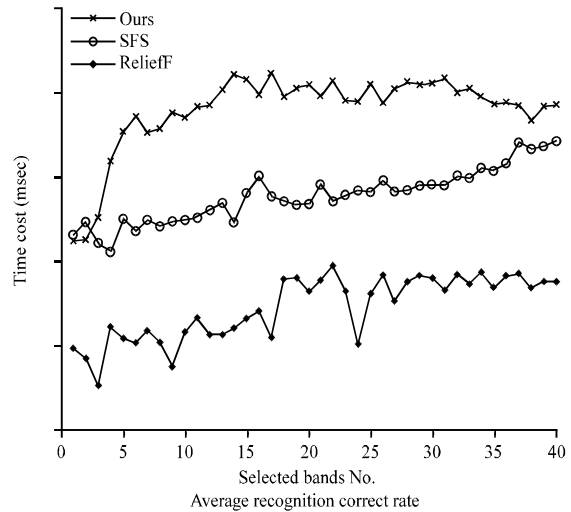

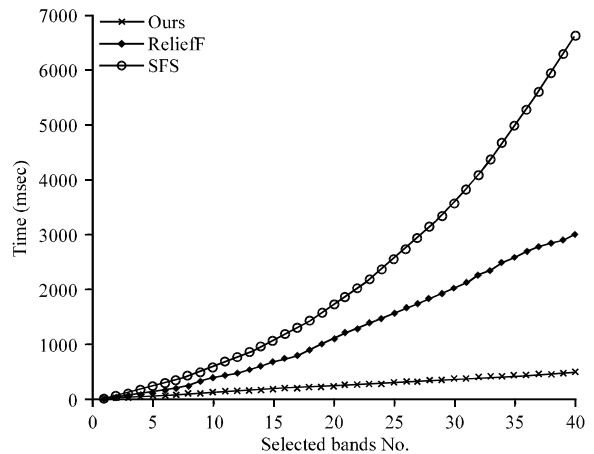Fig. 2: The time cost of ours, SFS and ReliefF

Fig. 3: Average recognition correct rates algorithm, using the Bayesian classification on the selected bands which selected by our algorithm, SFS, and ReliefF
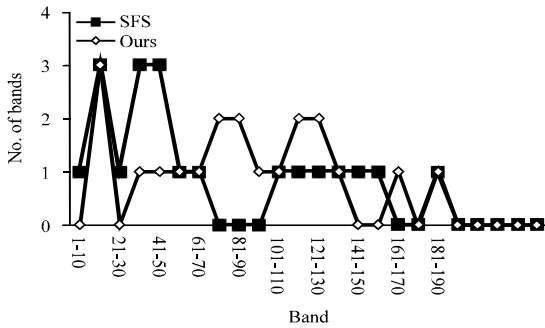
Fig. 4: Distribution of the selected bands

our algorithm, SFS and ReliefF. It is apparent that the time cost of our algorithm is smaller than the other algorithms and with the number of selected bands rising, the advantage of our algorithm in the time cost is more and more greater. When the number of selected band is 40, the time cost of our algorithm is about 17% of the ReliefF and about 8% of the SFS.

Figure 4 shows the distribution of the selected bands by our method and SFS. It shows that the selected bands are distributed evenly on the whole spectrum.

CONCLUSIONS

This study proposed a new bands selection algorithm for hyperspectral image using the hyperspectral derivative on Clifford manifold. Firstly, the properties of the hyperspectral image in the Clifford algebra are analyzed and then a new bands selection algorithm for hyperspectral image is represented. Compared with the existing algorithm, the proposed algorithm reduces the computational time considerably. However, there are still many further studies worth to do, including the analysis technologies with Clifford algebra theories, Clifford manifold theories, bands selection algorithm in high dimension space and so on.

REFERENCES

Batard, T., C. Saint-Jean and M. Berthier, 2009. A metric approach to nD images edge detection with Clifford algebras. J. Math. Imag. Vision, 33: 296-312.

Brackx, F. and N. de Schepper, 2005. The Clifford-Fourier transform. J. Fourier Anal. Applied, 11: 669-681.

Bruzzonne, L., F. Roli and S.B. Serpico, 1995. An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection. IEEE Trans. Geosci. Remote Sens., 33: 1318-1321.

Byrnes, J., 2006. Computational Noncommutative Algebra and Applications. Springer, Netherlands, pp: 197-225.

Hestenes, D. and G. Sobczyk, 1984. Clifford Algebra to Geometric Calculus. D. Reidel Publ. Co., Dordrecht/Boston.

Keshava, N., 2001. Best bands selection for detection in hyperspectral processing. Acoustics Speech Signal Process., 5: 3149-3152.

Kira, K. and L.A. Rendell, 1992. The feature selection problem: Traditional methods and a new algorithm. Proceedings of the Tenth National Conference on Artificial Intelligence, July 12-16, MIT Press, San Jose, CA. Cambridge, MA., pp: 129-134.

Landgrebe, D., 2000. Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data. In: Information Processing for Remote Sensing, Chen, C.H. (Ed.). World Scientific Publishing Co., River Edge.

Li, H., D. Hestenes and A. Rockwood, 2001. Generalized Homogeneous Coordinates for Computational Geometry. In: Geometric Computing with Clifford Algebra, Sommer, G. (Ed.). Springer, Heidelberg, pp: 27-60.

Mazzoni, D., M.J. Garay, R. Davies and D. Nelson, 2007. An operational MISR pixel classifier using support vector machines. Remote Sens. Environ., 107: 149-158.

Mianji, F.A., Y. Zhang and A. Babakhani, 2009. Optimum method selection for resolution enhancement of hyperspectral imagery. Inform. Technol. J., 8: 263-274.

Mianji, F.A., Y. Zhang, H.K. Sulehria, A. Babakhani and M.R. Kardan, 2008. Super-resolution challenges in hyperspectral imagery. Inform. Technol. J., 7: 1030-1036.

Misman, M.A., H.Z.M. Shafri and R.M.K.R. Ahmad, 2010. Effects of hyperspectral data transformations on urban inter-class separations using a support vector machine. J. Applied Sci., 10: 2241-2250.

Penrose, R. and W. Rindler, 1984. Spinors and Space-Time. Cambridge University Press, UK.

Rajpoot, K.M. and N.M. Rajpoot, 2004. Hyperspectral Colon Tissue Cell Classification. SPIE Medical Imaging (MI), San Diego, California.

Schlemmer, M., 2004. Fourier Transformation and Filter Design for Clifford Convolution. University of Kaiserslautern, Kaiserslautern.

Sommer, G., 1998. Geometric Computing with Clifford Algebras. Springer, Heidelberg.

Stein, D.W.J., S.E. Stewart, G.D. Gilbert and J.S. Schoonmaker, 1999. Band selection for viewing under water objects using hyperspectral sensors. Proc. SPIE, 3761: 50-50.

Swain, P. and R. King, 1973. Two effective feature selection criteria for multispectral remote sensing. Proceedings of the International Joint Conference on Pattern Recognition, October 30-November 1, 1973, Washington, DC., pp: 536-540.

Windham, W.R., D.P. Smith, M.E. Berrang, K.C. Lawrence and P.W. Feldner, 2005. Effectiveness of hyperspectral imaging system for detecting cecal contaminated broiler carcasses. Int. J. Poult. Sci., 4: 657-662.

Xia, Z., X. Sun, J. Qin and C. Niu, 2009. Feature selection for image steganalysis using hybrid genetic algorithm. Inform. Technol. J., 8: 811-820.

Yanshan, L., 2008. A digital image watermarking algorithm based on clifford algebra. Acta Elect. Sinica, 36: 852-855.