The Third AFSS(1998). 314-319

Band Feature Extraction of Normal Distributive Multispectral Image Data using Rough Sets.

Hwanmook Chung * Sunghyun Won **

* : Faculty of Electronic & Information Engineering, Catholic University of Taegu-Hyosung

** : Department of Computer Information Processing, Jisan College

Abstract

In this paper, for efficient data classification in multispectral bands environment, a band feature extraction method using the Rough sets theory is proposed. First, we make a look up table from training data, and analyze the properties of experimental multispectral image data, then select the efficient band usin indiscernibility relation of Rough sets theory from analysis results. Proposed method is applied to LANDSAT TM data on 2. June, 1992. Among them, normal distributive data were experimented, mainly. From this, we show clustering trends that similar to traditional band selection results by wavelength properties, from this, we verify that can use the proposed method that centered on data properties to select the efficient bands, though data sensing environment change to hyperspectral band environments.

1. Introduction

Processing techinques of remote sensed image data using computer have been recognized very necessary techniques to all social fields, such as, environmental observation, land cultivation, resource investigation, military trend grasp and agricultural product estimation, etc. Especially, accurate classification and analysis to remote sensed image data are important elements that can determine reliability of remote sensed image data processing systems, and many researches have been processed to improve these accuracy of classification and analysis. Traditionally, remote sensed image data processing systems have been processed 2 or 3 selected bands in multiple bands, in this time, their selection criterions are statistical separability or wavelength properties. But, it have be bring up the necessity of bands selection method by data distribution characteristics than traditional bands selection by wavelength properties or statistical separability. Because data sensing environments change from multispectral environments to hyper-spectral environments.

In this paper, for efficient data classification in multispectral bands environment, a band feature extraction method using the Rough sets theory is proposed. First, we make a look up table from training data, and analyze the properties of experimental multispectral image data, then select the efficient band using indiscernibility relation of Rough sets theory from analysis results. Proposed method is applied to LANDSAT TM data on 2, June, 1992. Among them, normal distributive data were experimented, mainly. From this, we show clustering trends that similar to traditional band selection results by wavelength properties, from this, we verify that can use the proposed method that centered on data properties to select the efficient bands, though data sensing environment change to hyperspectral band environments.

2. Processing of multispectral image data

(Fig. 1) shows general procedure of multispectral image data. Image data from satellite will be processed various intermediate procedures, and must be classification accuracy, finally. In this paper, we studies efficient band selection, that is, band feature extraction.

data aquisition from satellite

⇓

preprocessing

⇓

define of category class

⇓

selection of training field

⇓

band feature extraction

⇓

applying classification methdology to image data

⇓

evaluation of classification accuracy

(Fig. 1) General procedure of multispectral image data

3. Rough set theory

3.1 Feature of Rough set theory

The strong points from system using Rough sets theory are as follows.

First, it is easy that can develop the efficient algorithm for finding hidden patterns in data. Second, it is easy that make minimum set eliminating redundancy of pattern characteristic in data. Third, it can evaluate meaning or importancy. Fourth, it is easy that generate determinance rule set from data. Fifth, it is easy that understand processing procedure. Sixth, it is simple that analyze the gained result. Finally, it is very useful such as parallel processing system.

Most of these characteristics are related to analysis or evaluation to characteristics of data. Therefore, if we use the Rough sets theory, it is able to efficient processing from incomplete data, especially, classification of pattern characteristics, that is, it is very useful clustering.

3.2 Basic concept of Rough set theory

Assume that there is attribute set Q and each attribute of n elements.

$$Q = \{q_1, q_2, \cdots, q_n\} \qquad (1)$$

And, assume that there is set X to m objects that become classification object and their elements $x_1, x_2, \cdots, x_m$.

$$X = \{x_1, x_2, \cdots, x_m\} \qquad (2)$$

Also, let V_{q_j} is the set of q_j that express the attribute value.

$$V_{q_j} = \{\alpha, \beta, \cdots, \omega\} \qquad (3)$$

(where, $j = 1, 2, \cdots, n$)

Then, attribute value description function P_X to these is as follows.

(Def. 1) attribute value description function P_X

$$P_X : Q \rightarrow V : P_X(q_i) = P(X, q_i) \quad (4)$$

(where, $Q \rightarrow V$ is the mapping from attribute set Q to attribute value set V)

And, indiscernibility relation ind(Q) that can't discriminate two set of object is defined as follows.

(Def. 2) indiscernibility relation ind(Q)

If object x_i and x_j are indiscernibility relation to any attribute q_i, then we note that as follows.

$$P(x_i, q_i) = P(x_j, q_i)$$

$$P(x_i, x_j) \in ind(q_i)$$

$$R = ind(Q) \qquad (5)$$

(where, R is equivalence relation that two object sets x_i and x_j are indiscernibility relation each other)

Therefore, it shows as follows in case of x_i, x_j can't be discriminated by $P \subset Q$.

$$(x_i, x_j) \in ind(P)$$

$$ind(P) = \bigcap_{q_i \in P} ind(q_i) \qquad (6)$$

Here, if $P = Q$ and $(x_i, x_j) \in ind(q_i)$ then x_i and x_j become indiscernibility relation, partition by $ind(q_i)$ is quotient set, because that is equivalence relation.

$$X/ind(q_i) = \{[x_i] \mid x_i \in X\} \qquad (7)$$

3.3 Construction of equivalence class using Rough set theory

For classifying objects using indiscernibility relation of Rough sets theory, it must be defined attribute set and object set, first. Let attribute set Q and object set X be defined as like formula (1), (2). Then, attribute value $AV(x_i)$ that object x_i has is one of the q_js.

$$q_j\{AV(x_i)\} = \text{one of } \{\alpha, \beta, \cdots, \omega\} \qquad (8)$$

Therefore, indiscernibility relation object to attribute q_k can be gotten following formula.

$$X/ind(q_k) = \{x_i, x_j \mid q_k\{AV(x_i)\} = q_k\{AV(x_j)\}\}$$

(where, $i, j = 1, 2, \cdots, m$) (9)

If q_j, q_k are two attributes of classification criterion, indiscernibility relation object to them can be gotten following formula.

$$X/ind(q_j, q_k) = \{x_i, x_j \mid (q_j, q_k)\{AV(x_i)\} = (q_j, q_k)\{AV(x_j)\}\}$$
$$(\text{where, } i, j = 1, 2, \cdots, m) \quad (10)$$

indiscernibility relation object for all attribute Q can be gotten following formula.

$$X/ind(Q) = \{x_i, x_j \mid Q\{AV(x_i)\} = Q\{AV(x_j)\}\}$$
$$(\text{where, } i, j = 1, 2, \cdots, m) \quad (11)$$

4. Band feature extraction of multispectral image data using Rough set theory

4.1 Look up table of multispectral image data

(Fig. 2) is the look up table of spectral intensity that means the relation between classes and bands to n frames. At each frames, $\mathbb{C}$ means the set of classes from C_1 to C_m. That is, it means the class that can be belong to any class for example, forest, water and so on. Also, $\mathbb{B}$ is the set of bands from B_1 to B_7. Although the goal of this study is the band feature extraction in hyperspectral environment, but theoretical basis was made in multispectral environment. v_{111}, v_{27m} are spectral intensity that must be belong to special class(v_{111} is class 1, v_{27m} is class m) at special frame(v_{111} is first frame, v_{27m} is second frame), special band(v_{111} is band 1, v_{27m} is band 7).

(Fig. 3) is the table of mean(μ) and standard deviation(σ) of same class and band. μ_{11} is the mean of spectral intensity that is appeared to class 1(C_1) at band 1(B_1) for n frames, σ_{11} is the standard deviation of spectral intensity that is appeared to class C_1 at band B_1 for n frames.

frame 1

B \ C	C_1	C_2	$\cdots$	C_m
B_1	v_{111}	v_{112}	$\cdots$	v_{11m}
B_2	v_{121}	v_{122}	$\cdots$	v_{12m}
.	.			.
.	.			.
B_7	v_{171}	v_{172}	$\cdots$	v_{17m}

frame 2

B \ C	C_1	C_2	$\cdots$	C_m
B_1	v_{211}	v_{212}	$\cdots$	v_{21m}
B_2	v_{221}	v_{222}	$\cdots$	v_{22m}
.	.			.
.	.			.
B_7	v_{271}	v_{272}	$\cdots$	v_{27m}

$\cdots$

frame n-1

B \ C	C_1	C_2	$\cdots$	C_m
B_1	v_{n-111}	v_{n-112}	$\cdots$	v_{n-11m}
B_2	v_{n-121}	v_{n-122}	$\cdots$	v_{n-12m}
.	.			.
.	.			.
B_7	v_{n-171}	v_{n-172}	$\cdots$	v_{n-17m}

frame n

B \ C	C_1	C_2	$\cdots$	C_m
B_1	v_{n11}	v_{n12}	$\cdots$	v_{n1m}
B_2	v_{n21}	v_{n22}	$\cdots$	v_{n2m}
.	.			.
.	.			.
B_7	v_{n71}	v_{n72}	$\cdots$	v_{n7m}

where, set of classes $\mathbb{C} = \{C_1, C_2, \cdots, C_m\}$
set of bands $\mathbb{B} = \{B_1, B_2, \cdots, B_7\}$
v_{ijk} is trained spectral intensification that must be belong to class k at jth band of ith frame

(Fig. 2) Look up table of multispectral image data

B \ C	C_1	$\cdots$	C_m
B_1	$[\mu_{11}-\sigma_{11}, \mu_{11}+\sigma_{11}]$	$\cdots$	$[\mu_{1m}-\sigma_{1m}, \mu_{1m}+\sigma_{1m}]$
B_2	$[\mu_{21}-\sigma_{21}, \mu_{21}+\sigma_{21}]$	$\cdots$	$[\mu_{2m}-\sigma_{2m}, \mu_{2m}+\sigma_{2m}]$
.	.		.
.	.		.
.	.		.
B_7	$[\mu_{71}-\sigma_{71}, \mu_{71}+\sigma_{71}]$	$\cdots$	$[\mu_{7m}-\sigma_{7m}, \mu_{7m}+\sigma_{7m}]$

where, μ_{ij} is mean of spectral intensity that must be belong to class j at band i for n frames
σ_{ij} is standard deviation of trained spectral intensity that must be belong to class j at band i for n frames

(Fig. 3) Permitted error limitation of multispectral image data

4.2 Class reversion of remote sensed pixels

(Fig. 4) shows the table of spectral intensity for remote sensed data. Here, $\mathbb{P}$ is the set of remote sensed pixel P_1, P_2, $\cdots$, P_k. Also, pv_{11} means spectral intensity of pixels P_1 at band 1.

B \ P	P_1	P_2	$\cdots$	P_k
B_1	pv_{11}	pv_{12}	$\cdots$	pv_{1k}
B_2	pv_{21}	pv_{22}	$\cdots$	pv_{2k}
.		.		.
.		.		.
B_7	pv_{71}	pv_{72}	$\cdots$	pv_{7k}

where, set of sensed pixel $\mathbb{P} = \{P_1, P_2, \cdots, P_k\}$
pv_{ij} is spectral intensity of pixel j at band i

(Fig. 4) Spectral intensity of pixels for each bands

We make each pixel belong to special class as like (Fig. 6) from (Fig. 2), (Fig. 3). In this time, the reversion rule is as follows. First, if it include spectral intensity of pixel on range (Fig. 3), then belong to that class, otherwise, belong to nearest class. This algorithm is (Fig. 5).

```
for i = 1 to 7 by 1
 for j = 1 to k by 1
  for n = 1 to m by 1
   if (μin - σin ≤ pvij ≤ μin - σin) then Category = Cn
                                      exit
   end
   if (μin - σin > pvij or pvij > μin - σin) then
    for n = 1 to m by 1
     Category = Cn of min(μin, pvij)
    end
  end
end
```

(Fig. 5) class reversion algorithm

(Fig. 6) is the result that belong to special class for each sensed pixel by algorithm (Fig. 5).

B \ P	P_1	P_2	…	P_k
B_1	C_1	C_n	…	C_2
B_2	C_2	C_1	…	C_1
B_3	C_{n-1}	C_n	…	C_1
B_4	C_2	C_n	…	C_3
B_5	C_2	C_1	…	C_2
B_6	C_{n-1}	C_3	…	C_n
B_7	C_1	C_2	…	C_n

(Fig. 6) Class reversion result of remote sensed pixels

4.3 Generation and Analysis of equivalence class using Rough set theory

From (Fig. 6) was made by (Fig. 2) and (Fig. 3), we generate equivalence class using Rough sets theory as like (Fig. 7).

B \ P	P_1	P_2	…	P_k
B_1	C_1	C_n	…	C_2
B_2	C_2	C_1	…	C_1
B_3	C_{n-1}	C_n	…	C_1
B_4	C_2	C_n	…	C_3
B_5	C_2	C_1	…	C_2
B_6	C_{n-1}	C_3	…	C_n
B_7	C_1	C_2	…	C_n

(Fig. 7) Equivalence class of remote sensed pixels(for only P_1)

From this, extracting equivalence class is as follows.

$\mathbb{B}/ind(P_1)=\{[B_1,B_7],[B_2,B_4,B_5],[B_3,B_6]\}$ (12)

$\mathbb{B}/ind(P_2)=\{[B_1,B_3,B_4],[B_2,B_5],[B_6,B_7]\}$ (13)

$\mathbb{B}/ind(P_k)=\{[B_1,B_5],[B_2,B_3],[B_4],[B_6,B_7]\}$ (14)

$\mathbb{B}/ind(\mathbb{P})=\{[B_1],[B_2],[B_3],[B_4],[B_5],[B_6],[B_7]\}$ (15)

Therefore, band 1 and band 7 are indiscernibility relation to pixels P_1, each other, and band 2, band 4, band 5 are indiscernibility relation, too.

5. Experimentation and discussion of result

5.1 Experimental object

Experimental zone is LANDSAT TM data at near the Han River on 2, June, 1992. LANDSAT TM data has better spectral resolution than SPOT HRV data and, it is easy to get, our nation. Land cover size of LANDSAT TM per scene 170km(vertical)×185km(horizontal), the number of pixels are 5,965×6,920, this experimentation use the part of whole zone, so, 136×136 = 18,496 pixels were experimented. Although experimental zone has many classes, but we select four kind of class among them, that is, water, crop, urban, forest.

5.2 Distributive characteristics of training data

There are two kind of data for experimentation, that is training data and experimental data. Training data means data for training. Experimental data means that the data were sensed remotely at special zone for doing experimentation using suggested method.

Distributive characteristics of training data are as following <Table 1>~<Table 4>.

<Table 1> pattern characteristic of water

	band1	band2	band3	band4	band5	band6	band7
min S.I.	103	44	44	29	12	136	5
max S.I.	120	50	76	33	30	161	9
mean	109.4	46.9	48.9	30.1	16.4	138.5	7.0
S.D.	1.38	1.41	1.12	0.56	0.38	0.48	0.94
medium value	109	47	48	30	17	138	7

<Table 2> pattern characteristic of crop

	band1	band2	band3	band4	band5	band6	band7
min S.I.	103	47	48	46	21	156	8
max S.I.	128	52	88	60	156	176	27
mean	108.1	49.4	57.2	52.3	58.6	161.2	14.4
S.D.	0.67	0.99	1.78	2.80	5.31	1.77	4.28
medium value	108	49.5	57	52	60	161	14

<Table 3> pattern characteristic of forest

	band1	band2	band3	band4	band5	band6	band7
min S.I.	101	56	43	61	21	147	47
max S.I.	134	72	101	95	194	182	122
mean	109.6	64.5	65.9	82.7	106.6	167.3	91.3
S.D.	6.66	3.06	9.9	6.57	12.4	8.41	14.44
medium value	109	64	66	83	107	167	94

<Table 4> pattern characteristic of urban

	band1	band2	band3	band4	band5	band6	band7
min S.I.	99	41	39	74	19	143	21
max S.I.	126	53	103	111	198	177	53
mean	108.4	45.3	53.2	90.8	83.2	161.3	31.0
S.D.	3.13	3.09	5.14	8.66	6.78	3.47	9.51
medium value	108	44	53	90	83	161	27

5.3 Distributive characteristics of experimental data

In this section, we analyze distributive characteristics of the 136×136 = 18,496 pixels.

Pattern distribution characteristics of experimental data are as following <Table 5>~<Table 8>.

<Table 5> pattern characteristic of water among experimental data

	band1	band2	band3	band4	band5	band6	band7
min S.I.	99	41	41	25	12	136	3
max S.I.	131	67	89	116	144	170	92
mean	108.1	45.7	48.2	30.3	16.1	137.5	7.2
S.D.	3.11	3.03	3.13	8.99	9.98	2.34	6.99
medium value	109	46	48	30	16	138	7

<Table 6> pattern characteristic of crop among experimental data

	band1	band2	band3	band4	band5	band6	band7
min S.I.	99	43	40	42	18	149	6
max S.I.	153	79	106	127	194	183	160
mean	108.4	49.3	55.1	49.9	27.6	158.3	11.0
S.D.	3.13	3.20	5.04	9.16	12.67	3.11	10.44
medium value	108	49	55	50	27	158	11

<Table 7> pattern characteristic of forest among experimental data

	band1	band2	band3	band4	band5	band6	band7
min S.I.	97	40	39	29	17	141	7
max S.I.	134	71	103	126	196	181	118
mean	101.4	43.9	43.1	84.0	83.2	153.6	23.9
S.D.	2.33	3.09	4.14	8.66	11.78	3.47	9.51
medium value	101	44	43	84	83	154	24

<Table 8> pattern characteristic of urban among experimental data

	band1	band2	band3	band4	band5	band6	band7
min S.I.	99	42	41	26	17	144	8
max S.I.	146	77	106	124	206	184	182
mean	116.4	53.2	63.2	83.3	97.1	173.5	47.2
S.D.	3.13	3.09	5.09	8.66	16.31	3.33	9.51
medium value	116	54	64	83	97	173	47

5.4 Band feature extraction using Rough sets

We made an experiment for reasonability of our suggestion using simulation program. We got the 10 number of pixel for classification criterions. And, experimentations were iterated 1,000 times. <Table 9>~<Table 11> are experimentation results when it is the case of 3~5 pixels of classification criterion.

<Table 9> result(number of pixel : 3~5)

number of pixel : 3	clustering trend	[B_1,B_3]	[B_2,B_4,B_7]	[B_5]	[B_1,B_7]	others	total
	frequency	101	614	119	53	113	1,000
number of pixel : 4	clustering trend	[B_1,B_2]	[B_1,B_4,B_7]	[B_2,B_4,B_7]	[B_2,B_5]	others	total
	frequency	99	214	549	111	27	1,000
number of pixel : 5	clustering trend	[B_1,B_3]	[B_1,B_5]	[B_2,B_4,B_7]	[B_3,B_5]	others	total
	frequency	136	152	492	103	117	1,000

<Table 10> result(number of pixel : 6~7)

number of pixel : 6	clustering trend	[B_1,B_3]	[B_1,B_5]	[B_2,B_7]	[B_2,B_4]	others	total
	frequency	89	113	97	126	575	1,000
number of pixel : 7	clustering trend	[B_1,B_2]	[B_4,B_7]	[B_5,B_7]	[B_2,B_4]	etc	total
	frequency	76	99	65	89	671	1,000

<Table 11> result(number of pixel : 8~10)

number of pixel : 8	clustering trend	[B_2,B_4]	[B_2,B_7]	[B_1,B_3]	others	total
	frequency	72	91	65	772	1,000
number of pixel : 9	clustering trend	[B_2,B_4]	[B_2,B_7]	[B_3,B_5]	others	total
	frequency	79	88	53	780	1,000
number of pixel : 10	clustering trend	[B_2,B_4]	[B_4,B_7]	[B_1,B_3]	others	total
	frequency	66	58	43	833	1,000

5.5 Evaluation of experimentation result

<Table 12> is synthetic result of experimentation until now. To put it shortly, it is very hard to find the clustering trend when it

is the case of less than 2 pixels of classification criterion. Also, it can't cluster when it is the case of over than 10 pixels of classification criterion. But, it shows clustering trend with band 2, 4, 7 when it is the case of 3~5 pixels of classification criterion, clearly, alse, with band 1, 3, too. Increasing the number of pixels to 6~8, we can't find the clustering trend with three bands each other. Overall, it shows clustering trend with the band 1 and band 2, increasing the number of pixels to 9~10, we can find the clustering trend with band 2 and band 4.

<Table 12> Synthetic result of experimentation

number of pixel	less than 3	3~5	6~8	8~10	over 10
clustering trend	irregular clustering	$[B_1,B_3]$ $[B_2,B_4,B_7]$	$[B_1,B_5]$ $[B_1,B_3]$ $[B_2,B_7]$ $[B_2,B_4]$	$[B_2,B_4]$	irregular clustering

6. Conclusion

In this paper, we suggested new band feature extraction method using Rough sets theory for efficient band selection on multispectral environment. Suggested method used data classification and discernibility of Rough sets theory, we verified appropriateness and performance using LANDSAT TM data near the Han River, on 2, June, 1992. By the result, we can find that it can extract band feature, automatically. Maybe we consider that it can be applied efficient data analysis.

Although the suggested feature extraction method was experimented in multispectral environment, it is considered that good classification performance is appeared when spectral environment moves to hyperspectral environment.

After this, suggested method in this paper must be applied the case of abnormal distribution data environment.

References

[1] L. Jimenez and D. Landgrebe, "Supervised Classification in High Dimensional Space : Geometrical, Statistical and Asymptotical Properties of Multivariate Data," IEEE Trans. on Systems, Man, and Cybernetics, Jan., 1998.

[2] T. M. Lillesand and R. W. Kiefer, *Remote Sensing and Image Interpretation*, 2nd Ed., John Wiley & Sons, 1987.

[3] P. M. Mather, *Computer Processing of Remotely-Sensed Image*, John Wiley & Sons Inc., 1993.

[4] Z. Pawlak, "Rough Sets," International Journal of Computer and Information Science, 1982.

[5] Z. Pawlak, Rough Sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, 1991.

[6] P. Rosenberg, "Resolution, Detectability and Recognizability," PE&RS, Vol. 37, No. 12, pp. 1255-1258, 1971.

[7] R. A. Schowengerdt, *Techniques for Image Processing and Classification in Remote Sensing*, Academic Press, 1983.

[8] ERDAS Field Guide, ERDAS Inc., 1991.