MULTI-SOURCE INFORMATION INTEGRATION BASED REMOTE SENSING INFORMATION RETRIEVAL IN NORTH HEBEI PROVINCE

Jinguo YUAN [a,b], *, Zheng NIU [a], Wei WANG [b], Xiaoli SHI [b]

[a] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications，Chinese Academy of Sciences，P.O.Box 9718, Datun Road，Beijing 100101, China-yuanjinguo8@163.com, zheng_niu@263.net

[b] College of Resource and Environmental Sciences, Hebei Normal University, No.265, Yuhua East Road, Shijiazhuang 050016, China- wangwei@mail.hebtu.edu.cn, shixiaoli_2004@163.com

Commission VII Mid-term Symposium 2006

KEY WORDS: North Hebei province, Remote sensing data fusion, panchromatic and multispectral data fusion, normalized difference vegetation index, tasseled cap transformation, remote sensing image and DEM information fusion, supervised classification, Field survey

ABSTRACT:

North Hebei province lies in the ecotone from agriculture to animal husbandry and forest and is eco-fragile transitional region. Serous land degradation, especially grassland degradation occurs in this area, because of overgrazing and over-reclamation. There are also many disasters, such as drought, dust storm, forest fire, and so on. This area has characteristics of ecological degradation and economic poverty. Soil erosion and dust storm has important impacts on eco-safety of Beijing and Tianjin city. This paper used multi-source information integration method to retrieve land use/cover information using Landsat TM data as information source. Multi-source information integration method used in this paper included multi-source remote sensing data fusion method, visual interpretation and field survey and validation method. Multi-source remote sensing data fusion methods included panchromatic and multispectral data fusion, normalized difference vegetation index, tasseled cap transformation, remote sensing image and DEM information fusion, remote sensing information and GIS database fusion and supervised classification. The result shows that total classification accuracy in 2000 was 84.62% and total Kappa was 0.8026. Arable land increased from 34.3 percent in 1987 to 35.3 percent in 2000, forest land decreased from 38.0 percent in 1987 to 37.4 percent in 2000, and grassland decreased from 24.3 percent in 1987 to 23.9 percent in 2000. Arable land, forest land and grassland maintained 96.6 percent of total land use/cover types both in these two years.

1. INTRODUCTION

Land is an important component of global ecosystem; land use/cover change (LUCC) is one of the important research fields in global change. Research on land use/cover change can promote research process of global environment change and sustainable development. Remote sensing technology has characteristics of dynamic, macroscopical, real-time and multi-temporal, traditional land use/cover survey technology is incomparable to remote sensing.

There are many methods to retrieve land use/cover information, Liu Yalan used segmentation-based classification approaches for different spatial images (Liu et.al., 2002), Others used knowledge-based expert classification (Gao et. al., 2000a; Ton J, et al.,1991a; Wharton S W.A., 1987a), improved artificial neural network classification (Jia, 2000a; Xiong, et. al., 2000a) fuzzy logic classification, spatial data mining (Di et. al., 2000a), and so on. This paper used multi-source information integration method to retrieve land use/cover information using Landsat TM data as information source. Multi-source information integration method used in this paper included multi-source remote sensing data fusion methods, visual interpretation and field check method. Multi-source remote sensing data fusion methods included panchromatic and multispectral data fusion, normalized difference vegetation index, tasseled cap transformation, remote sensing image and DEM information fusion, remote sensing information and GIS database fusion and supervised classification.

2. STUDY AREA

The study area is located in the north of Hebei province of China, between 39º37’N and 42º38’N latitudes and between 113º49’E and 119º15’E longitudes. It contains Zhangjiakou and Chengde these two administrative districts, including two urban area and twenty-one counties, with an area of 7,6300km^2. This area is divided into three geographical subregions, that is, Bashang plateau region, basin region in Northwest Hebei Province and Yanshan Mountainous region in North Hebei Province (Figure 1). There are two climate types, that is, semi-humid and semi-arid, the former includes most of Chengde administrative districts, and the latter includes all Zhangjiakou administrative districts. This area belongs to ecologically fragile region in North China, and is transitional zone from farming to forest and grassland. Many disasters occur in this area, such as drought, dust storm and water erosion, which has great impacts

* Corresponding author

on the eco-safety of Beijing and Tianjin city. It is relatively undeveloped region, and many poverty counties in Hebei province lie in this area. It has characteristics of ecological degradation and economic poverty.

Figure 1. Location of study area

3. DATA PROCESSING

Remote sensing data were seven-band Landsat TM in 1987 and 2000 with TIFF, Raw and Generic Binary format. Nine scenes were needed in this study, data details were shown in Table 1. Each band remote sensing data was input using ERDAS IMAGINE 8.6 image processing system and Photoshop software. Seven-band TM data were composed to color image in ERDAS. The projection is Albers Conical Equal Area, parameters are: latitude of 1st standard parallel is 25°N, latitude of 2nd standard parallel is 47°N, longitude of central meridian is 105°E, spheroid is Krasovsky. Image geometric correction and image mosaic were made. Digital elevation model (DEM) was generated from 100m×100m ASCII format in ERDAS software. The resolution of images was 30 meters. 1:250,000 scale thematic data with E00 format, including transportation, water, and buildings information were transformed to Coverage in ARC/INFO GIS software. Other thematic maps, including soil, geology, soil erosion maps were digitized using ARCVIEW GIS software. Other accessorial information was 1:10,000 scale relief map and land use map in 1987.

Scene	WRS No.	Acquisition Date		Format	
		1987	2000	1987	2000
Keshiketeng	123-30	87.9	2000.10	Raw	Tiff
Guyuan	124-31	87.5.	2000.5	Generic binary	Tiff
Weichang	123-31	87.7	2000.7	Raw	Tiff
Chifeng	122-31	87.2	2000.8	Raw	Tiff
Yangyuan	125-32	87.9	2000.10	Raw	Tiff
Zhangjiakou	124-32	87.7	2000. 4	Generic binary	Tiff
Beijing	123-32	87.10	2000.7	Raw	Tiff
Tangshan	122-32	87.5	2000.5	Generic binary	Tiff
West Guyuan	125-31	87.10	2000.10	Generic binary	Tiff

Table 1. Remote sensing data of study area

4. MULTI-SOURCE INFORMATION INTEGRATION RETRIEVAL METHODS

4.1 Determination of Classification System

Land use/cover classification system of study area were determined according to resource and environment remote sensing investigation of China made by LIU Jiyuan (Liu, 1996), shown in Table 2.

First level	Second level	Definition
Arable Land	In plain In hill In mountain	Cropland
Forest Land	Forested land	Natural and manmade forest, greater than 30% crown cover
	Shrub land	Low forest and shrub, greater than 40% crown cover, less than two meters high
	Sparse forest land	10-30% crown cover
Grassland	High coverage	Natural and improved grassland, greater than 50% coverage, good water condition
	Middle coverage	Natural and improved grassland, 20-50% coverage, insufficient water
	Low coverage	Natural grassland, 5-20% coverage, short of water
Water body		Streams, lakes, reservoirs and irrigation works establishment
Buildings		Residential, industrial and transportation
Other		Bare rock, sandy areas and salination land

Table 2. Land use/cover classification system of study area

4.2 Multi-source remote sensing data fusion methods

Development trend of remote sensing technology is multi-platform, multi-temporal, multi-sensor, multi-spectral, high spatial resolution. Multi-source remote sensing data fusion is multi-layer information matched technology of different spatial resolution, radiometric resolution, spectral resolution and temporal resolution remote sensing data, remote sensing data and non-remote sensing data, including geometric correction and data fusion. Data after fusion is a set of new spatial information and composite image. Its principle is character information optimization; the purpose is to integrate multi-band information of single sensor or information from different sensors, to improve application of remote sensing data and identification ability of targets. This method can exert advantage of different remote sensing data source, make up for shortage of certain remote sensing data, decrease uncertainty, and improve classification accuracy and dynamic monitoring function (Jia, 1997a; J Zhou, 1998a)

Multi-source remote sensing data fusion is divided into three levels: pixel level, character level and decision-making level (Pan et. al., 2002a; Wang et. al., 2003a; Zhou et. al., 2003a; Zhu et. al., 2000a.).Remote sensing data fusion in this study took Fengning County as test area, because acquisition time of remote sensing data in this county was consistent with each other, and it was convenient for data processing and analysis.

4.2.1 Panchromatic and multispectral data fusion： To dig sufficiently the potential resource values of Landsat-7 TM data, fusion of panchromatic and multispectral data in 2000 was made (Figure 2) , the image after fusion had high spatial resolution of 15 meters and abundant color multispectral information, so as to increase the accuracy of image interpretation.

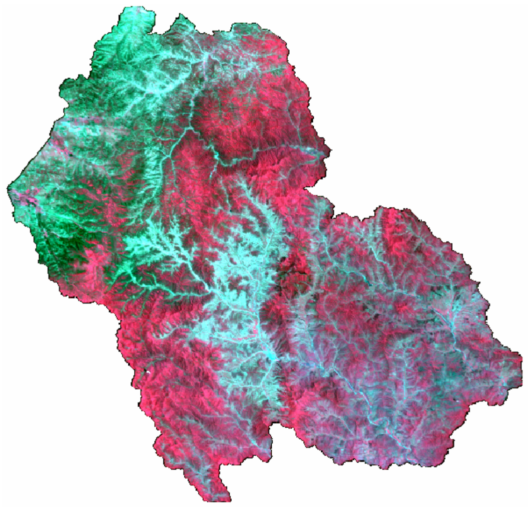

Figure 2. Landsat TM panchromatic and multispectral data fusion image in 2000

4.2.2 Vegetation index method： The principle behind vegetation index is that in the red-light region of the electromagnetic spectrum (0.67μm) chlorophyll causes considerable absorption of incoming sunlight, whereas in the near-infrared region of the spectrum (0.74～1.3μm) plant's spongy mesophyll leaf structure creates considerable reflectance. As a result, vigorously growing healthy vegetation has low red-light reflectance and high near-infrared reflectance, and hence, high NDVI. Vegetation index is applied extensively in vegetation, and is thought as the best indicator of vegetation growth condition, vegetation cover and environment change. It is highly related to LAI (Leaf Area Index), APAR (Absorbed Photosynthetically Active Radiation) and green biomass. Many vegetation indices have already been proposed, and Normalized Difference Vegetation Index (NDVI) was used in this study. NDVI value of Fenning County was shown in Table 3 and Figure 3. From Figure 3 we can see that area of vegetation index in 0.3～0.64 accounted for almost half of total area, vegetation growed better, especially forest; whereas in Bashang plateau, area of vegetation index in -0.13～0.1 dominated, desertification was serious.

NDVI value	Landscape
0.30～0.64	Dense forest, shrub forest, cropland
0.20～0.30	Sparse forest, middle coverage grassland
0.10～0.20	Low coverage grassland
-0.13～0.10	Water body, residential, bare land, sandy land, bare sock

Table 3. NDVI values of different targets

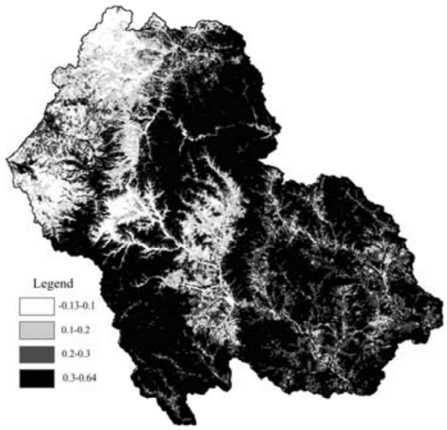

Figure 3. NDVI image of Fengning County in 2000

4.2.3 Tasseled cap transformation： Greenness image after tasseled cap transformation of Fengning County in 2000 was shown in Figure 4, the darker tone, the bigger greenness and the better vegetation, which was consistent with that in NDVI.

Figure 4.Greenness image after tasseled cap transformation

4.2.4 Remote sensing image and DEM information fusion： Remote sensing image and DEM information fusion in mountainous area can rectify distortion caused by relief, because elevation, slope and aspect map can be generated from DEM, vegetation distribution is related to elevation, slope and aspect (Liu et. al., 1996), shown in Table 4.

Elevation (meters)	Main vegetation	Main aspect	Physiognomy type
1000-1600	Meadow grassland, represented by *Leymus chinensis* , *Spiraea* dominates under layer plants.	---	Bashang Plateau
>2000	Meadow	---	Edge of Bashang Plateau in mountainous area
1600-2000	Coniferous forest, mainly *Larix* and *Picea asperata*	Northern	
1400-1600	*Betula phatyphylla* forest, mixed with *Populus davidiana*	Northern	
1200-1400	*Pinus tabulaeformis* forest and *Quercus mongolica* forest	Northern	
1000-1200	*Quercus mongolica* forest, mixed with *Populus davidiana*, *Betula phatyphylla* and *Pinus tabulaeformis*	Southern	
300-1000	Shrub, maily *Zizyphus jujuba var. sponosa*, *Vitex* and Corylus heterophylla Grassland, mainly *Bothriochloa ischaemum*, *Quercus mongolica* and *Quercus liaotungensis*	Northern	Hill, broad valley and Bashang Plateau

Table 4. Vegetation distribution in different elevation and aspect of study area

4.2.5 Remote sensing information and GIS database fusion: The function of GIS in data fusion is to help to classify as data layers. Traffic map, water system map, residential map, soil map and geology map of study area as ancillary data can improve classification accuracy. Spatial data fusion from different source, collection way, scale, spatial and temporal series in GIS is an effective way to solve spatial analysis. In GIS database, digital raster graph (DRG), digital line graph (DLG), digital elevation model (DEM), digital ortho map (DOM), also called “4D” data, are base geographical data. GIS is regarded as information fusion platform with the core of geographical information.

4.2.6 Supervised and unsupervised classification: In supervised classification process, computer system must be trained to recognize patterns in the data, pixels that represent patterns of land use/cover features are first selected. By identifying patterns, computer can identify pixels with similar characteristics. Training area (signature) selection is a key step. To retrieve land use/cover information, optimal band combination scheme need to be selected. There is correlation among seven bands of Landsat TM data, so band 4 (red), band 3 (green), and band 2 (blue) colors composite can be used mainly to retrieve land use/cover information. Band 5 (red), band 4 (green), and band 3 (blue) color composite scheme can be as reference. Band 1 (red), band 4 (green), and band 6 (blue) color composite scheme is helpful to retrieve residential area. In ERDAS IMAIGNE 8.6 software, using AOI polygon tool and seed growth properties, we selected representative objects to gather signatures of image in 2000. To improve classification accuracy, we selected image in 1987 and thematic maps, such as soil, geology, soil erosion maps and DEM as reference information. For each land use/cover types, many training areas could be selected, and then merged after classification. Signatures evaluation methods were spectral response curve according to pixel mean of signatures, histogram, alarms, separability, error matrix, ellipse in feature space image. After reselecting and optimizing signatures, supervised classification was made. Unsupervised classification method could also be used to combine to acquire remote sensing information.

4.3 Visual interpretation method

Remote sensing image is the real time record of electromagnetic spectral character of objects, image characteristics of different targets in the study area were shown in Table 5.

First level	Second level	Color Infrared Image Character
Arable Land	In plain	Blue gray or black green，obvious geometric character, regular texture, strip distribution
	In hill	Grass green or red white
	In mountain	Bright sage green or gray，distributed in river valley, irregular distribution
	Paddy field	dark cyan, light cyan, even tone, distributed in both sides of river valley, strip distribution
Forest Land	Forested land	Bright red, even tone，natural smooth boundary，irregular distribution
	Shrub land	Dark red grey, even tone and structure
	Sparse forest	Dark sage green，red speckle, distributed in mountainous area, coarse structure, irregular shape
Grassland	High coverage	Pink or dark cyan，even structure，strip distribution, clear boundary
	Middle coverage	White pin, white speckle, clear boundary
	Low coverage	Green gray, white speckle, irregular patch
Water body		Light blue to dark blue, obvious character, even tone
Buildings		Bright gray or bice, regular shape, coarse structure, obvious character
Other		White, gray, dark cyan

Table 5. Image characteristics of targets in study area

4.4 Field survey

Information retrieval could be different because of professional level and experience of interpreter, image quality and familiarity with the study area. And because of the complexity of objects and difference of acquisition date, there was maybe wrong interpretation result. Also there were objects not decided and in doubt. All of these needed field check and validation. Our research group went field survey and validation in June, 2004, and set 123 check points (Figure 5). Geographic coordinates, elevation and landscape of check points were recorded using Global Positioning System and digital camera. Coverage of field check points was overlaid on remote sensing image, and then image spectral characteristics of check points representing different objects were drawn, so as to improve interpretation accuracy. Spectral curves of the same objects sometimes were different, because of different acquisition time, growth condition or water content.

Figure 5. Field check points in each county of study area

5. RESULTS

Using ERDAS software, land use map was overlaid to the classified image in 2000. Three hundred randomly selected reference pixels on the classified image were used (fifty per classes). Based on field survey data and land use map, the actual class of each point can be determined. Table 6 showed the classification accuracy of result. Total classification accuracy was 84.62% and total Kappa was 0.8026。

Land use/cover type	Accuracy	Kappa
Cultivated land	80%	0.7753
Forest land	93.33%	0.9157
Grassland	81.25%	0.7725
Water	82.35%	0.7914
Buildings	75%	0.7415
Others	85.29%	0.7842

Table 6. Classification accuracy of result

Land use/cover maps (Coverage format) interpreted from remote sensing image in different time were overlaid using IDENTITY command in ARC/INFO software, land use/cover change from 1987 to 2000 was generated. Land use/cover pattern in this study area changed not greatly from 1987 to 2000 (Table 7). Arable land increased from 34.3 percent in 1987 to 35.3 percent in 2000, forest land decreased from 38.0 percent in 1987 to 37.4 percent in 2000, and grassland decreased from 24.3 percent in 1987 to 23.9 percent in 2000. Arable land, forest land and grassland maintained 96.6 percent of total land use types both in these two years.

Irrational land use action of human being dominated during 1987-2000, which is manifested in these aspects: Reclamation speed of land is greater than that of changing arable land into forest land and grassland, forest cut speed is greater than afforestation speed, and degradation speed of grassland is greater than fostering speed.

Land use/cover types	Arable land	Forest land	Shrub land	Sparse forest land	Other forest land	High coverage grassland	Middle coverage grassland	Low coverage grassland	Water body	Buildings	Unused land	Change to	Net change
Arable land	0	6901	5519	934	195	16233	7364	928	3176	11567	986	53803	76505
Forest land	8638	0	173950	32949	1060	37182	15851	257	552	144	9	270591	-94894
Shrub land	16567	122353	0	16209	51	70963	43782	818	79	272	458	271552	10896
Sparse forest land	892	7333	9285	0	115	6109	1717	249	1	57	0	25757	34333
Other forest land	1343	469	8	0	0	195	308	0	0	73	2	2399	-639
High coverage grassland	65800	30851	61267	6422	235	0	56432	6090	1103	786	2059	231045	-44349
Middle coverage grassland	23049	6688	29240	3452	104	51964	0	9146	1201	560	315	125719	10354
Low coverage grassland	1946	664	881	114	0	641	9109	0	5	68	87	13515	3975
Water body	6660	281	1287	9	0	1333	917	0	0	164	776	11427	-4870
Buildings	895	124	132	2	0	107	155	1	30	0	2	1448	12330
Used land	4518	34	878	0	0	1970	438	2	409	88	0	8338	-3644
Change from	130308	175698	282447	60090	1760	186696	136073	17490	6556	13778	4694		

Table 7 Land use/cover conversion matrix from 1987 to 2000 (Unit: hm^2)

References:

Di Kaichang, Li Deren, Li Deyi, 2000a. Study of remote sensing image classification based on spatial data mining techniques. *Journal of Wuhan Technical University of Surving and Mapping*, 25(1), pp. 42-48.

Gao Zhiqiang, Liu Jiyuan, 2000a. Application of remote sensing technology in land resources dynamic research. *Journal of Remote Sensing*, (3), pp. 27-30.

Jia Yonghong, 1997a. A method of image fusion to enhance spatial resolution of remotely sensed multispectral images. *Remote Sensing Technology and Applications*, 12(1), pp. 19-23.

J Zhou, 1998a. Wavelet transform method to merge Landsat TM and SPOT panchromatic data. *INT. J. Remote Sensing*, 19(4), pp. 743-757.

Jia Yonghong, 2000a. Application of artificial neural network to classification of multi- source remote sensing imagery. *Aviso of Surveying and Mapping*, 7, pp. 7-8.

Liu Jiyuan, 1996. *Resource and Environment Remote Sensing Macroscopical Investigation and dynamic Research of China*. Chinese Science and Technology Press, pp. 262-275.

Liu Lian, hang Yumin, Wang Shouyi, 1996. *Vegetation of Hebei Province*. Science Press, pp. 214-222.

Liu Yalan, Yan Shouyong, Wang Tao, et al., 2002a. A case study on segmentation-based classification approaches for remotely sensed imagery. *Journal of Remote Sensing*, 6(5), pp. 357-362.

Pan Yaozhong, Chen Zhijun, Nie Juan, Wang Xiushan, 2002a. Research on comprehensive monitoring approach in land use dynamic change using multi-source remote sensing data. *Advance in Earth Sciences*, 17(2), pp. 182-187.

Ton J, et al.,1991a. A knowledge-based segmentation of land images. *IEEE Trans. Geosci. Remote Sensing*, 29(3), pp. 222-231.

Wang Ping, Zhang Jixian, Lin Zongjian, Li Chunxia, 2003a. Extraction experiment of LUCC information based on fusion of multi−source RS data. *Aviso of Surveying and Mapping*, 4, pp. 14-17.

Wharton S W.A., 1987a. Spectral knowledge based approach for urban land cover discrimination. *IEEE Trans. Geosci. Remote Sensing*, 25(5), pp. 272-282.

Xiong Zhen, Tong Qingxi, Zhen Lanfen, 2000a. High-rank artificial neural network algorithm for classification of hyperspectral image data. *Journal of Image and Graphics*, 5(3), pp. 196-201.

Zhou Qianxiang, Jing Zhongliang, Jiang Shizhong, 2003a. Comment on fusion theory for different spatial and spectral remote sensing images. *Remote Sensing Technology and Applications*, 18(1), pp. 41-45.

Zhu Changqing, Wang Qian, Yang Xiaomei, 2000a. The fusion of SPOT panchromatic and multi- spectral remote sensing image data by the multi-band wavelet. *Acta Geodaetica et Cartographica Sinica*, 29(2), pp. 132-136.

Acknowledgements

The authors would like to greatly thank the National Natural Science Foundation of China (No: 40171001), the grant of the Knowledge Innovation Program of the Chinese Academy of Sciences (approved # KZCX1-SW-01-02) and National High Technology Research and Development Program of China (No: 2003AA131170) which founded this research.