THE INFLUENCES OF IMAGE CLASSIFICATION BY FUSION OF SPATIALLY ORIENTED IMAGES

Waileung LAU, Bruce A. KING, Zhilin LI
The Hong Kong Polytechnic University, Hong Kong SAR.
Department of Land Surveying and Geo-Informatics
97980389r@polyu.edu.hk, lsbaking@polyu.edu.hk, lszlli@polyu.edu.hk

KEY WORDS: Classification, Data Fusion, Remote Sensing

ABSTRACT

Image classification of either supervised or unsupervised approach is an essential tool to categorise the unidentified pixel in an image to the thematic or the spectral separable classes in the remote sensing science. The informational utility of single-image classification in somehow is limited by either the spatial or spectral resolution, due to the physical trade-off between the resolutions of imaging system. In order to integrate both high spatial and spectral resolution in a single image, the technique of image fusion may be employed. This paper investigates the influences of the multispectral image which is fused with the spatial-oriented image, on the thematic accuracy and the resultant clusters of supervised and unsupervised classification respectively. Through two examples of spatial-oriented images: SPOT panchromatic and scanned aerial images, two respective SPOT multispectral images were fused by intensity-hue-saturation (IHS), principal component analysis (PCA) and high pass filter (HPF) fusion methods. All the images were then classified under the supervised classification approaches of maximum likelihood classifier (MLC) and the unsupervised approaches of ISODATA clustering. Using the classified result of the parent (original multispectral) image as a benchmark, the integrative analysis of the overall accuracy and the numbers of clusters indicated a certain degree of improvement in the classification from using the fused images. The effect of various resolutions for image fusion is also presented. The validity and limitations of image fusion for image classification are finally drawn.

1 INTRODUCTION

Image classification is an image processing technique to categorise the unidentified pixels into themes or classes, based on its spectral response in an image. To achieve "greater quality" of classification, a spectral image with a larger number of narrow spectral bandwidths (high spectral resolution), is necessary. To gather image data with high spectral resolution, a sensor with large-sized instantaneous field of view (IFOV) is required to allow the necessary quantity of radiant energy to enter the detector elements. This results in a sensor with low spatial resolution. The opposite case occurs for high spatial resolution sensors. For example, an image coming from the spectrally oriented (seven-band) LandSat TM sensor has 30-meter resolution and the image from the spatially oriented (1-metre resolution) IKONOS-2 sensor is panchromatic. These fundamental sensor characteristics directly affect the thematic and spatial accuracy of classification of a single image. Image fusion is a technique whereby single image that has the characteristics of both high spectral and spatial resolution sensors can be generated.

Image fusion is an image processing technique that to synthesises image data coming from multiple imaging systems into a single image data source. Previous researches in this arena concentrated on developing fusion algorithms and their assessment of its advantage was limited to visual enhancement. The enhancement of spatial information in various degrees of fused images was seldom repeated. In particular, the effects of fusing spatial-oriented image in image classification have only been reported in Munechika et al (1993).

In Munechika's work, a pair of SPOT panchromatic and LandSat TM images were considered for the multisensor fusion process. In the testing of the classification performance of the fused images, only Gaussian maximum likelihood classification was executed for five types of land cover and evaluation consisted of comparison of the outcomes with ground truth data. The results indicated an enhanced accuracy of classification using the fused image over using the individual (parent) images. However, alternative types of images, levels of enhanced spatial resolution, fusion methods and levels of classification were not evaluated.

This paper investigates the influences on classification of the multisensor fusion of two (relatively) high spectral resolution SPOT multispectral (XS) images with the (relatively) high spatial resolution SPOT panchromatic (PAN) and scanned aerial image (AERIAL) respectively. Three fusion approaches - intensity-hue-saturation

(IHS), principal component analysis (PCA) and high pass filter (HPF) - were used to generate the fused images containing both high spatial and spectral resolutions for classification. The fused (child) images and the original multispectral (parent) image were classified under the supervised classification approaches of maximum likelihood (MLC). The classifications for USGS level-one natural land cover and level-two cultural land use were performed for multilevel-based analysis. Besides, those images are also processed with the unsupervised classification approach of ISODATA clustering. Instead of visual inspection, the classification results with respect to the overall accuracy from the confusion matrices and the number of clusters were assessed. Using the parent image as a benchmark, the influences of classification using the fused images were evaluated.

2 OVERVIEW OF COMMONLY USED IMAGE FUSION APPROACHES

Image fusion is a subset of data fusion to synthesise two or more sets of image data coming from different sources into a new image data set (Pohl and Genderen, 1998). Many algorithms of multisensor fusion have been developed and studied previously (eg. Carper et al., 1990; Chavez et al., 1991; and Varbel, 1996). In general, there are three commonly used approaches: IHS, PCA and HPF. As they are pixel-by-pixel process, they may be applied to general types of images.

The principle of the IHS approach is to transform the multispectral data of three spectral bands in RGB (red-green-blue) mode into IHS color space and replace the intensity component by the content of the panchromatic spatial image. Generally, I is the brightness, and H and S are the color (spectral) components of an image. The procedure of the IHS approach is to transform a multispectral (XS) image from RGB mode into IHS mode. The intensity component (I) is independent of the other two spectral components (H and S) and so can be directly replaced by the spatial (PAN) data, as it does not affect the spectral components. Afterwards, the reverse transformation converts the substituted and two spectral components back into RGB colour space for display and processing. After the reverse process, the synthesised image (as R’G’B’) possesses the spatial resolution of the panchromatic image and the spectral resolution of the multispectral image.

The principle of the PCA approach is similar to that of IHS. A multispectral image (Xr) is transformed into the principal component space. The principal components are simply the largest trend of grey levels in any two channels of the image. The spatial image (PAN) is then directly substituted into the first principal component (Xp), the one that contains largest trends among the three components. This is based on there being a higher similarity between a panchromatic image and first the principal component of a multispectral image than the other principal components, and so the deterioration or changing of spectral data in other principal components is minimised. The output image (Xr’) is generated by re-transforming the replaced channel and the others two components back into RGB space.

The HPF approach is a more direct method than either of the previous two. Spatial details are extracted from the high spatial resolution image (PAN) by high-pass filtering. The extracted spatial information is then independently inserted into each channel of the multispectral image. Due to the arithmetic process, the fused spatial detail increases the DN value of the pixels, and so the spatial detail is dominant in the fusion result. The fused image (XS 1’, XS 2’ and XS 3’) is generated with enhanced spatial detail.

3 EXPERIMENTAL TESTING

In this study, all the image processing operations (multisensor fusion and image classification) were carried out on a Digital Ultimate workstation model 533AU2 using PCI ImageWorks and Xpace version 6.01.

3.1 Test Data

The experiment in this study used two pairs of image: SPOT panchromatic (PAN) and multispectral (XSa) images and scanned aerial photograph (AERIAL) and SPOT multispectral image (XSb). The SPOT PAN image as Figure 1(a), was dated 92/01/31 and had a spectral bandwidth of 0.51-0.73μm. The image provides 10-meter spatial resolution and is used for enhancing the spatial information of the SPOT XSa multispectral image as Figure 1(b), obtained on 91/12/20. The XSa image provides three spectral bands: green (0.50-0.59μm), red (0.61-0.68μm) and near-IR (0.79-0.89μm) with 20-meter ground resolution. To diminish the image size and maintain geographic consistency during the fusion, the equivalent area of sub-scene are used. A large portion of the Kowloon Peninsula in Hong Kong is included. Features such as buildings, roads, airport runways and harbour infrastructures are clearly seen.

For another set of testing data, the AERIAL image as shown in Figure 1(c) was dated 94/03/20 and provides 3-meter spatial resolution. The fusion of this image is used for enhancing another subscene of SPOT XSb image as shown in Figure 1(d) obtained on 95/02/05 in higher spatial resolution (than first set of testing data). To maintain the geographic and temporal consistencies, another equivalent area of subscene is used. The area was located at the Shau Kei Wan at the eastern Hong Kong Island. Features such as buildings, harbour infrastructure, roads and vegetation covers are imagined.

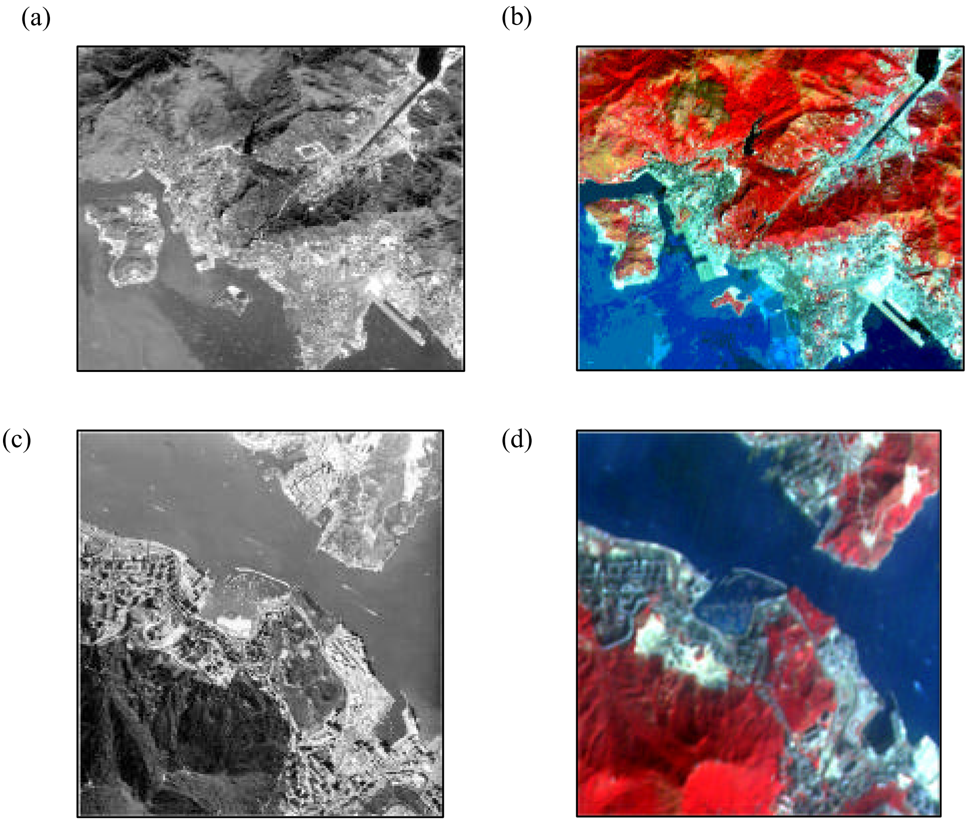


Figure 1. Respective pair of images for the multisensor fusion: (a) SPOT PAN, 92/01/31; (b) SPOT XSa, 91/12/20; (c) Scanned aerial image, 94/03/20; and (d) SPOT XSb, 95/02/05

3.2 Multisensor Fusion Processing

The experiment began with geometric processing. Each parent image was geometrically corrected to reduce the geometric inconsistency between images. Ten stations of the local triangulation network (Hong Kong 80 Grid System) were selected as ground control points (GCPs) for each image. They were evenly distributed through the study area and was accepted only if its root-mean-square (RMS) error was lower than the dimension of one pixel.

Following the geometric registration, the pair of images was “geo-locked”. This “lock” means that the two images are maintained at the same orientation and dimension and is achieved through pixel-by-pixel fusion. This process used the higher spatial resolution image (PAN or AERIAL) as the control image and the lower spatial image (XSa or XSb) was then registered to it. After the “geometric locking” process, fusion of the images was performed by the three approaches, IHS, PCA and HPF.

The XSa or XSb image was transformed into IHS and PC spaces, and the high spatial component (PAN or AERIAL image) replaced the intensity and first principal component respectively. The reverse function transformed the replaced and other components back into the RGB mode of the newly generated, fused XS image. For the HPF fusion, the spatial detail from panchromatic image (PAN or AERIAL) was extracted by differential high pass filtering and directly integrated into the XSa or XSb image to form HPF-fused image.

3.3 Image Classification Processing

In the experiment, supervised and unsupervised classification approaches was considered. For the training stage of supervised classification, the U.S. Geological Survey (USGS) classification systems for remote sensing (Anderson et al., 1972) defines natural land cover and cultural land use as two different levels of classification. Level-one classification, natural land cover, relates to water and the land surfaces that are barren or covered by the vegetation or artificial constructions. The category is defined in more general and simple terms such as water urban, agriculture and forestland classes. The information from satellite imagery with little ancillary data is adequate to identify these land classes. For level-two classification, cultural land use, the impact of man's activities is directly assessed. One piece of land surface identified as a single class in level one can be categorised into several classes in level-two. For example, water bodies may be sub-divided into lake, reservoir and stream sub-categories. Instead of just the spectral response, more spatial details and patterns in an image are needed for level-two classification.

CLASS	DESCRIPTION
LEVEL 1	
WATER	Water coverage includes ocean, streams and reservoirs
VEG	Vegetated coverage includes woodland, grassland and swamp.
URBAN	Artificial coverage includes residential, commercial, industrial and transportation.
BARREN	Uncovered coverage include beach, rock outcrop and construction sites
LEVEL 2	
SEA	Predominant and large bodies of water separate the land mass
INLAND	Water covered area within the land mass includes reservoirs, streams and estuaries.
GRASS	Shrub and brush covered area
WOOD	Forest covered area.
MXURBAN	Intensive built-up area includes residential, industrial and commercial zone.
TRANS	Linear routes as clearcuts and boundaries outline the other land use.
SANDY	Un-vegetated sands dunes dominated along the coast.
OPEN	Temporary barren land included construction sites and bare space.

Table 1 Description of designed categories in level one and two.

Using Anderson's classification system, four categories - water, vegetation, urban and barren land - were defined for level one classification. Another eight categories - sea, inland water, grassland, woodland, mixed urban area, transportation, sandy area and open space - were defined for level two classification. The descriptions of all categories are summarised in Table 1. For the supervised classification, one approach – maximum likelihood (MLC) was applied in this study. After the training process, all the child and parent multispectral images were classified into the two levels through three selected approaches.

For unsupervised classification, no *a-priori* knowledge is required. Unidentified themes are automatically generated by the grouping of statistically similar pixels identified through pixel-by-pixel examination. The resultant clusters represent the spectrally separable classes of an image. In the unsupervised classification – ISODATA clustering (ISOCLUS) was used. The self-trained ISOCLUS requires the defining number of resultant clusters in an image. To test all the images (parents and fused), each image was asked to generate 20 clusters and 256 clusters independently after the process in this study. Due to the variation in the quantity of spectral information in each image, the actual number of clusters identified in an image is not necessarily equal to the defined numbers (20 & 256). Using the result of the parent image as a benchmark, the change in spectral discrimination of the fused images is reflected in the number of classes found in them.

4.4 Classification Results

Instead of visual inspecting the classified images and providing quantitative analysis, overall accuracy from confusion matrix was used (Congalton and Green, 1999). Table 2 lists the overall accuracy and the percentage of correctly classified pixels in individual classes of the level-one classification result using three child-images and parent spectral image. The comparative analyses of the accuracy in each class of either fusion cases are illustrated in Figure 2 (a) and (b).

For the level two classification, a second, and independent data sets was obtained for the training stage. The classification results are summarised in Table 3. The difference of overall accuracy and individual classification accuracy of each class in compared of parent spectral image are also illustrated in Figures 3 (a) and (b).

Thematic Accuracy of Classification Results (%)				
Accuracy Assessment	XSa	IHS-fused	PCA-fused	HPF-fused
MLC				
WATER	99.10	100.00	95.10	96.50
VEG	100.00	95.90	95.40	93.30
URBAN	92.90	88.80	86.30	76.10
BARREN	94.40	91.90	86.90	73.40
Overall Accuracy	93.45	91.66	88.78	80.51

(a)

Thematic Accuracy of Classification Results (%)				
Accuracy Assessment	XSb	IHS-fused	PCA-fused	HPF-fused
MLC				
WATER	69.90	98.10	79.10	89.20
VEG	64.90	91.40	70.80	91.90
URBAN	71.80	52.90	64.40	62.70
BARREN	6.20	5.00	68.60	0.10
Overall Accuracy	50.48	57.88	71.14	56.13

(b)

Table 2. The thematic accuracy of natural land cover classification results: (a) SPOT XSa + SPOT PAN; (b) SPOT XSb + AERIAL images

Thematic Accuracy of Classification Results (%)				
Accuracy Assessment	XSa	IHS-fused	PCA-fused	HPF-fused
MLC				
SEA	66.30	83.80	88.30	65.50
INLAND	92.00	73.60	74.70	85.10
GRASS	64.10	83.20	48.30	86.20
WOOD	64.40	72.60	56.40	79.50
MXURBAN	59.50	65.30	45.90	16.60
TRANS	37.90	29.20	45.10	35.10
SANDY	51.00	89.40	57.00	80.80
OPEN	79.20	59.50	30.90	83.50
Overall Accuracy	63.55	67.54	53.47	64.55

(a)

Thematic Accuracy of Classification Results (%)				
Accuracy Assessment	XSb	IHS-fused	PCA-fused	HPF-fused
MLC				
SEA	74.20	84.60	51.80	66.70
INLAND	25.60	28.20	20.60	46.20
GRASS	9.00	82.30	14.80	8.10
WOOD	91.20	95.10	87.30	72.30
MXURBAN	65.50	20.00	55.40	24.80
TRANS	15.80	34.30	8.90	30.20
SANDY	0.00	2.20	0.00	0.00
OPEN	11.50	4.30	27.30	24.70
Overall Accuracy	43.48	44.58	39.11	34.59

(b)

Table 3. The thematic accuracy of cultural land use classification results: (a) SPOT XSa + SPOT PAN; (b) SPOT XSb + AERIAL images

For the unsupervised clustering, the number of resultant clusters is shown in Table 4.

	Number of Resultant Clusters	
Image	ISOCLUS (20)	ISOCLUS (256)
PAN	6	55
XSa	12	137
IHS	18	99
PCA	19	254
HPF	19	227

(a)

	Number of Resultant Clusters	
Image	ISOCLUS (20)	ISOCLUS (256)
AERIAL	20	255
XSb	12	199
IHS	23	250
PCA	19	254
HPF	19	174

(b)

Table 4. Number of clusters in ISOCLUS: (a) SPOT XSa + SPOT PAN; (b) SPOT XSb + AERIAL images

5 ANALYSIS OF EXPERIMENTAL RESULTS

In the level-one classification (Table 3a), the overall accuracy of the parent XSa image was decreased by an average of 6.5 percent when it is classified using the fusion of SPOT PAN image. For some classes, such as BARREN in MLC (Figure 2a), the decreased classification accuracy was greater than 2.5 percent in using IHS-fused image and up to 21 percent in using HPF-fused image. Only one example, WATER, in using the IHS-fused image is resulted the increased accuracy about 0.9 percent. The significant improvement of classification accuracy was resulted in the fusion of AERIAL image (Table 3b). The overall accuracy was increased about 6.3 percent in the averaging. Using the PCA-fused image, the increasing accuracy was up to 20.67 percent. For two classes: WATER and VEG (Figure 2b), the individual accuracy was upgraded to 18.9 and 19.8 percents (on average) respectively when using those fused images. The individual accuracy was only degraded in the URBAN class (Figure 2b) when using the fused images. The accuracy was decreased about 11.8 percent on average. In the comparative analysis of two fusion cases, the overall classification results of fusing PAN image (about 88.6 percent on average) was accurate than that of fusing AERIAL image (about 58.9 percent on average).

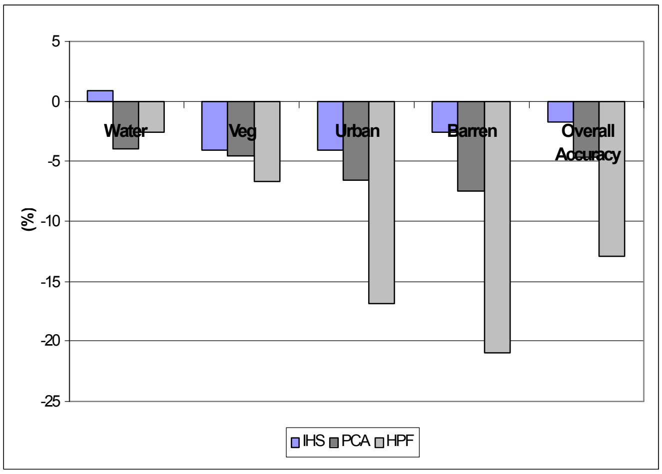
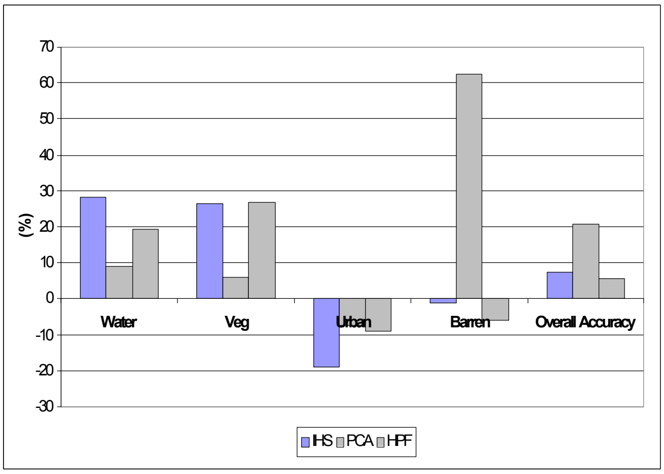


Figure 2(a). A Comparison of Accuracy of Each Class in level-one MLC from SPOT XSa + SPOT PAN Images



(b). A comparison of Accuracy of Each Class in level-one MLC from SPOT XSb + AERIAL Images

The accuracy of classification results using the fused images is significantly improved in fusing of AERIAL image. The fused spatial information enhances the details of spatial features of parental spectral image. However, the spatial information fused from a SPOT PAN image in 10-meter does not improve the accuracy of classification by MLC approach. The reason of degraded accuracy may be came from the distorted spectral characteristics. By the visual inspection of the fused images, the spectral characteristics of the parent spectral image were distorted after the multisensor fusion. For the homogenous natural land cover areas, which are recognised by the high purity of spectral characteristics, the fused images with changed characteristics in larger extent are inappropriate for the classification. The distortion of spectral characteristics was also occurred during the fusion of AERIAL image, but the significant enhancement of spatial details may be over the negative effects of distortion.

In the level-two MLC classification of first fusion case (Table 4a), the higher overall accuracy of results using the IHS and HPF fused images over the parent spectral image were achieved by up to 4.0 and 1.0 percent respectively. The improvement of overall accuracy was significant (Table 4a). For some classes, such as GRASS and SANDY (Figure 3a), the enhanced accuracy of the IHS-fused image is up to 19.1 and 38.4 percent respectively. All the classes (Figure 3a), except INLAND, showed the significantly enhancement in the accuracy of the classification of the fused images. However, the PCA-fused image, with a large degree of distorted spectral characteristics, produced a 10.1 percent degradation of the overall accuracy of classification. Certain improvements of accuracy in using PCA-fused image were achieved, such as SEA and TRANS (Figure 4a). The accuracy of other classes decreased, such as MXURBAN and OPEN classes (Table 3a), the deterioration level being 13.6 and 48.3 percent respectively.

In the comparison of the level-two MLC classification in fusing an AERIAL image (Table 3b), the overall accuracy of results of was only improved by using IHS-fused image about 1.1 percent. For example, in some classes such as GRASS (Figure 3b), the individual accuracy was increased up to 73.3 percent. But the accuracy of two classes: MXURBAN and OPEN (Figure 3b), were still degraded to 45.5 and 7.2 percent respectively. In using the PCA- and HPF-fused images, the overall accuracy was decreased on average by 6.6 percent. Reference to Figure 3(b), the accuracy of five classes (e.g. SEA; WOOD and MXURBAN) were degraded in using PCA-fused image (about 22.4; 3.9 and 10.1 percent respectively); and four classes (e.g. SEA; GRASS and WOOD) were also degraded in using HPF-fused image (about 7.5; 0.9 and 18.9 percent respectively). The degradation of spectral images after the multisensor fusion was obviously in the classification. However, the individual accuracy of two classes was still improved after using the PCA-fused images. For the classes of GRASS and URBAN, the range of improvement was from 5.8 to 15.8 percent (Figure 3b). Three classes: INLAND, TRANS and OPEN were improved in range of 13.2 to 20.6 percent when using the HPF-fused image. For the overall level-two classification results, the accuracy from the fused images in level two was lower than that of level one. As those improved classes were categorised by the sub-division of classification level, the intention of using fused spatial information to improve the level-two classification accuracy can be upheld.

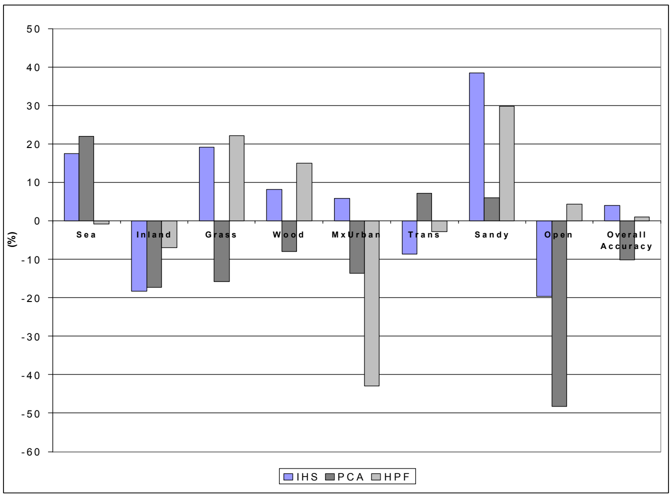
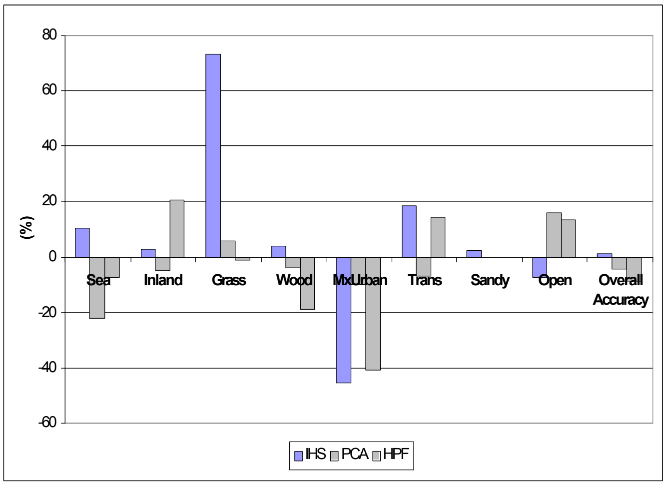


Figure 3(a). A Comparison of Accuracy of Each Class in Level-two MLC from SPOT XSa + SPOT PAN Images



(b). A Comparison of Accuracy of Each Class in Level-two MLC from SPOT XSb + AERIAL Images

According to the unsupervised classification results (Table 4), numbers of clusters was increased in using the fused images. For all the fused images in the testing of 20 clusters, all the fused images in two cases improved the clustering results. For example, the resultant clustering in two cases was increased up to 58.3 and 91.7 percent respectively in compared of parental spectral image. However, in the testing of 256, the cluster numbers from an IHS-fused image were decreased to 27.7 percent in first fusion case (Table 4a). And, in the second fusion case (Table 4b), using a HPF–fused image declined the cluster number to 12.6 percent. Only the PCA-fused image enhanced the clustering in all cases. From the increased clusters (20), the spectral discrimination of an image was enhanced after multisensor fusion. The spectral information of parent spatial image was imported to the child images.

6 DISCUSSION AND CONCLUSION

Land cover classification works with the spectral information of an image. The presented fusion experiments showed that the spectral characteristics of a SPOT XS image were distorted after the fusion with a high-spatial resolution SPOT PAN and AERIAL images. This implies that the successful image fusion, which generates an image with both spatial and spectral resolutions, cannot avoid changing the spectral characteristics of the parent multispectral image. Since the purity of spectral characteristics was degraded after the fusion, the use of multisensor fusion is invalid to enhance the thematic accuracy in the case of the natural land cover classification, which requires greater quantity of spectral information than spatial details. In the level-one classification results, the overall accuracy using the fused images of SPOT PAN deteriorated. However, the higher level of fusing spatial information from AERIAL image improved the overall accuracy of classification. For level-two classification that requires more spatial details, the complicated patterns of features, multisensor fusion has potentials to improve the accuracy of results. The classification results of MLC in two cases demonstrated the improvement of thematic accuracy after using the multisensor fusion method of IHS. Among three selected fusion approaches, the IHS fusion approach for the level-two classification produced higher accuracy results than the parent image. If higher level land cover classification is required from multispectral images, unless the image has high spatial resolution, multisensor fusion with an image of higher spatial resolution is recommended. Furthermore, for the unsupervised clustering, more spectral classes can be determined from the fused images after the PCA approach. However, the increasing clusters can be caused by either the increasing spatial information from fused image or the noise generated from the fusion. The additional evaluation of the noise generated in the fusion is necessary.

ACKNOWLEDGMENTS

This study was funded by the Hong Kong Polytechnic University (Project Account: G-V544).

REFERENCES

Anderson, J. R., E. E. Hardy and J. T. Roach (1972). “A Land-Use Classification System for Use with Remote Sensor Data.” *U. S. Geological Survey Circular*, No. 671.

Carper, W. J, T. M. Lillesand and R. W. Kiefer (1990). “The Use of Intensity-Hue-Saturation Transformations for Merging SPOT Panchromatic and Multispectral Image Data.” *Photogrammetric Engineering & Remote Sensing*, 56(4): 459-467.

Chavez, P. S., Jr., S. C. Sides and J. A. Anderson (1991). “Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic.” *Photogrammetric Engineering & Remote Sensing*, 57(3): 295-303.

Congalton, R. G. and K. Green (1999). *Assessing the Accuracy of Remote Sensed Data: Principle and Practices.* CRC Press, Inc., Denvers, 137p.

Munechika, C. K., J. S. Warnick, C. Salvaggio and J. R. Schott (1993). “Resolution Enhancement of Multispectral Image Data to Improve Classification Accuracy.” *Photogrammetric Engineering & Remote Sensing*, 59(1): 67-72.

Pohl, C. and J. L. Van Genderen. (1998). “Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Application”. *International Journal of Remote Sensing*, 19(5): 823-854.

Varbel, J. (1996). “Multispectral Imagery Band Sharpening Study.” *Photogrammetric Engineering & Remote Sensing,* 62(9): 1075-1083.