Remote Sensing in Transition, Goossens (ed.)
© 2004 Millpress, Rotterdam, ISBN 90 5966 007 2

Detection of buildings from different satellite images by using image enhancement techniques

M. Çetin
Gebze Institute of Technology, Gebze, Kocaeli, Turkey
mcetin@gyte.edu.tr

N. Musaoğlu
Istanbul Technical University, Civil Engineering Faculty, Remote Sensing Division, 34469 Ayazağa, İstanbul, Turkey
nmusaoglu@ins.itu.edu.tr

Keywords: detection of buildings, data fusion, high-pass filter

ABSTRACT: The province of İstanbul and its immediate surroundings have been facing serious environmental problems within the past forty years due to migration from the rural areas, industrial and tourism-oriented investment. The city has about 20% of the total population of the country while it makes up only 0.74 % of the total surface area of Turkey. The conventional methods of urban data collection and analysis cannot deliver necessary information in a timely and cost effective fashion. To meet the demand for current and accurate data remotely sensed images are increasingly used as a source of data. As the resolution of remotely sensed data increases, more detailed information can be obtained.
In this study, the Tuzla region, located in the east coastline of İstanbul, was examined. At the beginning this region developed as a small village having flat, productive agricultural lands and had summer residential area characteristics, then it developed rapidly because of regional investments.
As for the satellite images in the study, SPOT satellite images dated 1999, IRS1C dated 2000, and LANDSAT 7 TM dated 2001 were used. For detection of buildings from satellite images, many image fusion algorithms and a high pass filtering technique have been used. At the end of the study, a comparison of these methods has been made.

1 INTRODUCTION

Without up-to-date information on urban housing development, effective urban planning is hardly possible. Buildings form one of the most important classes in urban land-cover and land-use classification. Development and distribution of buildings in a city are essential information for urban environmental investigation and urban planning. Satellite remote sensing has demonstrated a large potential to obtain this urban information (Zhang 2001).

Urban mapping by means of satellite images is a promising market for Earth observation in the next decade. For such application, on the one hand, a high spatial resolution is necessary for an accurate description of the forms and structures within the cities. On the other hand, the different types of land use are better mapped if high spectral resolution images are available. Thus it is necessary to have both high resolutions for a better knowledge of urban areas.

Methods have been developed to extract useful information from remotely sensed data. Land cover maps may be produced by multispectral image classification. As the resolution of remotely sensed data increases, more detailed information can be obtained. Spectral information extracted from multispectral data is useful for detecting urban expansion. With both multispectral and PAN data urban housing development can be mapped by performing spatial feature post-classification. This paper presents a set of customized image filters, combined with multispectral information for building detection in Istanbul, Turkey, using data from Landsat7 ETM+ and IRS1-C PAN.

2 STUDY AREA

The test area was the small part of urban area of Istanbul, Turkey, which covers over 20 km X 20 km (Fig. 1). Land use in Istanbul is very cosmopolitan and irregular. Most buildings in the area are small and form closely. The streets in the city are narrow. Most green areas in the city centre are very small, apart from some parks. The buildings in the central part of the city are semi-detached or detached in the surrounding parts of the city.

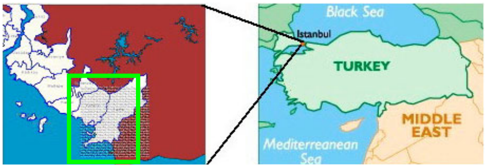

Figure 1. Study area

Because of the economic growth the development of the city construction reached the most dynamic period since 1970 in the last decade. From 1980 to 1995 twice as many residential buildings were built compared with the past years. Therefore it is necessary to rapidly detect the housing development in the urban area of Istanbul using satellite data.

3 DATASETS AND PREPROCESSING

The Landsat7 ETM+ data, SPOT XS and the IRS1-C PAN data were used for this application. In order to register these three datasets together, the Landsat7 ETM+ and SPOT XS data scenes were registered to the IRS1-C PAN scene using a first order polynomial function. A sub scene of 4254x4038 pixels covering the south-eastern part of the city was selected for the building detection study. Sub images of the ETM+, XS and IRS1-C PAN data are shown in Fig. 2, Fig. 3 and Fig. 4. Generally, the houses are small and attached to each other within the old city blocks. Then these buildings were not recognizable on the imagery.

Figure 2. Landsat TM (321) image of study area

Figure 3. Spot XS (432) image of study area.

Figure 4. Irs1-C Pan image of study area.

In the pre-processing, TM and SPOT images were rectified using the nearest neighbour method to the coordinate system of the IRS1-C PAN image. These data have been co-registered to an accuracy of less than 0.4 pixel. Spatial information on the IRS1-C PAN image was improved with high-pass filtering.

4 COMBINING HIGH- AND LOW-RESOLUTION IMAGE DATA

Landsat TM data contain a high level of spectral information, but their spatial resolution is too coarse to interpret urban buildings. IRS1-C pan data are well known for their detailed spatial information, buildings can be interpreted visually, but their spectral information is not sufficient for classifying buildings and other classes automatically. Therefore, it is necessary to fuse the spectral and spatial information of the two datasets to achieve a detailed detection (Zhang 2001).

4.1 *Data fusion techniques*

Several methods have been devised to combine remotely sensed data. While these procedures may offer improvements in spatial resolution, the fused images often exhibit residual noise. The methods may adversely change the information content of the resulting image product. In this study simple band arithmetic techniques (Brovey, Multiplication), component substitution techniques (IHS transformation, principal component substitution PCS) and classical filter techniques (high frequency addition method HFA, high frequency modulation method HFM) are used for the data fusion.

4.1.1 *IHS*

The IHS transform, an alternative colour space, intensity (I), hue (H) and saturation (S) as the three seperate, orthogonal and easily perceived colour attributes (Sunar and Musaoglu, 1998). This numerical procedure was developed to convert a three-band RGB (red-green-blue) display into its fundamental physiological (IHS) elements of human color perception (Grasso 1993).

4.1.2 *Brovey transform*

The Brovey Transform was developed to visually increase contrast in the low and high ends of an image's histogram. Consequently, the Brovey Transform should not be used if preserving the original scene radiometry is important. However, it is good for producing RGB images with a higher degree of contrast in the low and high ends of the image histogram and for producing visually appealing images (Erdas Field Guide 1999).

4.1.3 *Multiplicative*

The algorithm is derived from the four-component technique of Crippen. In this paper, it is argued that of the four possible arithmetic methods to incorporate an intensity image into panchromatic image (addition, substraction, division and multiplication), only multiplication is unlikely to distort the colour (Erdas Field Guide 1999).

4.1.4 *PCS*

The procedure to merge the TM and PAN data using the PCS method is similar to that of the IHS method. The TM data, either three or all six reflective TM bands, are used as input to a principal component analysis procedure. As with the IHS method, the PAN data are stretched to have approximately the same variance and average as the first principal component. The results of the stretched PAN data replace the first principal component image before the data are retransformed back into the original space. This assumption is made because the first principal component image will have the information that is common to all the bands used as input to PCS, while spectral information unique to any of the bands is mapped to the other components (Chavez 1991*)*.

4.1.5 *HFA*

HFA is a filter technique in spatial domain to extract the panchromatic channel high frequencies, a degraded or low-pass-filtered version of the panchromatic channel has to be created. This low-resolution version should fit to the resolution level of the multispectral image. Subsequently, the high frequency addition method (HFA) extracts the high frequencies using a subtraction procedure and adds them to the multispectral channels via addition:

$$\text{multi}_j^{high} = \text{multi}_j^{low} + (\text{pan}^{high} - \text{pan}^{low}) \quad (1)$$

where [multi $_j^{low}$] is the original low resolution multispectral image, [pan^{low}] the degraded panchromatic band, [pan^{high}] the original (high resolution) panchromatic band, [multi $_j^{high}$] the fusion result and [j] the channel index (Hill, 1999).

4.1.6 *HFM*

The other high frequency modulation method (HFM, or Sparkle) extracts the high frequencies via division and adds them to each multispectral channel via multiplication:

$$\text{multi}_j^{high} = \text{multi}_j^{low} * \text{pan}^{high} / \text{pan}^{low} \quad (2)$$

Because of the multiplication operation, every multispectral channel is modulated by the same high frequencies. Thus, the HFM technique allows a straightforward fusion of both datasets (Hill, 1999).

4.2 *Comparison of the spectral and spatial effects of the merged TM- IRS1C and TM-SPOT images*

4.2.1 *Visual comparison*

For comparison of the merging methods, a small section was cut out from each of the original TM, SPOT XS images. The Multiplicative, HFA and HFM methods reproduce the spectral information of the original TM image better than the other methods tested. Especially in homogeneous areas such as the colour of the water areas and the green areas is also almost identical in these images. But HFA and HFM methods are affected for combining remotely sensed image data having large (6 x) differences in spatial resolution. The colour effects of the PCS, IHS method and the Brovey method are different and more distorted for the test areas. The colour contrasts of the Brovey and IHS methods are weaker than that of the PCS method for the test area (Fig. 5 and Fig. 6).

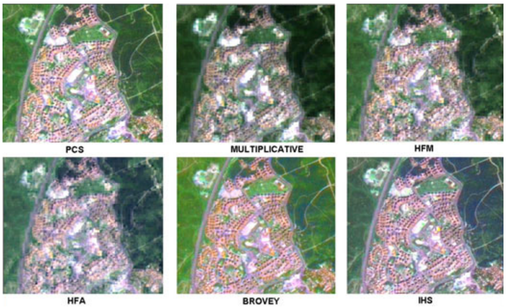

Figure 5. HFA and HFM methods are affected for combining images (Landsat TM&IRS1-C PAN) having large (6 x) differences in spatial resolution.

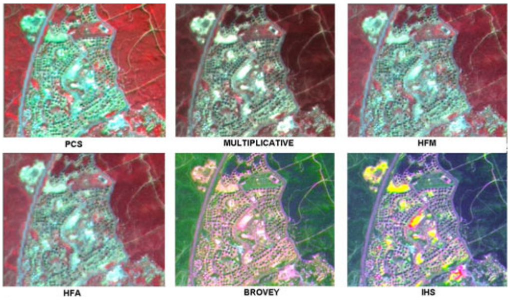

Figure 6. HFA and HFM methods are not affected for combining images (SPOT XS&IRS1-C PAN) having large (4 x) differences in spatial resolution.

In a comparison of spatial effects, it can be seen that the results of the six methods display the different details. Houses, streets, green fields, roads and so on can be clearly recognized in the merged images using all three methods. The Brovey, IHS and PCS methods (Fig. 7) reproduces the spatial information of the original IRS1-C image better than the other methods. The main roads of the city (i.e. streets whose width is larger than 4 m) are clearly visible. We can clearly separate the different textures of the city (green parks, building areas). This spatial resolution does not allow the observation of the small details within the city.

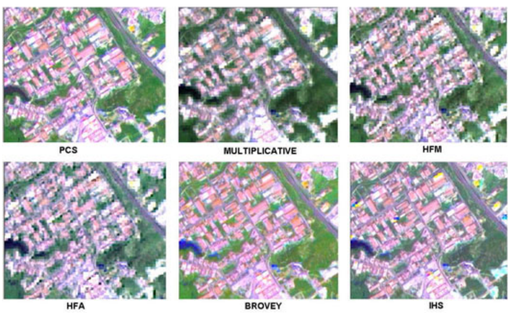

Figure 7. The visual comparison of the merging methods from TM&IRS1-C PAN

4.2.2 *Statistical comparison*

Table 1. Correlation coefficients between TM images and IRS1-C PAN.

	TM1	TM2	TM3	TM4	TM5	TM7
PAN	0.72	0.85	0.85	0.76	0.87	0.86

Simple band-arithmetic techniques, such as multiplication, IHS transformation, principal component substitution (PCS) and the Brovey algorithm may produce spectral distortions in the result, which depend on the degree of global correlation between the panchromatic and the multispectral channels to be enhanced. For the test area TM (3,5,7) and SPOT XS (3,2,1) bands have highest correlation coefficients with the IRS1-C PAN were used in the merging with IHS and Brovey transformation techniques.

Table 2. Correlation coefficients between SPOT XS images and IRS1-C PAN

	XS1	XS2	XS3	XS4
PAN	0.85	0.89	0.89	0.82

The set of criteria, proposed by Wald (1997), was used to assess the performance of the data fusion methods. It is;

1. Bias, as well as its value relative to the mean value of the original image. Recall that the bias is the difference between the means of the original image and of the synthetic image. Ideally, the bias should be null.

2. The difference in variances, as well as its value relative to the variance of the original image. This difference expresses the quantity of information added or lost during the enhancement of the spatial information. The difference will be negative because the variance of the synthetic image will be larger than the original variance. In opposite case, the difference will be positive. Ideally, the variance difference should be null.

3. The correlation coefficient between the original and synthetic images. It shows the similarity in small size structures between the original and synthetic images. It should be as close as possible to 1.

4.The standard deviation of the difference image, as well as its value relative to the mean of the original image. It globally indicates the level of error at any pixel. Ideally, it should be null.

In the multiplicative, HFM and HFA methods, the differences between the means of the fused images and original images are close to zero (Table 3a, Table 3b, Table 4a and Table 4b). Other methods are not sufficient for this condition.

The HFM and HFA methods have the minimum and negative values in this criteria (Table 5a and Table 5b). Other methods are larger than both methods and they have positive values.

Table 3a. The mean values of the fused images (Landsat TM&IRS1-C) and original images (Landsat TM).

	TM	IHS (753)	Brovey (753)	PCS	Multip	HFM	HFA
TM1	92.65			88.74	92.00	92.79	92.45
TM2	74.64			68.58	75.30	74.78	74.44
TM3	72.86	51.45	27.03	63.78	73.70	73.03	72.64
TM4	61.44			53.74	61.60	61.44	61.28
TM5	76.64	54.20	25.56	63.89	76.60	76.74	76.41
TM7	56.04	38.47	18.34	45.70	56.10	56.12	55.82

Table 3b. The mean values of the fused images (SPOT XS&IRS1-C) and original images (SPOT XS).

	XS	IHS (321)	Brovey (321)	PCS	Multip	HFM	HFA
XS1	108.38	50.87	23.49	58.23	108.73	107.72	108.17
XS2	116.14	47.75	25.93	93.97	116.66	115.56	116.00
XS3	110.17	49.27	21.51	75.36	109.99	109.60	110.04
XS4	82.00			49.00	80.98	81.16	81.54

Table 4a. The bias values of the fused images (Landsat TM&IRS1-C) and original images (Landsat TM).

	PCS	Multip	HFM	HFA	IHS	Brovey
TM1	4.1	0.8	0.0	0.4		
TM2	6.4	-0.4	0.1	0.5		
TM3	9.5	-0.4	0.3	0.7	21.9	46.3
TM4	7.8	-0.1	0.1	0.3		
TM5	13.3	0.5	0.4	0.7	22.9	51.6
TM7	10.9	0.5	0.5	0.8	18.1	38.3

Table 4b. The bias values of the fused images (SPOT XS&IRS1-C) and original images (SPOT XS).

	PCS	Multip	HFM	HFA	IHS	Brovey
XS1	50.15	-0.35	0.66	0.21	57.51	84.89
XS2	22.17	-0.52	0.58	0.14	68.39	90.21
XS3	34.81	0.18	0.57	0.13	60.90	88.66
XS4	33.00	1.02	0.84	0.46		

Table 5a. The differences in variances between the fused images and original images.

	PCS	Multip	HFM	HFA	IHS	Brovey
TM1	81.1	30.6	-207.2	-151.2		
TM2	220.5	294.2	-189.8	-151.8		
TM3	544.3	911.1	-236.5	-158.1	605.2	754.3
TM4	304.5	1031.2	-145.6	-138.1		
TM5	1133.6	1994.7	-246.2	-121.7	1192.3	1220.2
TM7	750.5	1284.0	-183.9	-132.8	894.7	823.2

Table 5b. The differences in variances between the fused images and original images.

	PCS	Multip	HFM	HFA	IHS	Brovey
XS1	4364.82	4606.18	-124.26	-58.22	4258.96	5466.84
XS2	737.45	647.16	-70.66	-20.29	771.71	1186.91
XS3	1881.00	2065.37	-93.19	-31.26	1376.63	2683.43
XS4	1849.32	2058.64	-51.43	-30.19		

The multiplicative, HFM and HFA methods have the maximum correlation coefficient values. The Brovey method has the minimum values (Table 6a and Table 6b).

Table 6a. Correlation coefficients between fused images (Landsat TM&IRS1-C) and Landsat TM.

	Brovey 753	Brovey 321	HIS 321	HIS 753	Multip	PCS	HFM	HFA
TM1	-	0.62	0.62	-	0.90	0.93	0.82	0.84
TM2	-	0.84	0.81	-	0.95	0.90	0.91	0.91
TM3	0.64	0.92	0.91	0.85	0.95	0.90	0.95	0.95
TM4	-	-	-	-	0.93	0.93	0.96	0.95
TM5	0.90	-	-	0.90	0.94	0.90	0.96	0.97
TM7	0.93	-	-	0.90	0.95	0.90	0.96	0.96

Table 6b. Correlation coefficients between fused images (SPOT XS&IRS1-C) and SPOT XS.

	Brovey 321	HIS 321	Multip	PCS	HFM	HFA
XS1	0.96	0.91	0.94	0.90	0.98	0.99
XS2	0.58	0.64	0.95	0.85	0.95	0.97
XS3	0.89	0.89	0.95	0.85	0.97	0.99
XS4	-	-	0.92	0.88	0.98	0.98

The HFM and HFA methods have the minimum value in the standard deviations of the difference images. The Brovey method has the maximum value (Table 7).

Table 7a. The standard deviations of the difference images.

	HIS (753)	Brovey (753)	PCS	Multip	HFM	HFA
TM1			6.46	7.48	12.46	10.99
TM2			10.54	10.01	11.23	10.96
TM3	19.39	31.25	16.21	18.34	11.98	10.84
TM4			13.63	22.62	10.09	10.96
TM5	22.96	37.67	22.76	30.03	13.34	10.79
TM7	19.45	31.36	18.51	23.48	10.77	10.73

Table 7b. The standard deviations of the difference images.

	His 321	Brovey 321	PCS	Multip	HFM	HFA
XS1	41.21	53.17	44.56	47.21	12.03	7.05
XS2	22.58	27.51	19.96	14.08	10.89	7.11
XS3	27.06	37.26	31.61	27.80	11.60	7.04
XS4			29.61	31.50	8.77	7.13

5 CLASSIFFICATION AND ANALYSIS

There are a few procedures involved in the proposed building detection method. 1. Image fusion: to make use of higher spatial resolution of the IRS1-C PAN data and higher spectral resolution of the Landsat-7 data. 2. Multi-spectral analysis: to separate open water and vegetation area, construction sites and roads from other built-up areas where buildings stand.

3. Building detection using a set of filters and post-processing. Methods have been developed for image fusion. The most popular approach is the IHS-RGB transformation. However, IHS-RGB transformation can cause spectral distortion. Since spectral distortion is not desired for multispectral classification, high frequency modulation method (HFM) was adopted in this study for multispectral analysis instead. This method calculates the ratio between lower resolution imagery and a smoothed (low-pass filtered) higher resolution imagery, and then multiplies this ratio with the higher resolution imagery. Thus, it tends to preserve the spectral properties and contrast of the lower resolution imagery (Zhang 2002). The HFM fusion was performed on ETM+ bands 1-5,7 and IRS1-C PAN to produce six merged channels.

In urban areas buildings, green vegetation and water surfaces are fundamental components. Thus, for a detailed detection of urban housing development, the other two relevant classes—green vegetation and water areas—were also classified. Since the unsupervised clustering method is better suited for classifying heterogeneous classes in high-resolution satellite images than supervised classification (Zhang 2001). Multispectral IsoDATA clustering was then performed on the seven HFM fused channels. The clusters were labeled and grouped into water, vegetation and buildings classes. In the unsupervised classification the image data were first subdivided into 50 clusters. The resulting classes were reduced to three classes extracting through additional interpretation of the 50 clusters. The road networks could be confused with buildings. Roads were best distinguished on the IRS1-C PAN image. This band was carefully examined and a few grey levels on the continuous histogram of IRS1-C PAN were identified as 'roads'. This road map included most parts of major roads. High-pass filter was applied to the IRS1-C PAN imagery to detect the buildings. Then a threshold was applied to each channel. Pixels with a value above the threshold were selected as potential building pixels. Then the building pixels resulted from the filtered channel were combined into one map. In addition, pixels within vegetation area, water area, construction sites from the multispectral analysis result were masked out.

The advantage of the fused data compared with the TM data can clearly be seen in the classified images. In the results of the original TM data, green and construction areas in the city cannot be recognized, whereas in the results of the fused data, they are clearly recognizable. On the other hand, the original TM data are more suitable for classification of areal classes such as vegetation areas and water areas, because the pixel values in TM data are more homogenous than in fused data and the results are more continuous (Zhang 2001). Therefore the areal classes were classified from the fused TM–IRS1-C PAN data (Fig. 8).

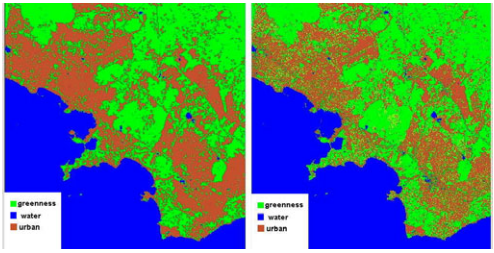

Figure 8. Classified images from the original TM (left) and fused HFM (right) images.

In the accuracy assessment of the building extraction, 100 random points were assessed in study area area. The nearly true-color image fused with TM and IRS1-C pan data and aerial photos were used as reference. The visible buildings in the image were checked. The average user accuracy of the extracted buildings reached a level of 81% after the high-pass filtering, while it was only 52 % using multispectral classification.

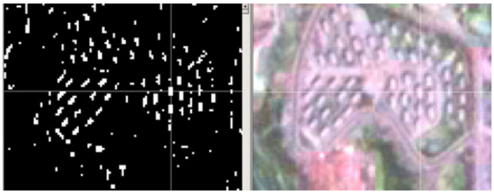

Figure 9. Detected regular shaped buildings (left) and IHS image (right).

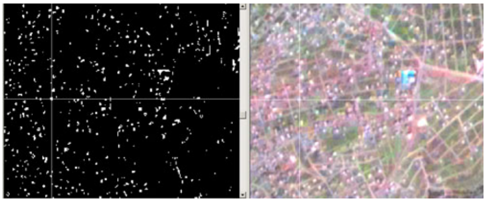

Figure 10. Detected irregular shaped buildings (left) and IHS image (right).

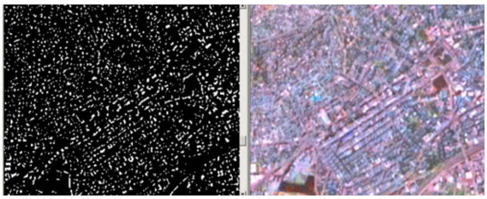

Figure 11. Detected buildings attached to each other (left) and IHS image (right).

6 RESULTS

The detection results (Fig.9, Fig.10 and Fig.11) in this study demonstrated the ability to classify large buildings in urban areas after the fusion of TM multispectral with IRS1-C pan data.

Performing the spatial feature post-classification the results of the multispectrally extracted classes—buildings, water areas and green areas—can be improved significantly. In comparison to the conventional spectral/temporal classification algorithm using multispectral data, the results in this study are much more detailed and accurate. Not only were the newly developed built-up areas detected, but also the new buildings and infrastructure developments in the new areas could be recognized.

REFERENCES

Chavez, P.S., Stuart, Jr., Sides, C. And Anderson, J.A. 1991. Comparison Of Three Different Methods To Merge Multiresolution And Multispectral Data: Landsat TM And SPOT Panchromatic. *Photogrammetric Engineering & Remote Sensing,* Vol. 57 (No. 3): 295-303.

Couloigner, I., Ranchin T., Valtonen V.P. & Wald, L. 1998. Benefit Of The Future SPOT-5 And Of Data Fusion To Urban Roads Mapping. *Int. J. Remote Sensing.* Vol. 19 (No. 8): 1519 –1532.

Erdas Field Guide Fifth Edition. 1999. Erdas Inc. Atlanta, Georgia.

Grasso, D.N., 1993. Applications Of The IHS Color Transformation For 1:24,000-Scale Geologic Mapping: A Low Cost SPOT Alternative. *Photogrammetric Engineering & Remote Sensing.* Vol. 59 (No. 19): 73-80.

Hill, J., Diemer, C., Stöver, O. & Udelhoven, Th. 1999. A Local Correlation Approach For The Fusion Of Remote Sensing Data With Different Spatial Resolutions In Forestry Applications. *International Archives Of Photogrammetry And Remote Sensing.* Vol. 32, Part 7-4-3 W6, Valladolid, Spain.

Sunar, F., Musaoglu, N. 1998. Merging Multiresolution Spot P and Landsat TM Data: The Effects and Advantages, *International Journal of Remote Sensing,* Vol.19, No: 2, pp. 219-224.

Zhang, Q., Wang, J. & Shi, P. 2002. Detection Of Buildings From Landsat-7 ETM+ And SPOT Panchromatic Data In Beijing, China. *IEEE.* 2977-2979.

Zhang, Y. 1999. A New Merging Method And Its Spectral And Spatial Effects. *Int. J. Remote Sensing*. Vol. 20 (No. 10): 2003-2014.

Zhang, Y. 2001. Detection Of Urban Housing Development By Fusing Multisensor Satellite Data And Performing Spatial Feature Post-Classification. *Int. J. Remote Sensing*. Vol. 22 (No. 17): 3339-3355.

Wald, L., Ranchin, T. & Mangolini, M. 1997. Fusion Of Satellite Images Of Different Spatial Resolutions: Assessing The Quality Of Resulting Images. *Photogrammetric Engineering & Remote Sensing.* Vol. 63(No. 6): 691-699.