Satellite imagery fusion methods to improve urban land cover classes identification

Caprioli M., Tarantino E.

in

Camarda D. (ed.), Grassini L. (ed.).
Local resources and global trades: Environments and agriculture in the Mediterranean region

Bari : CIHEAM
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 57

2003
pages 421-428

SATELLITE IMAGERY FUSION METHODS TO IMPROVE URBAN LAND COVER CLASSES IDENTIFICATION

M. Caprioli, E. Tarantino
Polytechnic of Bari, Italy

ABSTRACT

The growing use of Geographic Information Systems (GIS) has led to new research opportunities in the application of satellite imagery to urban analysis. The information content of such images is a function of the combined influence of the radiometric, spatial, and spectral resolution of the sensor. The different bands of satellite sensors are recorded synchronously so that their pixels may be precisely matched and compared with their counterpart pixels in other bands. This means that we can use spectral ("colour") differences to identify urban features to the extent that colours are diagnostic sort of a coarse spectroscopy from space.

Spectral sensor technology, however, coupled with the complexity of ground features in urban areas, can make visual interpretations of satellite imagery both labour intensive and uncertain. Moreover, the informational utility of a multispectral image for land cover classification is often limited by the spectral and spatial resolution of the imaging system. No currently existing single system offers both high spatial and high spectral resolution. Furthermore, if the same techniques that were developed for earlier lower resolution satellite imagery are used on high-resolution imagery (such as maximum likelihood classification), the results can create a negative impact. Lower resolution data are not affected greatly by artefacts, such as shadows, and they also “smooth” out variations across ranges of individual pixels, allowing statistical processing to create effective land cover maps. Individual pixels in higher resolution data can represent individual objects and contiguous pixels in an image can vary dramatically, creating very mixed or “confused” classification results.

This paper proposes a two-stage classification procedure that effectively reduces negative impacts related to spectral ambiguity and spatial complexity of land cover classes of high resolution imagery in urban environments.

In order to achieve both high spatial and spectral resolution in a single image, image fusion is employed and its influences on thematic accuracy of land cover classification, through an example using IKONOS panchromatic and multispectral images, is examined. Three fused images were generated using 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). Using the classified result of the original multispectral image as a benchmark, integrative analysis of overall accuracy, with the degree of improvement in the classification from using the fused images, are executed. Validity and limitations of image fusion for land cover classification are finally drawn.

Keywords: urban environment; satellite imagery; data fusion; classification.

1. INTRODUCTION

The advent of recent multispectral imagery with very fine spatial resolution has increased our ability to map land use in geometric detail and accuracy (Aplin *et al.*, 1997) for local scale investigations. Such improved satellite sensor technology; improved spatial data handling and analysis techniques, the increasing capabilities of GIS software, the widespread availability of training in GIS and image processing, and the increasing speed, memory, and storage capacity of personal computers have all made previously difficult image processing and data integration possible for users and analysts.

The use of very fine spatial resolution images, however, brings with it some major problems. First, the mass of corresponding information creates difficulties in terms of image storage, data exchange and processing time; and, second, the majority of the highest resolution images (1 m) are presently recorded in panchromatic mode only. The problems concerning data volume might appear to be of secondary importance, given rapid and continuing improvements in computer technology. The solutions suggested for solving the latter problem revolve around the use of various methods of data fusion, by merging the higher resolution panchromatic data with lower resolution multispectral data (Pohl *et al*., 1998). Additionally, one may observe that the fusion issue is, and remains, of prime importance in diachronic analyses using images recorded by different generation and typology sensors (i. e. with different spatial resolutions).

Spectral sensor technology, however, coupled with the complexity of ground features in urban areas, can make visual interpretations of satellite imagery both labour intensive and uncertain. Moreover, informational utility of recent multispectral images for land cover classification is often limited by the spectral and spatial resolution of the imaging system. Lower resolution data are not affected greatly by artefacts, such as shadows, and also “smoothe” out variations across ranges of individual pixels, allowing statistical processing to create effective land cover maps. Furthermore, lines of communication (such as roads) can be identified and extracted as a single line. In high-resolution images there is also the issue of linear feature extraction and roads comprise a variety of information: the road markings, the road itself, the kerb with its shadows, the pavement or sidewalk, overhanging trees and parked cars (Caprioli *et al.*, 2001).

This paper proposes a two-stage classification procedure that effectively reduces negative impacts related to spectral ambiguity and spatial complexity of land cover classes of urban areas in high resolution imagery. In order to achieve both high spatial and spectral resolution in a single image, image fusion is employed and its influences on thematic accuracy of land cover classification, through an example using IKONOS panchromatic and multispectral images, is examined. Three fused images were generated using 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). Using the classified result of the original multispectral image as a benchmark, the integrative analysis of the overall accuracy indicated a certain degree of improvement in the classification from using the fused images. Validity and limitations of image fusion for land cover classification are finally drawn.

2. REMOTE SENSING OF URBAN ENVIRONMENTS

There is a long history of using remote sensing as a data source for urban management information. In 1858, Gaspard Felix Tournachon (later known as "Nadar") took the world's first aerial photographs (of Paris and its surrounding countryside) from a hot air balloon (Lillesand *et al.*, 2000). Since then planners, among others, have come to recognise the value of this "bird's eye view" for discovering the distinctive spatial patterns and forms that characterise urban spaces. There is a richly illustrated literature of cityscapes and urban patterns using aerial photos and, more recently, images from other airborne sensors such as radars, scanners, and even video cameras. However, applications of data obtained from satellite platforms at 300 to 600 miles above the earth have remained limited for urban areas for a variety of reasons, including the low resolution of the images, the complexity of ground features in urban areas, and the technological differences with conventional photography. Until recently, low resolution of satellite images has inhibited their use for urban analysis. Planners are used to seeing roads, buildings, and other small structures in aerial photos, and many urban analyses require mapped data to be at that level of detail, say scales of 1:500 to 1:10.000. The world's first civilian Earth Resources Technology Satellite (ERTS-1), launched in the early 1970s, produced pictures which were too coarse to show details of the built environment. With pixels 80 meters on one side, the pictures from ERTS-1 (later renamed "Landsat") could not be used to locate houses, streets, or individual plots. However, as space sensor technology improved, pixel sizes decreased. The latest Landsat (#7), launched in April 1999, produces images with pixel sizes 15 meters square in the panchromatic (black and white) band, 30 meters square in the six visible and near-infrared bands, and 60 meters square in a single heat-sensitive thermal band. These Landsat data are available in image and digital formats from U.S. government and commercial sources at modest prices. Moreover, several private companies operate their own satellites and sell images with 1 to 10 meters resolutions (e.g., SPOT Image Corporation, OrbImage, Space Imaging Corporation). This third generation of satellite sensors is likely to stimulate the development of urban remote sensing (Fritz, 1999). To emphasize the significance of these new data sources, it is worth bearing in mind that this level of spatial resolution corresponds to scales of analysis between 1:10.000 and 1:25.000 (ignoring the effects of relief distortion etc.), that is to say, scales typical of projects dealing with urban planning. Although , their prices are higher than Landsat images, they are still less than the cost of commissioning new aerial photography.

Satellite imagery technology is very different from conventional photography. Landsat's sensors record wavelengths of reflected light ranging from 0.45 mm to 12.5 mm-the visible, near-infrared, short-wave infrared, and thermal infrared portions of the spectrum. The number of spectral bands in civilian imaging satellites has increased from four in the first Landsats to hundreds in today's hyperspectral satellite sensors. These different bands are recorded synchronously so that their pixels may be precisely matched up and compared with their counterpart pixels in other bands. This means that we can use spectral ("colour") differences to identify urban features to the extent that colours are diagnostic sort of a coarse spectroscopy from space. Spectral sensor technology, however, coupled with the complexity of ground features in urban areas, can make visual interpretations of satellite imagery both labour intensive and uncertain. To implement image interpretation, commercial or free software can be used to "train" computers to recognise spectral signatures from sample locations. The software program will then read through an entire satellite dataset to tag other pixels that have similar characteristics (Gong *et al.*, 1990). This procedure works well with images of agricultural crops or forests in rural areas, but computers have had trouble distinguishing between different urban features. For example, roofs of building are small compared to satellite pixels and have many colours and shadings due to different materials and orientations towards the sunlight. As a result, this approach has had limited success in dealing with data for urban environments (Gao *et al.*, 1998; Harris *et al.*, 1995; Ridd, 1995).

New high resolution images are being used to update urban information (Ehlers *et al.*, 1990; Westmoreland *et al.*,1992; Caprioli *et al.*, 2001), but optimal results still usually need human visual interpretation of the computer-enhanced images. More to the point, even today satellite imagery is seldom employed for making population estimates or for supporting housing surveys, at least not without accuracy checks using aerial photography. For these reasons, much of the research in satellite data interpretation has focused on agricultural and forested areas, where the spectral responses from large fields with homogeneous types of vegetation provide relatively uncomplicated landscapes.

3. DATA FUSION APPROACHES

To achieve accurate classification a spectral image with a larger number of narrow bandwidths spectral bands (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 a 30-meter resolution and the image from the spatially oriented (1-meter resolution) IKONOS-2 sensor is panchromatic. These fundamental sensor characteristics directly affect the thematic and spatial accuracy of classification of a single image. Multisensor data fusion can help to obtain better results in this field of application because it is a technique whereby single images that have the characteristics of both high spectral and spatial resolution sensors can be generated.

Data fusion means a very wide domain and it is quite difficult to provide a precise definition. Several definitions in data fusion have been proposed. Pohl and Van Genderen (1998) defined image fusion as “the combination of two or more different images to form a new image by using a certain algorithm” which is restricted to image. Hall and Llinas (1997) said data fusion techniques “combine data from multiple sensors, and related information from associated databases, to achieve improved accuracy and more specific inferences that could be achieved by the single sensor alone”. This definition focuses on information quality and fusion methods. These definitions imply that purpose of data fusion should hopefully be information obtained to improve image visualisation and interpretation.

There are several fusion approaches; generally fusion can be divided in three main categories based on the stage at which the fusion is performed namely: pixel-based, feature-based and decision-based. Among these approaches, only the pixel-based methods has been considered in this study. In pixel-based fusion, the data are merged on a pixel-by-pixel basis. Feature-based approaches always merge the different data sources at the intermediate level. Each image from different sources is segmented and the segmented images are fused together in. Decision-based fusion, the output of each of the single source interpretations are combined to create a new interpretation.

Previous research into this theme concentrated on developing fusion algorithms and their assessment of their advantage was limited to visual enhancement. Experiments using integrated information of fused images were seldom repeated. In particular, the effects of multisensor fusion in image classification have only been reported in Munechika *et al.* (1993). Here, a pair of SPOT panchromatic and Landsat TM images was used in the fusion process. In testing the classification performance of the fused images, 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 individual images. However, alternative types of images, fusion methods, levels of classification and classifiers were not evaluated.

In this paper we applied intensity-hue-saturation (IHS), principal component analysis (PCA) and high pass filter (HPF) fusion methods, in order to test the most suitable to improve land cover classification of an urban area.

IHS colour transformation separates spatial (I) and spectral (H, S) information from a standard RGB image. There are two ways of applying IHS techniques in image fusion: direct and substitutional. The first refers to the transformation of three image channels assigned to I, H and S. The second transforms three channels of the data set representing RGB into the IHS colour space which separates the colour aspects in its average brightness (intensity). This corresponds to surface roughness, its dominant wavelength contribution (hue) and its purity (saturation). Then, one of the components is replaced by a fourth image channel which is to be integrated. Reverse transformation from IHS to RGB converts the data into its original image space to obtain the fused image.

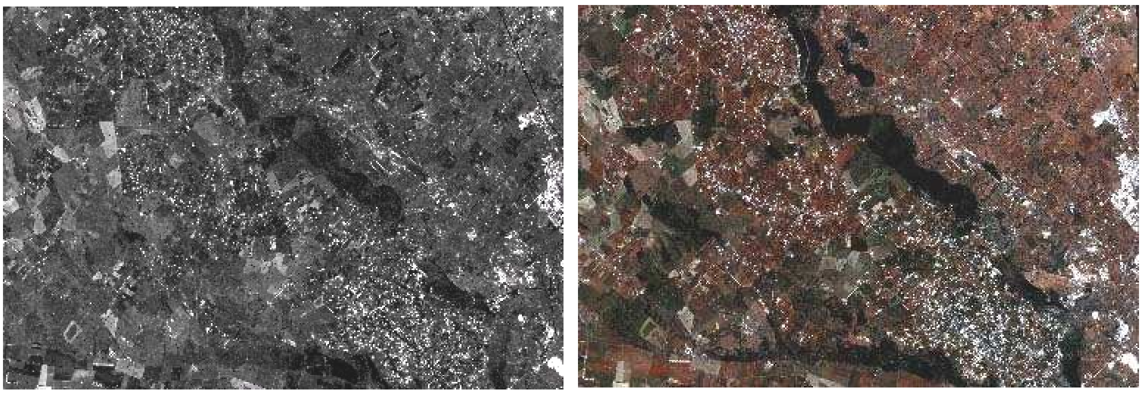
The purpose of applying PCA is to reduce the dimensionality of input data with high correlation into a smaller number of output channels. This fact is well documented for all of the optical sensors currently available. In PCA most input information is transformed into the first component and information content decreases by increasing of the number of PCA components.

In the HPF method, the high spatial resolution image is filtered with a small high-pass-filter resulting in the high frequency part of the data which is related to the spatial information. This is pixel wise added to the low resolution bands (Pohl *et al.*, 1998).

4. EXPERIMENTAL TESTING

4.1 Test Area and Data Used

The experiment in this study used a pair of IKONOS panchromatic (Figure 1a) and multispectral (Figure 1b) images (Acquisition Date/Time: 2000-05-12 / 09:14). To reduce the size of image to be manipulated and, to maintain geographic consistency during fusion, equivalent sub-scenes were defined.



Figg. 1a-1b. Original image data:
(a) IKONOS panchromatic image; (b) IKONOS multispectral image.

4.2 Fusion Processing

Examination of the correlation between intensity data derived from the multispectral images and the panchromatic data underscores the need for additional processing before directly substituting panchromatic data for the intensity component of the merged product (Carper, Lillesand and Kiefer, 1990).

To facilitate integration and geo-locking of the IKONOS panchromatic and multispectral imagery it was necessary to register both data sets in a single map coordinate system, in this case Gauss-Boaga map projections and Roma40 datum. A high (3rd) order polynomial transformation for rectification (RMS value lower than 1 pixel) was executed, in order to consider the relief of certain sub-areas. In this process identification of 30 common Ground Control Points (GCP), regularly distributed in the study area and obtained by recognising permanent and detailed elements at imagery scale, was planned. After the "geometric locking" process, fusion of the images was performed by the three approaches, IHS, PCA and HPF.

The multispectral image was transformed into IHS and PC spaces, and the high spatial component (panchromatic 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 multispectral image. For the HPF fusion, spatial detail from panchromatic image was extracted by differential high pass filtering and directly integrated into the multispectral image to form HPF-fused image. The three images generated by IHS, PCA and HPF fusion process are shown in Figure 2.

Fig. 2. IKONOS panchromatic and multispectral images:
(a) IHS-fused image; (b) PCA-fused image; (c) HPF-fused image.

The aim of drawing attention improving land cover classification and the spatial resolution of the IKONOS images permitted individuation of four categories agriculture, vegetation, urban and barren land for level one classification. Another seven categories sown ground, olive-groves, grassland, woodland, transportation, mixed urban areas and open space were defined for level two classification. The descriptions of all categories are summarized in Table 1.

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

CLASS	DESCRIPTION
LEVEL 1	
Agriculture	Cultivated coverage includes sown ground and olive-groves
Vegetation	Vegetation coverage includes woodland, grassland and swamp
Urban	Artificial coverage includes residential, commercial, industrial and transportation
Barren	Uncovered coverage includes rock outcrop, country roads and construction sites
LEVEL 2	
Sown ground	
Olive-Groves	Include also fruitful trees
Grassland	Shrub and brush covered areas
Woodland	Forest covered areas
Transportation	Linear routes such as asphalt roads and boundaries outline the other land use
Mixed urban	High density include residential, commercial and industrial buildings
Open	Temporary barren land include construction sites, country roads and bare space

In acquiring ground truth data, the size and delineated units were designed with reference to Congalton (1991). For the sample size of training data, a minimum of 50 samples for each category was used. For each class, a stratified random sampling technique was used to capture the location of each training set. The delineated locations were distributed over the entire image. Based on ancillary information, more than 30 sites were delineated for each of the two classification cases. Two sets of data were identified for each class, one set was used for classification training and the other to assess the thematic accuracy.

After the training process, all the initial and transformed multispectral images were classified into the two levels through the three selected approaches.

The basic procedure for this study was to conduct a digital classification of the study area using standard processing techniques applied to various combinations of manipulations of the original data set, the original data set being a spatially coregistred set of five IKONOS bands (multispectral and panchromatic data), all resampled to 1 m pixels. Spectral signatures were extracted for the various cover types (Table 1) using supervised training site procedures. The field visits were used to identify valid training sites. After signature extraction, a maximum likelihood decision rule was employed to classify the data set and a contingency table created for accuracy assessment. The contingency table was developed from a set of truth sites, separate from the training areas, also derived from the field effort. This study area did not allow similar size truth data for each cover type.

Table 2. The thematic accuracy of land cover classification results level one.

	Original Multispectral Data	IHS-fused	PCA-fused	HPF-fused
Agriculture	96.30%	92.10%	90.50%	89.70%
Vegetation	90.60%	85.20%	83.30%	82.70%
Urban	85.20%	81.60%	80.10%	75.40%
Barren	62.30%	60.70%	59.00%	57.80%
Overall Accuracy	83.60%	79.90%	78.22%	76.40%

Table 2 lists overall accuracy and the percentage of correctly classified pixels in individual classes of level one classification result using three initial images and transformed spectral image.

For level two classification, a second, independent data sets was obtained for the training stage. Classification results are summarized in Table 3.

Table 3. The thematic accuracy of land cover classification results level two.

	Original Multispectral Data	IHS-fused	PCA-fused	HPF-fused
Sown ground	95.30%	96.70%	80.20%	85.70%
Olive-Groves	85.80%	89.30%	74.10%	86.60%
Grassland	64.10%	83.20%	52.30%	71.50%
Woodland	64.40%	73.20%	58.70%	75.60%
Transportation	36.60%	40.80%	43.10%	39.60%
Mixed urban	76.70%	84.10%	73.20%	69.20%
Open	78.20%	80.20%	40.90%	78.40%
Overall Accuracy	71.58%	78.21%	60.35%	72.37%

5. ANALYSIS OF EXPERIMENTAL RESULTS

In level one classification (Table 2), the mean accuracy of the fused image was 5.3 percent less than the original multispectral image. The degradation of accuracy is significant in classification using fused images. The spatial information obtained from the panchromatic image did not improve the accuracy of classification of the multispectral image by MLC approach. The reason for degraded accuracy may have come from the distorted spectral characteristics. By the visual inspection of the fused images, the spectral characteristics of the multispectral 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 are inappropriate for the classification.

In the level two MLC classification (Table 3), the higher overall accuracy of results using the IHS and HPF fused images over the original spectral image were achieved by up to 6.6 and 0.8 percent respectively. The improvement of overall accuracy was significant.. All the classes showed significant enhancement in accuracy of the classification of the fused images. However, the PCA-fused image, with a large degree of distorted spectral characteristics, produced a 11.2 percent degradation of the overall accuracy of classification. As these improved classes were categorised by the sub-division of classification level, the intention of using fused spatial information to improve level two classification accuracy can be upheld.

6. CONCLUSION AND DISCUSSION

Land cover classification works with spectral information of an image. The fusion experiments presented showed that the spectral characteristics of an IKONOS multispectral image were distorted after fusion with an high-spatial resolution IKONOS panchromatic image. This implies that successful image fusion, which generates an image with both spatial and spectral resolutions, cannot avoid changing the spectral characteristics of the transformed multispectral image. Since purity of spectral characteristics was degraded after the fusion, the use of multisensor fusion cannot enhance the thematic accuracy in the case of the natural land cover classification, which requires a greater quantity of spectral information than spatial details. In level one classification results the overall accuracy using the fused images deteriorated.

For level two classification that requires more spatial details, the complicated patterns of features, multisensor fusion can potentially improve accuracy of results. The classification results of MLC demonstrated the improvement of thematic accuracy after using the multisensor fusion methods of IHS and HPF. Among the three selected fusion approaches, IHS fusion approach for level two classification produced higher accuracy results than the transformed image.

If higher level land cover classification for urban areas is required from multispectral images, unless the image has high spatial resolution, multisensor fusion with an image of higher spatial resolution is recommended.

REFERENCES

Aplin P., Atkinson P. M. and Curran, P. J. (1997). Fine spatial resolution satellite sensors for the next decade. *International Journal of Remote Sensing*, 18, pp. 3873-3881.

Caprioli M. and Tarantino E. (2001), "Linear Elements in the Classification of High Spatial Resolution Satellite Imagery", in Proceedings of International Symposium on "Geodetic, Photogrammetic and Satellite Technologies Development and Integrated Applications", Sofia, November, 8-9, 2001, pp. 244-250.

Caprioli M. and Tarantino E. (2001), "Accuracy assessment of per-field classification integrating very fine spatial resolution satellite sensors imagery with topographic data", *Journal of Geospatial Engineering*, 3(2), pp. 127-134.

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

Congalton R.G. (1991), "A review of Assessing the Accuracy of Classifications of Remote Sensed Data", *Remote Sensing of Environment*, 20(37), pp. 35-46.

Ehlers M., Jadkowski M.A., Howard R.R., Brostuen D E. (1990), "Application of SPOT data for regional growth analysis and local planning", *Photogrammetric Engineering and Remote Sensing*, 56(2), pp. 175-180.

Fritz L.W. (1999), "High resolution commercial remote sensing satellites and spatial information", ttp://www.isprs.org/publications/highlights/highlights0402/fritz.html.

Gao J., Skillcorn D. (1998), "Capability of SPOT XL data in producing detailed land cover maps at the urban-rural periphery", *International Journal of Remote Sensing*, 19(15), pp. 2877-2891.

Gong P., Howarth P.J. (1990), "An assessment of some factors influencing multispectral land-cover classification", *Photogrammetic Engineering and Remote Sensing*, 56(5), pp. 597-603.

Hall D. and Llinas J. (1997), "An introduction to multisensor data fusion", *Proceedings of the IEEE*, 85(1), pp. 6-23.

Harris P.M., Ventura S.J. (1995), "The integration of geographic data with remotely sensed imagery to improve classification on an urban area", *Photogrammetric Engineering and Remote Sensing*, 61(8), pp. 993-998.

Lillesand T.M. and Kiefer R.W. (2000), *Remote Sensing and Image Interpretation*, 4th edition, New York, Wiley & Sons.

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

Pohl C. and Van Genderen J.L. (1998), "Multisensor image fusion in remote sensing: concepts, methods and applications", *International Journal of Remote Sensing*, 19(5), pp. 823-854.

Ridd M.K. (1995), "Exploring a VIS (vegetation-impervious surface soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities", *International Journal of Remote Sensing*, 16(12), pp. 2165-2185.

Saraf K. (1999), "IRS-1C-LISS-III and Pan data fusion: an approach to improve remote sensing based mapping techniques", *International Journal of Remote Sensing*, 20(10), pp. 1929-1934.

Westmoreland S. and Stow D.A. (1992), "Category identification of changed land-use polygons in an integrated image processing/geographic information system", *Photogrammetic Engineering and Remote Sensing*, 58, pp. 1593-1599.