URBAN LAND COVER MAPPING USING HYPERSPECTRAL AND MULTISPECTRAL VHR SENSORS: SPATIAL VERSUS SPECTRAL RESOLUTION

Fabio DELL'ACQUA, Paolo GAMBA, Gianni LISINI

Department of Electronics, University of Pavia, ITALY
{fabio.dellacqua, paolo.gamba, gianni.lisini}@unipv.it

Working Group III/5

KEY WORDS: Urban remote sensing, hyperspectral imaging, very high resolution sensors.

ABSTRACT

In this paper a discussion about the effects of very high resolution (VHR) in both the spatial and the spectral dimensions for urban land cover mapping is proposed. We confirm that VHR spatial sensors are unable to discriminate between some urban materials, since they often come from the same chemical family. However, VHR in the spectral sense sometimes carries too much information, since it differentiates covers made by the same material but with different age or illumination conditions. Finally, after out test with a supervised classifier, we stress the importance to use the context for a more accurate mapping, and offer one simple way to improve the classification using *a priori* knowledge about the geometric constraints of the segmented map.

1 INTRODUCTION

Land cover mapping (and consequently land use mapping) in urban areas relies, at the level of detail required by most urban planners, on very high spatial resolution (VHR) sensor. Recently, the availability or the development of new airborne and satellite sensors with very high spectral resolution allowed the comparison of data sets from these different sources for urban material detection, recognition and characterization.

Indeed, hyperspectral data have been used for many urban applications. Urban material characterization Marino *et al.* (2001) has been considered by means of the MIVIS sensor, allowing not only to provide a precise definition of the roof materials, but also to build a database of dangerous roof tops for subsequent substitution and destruction. Similarly, vegetation canopies and tree in urban forests provide information extremely interesting for energy exchange analysis, air quality (micro-climate) and hydrology. Their distribution, height, stress and species have been already derived from low altitude AVIRIS measurements in Xiao *et al.* (1999). As a further example, detection of land cover and sealed parts in an urban area can be accomplished using the multi-technique approach proposed in Heiden *et al.* (2003), where HyMap data have been analyzed for a better characterization of urban environments. Finally, a detailed analysis of the spectral properties required by sensors to obtain interesting results in urban areas is given in Herold *et al.* (2003), where it was shown that AVIRIS's fine sampling of the electromagnetic spectrum provides a much better result than multi-spectral Ikonos (simulated) data, because the latter lack information on significant portion of the infrared wavelength range. However, no discussion of the spatial resolution issue is provided in the paper, since Ikonos is simulated at the same resolution than AVIRIS.

One final point to be discussed is therefore if the combination of spatial and spectral VHR leads to significantly better results than one of the two characteristics taken alone, and to which extent, if any. So, this paper is devoted to the comparison of two different hyperspectral data sets, coming from the ROSIS and DAIS sensors, and a Quickbird image.

The two hyperspectral sensors provide the best spatial available up to date (2.6 and 1.2 m posting, respectively), while Quickbird mounts the finest multi-spectral satellite sensor available (2.8 m). The aim of this paper is therefore to improve the knowledge of the present possibility for urban land cover mapping right now, considering the effect of both the spectral and the spatial properties of the data. Moreover, it may provide useful information for future implementation and choice in sensor design targeted to urban applications.

2 NEIGHBOURHOOD-AWARE SPECTRAL CLASSIFICATION

The classification tool used for comparison is based on the technique recently developed in Gamba *et al.* (2003) for multi-classification fusion of very fine spatial and spectral data. A simple scheme of the processing steps is as follows:

- a feature extraction procedure is applied to the data in order to reduce dimensionality;
- supervised classifications are performed, comparing the results of different spectral and spatial classifiers;
- all the classification maps are combined using majority voting or Linear/Logarithmic Opinion Pools;
- a further geometric refinement step is introduced, and spatial information from the neighborhood of each pixel is taken into account to improve the results.

(a) (b) (c)

Figure 1: First test area: a part of the town center, as imaged (false colors) by the DAIS (a), ROSIS (b) and Quickbird (c).

Several feature extraction approaches have been proposed for the first step. One of the best known is discriminant analysis feature extraction (DAFE), (Landgrebe, 2003), which is a method which is intended to enhance separability. In DAFE, a within-class scatter matrix, Σ_W, and a between-class scatter matrix, Σ_B, are defined:

$$\Sigma_W = \sum_i P(\omega_i)\Sigma_i \tag{1}$$

$$\Sigma_B = \sum_i P(\omega_i)(M_i - M_0)(M_i - M_0)^T \tag{2}$$

$$M_0 = \sum_i P(\omega_i)M_i \tag{3}$$

where M_i is the mean vector for the $i-th$ class, Σ_i is the covariance matrix for the $i-th$ class, and $P(\omega_i)$ is the prior probability of the $i-th$ class. The criterion for optimization may be defined as

$$J = \mathrm{tr}(\Sigma_W^{-1}\Sigma_B) \tag{4}$$

where tr( ) denotes the trace of a matrix. New feature vectors are selected to maximize the criterion. The necessary transformation from X to Y is found by taking the eigenvalue-eigenvector decomposition of the matrix $\Sigma_W^{-1}\Sigma_B$ and then taking the transformation matrix as the normalized eigenvectors corresponding to the eigenvalues in decreasing order. However, this method does have some shortcomings. For example, since discriminant analysis mainly utilizes class mean differences, the feature vectors selected by discriminant analysis are not reliable if mean vectors are near to one another. Since the lumped covariance matrix is used in the criterion, discriminant analysis may lose information contained in class covariance differences. Also, the maximum rank of Σ_B is $M-1$ since Σ_B is dependent on M_0. Usually Σ_W is of full rank and, therefore, the maximum rank of $\Sigma_W^{-1}\Sigma_B$ is $M-1$. This indicates that at maximum $M-1$ features can be extracted by this approach. Another problem is that the criterion function above generally does not have direct relationship to the error probability.

In the early nineties, decision boundary feature extraction (DBFE) was proposed by Lee and Landgrebe (1997). They showed that discriminantly informative features and discriminantly redundant features can be extracted from the decision boundary itself. They also showed that discriminantly informative feature vectors have a component which is normal to the decision boundary at least at one point on the decision boundary. Further, discriminantly redundant feature vectors are orthogonal to a vector normal to the decision boundary at every point on the decision boundary. Thus, a *decision boundary feature matrix* (DBFM) was defined in order to extract discriminantly informative features and discriminantly redundant features from the decision boundary. It can be shown that the rank of the DBFM is the smallest dimension where the same classification accuracy can be obtained as in the original feature space. Also, the eigenvectors of the DBFM, corresponding to non-zero eigenvalues, are the necessary feature vectors to achieve the same classification accuracy as in the original feature space.

The choice to perform different classifications and combine their outputs depends on their different behaviors. It is extremely difficult to find rules for choosing the best classifier in any situation. It's better to combine their maps as in Briem *et al.* (2002). Here we refer to the well-known Majority voting and Opinion Pools approaches (Benediktsson and Kanellopoulos, 1999). Majority voting is a very standard procedure, and provides a classification map where each pixel is assigned to the class to which it belongs in the majority of the maps to be jointly considered. The Linear Opinion Pool (LOP) and the logarithmic opinion pool (LOGP) procedures require instead a training set, which inherently defines the output classes to be considered in the final classification map. For any possible pattern of the output classes in the training set we compute $p(i,j)$, the probability that the pixels of the i-th class of the training set are associated with the j-th pattern in the multiple classification set. Finally, the output class is defined by maximizing $C_i^{LOP}(\mathbf{X}) = \sum_{j=1}^{n} \lambda_j p(i,j)$ or $C_i^{LOGP}(\mathbf{X}) = \sum_{j=1}^{n} p(i,j)^{\lambda_j}$ where $\mathbf{X}$ is the input classification pattern, n the number of classified images, and λ_j the weight of

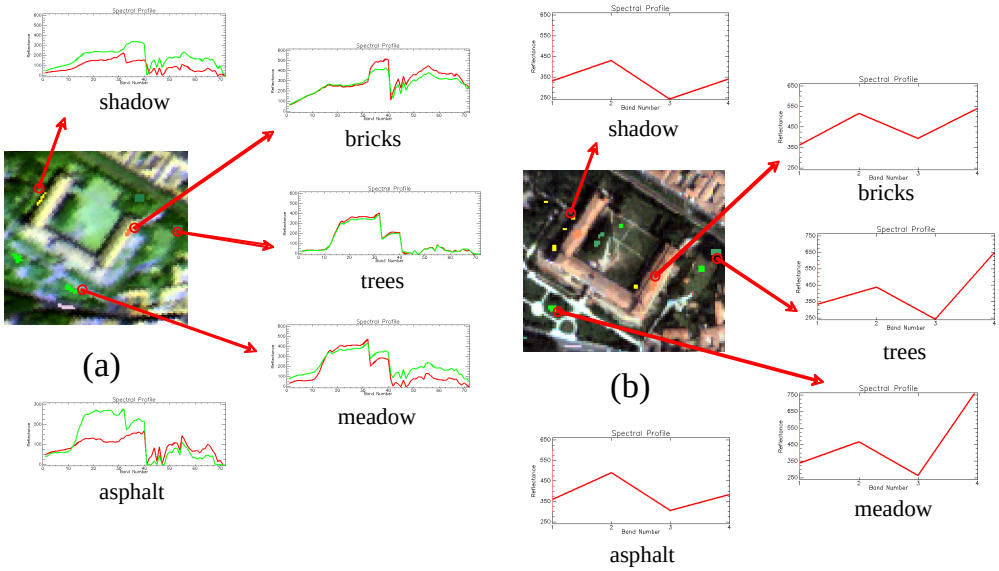

Figure 2: Spectra of some covers in DAIS data (a), and in Quickbird data (b).

the j-th map (default is $\lambda_j = 1$). Of course, there could be some patterns that are not present in the training set. For these patterns, the output class assignment follows majority voting.

The combined maps come from different classifiers, some based on spectral information (Maximum Likelihood, Fisher Linear Discriminant, Fuzzy ARTMAP), others on spatial information also (ECHO, Fuzzy ARTMAP with spatial reclassification). The last classifier is a refined version of the fuzzy ARTMAP classification chain introduced in Gamba and Houshmand (2001); Gamba and Dell'Acqua (2003). The output of a pixel-by-pixel classification is used by a second classifier whose input is a vector representing the mapping class percentages in a window around each pixel plus a random value to enhance their discriminability.

After multiple map combination, we applied a further reclassification step, again by means of the fuzzy ARTMAP classifier, which usually provide better and more precise results, since it corrects random errors due to randomly correlated noise, i. e. errors, in the maps.

Even if the methodology has been fully presented in previous papers, the usefulness is this research is in comparing the best results on the same test area and the same test and training sets for different data sets. An accurate test on the spectral and spatial properties of the training sample may help not only to understand the different performances, but also to extrapolate, beyond the particular class legend we used, which may be the best choice for urban land cover mapping with these sensors.

Moreover, in this paper we introduce a final step after the above mentioned classification chain, aimed at a further geometric refinement of the classification. This step is based on the usual methodology applied to panchromatic VHR images, and based on the experience gained with photogrammetric data processing. Indeed, as correctly stated in recent review papers, (see Guindon (2001) and Gamba *et al.* (2003), for instance), that it is time now to start considering that the photogrammetric-based and the classification-based views should be integrated into a general framework. One possible idea, which is somehow a generalization of Jin and Davis (2004), is to match a top-down segmentation step with a bottom up approach based on geometric features and cues grouping. In the top-down step, an initial partitioning of areas of interest within a scene is accomplished using simple supervised classification schemes. A reduced resolution image may be considered, and textural features may be used to complement the original information. In the bottom-up step, each part of the scene and its surroundings are searched for basic patterns of covers, and corrections to misclassifications are made based on the context. Alternatively, geometric features are extracted to recognize objects instead of classification segments, and the class of each object is chosen by majority voting.

Following these line guides, in this work we apply a post-processing step to the classified land cover map, aimed at exploiting some of the constraints for urban covers. For instance, we may assume that roof cover pixel belong to geometric structures delimited by orthogonal edges. This assumptions leads to a correction procedure that searches for segmented regions belonging to a roof cover class and extracts their boundary. Successively, these boundaries are matched with a piecewise linear path with 90° angles. Finally, the corresponding regions are re-drawn on the original map.

3 EXPERIMENTAL RESULTS

We offer results about the town of Pavia, Northern Italy. Four DAIS and ROSIS flight lines over the area, kindly provided by the German Space Agency in the framework of the HySens project, cover the town and its immediate surroundings. The Digital Airborne Imaging Spectrometer (DAIS) is a multi-band systems with 80 bands in the visible and infrared wavelength range (from 0.4 to 12.6 μm),

(a) (b) (c)

Figure 3: Second test area: the Engineering School: classification maps obtained using the DAIS (a), ROSIS (b) and Quickbird (c) data.

while the Reflective Optics System Imaging Spectrometer (ROSIS) is a multi-band sensor focused only on visible and near infrared frequency bands. The latter comprises 32 frequency bands from 0.45 to 0.85 μm, showing higher potential for vegetation and green area mapping than the former one, which, on the contrary, has a broader range of applications, since it provides information up to the thermal infrared. To these data sets we added elaborations made on a Quickbird panchromatic and multi-spectral images (4 bands, visible and near infrared), acquired in the same period (July 2002).

The fine spectral resolution of some of these data sets couples well with a very high spatial resolution, from 2.6 m for DAIS to 1.2 m for ROSIS to 0.7 m for pan-sharpened Quickbird data.

In the urban test area of interest, the following results show and compare algorithm performances on two different test areas, one in the center of the town, considering typical urban land covers, and one in the surroundings, where both urban and rural covers exist. This choice allows analyzing the overall and class accuracy values with respect to the spectral and spatial properties of different legends, and also to focus on the effects due to typical problems in dense urban area, such as masking by shadows. The first area is depicted in fig. 1: please note that the areas are not exactly the same, due to differences in the fight lines. For example, due to the finer spatial resolution and the subsequent reduced swath of the ROSIS sensor, only two parts of the scene are shown in the ROSIS false color image, which belong to two non-overlapping lines of flight.

All the supervised classifiers used exploit non-overlapping training and a test sets composed by nearly 100 pixels for each class, and referring to 7 classes of urban covers, plus water. Covers include trees, meadows and bare soil, asphalt and bitumen, parking special cover and brick roof. To discard problems with shadowed areas, a special class for shadow pixels was added.

The highest achievable values for the overall accuracy among all the tests made are shown in Table 1. Note that these values do not take into account the final geometric refinement, since this may introduce some bias in the results due to its dependence on spatial resolution only. This is the reason why it will be discussed separately, at the end of this section.

The accuracy values in the Table show that for this first test area, but also globally speaking, spatial re-classification using ARTMAP neural networks is by far the best option. Moreover, these values stress the fact that high spatial resolution is not enough, and lower spatial posting is well-balanced by multi-spectral capability. As a matter of fact, DAIS maps are always better than Quickbird ones. Moreover, it should be mentioned that in Quickbird maps some of the covers are not actually present. For instance, bitumen and asphalt, being virtually indistinguishable due to the low spectral resolution, are jointly considered. Finally, ROSIS mapping accuracy is based of course on both the fine spatial and spectral resolution.

To give an idea of the different problems coming from spatial and spectral high resolution, we offer in fig. 2 the spectra of five land covers around the area of the castle of Pavia, as they can be deduced by the Quickbird and DAIS data sets. The two subsets show also the difference in spatial details, and thus allow comparing the global effect of both resolutions on the mapping capabilities of the two sensors.

As for the second test area, the Engineering School of the University of Pavia, the land cover legend used was partially different. We considered 9 covers, i. e. trees, meadows and bare soil, asphalt and bitumen, gravel, parking special cover and metallic roofs. Even in this case the lower spectral resolution of Quickbird data did not allow discriminating bitumen against asphalt, and the two classes were jointly considered. Again, a shadow class was artificially introduced, to reduce misclassified pixels. The classification maps for this areas are shown in fig. 3.

Overall accuracy (%)	*DAIS*	*ROSIS*	*Quickbird*
Center	97.5	99.3	91.8
Engineering School	87.5	91.4	80

Class. algorithm	*DAIS*	*ROSIS*	*Quickbird*
Center	ARTMAP (DAFE)	Spatial ARTMAP	Spatial ARTMAP
Engineering School	Majority voting	Spatial ARTMAP	Spatial ARTMAP

Table 1: Comparison of the best overall classification accuracy values and classification methods for land cover maps obtained from data by each of the three sensors and referring to the same test areas.

A look at the maps reveals that there are some misclassifications and pixels assigned to the wrong (even if spectrally similar) class. We have, for instance, some exchanges between pixels in the bitumen and gravel classes. As a matter of fact, the gravel on top of some of the buildings is posed on a bitumen substrate, so that the spectral response is somehow mixed. Similarly, bitumen and asphalt are also exchanged in some areas. ROSIS data show a better performance, but suffer from a stronger dependency on the training set. Pure pixels are mandatory and a few mixed spectral response may change dramatically the per-class accuracy values.

3.1 Geometric refinement of the classified maps

As introduced above, a final step was applied to some of the classes. The procedure is based on a boundary extraction for each of the segments of the class of interest, followed by a piece-wise linear reconstruction of this boundary and, finally, a step that poses constraints on the angle between subsequent segments. There are therefore three parameters ruling the procedure: the maximum length of the segments and the maximum distance from the original boundary determine how approximate is its piece-wise linear representation. The tolerance on the angle to be "corrected" is used to choose which are the segments whose direction is to be re-assigned in a clock-wise path along the piece-wise linear approximation.

An example of block refinement through geometric constraints is shown in fig. 4(a). Similarly, the segments of the brick roof class in parts of the Quickbirdclassification map of the whole town, before and after the geometric correction, are showed in fig. 4(b) and (c), respectively. We may observe an improvement in the shape of the objects, that become more recognizable. However, the shapes are not as regular as expected, as a consequence of the imprecise piece-wise linear representation of their boundaries. This in turn is due to the small number of pixels defining the boundaries. We may observe more visible improvements when considering, in another part of the town, near the Engineering school, the same approach as applied to the "road" (actually, asphalt) class, as shown in fig. 4(d) and (e). In this example the advantage of using geometrical constraints on the shape of the objects is more evident, and results in a regularization of their boundaries.

4 CONCLUSIONS

The present work shows that, with respect to urban and urban/rural land cover mapping, very high resolution in the spectral sense is more valuable than in the spatial sense. It may be better to have more bands recorded by a sensor that a more detailed image of the scene.

The main reason for this behavior is the similarity of many covers due to their very similar chemical components. Indeed, once a sufficient spatial resolution is achieved, so that in the image there are mostly spectrally pure pixels, the geometric properties of the objects may be improved exploiting *a priori* information about their shape. This information is, in fact, valuable but already available for urban areas and objects herein.

Instead, the correctness of assignment of an object to the right land cover class (and possibly the right land use class) is based on the possibility to recognize the spectra of the materials, which in turn depends on the number and positions of the imaged wavelengths.

ACKNOWLEDGMENTS

The authors want to thank Francesco Grassi for his help in performing some of the classification tests. The classifications using Quickbird data were performed at the University of Siena. Thanks to Andrea Garzelli for running the tests and providing the mapping results.

References

C.M. Marino, C. Panigada, L. Busetto: "Airborne hyperspectral remote sensing applications in urban areas: asbestos concrete sheeting identification and mapping", Proc. of the 1st IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, Italy, 8-9 Nov. 2001, pp. 212-216.

Q. Xiao, S. L. Austin, E. G. McPherson, and P. J. Peper: "Characterization of the structure and species composition of urban trees using high resolution AVIRIS data", Proc. of the 1999 AVIRIS Workshop, Pasadena, CA, 1999, unpaginated CD-ROM.

U. Heiden,K. Segl, S. Roessner, H. Kaufmann: "Ecological evaluation of urban biotope types using airborne hyperspectral HyMap data", Proc. of the 2nd IEEE/ISPRS

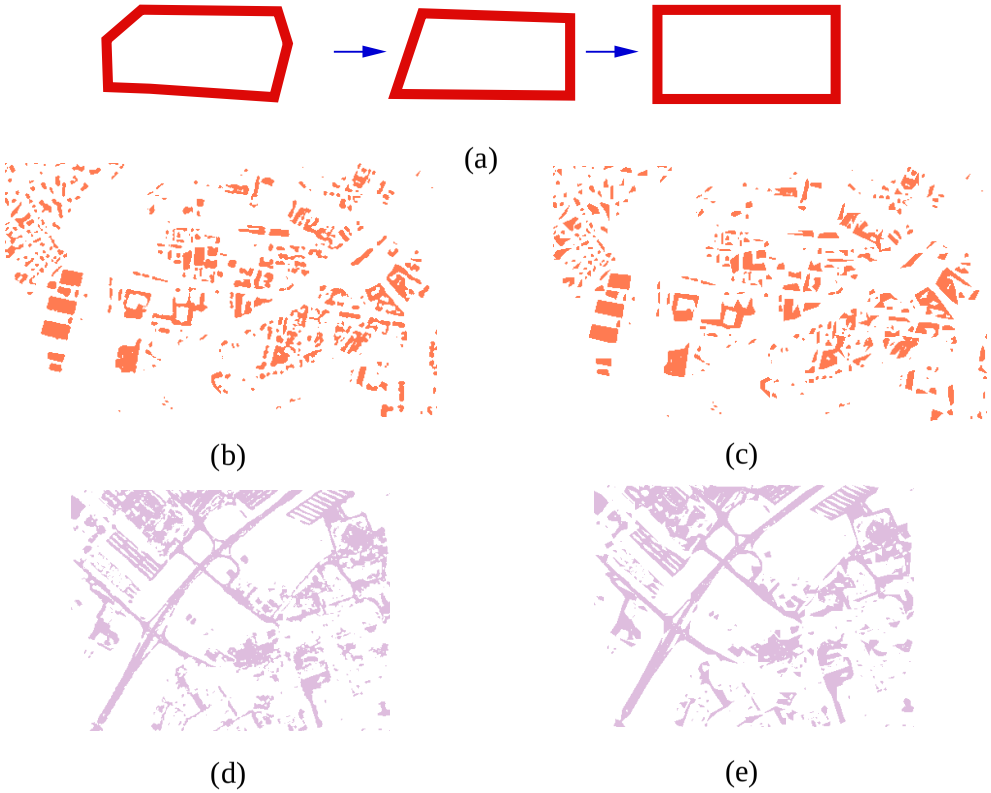

Figure 4: (a) an example of geometric refinement of a classified segment of the scene; the same approach is applied to the segments in (b) to obtain (c), and in (d) to obtain (e).

Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Berlin, Germany, 22-23 May 2003, pp. 18-22.

M. Herold, M.E. Gardner, D.A. Roberts: "Spectral resolution requirements for mapping urban areas", *IEEE Trans. Geosc. Remote Sens.* Vol. 41, n. 9, pp. 1907-1919, Sept. 2003.

P. Gamba, F. Dell'Acqua, A. Ferrari: "Exploiting spectral and spatial information for classifying hyperspectral data in urban areas", Proc. of *IGARSS'03*, July 2003, Vol. I, pp. 464-466.

D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley and Sons, Hoboken, New Jersey, 2003.

C. Lee and D. A. Landgrebe, "Decision Boundary Feature Extraction for Neural Networks, " *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 8, n. 1, pp. 75-83, 1997.

G.J. Briem, J.A. Benediktsson,and J.R. Sveinsson, "Multiple classifiers applied to multisource remote sensing data" *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 40, n. 10, pp. 2291-2299, Oct. 2002.

J.A. Benediktsson and I. Kanellopoulos: "Classification of multisource and hyperspectral data based on decision fusion," *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 37, n. 3, pp. 1367-1377.

P. Gamba, B. Houshmand: "An efficient neural classification chain for optical and SAR urban images", *International Journal of Remote Sensing*, Vol. 22, n. 8, pp. 1535-1553, May 2001.

P. Gamba, F. Dell'Acqua: "Improved multiband urban classification using a neuro-fuzzy classifier", *International Journal of Remote Sensing*, Vol. 24, n. 4, pp. 827-834, Feb. 2003.

B. Guindon, "Computer-based aerial image understanding: a review and assessment of its application to planimetric information extraction from very high resolution satellite images", Canadian Journal of Remote Sensing, Vol. 23, No. 1.

P. Gamba, O. Hellwich, P. Lombardo: Editorial to the Theme Issue on "Algorithms and Techniques for Multi-Source Data Fusion in Urban Areas", ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, n.1-2, pp. 1-3, May 2003.

Jin, X. and C.H. Davis, "Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information", EURASIP Journal on Applied Signal Processing, Special Issue on Advances in Intelligent Vision Systems: Methods and Applications, 2004.