A Watershed-Based Segmentation Technique for Multiresolution Data

Giuseppe Masi[1], Giuseppe Scarpa[1], Raffaele Gaetano[2], and Giovanni Poggi[1]

[1] DIETI, University Federico II of Naples
`firstname.lastname@unina.it`
[2] TSI Department,TELECOM-ParisTech
`gaetano@telecom-paristech.fr`

Abstract. A new watershed-based technique is proposed for the segmentation of multiresolution remote-sensing images. These images are composed by a high-resolution panchromatic band and a low-resolution multispectral set. To achieve a segmentation with the high resolution of the panchromatic image and the high accuracy granted by the spectral information, the two components are processed jointly, using both spectral and morphological properties. In addition, a fully automatic marker generation procedure is introduced to reduce the oversegmentation typical of watershed methods. Experiments on WorldView-2 multiresolution images demonstrate the potential of the technique.

1 Introduction

Image segmentation is a low-level processing task of critical importance in several applicative domains, like remote sensing [1–5] or medicine [6, 7], to mention just a few. In remote sensing, in particular, segmentation of multiresolution (MR) images is becoming a very relevant topic. In fact, because of technological limitations, satellite sensors cannot reach both high spatial and spectral (large number of bands) resolution. MR sensors, such as Ikonos, GeoEye, or World-View, overcome this problem by providing a single high-resolution (typically, below 50 cm) panchromatic (PAN) image complemented by a low-resolution multispectral (MS) image composed of 4-8 bands, relying then on signal processing, like pansharpening techniques [8], to recover a full multispectral data cube. Pansharpening, however, cannot really preserve all desired information, and high accuracy can be only guaranteed by direct processing of the original data. Here we propose a new segmentation technique specifically designed for multiresolution images and based on the watershed transform [9], a widespread technique particularly suited for the preservation of fine details.

To obtain a watershed segmentation, a topographic surface is first associated with the image. Then it is progressively filled with "water", and each time two basins meet a dam is built between them in order to avoid their merging. Once the surface is completely flooded the process stops and each water basin is regarded as an image segment. In the context of edge-based segmentation, one can apply watershed by using the distance transform to define the topographic

surface. Watershed will therefore provide a full segmentation where all original edges are guaranteed to be part of the final region boundaries, thus completing the original edge map.

Unfortunately, watershed produces many more segments than desired, because of noise and intrinsic image textures, calling for the adoption of some suitable merging strategy to reduce them, a computation-intensive step, and a possible source of errors. Several approaches have been proposed in the literature to limit over-segmentation. One can identify shallow basins, likely due to noise, and merge them with some suitably chosen neighbor [5]. Inconsistent region separations can be singled out and removed by analyzing the image at a larger scale of interaction [10]. Better yet, one can intervene before running the watershed by introducing some "markers" in the topographic surface, which force separate basins to be treated as a single entity. Markers can be put manually from an operator, a tedious and low-precision task or, more interestingly, through a specific automatic procedure, e.g. [1, 11]. This latter approach was followed in [11], leading to the Edge, Mark and Fill (EMF) algorithm, suited for single-band images. In this work we generalize the EMF strategy to the case of color/multispectral, and in particular multiresolution images, providing new tools to further reduce oversegmentation.

After briefly recalling EMF in next Section, the proposed segmentation algorithm is described in Section 3, while Section 4 shows experimental results. Concluding remarks are given in Section 5.

2   Edge, Mark and Fill Segmentation

The Edge, Mark and Fill technique is a marker-controlled watershed applied to the distance transform, which embeds a suitable morphological process for a fully-automatic generation of markers.

To gain insight into the EMF algorithm consider the example of Fig.1. In (a) the image to be segmented is shown together with the contours, in red, produced by a suitable edge detector. The distance transform, D, is shown in (b): its opposite $-D$ is used as topographic surface (DEM) for the watershed. In order to produce effective markers, EMF computes the crest lines of D, shown in (c) together with the edge map, which represent a support for the skeletons of the closed regions singled out by the watershed. The local maxima of D on the crest lines are then extracted, the red points in (d), and used as marker seeds. These seeds, say s_k, are dilated with a circular morphological profile of radius $D(s_k) - \epsilon$, where the margin ϵ prevents crossing the edges. Overlapping dilated seeds are merged together as shown in (d). The segmentation provided by watershed with the DEM $-D$ and the above markers is shown in (e), while (f) shows the difference between the initial edge map (a) and the final (closed) contour map (e). Removed edges are represented in red, while additional edges needed to close open segments are shown in green. For further details about EMF the reader is referred to [11].
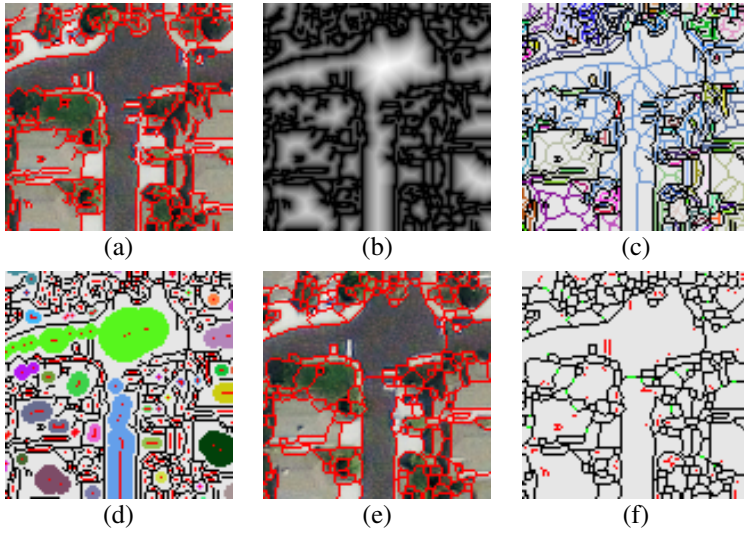
Fig. 1. Intermediate results of the EMF Algorithm

Experimental evidence shows that EMF reduces oversegmentation with no loss of valuable information, and preserving the high local accuracy of edge-based approaches. However, it keeps generating too many segments in textured and noisy areas, also because it works on a single-band image on the basis of purely morphological information. On one hand, this is a strength of EMF, which can deal with gray-level images or, more in general, with any given edge map, with no need for further information. On the other hand, when further information *is* available, as with MR images, it must be taken into account. Therefore, in next Section, we describe an enhanced version of EMF for MR images.

3   Multiresolution EMF

The oversegmentation typical of watershed is mostly due to edge map imperfections, with meaningless edges originated by minor texture variations, or edges which, because of noise, depart from their correct position. In both cases, the effect is a proliferation of small regions, with very similar spectral content, which should be merged together. Our primary goal, here, is to take into account the available spectral information to reduce oversegmentation from the beginning. In addition, with reference to MR images, we want to exploit effectively the different cues coming from the two image components, the single-band but high-resolution PAN, and the low-resolution spectrally rich MS.

In Fig.2 a high-level flowchart of the MR-EMF algorithm is shown. Unlike in basic EMF, the work-flow is split in two channels, in order to process the two image components. Edge detection is carried out in both channels, by means of
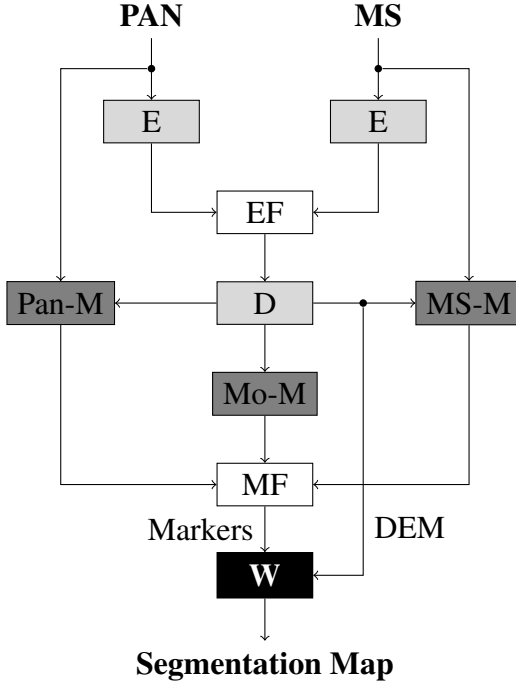
Fig. 2. Flowchart of MR-EMF. E: Edge Detection, EF: Edge Fusion, D: Distance transform, Pan-M: PAN Marker generation, MS-M: MS Marker generation, Mo-M: Morphological Marker generation, MF: Marker Fusion, W: Watershed.

a standard Canny detector [12], with the results combined to produce a single high-resolution edge map (here and in the following, for the sake of simplicity, we neglect all up/down-sampling and registration operations needed to combine data of the two channels). Two new sets of spectral markers are generated, one for each channel, besides the morphological markers used in EMF, and all markers are eventually combined at high resolution to guide the watershed transform.

Let us now analyze the algorithm structure in more detail. Intuitively one could choose to consider only the PAN data to generate geometrical markers, and only MS data for spectral markers, given the superior geometric resolution of the former and spectral resolution of the latter, thus renouncing to the full-symmetric approach depicted in Fig.2. However, there are good reasons to keep a balanced approach.

First of all, it can happen that edges visible in the MS component are not visible at all in the single-band PAN. Therefore, both edge maps must be taken into account. Of course, when contours are detected in both channels, preference must be given to the more accurate PAN data, neglecting the others. Experiments confirm that MS-edges help improving results. This is all the more true
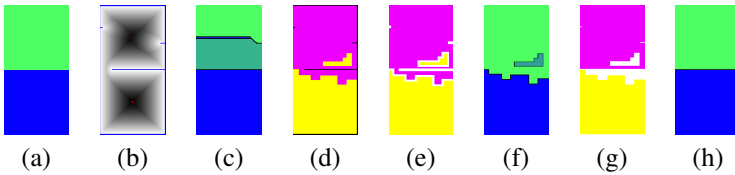
Fig. 3. Marker erosion and application: ground-truth (a); edge map (blue) with related distance transform (b); unmarked watershed (c); external spectral clustering (d) and its erosion, pre-markers, (e) by edge subtraction; (pre-)marker-controlled watershed (f); final markers (g) and related watershed (h)

for spectral marker generation, where both PAN and MS data are essential for a correct use of spectral information.

3.1   Generation of Spectral Markers

The generation of spectral markers is, in principle, very simple. Since our goal is to recognize and merge together homogeneous segments that share the same spectral properties, we only have to cluster data in the spectral domain and use the resulting connected components as markers. When a marker crosses the border between two segments, because pixels on both sides are spectrally homogeneous, the segments will be eventually merged. This vision, however, is overly simplistic as it relies on perfect segment homogeneity. In the actual cases, instead, within each segment, it can happen to find pixels assigned to different clusters, which risk to produce undesired results. This is basically due to the obvious inconsistency between the edge detection, which is a local process, and the clustering, which is a global one. Hence it is necessary to reshape the detected clusters before using them as markers.

Marker Erosion Process. In order to explain the proposed solution, we consider the toy example of Fig.3. Let (a) be the ideal segmentation (ground-truth), and (b) the detected edge map (blue) with corresponding distance function (gray scale). The given morphology gives rise to the wrong (over-) segmentation (c), if a simple unmarked watershed is applied. Now, assume to have a spectral-based clustering like that shown in (d), where three connected components can be distinguished. Edge map subtraction gives the markers in (e) which lead to an unsatisfactory marker-controlled segmentation (f) with both over- (additional region) and under-segmentation (lost edges). We therefore apply an additional erosion process to obtain more reliable markers. For each pixel p of a pre-marker M in (e) the attracting minimum in the DEM, say p_m, is localized. If $p_m \in M$ than p is kept, otherwise it is removed from M. As result, pre-markers which do not cover any minimum of the DEM will simply disappear because unreliable. Those who cover one or more minima will be constrained to the their attraction

basins. The final markers obtained after the erosion, shown in (g), lead to the desired segmentation shown in (h).

In particular, the algorithm used in our method to single out the clustering (d) is a simple mean shift procedure [13].

PAN *vs* MS Image Domain Partition. To apply this tool to MR images, a preliminary key observation is necessary, that the reliability of the spectral information provided by the MS component depends on the size of the segments to characterize. In fact, since MS pixels are much larger than PAN pixels (typically 4×4 times) they can happen to average together high-resolution pixels with different spectral characteristics, typically near segment boundaries. As a consequence, "pure" MS pixels can be found only in the segment interior, provided this is large enough. Therefore, for small or elongated segments, the MS spectral characterization is highly unreliable and, below a given segment size, it is more appropriate to use the PAN data.

Based on this consideration, we perform a preliminary partition of the image in two regions where "thick" and "thin" segments are found, using MS-based spectral markers in the former region and PAN-based in the latter. To this end, a simple thresholding on the distance function may seem to work well. Given a suitable threshold d, objects less than d pixel wide (along the shortest dimension) are better characterized in the PAN domain, while the others are reliably featured by MS data. The region $\widetilde{M} = D < d$, which is nothing but a dilation of radius d of the edge map, will certainly cover entirely the thin objects. However, it will also overlap with the contours of thick elements, partially covering them. To avoid this inconvenient, we apply to $\widetilde{M}$ the same region erosion mechanism described above to reshape the spectral markers. In other words, elements of $\widetilde{M}$ whose attracting minimum does not belong to $\widetilde{M}$ are discarded. Once singled out the mask for thin object, the PAN domain, its complement is used as MS domain.

An example of this process for a real-world MR image is shown in Fig.4. In (a) is the PAN component of the image to segment, and in (b) the composite MS-PAN edge map. (c) shows the thin-thick mask obtained on the basis of the distance transform of (b). The final PAN and MS markers are shown in (d) in green and red, respectively.

4   Experimental Results

The segmentation and classification accuracy assessment in remote sensing is a challenging problem as testified by recently developed benchmarking systems, like [14], specifically thought for segmentation of remote sensing data. Unfortunately no one such solutions deals with multiresolution data.

The lack of multiresolution data with suitable ground-truth (not just the interior of large segments) prevents a reliable quantitative assessment of segmentation accuracy, as well as a meaningful comparison with competing algorithms, like the TFR [15, 16] or the technique proposed in [17]. Therefore, we confine
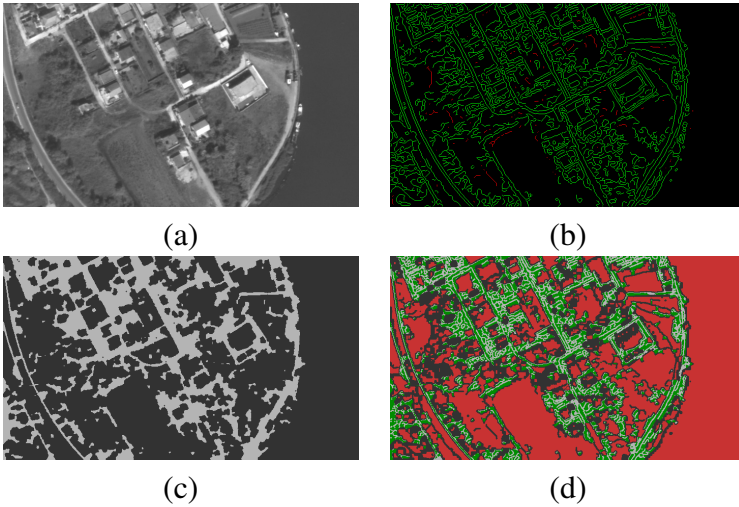
(a)   (b)

(c)   (d)

Fig. 4. Spectral markers generation: PAN component (a); edge map (b); PAN (light) and MS (dark) domains (c); PAN (green) and MS (red) markers (d)

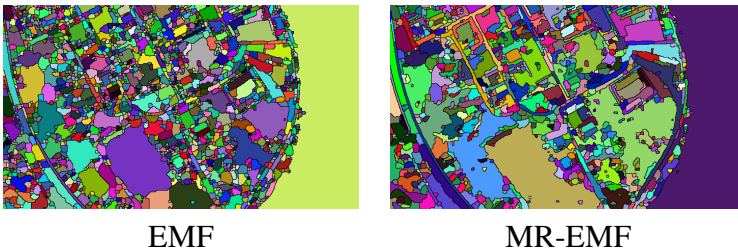EMF   MR-EMF

Fig. 5. Segmentation maps obtained with EMF (left) and with MR-EMF (right)

our analysis, for the time being, to the visual inspection of results obtained on a large dataset of ©WorldView-2 images, comprising a 8-band MS component with 2-meter resolution and a PAN component with 50-cm resolution. We found experimentally that the best value for d, the threshold used to generate the PAN-MS domain partition, is in the range 6-8 (high-resolution) pixels, not far from the size of a MS pixel. In particular we selected $d = 6$ for the experiments.

The first experiment presented concerns the image used in the running example of Fig.4. In Fig.5 we show the marker-controlled watershed segmentation with morphological markers alone (EMF), and with superposition of spectral markers (proposed MR-EMF). Although a significant simplification of the map is evident, the achieved gain can be better appreciated by looking at the few enlarged details shown in Fig.6. Segmentation contours (in red) are superimposed to the RGB subset of the MS component. On the left column are the results of the unmarked
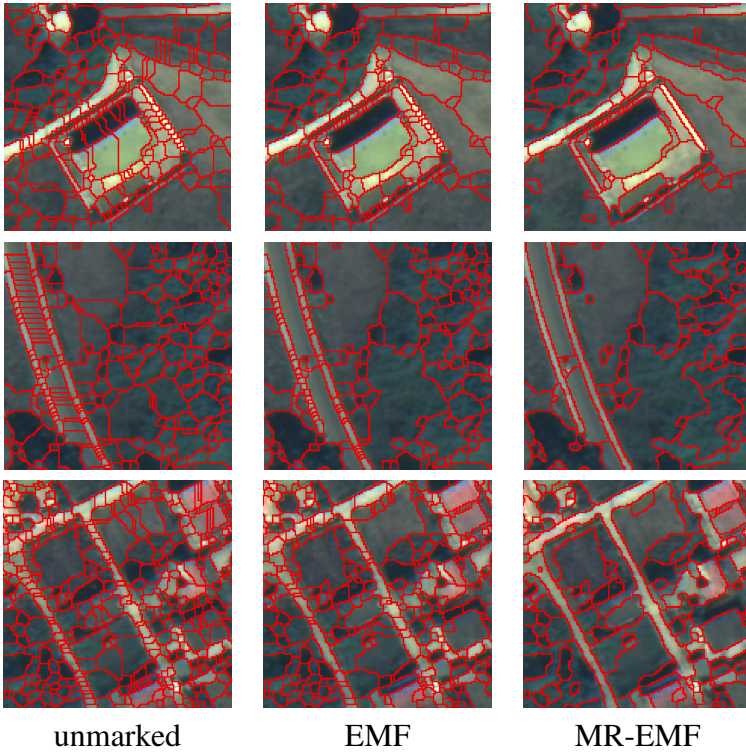
Fig. 6. Segmentation details for the running experiment of Fig.4: unmarked watershed (left), EMF (center), and MR-EMF (right)
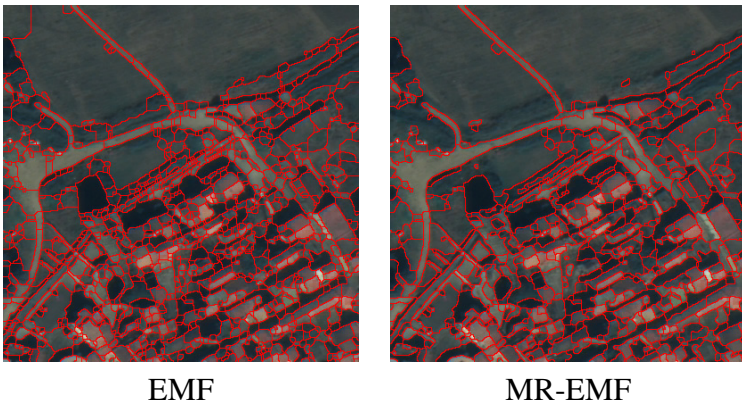

Fig. 7. Segmentation resultrs provided by EMF (left) and MR-EMF (right)

watershed. In the middle and right columns the results of EMF and MR-EMF, respectively. As can be easily appreciated by visual inspection the accuracy improves considerably moving toward the proposed solution. Over-segmentation is largely reduced, with no apparent loss in discriminating capabilities.

Similar considerations hold for the second experiments shown in Fig.7 which confirms a general improvement with respect to the basic EMF.

5   Conclusions

The proposed watershed-based segmentation algorithm for multiresolution images guarantees the high local accuracy of all edge-oriented techniques and a limited over-segmentation. A key step is the automatic generation of morphological and spectral markers, obtained by properly combining cues coming from both PAN and MS components and weighting their relative reliability.

Preliminary experiments have shown very promising results, with a large reduction of the over-segmentation and no apparent loss of details. In future work we will consider replacing the edge map fusion with a high resolution edge detector that works jointly on PAN and MS data, and designing an *ad hoc* clustering algorithm to improve upon the basic mean-shift.

References

1. Xiao, P., Feng, X., Zhao, S., She, J.: Multispectral ikonos image segmentation based on texture marker-controlled watershed algorithm. In: SPIE 6790, MIPPR (2007)
2. Cagnazzo, M., Poggi, G., Verdoliva, L.: Region-based transform coding of multispectral images. IEEE Transactions on Image Processing 16, 2916–2926 (2007)
3. Parrilli, S., Poderico, M., Angelino, C.V., Scarpa, G., Verdoliva, L.: A nonlocal approach for SAR image denoising. In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2010, pp. 726–729 (2010)
4. Cagnazzo, M., Parrilli, S., Poggi, G., Verdoliva, L.: Improved Class-Based Coding of Multispectral Images With Shape-Adaptive Wavelet Transform. IEEE Geoscience and Remote Sensing Letters 4(4), 566–570 (2007)
5. Li, P., Guo, J., Song, B., Xiao, X.: A multilevel hierarchical image segmentation method for urban impervious surface mapping using very high resolution imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4(1), 103–116 (2011)
6. Bova, N., Ibanez, O., Cordon, O.: Image Segmentation Using Extended Topological Active Nets Optimized by Scatter Search. IEEE Computational Intelligence Magazine 8(1), 16–32 (2013)
7. D'Elia, C., Marrocco, C., Molinara, M., Poggi, G., Scarpa, G., Tortorella, F.: Detection of microcalcifications clusters in mammograms through TS-MRF segmentation and SVM-based classification. In: 17th International Conference on Pattern Recognition, ICPR 2004, vol. 3, pp. 742–745 (2004)
8. Alparone, L., Wald, L., Chanussot, J., Thomas, C., Gamba, P., Bruce, L.M.: Comparison of Pansharpening Algorithms: Outcome of the 2006 GRS-S Data-Fusion Contest. IEEE Transactions on Geoscience and Remote Sensing 45(10), 3012–3021 (2007)

9. Beucher, S., Lantuejoul, C.: Use of Watersheds in Contour Detection. In: International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation, Rennes, France (September 1979)
10. Arbeláez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(5), 898–916 (2011)
11. Gaetano, R., Masi, G., Scarpa, G., Poggi, G.: A marker-controlled watershed segmentation: Edge, mark and fill. In: IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2012, pp. 4315–4318 (2012)
12. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI 8(6), 679–698 (1986)
13. Comaniciu, D., Meer, P.: Mean Shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(5), 603–619 (2002)
14. Mikes, S., Haindl, M., Scarpa, G.: Remote sensing segmentation benchmark. In: 7th IAPR International Workshop on Pattern Recognition in Remote Sensing, PRRS 2012, Tsukuba Science City, Japan (November 2012)
15. Scarpa, G., Haindl, M.: Unsupervised texture segmentation by spectral-spatial-independent clustering. In: 18th International Conference on Pattern Recognition, ICPR 2006, vol. 2, pp. 151–154 (August 2006)
16. Gaetano, R., Scarpa, G., Poggi, G.: Hierarchical texture-based segmentation of multiresolution remote-sensing images. IEEE Transactions on Geoscience and Remote Sensing 47(7), 2129–2141 (2009)
17. Yuan, J., Wang, D.L., Li, R.: Remote Sensing Image Segmentation by Combining Spectral and Texture Features. IEEE Transactions on Geoscience and Remote Sensing (to appear)