Different implementations of the ARSIS concept to fulfill users needs

Thierry Ranchin, Lucien Wald

To cite this version:
Thierry Ranchin, Lucien Wald. Different implementations of the ARSIS concept to fulfill users needs. 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, May 2003, Berlin, Germany. pp.299-304. hal-00395055

HAL Id: hal-00395055
https://hal.archives-ouvertes.fr/hal-00395055
Submitted on 14 Jun 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Different Implementations of the ARSIS Concept to Fulfill Users Needs

Thierry Ranchin, Member, IEEE, and Lucien Wald

Abstract—There is a need for synthetic products that simulate satellite multispectral data at a higher spatial resolution. The users needs are various and somehow opposite: needs of high spatial resolution for accurate geometrical description of objects and needs for high spectral resolution for accurate spectral content and subsequent classification process. The ARSIS concept (from its French name "amélioration de la résolution spatiale par injection de structures" that means improvement of the spatial resolution by structure injection) permits to construct such synthetic products by fusion of two sets of data, namely panchromatic images at high spatial resolution and multispectral images at lower spatial resolution. It offers a framework wherein different models can be combined to form different solutions to better satisfy the users needs. An example is given of an Ikonos image of the city of Hasselt, Belgium.

Index Terms—High spatial resolution, Ikonos, Image fusion, Multiscale analysis, Remote sensing.

I. INTRODUCTION

There are several examples of application in image fusion where two sets of images are available: one with the desired quantity / measurement but with a low spatial resolution, hereafter noted $B_h$, and the other hereafter noted $A_h$, having a high spatial resolution, $h$, but describing another quantity. If these quantities can be associated, there is a desire to fuse both sets of images in order to synthesize a new set, noted $B_h^s$, that are representations of the quantity $B$ at the spatial resolution $h$.

Remote sensing images are current examples. They exhibit usually either high spectral resolution and low spatial resolution, or low spectral resolution (broadband) and high spatial resolution. The high spatial resolution is necessary for an accurate description of shapes, features and structures. The different objects are better identified if high spectral resolution images are used.

The commercial companies are tackling the need for fusion products combining the high spatial and the high spectral resolutions using well-known procedures and deliver such products in an operational way. These fusion products are synthetic images aiming at simulating what a sensor having the same spectral bands but the highest spatial resolution would observe.

The properties of these synthetic images $B_h^*$ have been established [1]:

1) Any synthetic image $B_h^*$ once degraded to its original resolution $l$, should be as identical as possible to the original image $B_l$.

2) Any synthetic image $B_h^*$ should be as identical as possible to the image $B_h$ that the corresponding sensor would observe with the highest spatial resolution $h$, if existent.

3) The multispectral (or multi-modality) set of synthetic images $B_h^*$ should be as identical as possible to the multispectral (or multi-modality) set of images $B_h$ that the corresponding sensor would observe with the highest spatial resolution $h$, if existent.

II. THE ARSIS CONCEPT

In the different methods for production of synthetic images by fusion of two sets of modalities, some are relevant to the ARSIS concept [2, 3]. The ARSIS concept (from its French acronym: "amélioration de la résolution spatiale par injection de structures", which means “spatial resolution enhancement by injection of structures”) has been designed in a generic way, in order to fulfill the properties previously defined, whatever the mathematical tools used for its implementation. The approach is not new, but was never expressed as a concept before the common work of Ecole des Mines de Paris and Aérospatiale [4, 5, 6].

Figure 1 presents the general scheme of the ARSIS concept in the case where $A$ is made of one modality. Inputs to the fusion process are the modality $A$ at high spatial resolution ($A_h$, resolution $n1$) and the multi-modality images $B$ at low spatial resolution ($B_n$, resolution $n2$). The first step consists in the use of a model for the hierarchical description, the modeling and the synthesis of structures. This model allows the computation of approximations of image $A$ and the
hierarchical description of existing structures (also called
details) in this image (step 1 in Figure 1). The model used
should allow image reconstruction without loss. It is called
MultiScale Model (MSM). The same operation is applied on
image \( B \) (step 2 in Figure 1).

This hierarchical description of structures given for each
decomposition is used for parameter adjustment of a model
linking the known structures of sets \( A \) and \( B \) (step 3).
Numerous solutions exist for this model. It is called Inter-
Modality Model (IMM) in Figure 1.

From this model, another model called High Resolution
Inter-Modality Model (HRIMM) is derived. It performs the
transformation of the parameters of the Inter-Modality Model
with the change in resolution and allows the computation,
from the known structures from sets \( A \) and \( B \), of the searched
structures of image \( B \) (step 4). Finally, the inversion of the
model for the hierarchical description allows the synthesis of
image \( B^\# \) with the spatial resolution of image \( A \) (step 5).

Fig. 1. General scheme of application of the ARSIS concept. See text for more explanation.

III. IMPLEMENTATION OF THE ARSIS CONCEPT

A. The MSM Model

The most used MSM model for implementing the ARSIS
concept is the multiresolution analysis combined with wavelet
transform. A group of authors \([2, 7, 8]\) uses the algorithm
proposed by Mallat \([9]\). Others \([10, 11]\) use that proposed by
Dutilleux \([12]\), called here undecimated wavelet transform
(UWT).

Other tools exist for this MSM model. Laplacian pyramids
\([13]\) can be used for the hierarchical description of
information as well as gaussian pyramids \([14]\). \([15]\) defined
filters for the High Pass Filtering (HPF) method. \([16]\) used
iterative filters banks in place of wavelet transform. \([11, 17, 18]\)
propose the use of generalized laplacian pyramids (GLP).
Other solutions exist such as morphological pyramids \([19]\) or
Fourier transformations.

B. The IMM Model

Numerous solutions exist for the Inter-Modality Model
(IMM). \([2]\) proposed two global models of transformation
called model 1 and model 2, both performing on the detail
images. Model 1 (M1) is an identity model. The missing
details in image \( B \) are identified to those known in image \( A \).
Model 2 (M2) adjusts the mean and variance of the two
probability density functions (pdf) of the structure images of
\( A \) and \( B \) at resolution n°3.

\([11]\) describe two local IMM models. The first one (RWM),
used in this paper, describes the local transformation between
the structure descriptions of \( A \) and \( B \) given by the MSM
model at resolution n°3 (Figure 1). The second one (AABP)
is used similarly, but applies on approximations given by the
MSM model.

C. The HRIMM Model

The choice of the High Resolution Inter-Modality Model
(HRIMM) is not obvious. Influence of the spatial resolution
on the set of information extracted from satellite imagery is
well known but not well modeled. To our knowledge, no
published fusion method paid particular attention to this point
and the HRIMM is often set identical to the IMM, except \([20]\)
who performed a multiscale synthesis of the parameters of
their IMM from resolution n°3 to resolution n°2.

IV. SOME EXAMPLES OF IMPLEMENTATION OF THE ARSIS
CONCEPT

In this part is presented different implementations of the
ARSIS concept, applied to the fusion of images provided by
the Ikonos satellite. These implementations differ due to the
choice of the different MSM, IMM and HRIMM models,
taking into consideration their mutual influences.

In order to limit the combination of models, only one MSM
model was chosen: a wavelet transform combined with a
multiresolution analysis and called UWT [12]. This selection influences greatly the choice of the IMM and HRIMM models, because of the modeling of the structures as variations around a zero value. More precisely, the models should not introduce bias in the modeling of the missing structures.

The following part describes four methods for synthesis of fusion products in the case of Ikonos data.

A. Description of the Four Methods

The four methods use the same MSM models, but combine different IMM and HRIMM models. For the Ikonos imagery, resolution n\(^{1}\) is 1 m (pixel size of panchromatic image called PAN), resolution n\(^{2}\) is 2 m, resolution n\(^{3}\) is 4 m (pixel size of multispectral images called MS), and resolution n\(^{4}\) is 8 m. With respect to the description in Figure 1, where the resolution is defined in a generic way, resolution n\(^{2}\) of this scheme corresponds to resolution n\(^{3}\) (4 m) for Ikonos imagery case.

The first method is called UWT-RWM. The IMM model used is the RWM model. This model is described in great details in [11]. It is computed at resolution n\(^{4}\) (8 m). The HRIMM model is identical to the IMM model, and applied to resolutions n\(^{2}\) and n\(^{3}\) (resp. 2 and 4 m).

The second method is called UWT-ATS3. It was developed in the framework of the ATS3-E3 action dedicated to benchmarking of image fusion products, held by the French Space Agency (CNES) in 2000-2002. The IMM model combines the models M2 and RWM as follows. Both models are computed at resolution n\(^{4}\) (8 m). The HRIMM model consists in the application of the RWM model at resolution n\(^{3}\) (4 m), then on application of the M2 model combined with a denoising algorithm at resolution n\(^{2}\) (2 m) to provide the fused product at 1 m. The denoising algorithm gives the value zero to the structures that are less than a given threshold (here 0.1). It is inspired from the work of [21] on shrinkage.

The third method is called UWT-M1. In this case the IMM and HRIMM models are the identity model.

The last one proposed in this paper is called UWT-M2. The IMM model is model M2 computed at resolution n\(^{4}\) (8 m). The HRIMM model is in this case identical to the IMM model.

B. Description of the Case Study

The four methods are applied to a set of images from the Ikonos satellite. The GIM Company courtesy provided us this set for testing. This set is composed of a panchromatic image PAN at the spatial resolution of 1 m and of four multispectral MS images at the spatial resolution of 4 m. The spectral bands of the images are detailed in Table 1. The geographical area is the city of Hasselt in Belgium. The images were acquired the 28th of April 2000 at 10:39 UT. The original data are delivered with a dynamic range of 11 bits (gray values). Table 1 reports the mean value and standard deviation of each band, in gray values. It also provides the correlation coefficient between each band and the panchromatic image resampled at 4 m. The application of the methods leads to the synthesis of MS images at the spatial resolution of 1 m with the same dynamic range in 11 bits.

<table>
<thead>
<tr>
<th>Spectral Band</th>
<th>PAN (450-900)</th>
<th>Blue (450-520)</th>
<th>Green (520-600)</th>
<th>Red (630-690)</th>
<th>Near Infrared (760-900)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Value</td>
<td>352</td>
<td>328</td>
<td>339</td>
<td>249</td>
<td>467</td>
</tr>
<tr>
<td>Standard Dev.</td>
<td>55</td>
<td>26</td>
<td>39</td>
<td>48</td>
<td>124</td>
</tr>
<tr>
<td>Corr. Coeff.</td>
<td>1.00</td>
<td>0.271</td>
<td>0.409</td>
<td>0.323</td>
<td>0.875</td>
</tr>
</tbody>
</table>

Figure 2 shows an excerpt of the PAN image at 1 m. On the left, there is a river, crossed by two bridges, with small boats and several barges. Along the river bank is a main street. Several cars are visible. This area is mainly an industrial district with large buildings, surrounded by numerous trees. On the top right is a stadium. Its lawn is partly degraded. Figure 3 presents an extract of the same area in near infrared (NIR) band with a spatial resolution of 4 m. This NIR band is generally the most difficult to synthesize and is a good comparison case for the different methods.
C. Visual Results

Figures 4, 5, 6 and 7 are the NIR images synthesized respectively by the UWT-RWM, UWT-ATS3, UWT-M1 and UWT-M2 methods. The look-up tables of each image are similar, allowing visual comparison of the different products.

The visual quality of each product is very different. The UWT-RWM image (Fig. 4) presents a smooth aspect. The objects are difficult to distinguish. One can observe a granular aspect (see e.g. the stadium). The objects are not enough contrasted. A kind of halo seems to surround many objects in the UWT-RWM image; see in particular the boats and the stadium. The UWT-ATS3 image (Fig. 5) is of better visual quality. The objects are easier to distinguish. But the visual quality is not satisfactory. The UWT-M1 image (Fig. 6) presents a visual quality similar to the UWT-ATS3 one, but a bit inferior. The UWT-M2 image (Fig. 7) offers the best definition of objects. It allows a better recognition and identification of the objects in the image.

D. Quantitative Results

The protocol of quantitative evaluation of the fused products defined by [1] was applied. All the computations were made on the full dynamic range of images (11 bits). The dynamic range of the analyzed information is not at all perceptible on the visual representation proposed in this paper. The human eye is only sensitive to a few colors. The different quantitative results are not detailed in this paper. It was found that the three properties detailed in part 1, are respected with different levels of accuracy for the four methods. This is in accordance with a similar study achieved on SPOT images [22].

As an illustration, Table 2 deals with the test of the second property and reports some statistics on the relative discrepancies between the original images $B_d$ and the images $B_{*k}$. The differences are computed on a pixel basis and one image of differences is obtained per spectral band $k$. From each image of differences, the mean value (bias), standard deviation and root mean square error (RMSE) are computed. In Table 2, these quantities are expressed in percent, relative to the mean radiance value of the original image $B_d$. The ideal value for these parameters is 0. In addition, the difference between the variance of the original image $B_d$ and that of $B_{*k}$ is computed. It is expressed in percent, relative to the variance of the original image. Ideally, this value should be zero. The correlation coefficient between the original image $B_d$ and $B_{*k}$ is also computed. The ideal value is 1.

For all methods, the bias is very small. This is in accordance with the first property. The standard deviations and RMSEs are small for all methods. The fused products UWT-RWM and UWT-ATS3 do not contain enough variance (i.e. fine structures). For the blue and green bands, the UWT-M1 and UWT-M2 add too much fine structures (difference in variance is negative). This fact is confirmed by the low correlation coefficient in these cases. This observation meets the visual analysis performed on the 1 m products.

The UWT-RWM method is the method given the best achievable results whatever the spectral band (second property), as well as on the multispectral set of products (third property, results not shown here). The UWT-M2 method gives similar results. However, the addition of fine structures is usually more annoying for further processing than a lack of fine structures of similar intensity (difference in variance positive). Both can be considered as satisfactory. Then comes the UWT-ATS3 method. The UWT-M1 method is that giving the worst results of the four, the two latter being fairly unsatisfactory.

<table>
<thead>
<tr>
<th>Band</th>
<th>Bias</th>
<th>Standard Dev.</th>
<th>RMSE</th>
<th>Diff. in Variance</th>
<th>Correlation Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>UWT-RWM 0.01</td>
<td>3</td>
<td>3</td>
<td>23</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>UWT-ATS3 0.02</td>
<td>3</td>
<td>3</td>
<td>23</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>UWT-M1 0.02</td>
<td>4</td>
<td>4</td>
<td>-20</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>UWT-M2 0.04</td>
<td>7</td>
<td>7</td>
<td>-22</td>
<td>0.84</td>
</tr>
<tr>
<td>Green</td>
<td>UWT-RWM 0.01</td>
<td>5</td>
<td>5</td>
<td>24</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>UWT-ATS3 0.03</td>
<td>5</td>
<td>5</td>
<td>26</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>UWT-M1 0.03</td>
<td>6</td>
<td>6</td>
<td>-8</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>UWT-M2 0.04</td>
<td>7</td>
<td>7</td>
<td>-8</td>
<td>0.92</td>
</tr>
<tr>
<td>Red</td>
<td>UWT-RWM 0.01</td>
<td>8</td>
<td>8</td>
<td>23</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>UWT-ATS3 0.03</td>
<td>8</td>
<td>8</td>
<td>25</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>UWT-M1 0.03</td>
<td>10</td>
<td>10</td>
<td>-4</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>UWT-M2 0.08</td>
<td>7</td>
<td>7</td>
<td>1</td>
<td>0.92</td>
</tr>
<tr>
<td>NIR</td>
<td>UWT-RWM 0.03</td>
<td>8</td>
<td>8</td>
<td>9</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>UWT-ATS3 0.04</td>
<td>10</td>
<td>10</td>
<td>23</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>UWT-M1 0.04</td>
<td>9</td>
<td>9</td>
<td>24</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>UWT-M2 0.10</td>
<td>5</td>
<td>5</td>
<td>-1</td>
<td>0.98</td>
</tr>
</tbody>
</table>

V. Conclusion

The results of the comparison between the different methods enhance the difficulty to answer to users needs with a single method. The UWT-M1 method answers to a need of easy implementation and speed of use and production. The UWT-RWM and UWT-ATS3 methods answer to the spectral quality need of the fused products and are more adapted to further processing such as classifications [23, 24]. The UWT-M2 method is a good compromise between the visual and spectral qualities of products.

The ARSIS concept is an excellent framework for the development of accurate methods that can be tailored to
specific users needs and can be assembled within a toolbox. It is also an open framework with still many places for the development of different cases of applications and approaches for implementation.

Different methods can be developed based on this concept, depending upon the multiscale description and synthesis model MSM, the Inter-Modality Model relating the content of both representations $A$ and $B$ and the High Resolution Inter-Modality Model transforming the parameters of the IMM when increasing the spatial resolution. Though all examples of models given above are dyadic cases, that is the ratio $l/h$ is a power of 2, the ARSIS concept is applicable to any value of the ratio, provided one can find appropriate filters for the analysis and synthesis steps of the multiresolution analysis [16, 17, 18].

A dependency between the different models used in the ARSIS concept exists. The availability of a set of solutions for each model of the concept (MSM, IMM and HRIMM) allows the adaptation of the fusion process to users needs. They can be chained differently in order to obtain a result satisfying the targeted application. However, this should be done coherently in order to respect the mathematical and physical hypotheses of each model. Hence, the availability of a toolbox comprising all possible solutions, should be linked with a training of the users. This will allow them to fully understand the results of their choice on the quality of the synthetic products obtained by image fusion.

REFERENCES


Thierry Ranchin (M’01) a PhD degree from the university of Nice Sophia Antipolis, France in the field of applied mathematics. After a post-doctoral fellow in a company in Tromso, Norway, he joined the remote sensing group of Ecole des Mines de Paris in the fall of 1994. He was an invited scientist from the University of Jena, Germany in 1998. He has a patent about sensor fusion and more than 90 publications, communications in international symposia, or articles in journals with peer review committees in the field of remote sensing of the Earth system and in the field of image processing. Dr. Ranchin received the Autometrics Award in 1998 and the Erdas Award in 2001 for articles on data fusion.

Lucien Wald graduated in theoretical physics in Marseille and Paris, France, in 1977. He received the Ph.D. degree in 1980 (Paris, France) and the Doctorat
d’État ès Sciences in 1985, both on the applications of remote sensing to oceanography.

He is a Professor at École des Mines de Paris since 1991 and is supervising Ph.D. students. He is currently the Head of the Remote Sensing Group, and is focusing his own research in applied mathematics and energy.

Prof. Wald received the Autometrics Award in 1998 and the Erdas Award in 2001 for articles on data fusion. In 1996, he was the recipient of the Médaille Blondel, a famous French reward for outstanding career in information technologies.

Fig. 4. An excerpt of the image synthesized at 1 m by the UWT-RWM method.

Fig. 5. As Fig. 4, but for the UWT-ATS3 method.

Fig. 6. As Fig. 4, but for the UWT-M1 method.

Fig. 7. As Fig. 4, but for the UWT-M2 method.