A TWO-STEP DECISION FUSION STRATEGY: APPLICATION TO HYPERSPECTRAL AND MULTISPECTRAL IMAGES FOR URBAN CLASSIFICATION

Walid Ouerghemmi\textsuperscript{a}, Arnaud Le Bris\textsuperscript{a}, Nesrine Chehata\textsuperscript{b}, Clément Mallet\textsuperscript{a}

\textsuperscript{a} Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, F-94160 Saint-Mandé, France - walid.ouerg@gmail.com, (arnaud.le-bris, clement.mallet)@ign.fr
\textsuperscript{b} EA 4592 Géoressources & Environnement, Bordeaux-INP/Université Bordeaux Montaigne - nesrine.chehata@ensegid.fr

KEY WORDS: Decision fusion, Optimization, Graphical model, Hyperspectral, Multispectral, Urban classification

ABSTRACT:

Very high spatial resolution multispectral images and lower spatial resolution hyperspectral images are complementary sources for urban object classification. The first enables a fine delineation of objects, while the second can better discriminate classes and consider richer land cover semantics. This paper presents a decision fusion scheme taking advantage of both sources classification maps, to produce a better classification map. The proposed method aims at dealing with both semantic and spatial uncertainties and consists in two steps. First, class membership maps are merged at pixel level. Several fusion rules are considered and compared in this study. Secondly, classification is obtained from a global regularization of a graphical model, involving a fit-to-data term related to class membership measures and an image based contrast sensitive regularization term. Results are presented on three datasets. The classification accuracy is improved up to 5\%, with comparison to the best single source classification accuracy.

1. INTRODUCTION

Mapping urban environments requires very high spatial resolution (VHR) optical images (<5 m). Indeed, such spatial resolution is necessary to individualize and precisely delineate urban objects and to consider sharper geometrical details (e.g. Herold et al, 2003; Cleve et al, 2008). However, high spatial resolution sensors have generally a poor spectral configuration (i.e. usually three or four bands, RGB/RGB-NIR), limiting their ability to discriminate fine classes and limiting classification performances (Thomas et al, 2003; Carleer, 2005; Yu et al, 2006) compared to superspectral or hyperspectral (HS) sensors. Unfortunately, these latter generally exhibit a lower spatial resolution. To overcome the weakness of both sensors, multispectral (MS) and HS imageries could be jointly integrated to benefit from their complementary characteristics, so as to have: 1) rich geometrical and textural details to finely delineate objects, and 2) rich spectral information to efficiently separate the classes. Thus, the fusion of such sensors should enhance the classification performance at the highest spatial resolution. However, such fusion scheme would have to cope with both spatial and semantic uncertainties.

Fusion of heterogeneous datasets (i.e. different spatial and spectral resolution), have been widely investigated in the literature (e.g. Pohl and Genderen, 1998; Schmitt and Zhu, 2016). The fusion procedure can be carried at three distinct levels:

1) \textbf{observation level}: the most popular method for that purpose is the Pan-sharpening. It consists in merging high resolution panchromatic (PAN) image with low resolution MS one to produce high resolution MS image. A review of such methods can be found in (Loncan, 2015).

2) \textbf{feature level}: (e.g. Fauvel et al, 2007; Wegner et al., 2011; Ban and Jacob, 2013) consisting in applying a single classification using features extracted from both sources. For instance, Wegner et al. (2011) proposed a conditional random field (CRF) model for building detection using InSAR and orthophoto features.

3) \textbf{decision level}: (e.g. Djafari, 2003; Fauvel et al, 2006); consisting in merging different classification maps that are calculated from heterogeneous datasets. This paper aims at designing a generic fusion method that could be applied to other sources than HS/MS, involving poor spatial resolution but rich semantic information sensor and very high spatial MS sensor. Thus, the fusion at observation level is not conceivable. Only the fusion at decision level is investigated in this paper. The challenge here is to merge complementary single’s expert classifications to enhance them, in a case where training datasets are quite poor.

Fusion methods operating at the decision level can be applied in two situations, considering either they try to merge multiple classifiers applied to a same dataset, or multiple datasets. Concerning the second group, the aim is to take advantage of the information provided by multiple data sources. Fusion methods consisting in using all probabilistic outputs of several single source classifications as a new feature set for a classifier as in (Ceamanos et al., 2010), must be avoided here because of their strong requirements in terms of training dataset to prevent overfitting. Thus, this study will rather consider decision fusion rules, such as the ones based on probabilistic, fuzzy, possibilistic fusion or evidence theory. Benediktsson and Kanellopoulos (1999) combined neuronal and statistical Maximum Likelihood classifiers using several consensus theory rules (i.e. majority voting, complete agreement, CONSNN-NN) to classify MS and HS images. A characterization of the spatial organization of SAR images elements is investigated by Tupin et al. (1999) by merging the responses of multiple low-level detectors applied to the same image. The fusion is carried by Dempster-Shafer (Shafer, 1979) evidence theory rules. Fauvel et al. (2006) investigated the use of fuzzy decision rules to combine the classification results of a conjugate gradient neural network and a fuzzy classifier over an IKONOS image. Paiatkriangkrai et al (2015) combined convolutional neural networks and random forest classifiers using a multiplication scheme, the final result was regularized by a conditional random field (CRF). Jeon (1999), investigated the fusion of multi-temporal thematic mapper images, using decision fusion based methods (i.e. Jointly likelihood and weighted majority fusion). Waake and Benediktsson (2007) applied SVM to classify SAR and MS independently, the probabilities maps were then fused using different strategies (i.e. SVM’s, majority voting, and absolute maximum rule).
Hervieu et al. (2016) combined the classification results of HS and MS images, within a graphical energetic CRF model optimized using a graph-cut algorithm.

This paper proposes a fusion scheme at the decision level to merge a low resolution LR-HS and a VHR-MS images. The method aims at dealing with both semantic and spatial uncertainties, and is based on two steps; decision fusion at pixel level and classification optimization through a global regularization framework. Several decision fusion rules are tested including fuzzy, Bayesian, margin and Dempster-Shafer based rules. The fusion is optimized in as second step by a graph-cut algorithm including a contrast sensitive spatial regularization term.

2. PROPOSED METHOD

The method is based on three main steps: a) classification of a HS and a MS images and generation of the posterior probabilities, b) fusion of the posterior probabilities at the decision level, and c) classification optimization (Figure 1). First, two input images are classified. A gaussian kernel SVM classifier (Vapnik, 1999) was used here, but other supervised classifiers could be used (e.g. random forest). The posterior probabilities are retrieved with the Platt’s technique (Platt, 2000). At the end, a posterior class probabilities map is generated for each classification map.

Secondly, a decision fusion is applied to these posterior class probability maps. A variety of rules were tested. These rules could be divided into 4 classes; fuzzy decision rules (i.e. Min, Max, Compromise, Prioritized, Accuracy dependent), Bayesian combination (i.e. Sum and Product based rules), evidence theory (i.e. Dempster-Shafer rule), margin theory (i.e. Margin-Max rule). These decision fusion rules will be described hereafter. They enable to combine two class probability maps into a more accurate one at the highest resolution. The last step consists in performing a global regularization of the classification map obtained at step 2, so as to deal with spatial uncertainties between both sources. A graphical model was used, involving a fit-to-data term and a contrast sensitive regularization term. This formulation has been used successfully for many purposes related to image fusion (e.g. Kolmogorov and Zabih, 2004).

3. DECISION FUSION

The decision fusion rules used in this paper are exclusively based on the class membership probabilities generated at pixel level by the classifier. The fusion is carried out pixel by pixel, combining the class membership probabilities of each source. For this study, ten decision rules have been tested; these rules are derived from different probabilistic, possibilistic and evidential theories.

3.1. Fuzzy rules

Introduced first by Zadeh (1965), fuzzy rules become popular tools which were often used to process uncertain data (e.g. Dubois, 1992; Fauvel, 2006).

3.1.1 Theory and general properties: Let us consider a reference set \( L \) of classes, then a fuzzy set \( A \) in \( L \), is a set of ordered pairs:

\[
A = \{ (x, P_A(x)|x \in L) \}.
\]

(1)

where \( P_A: L \rightarrow [0,1] \) is called the membership probability (also called membership function) of \( A \) in \( L \), it ranges a nonnegative real numbers, whose supremum is finite. The membership probability of the fuzzy set is a crisp function (i.e. real valued). The intersection of two fuzzy sets \( P_A \) and \( P_B \) is given by the minimum of their membership probabilities:

\[
\forall x \in L \quad (P_A \cap P_B)(x) = \min(P_A(x), P_B(x)).
\]

(2)

The union of two fuzzy sets \( P_A \) and \( P_B \) is given by the maximum of their membership probabilities:

\[
\forall x \in L \quad (P_A \cup P_B)(x) = \max(P_A(x), P_B(x)).
\]

(3)

The complement of a fuzzy set \( P_A \) is given by:

\[
\forall x \in L \quad P_A(x) = 1 - P_A(x).
\]

(4)

3.1.2 Measure of conflict between two sources: Let us consider two sources \( A \) and \( B \), with the corresponding membership probabilities \( P_A(x) \) and \( P_B(x) \), the conflict between these sources is quantified using Dubois & Prade measure (Dubois and Prade, 1994) \( 1 - K \), where:

\[
K = \sup_x \min(P_A(x), P_B(x)).
\]

(5)

3.1.3 Confidence measure: To reduce the influence of unreliable information within each fuzzy set, a weight \( w_i \) called pointwise accuracy as proposed by Fauvel et al. (2006). Considering a multi-image classification case, let us consider the fuzzy set \( P_i(x) \), with \( i \) the number of source/classification images and \( x \) a pixel belonging to a source \( i \). Rationally, a classifier is considered reliable if regarding a pixel \( x \), one class has high membership and the others are low, conversely, if more than one class presents a high membership, the fuzzy set will present a high degree of fuzziness, and the classifier will be considered unreliable regarding the pixel \( x \). Starting from this assumption, one can weight each fuzzy set by \( w_i \) to reduce the influence of unreliable information as follows:

\[
w_i = \frac{\sum_{k=1}^{n} H_{\alpha QE}(P_k)}{(n - 1) \sum_{k=1}^{n} H_{\alpha QE}(P_k)},
\]

(6)

where \( n \) is the number of sources, and \( H_{\alpha QE}(P_k) \) the fuzziness degree of source \( k \). \( H_{\alpha QE} \) is a measure fuzziness called the \( \alpha \)-quadratic entropy (QE) (Pal and Bezdek, 1994).

Let us consider two sources \( A \) and \( B \) with class membership probabilities \( P_A \) and \( P_B \). The pointwise measure \( w_i \) was
integrated in the fusion rules calculation, to favor the most reliable source; each membership is multiplied by the pointwise measure as follows: \( p_A(x) = w_A \cdot P_A(x), \) \( p_B(x) = w_B \cdot P_B(x) \), where \( P_A, P_B \) are the original membership probabilities, and \( w_A, w_B \) are respectively the corresponding pointwise measures.

### 3.1.4 Fuzzy rules:

Five fuzzy operators were considered for fusion

1) A conjunctive T-norm \( \min \) operator:

\[
P_{\text{fusion}}(x) = \min(p_A(x), p_B(x)).
\]

2) A disjunctive T-norm \( \max \) operator:

\[
P_{\text{fusion}}(x) = \max(p_A(x), p_B(x)).
\]

3) A Compromise operator (as proposed by Dubois and Prade, 1992) following equation (9):

\[
P_{\text{fusion}}(x) = \begin{cases} 
\max(T_2, \min(T_1, (1-K))) & \text{if } (1-K) \neq 0 \\
\min(T_1, \max(P_A(x), P_B(x))) & \text{if } (1-K) = 1 
\end{cases}
\]

where \( T_1 = \min(p_A(x), p_B(x)) / K, \) \( T_2 = \max(p_A(x), p_B(x)) \).

- When the conflict between \( A \) and \( B \) is low (i.e. \( (1-K) \approx 0 \)), the operator behavior is conjunctive.
- When the conflict between \( A \) and \( B \) is high (i.e. \( (1-K) \approx 1 \)), the operator behavior is disjunctive.
- When the conflict is partial (i.e. \( 0 < (1-K) < 1 \)), the operator behaves in a compromise way.

Let us consider the Compromise rule (9), where \( T_1 \) is normalized to a membership belonging to \([0, 1]\) (i.e. division by \( K \)), unlike \( \min(T_2, (1-K)) \). Indeed, \( T_2 \) will be favored at \( P_{\text{fusion}}(x) \) level. To deal with this, the next alternative Compromise rule was proposed. Let us consider \( C_{\text{best}}1 = \arg\max_{c \in \mathcal{CL}} P_{\text{fusion}}(x) \), and \( C_{\text{best}}2 = \arg\max_{c \in \mathcal{CL}} C_{\text{best}}1 P_{\text{fusion}}(x) \), the intra-membership conflict \( f_c \) is measured as follows: \( f_c = \text{abs}(C_{\text{best}}1 - C_{\text{best}}2) \), a threshold \( t_c = 0.25 \) is proposed for the decision process:

\[
\text{If } f_c < t_c \ #\existence\ of\ intra-\ membership\ conflict \# \ \\
d o\ (P_{\text{fusion}}(x) = \max(p_A(x), p_B(x))) \ \\
else \ #\noexistence\ of\ intra-\ membership\ conflict \# \ \\
do\ (P_{\text{fusion}}(x) \text{ following equation (9)})
\]

4) Prioritized operators (Dubois and Prade, 1994) \( \text{Prior 1} \) (Equation 10) and \( \text{Prior 2} \) (Equation 11) rules are considered:

\[
P_{\text{fusion}}(x) = \max(p_A(x), \min(p_B(x), K)).
\]

\[
P_{\text{fusion}}(x) = \min(p_A(x), \max(p_B(x), (1-K))).
\]

For both operators, when the conflict between \( A \) and \( B \) is high (i.e. \( K = 0 \)), \( p_A \) contradicts \( p_B \), and only the information provided by \( p_A \) is taken into consideration. \( p_B \) is considered as a specific piece of information. Contextual dependent operators are more adequate to deal with conflictual situations than conjunctive and disjunctive operators (i.e. Bloch, 1996; Fauvel, 2006) which are ill suited to handle the conflict.

5) An Accuracy Dependent (AD) operator (Fauvel, 2006), integrating both local and global confidence measurements:

\[
P_{\text{fusion}}(x) = \max\left(\min\left(w_i P_i(x), f_i^c(x)\right), i \in [1,n]\right)
\]

where \( f_i^c \) is the global confidence of source \( i \) regarding class \( j \), \( P_i^c(x) = p_i^c(x) \), a normalization factor (see section 3.1.3). This operator ensures that only reliable sources are taken into consideration for each class, via the predefined coefficients \( f_i^c \). The idea seems interesting, nevertheless, the final result will depend on the classifier reliability and also on the ground truth availability since it is necessary to generate \( f_i^c \) term.

### 3.2 Bayesian combination

In addition to fuzzy rules we used simpler Bayesian Sum and Product combinations of membership probabilities (e.g. Bloch, 2006). Each membership is multiplied by the pointwise measure as for (see section 3.1.3). This permits to evaluate the limit of such operators compared to more complex combinations. The fusion is carried using a Bayesian Sum (Equation 13) and product (Equation 14) operators as follows:

\[
P_{\text{fusion}}(x) = p_A(x) + p_B(x).
\]

\[
P_{\text{fusion}}(x) = p_A(x) \cdot p_B(x).
\]

### 3.3 Margin-based rule (Margin-Max)

Let’s consider two sources \( A \) and \( B \), where \( \mathcal{S} = \{A, B\} \), and different classes \( \mathcal{L} = \{c_i\}_{i \in [1,n]} \). Let \( p_s^c(x) \), be the pointwise membership probability of point \( x \) to a class \( c \), according to a source \( s \). The margin of source \( s \) in \( x \) is:

\[
\text{margin}^c_s(x) = p_s^c(\text{chest}1(x)) - p_s^c(\text{chest}2(x)),
\]

where \( \text{chest}1 = \arg\max_{c \in \mathcal{CL}} P_s^c(x) \) and \( \text{chest}2 = \arg\max_{c \in \mathcal{CL}\{\text{chest}1\}} P_s^c(x) \).

Once the margin defined, different fusion rules could be defined based on it. For this study, a \( \text{Max-Margin} \) fusion method is tested. The aim is to combine the sources, conserving in each pixel the most confident source. To calculate the combined membership probabilities of two sources \( A \) and \( B \), where \( \mathcal{S} = \{A, B\} \), and different classes \( \mathcal{L} = \{c_i\}_{i \in [1,n]} \).

\[
P_{\text{fusion}}^c(x) = P_{\text{best}}^c(x),
\]

where \( S_{\text{best}} = \arg\max_{c \in \mathcal{CL}} \text{margin}^c_s(x) \).

### 3.4 Dempster-shafer (DS) evidence theory based rule

In DS formalism, the information from a source \( s \) for a class \( c \) is represented by a mass function \( m_s \), \( m_s \in [0,1] \) (Shafer, 1976). The evidence theory implies the use of simple classes \( c \in \mathcal{L} \) as well as compound classes (i.e. \( c' = c_q \cup \ldots \cup c_h \)).

Composed classes were here limited to the union of -at most- 2 simple classes (i.e. \( c' = c_q \cup c_h \)), let’s note \( \mathcal{L}' \) the new classes ensemble. Masses are associated to each class as:

\[
m(\emptyset) = 0
\]

-Simple classes: \( \forall c \in \mathcal{L}, v_{\text{pixel}}x, \) and \( \forall s \in \mathcal{S}, m_s^c(x) = P_s^c(x), \) where \( m \) is the mass affected to class \( c \) by source \( s \), and \( P \) is a pointwise membership probability of the considered class.

-Compound classes:

The compound classes masses were here generated as follows: \( \forall c_1, c_2 \in \mathcal{L}, \forall v_{\text{pixel}}x, \) and \( \forall s \in \mathcal{S} \)
The function $E$ is a probabilistic function of the posterior probability $P_{\text{fusion}}$ for a classification map $C$, the energy term is written following Equation (20):

$$E(P_{\text{fusion}}, C_fusion) = \sum_{x \in \text{img}} E_{\text{data}}(C(x)) + \lambda \sum_{x \neq y} E_{\text{reg}}(C(x), C(y)), \quad (20)$$

where:

$$E_{\text{data}}(C(x)) = f(P_{\text{fusion}}(C(x))),$$

$$E_{\text{reg}}(C(x) = C(y)) = g(P_{\text{fusion}}(C(x)), C_fusion),$$

$$E_{\text{reg}}(C(x) \neq C(y)) = h(P_{\text{fusion}}(C(x)), C_fusion).$$

$\lambda \in [0, \infty[$ is a tradeoff parameter between data and regularization terms, $N$ is the 8 connectivity neighbors.

$E_{\text{data}}$ is a fit-to-data attachment term, function of the probability map $P_{\text{fusion}}$ which models the result of the classification fusion, and defined by the function $f$, as follows:

$$f(t) = -\log(t) \text{ with } t \in [0, 1]. \quad (21)$$

The function $f$ ensures that if the probability for a pixel $x$ to belong to class $C(x)$ is close to 1, $E_{\text{data}}$ will be small, and will not impact the total energy $E$. Conversely, if the probability for a pixel $x$ to belong to class $C(x)$ is low, $E_{\text{data}}$ will be near its maximum, and will penalize such configuration. $E_{\text{reg}}$ is a regularization term defining the interactions between a pixel $x$ and its 8 neighbors, and enabling to smooth the initial classification map $C_{\text{fusion}}$, by favoring the neighboring pixels to belong to the same class (i.e. minimizing energy). A Potts model (Schindler, 2012) is a typical configuration defined as follows:

$$E_{\text{reg}}(C(x) = C(y)) = 0, \quad (22)$$

$$E_{\text{reg}}(C(x) \neq C(y)) = 1.$$
bands from 430 to 860 nm. Pavia University is a 335 x 605 image, Pavia Center is a 715 x 1096 image, and both have a ground sample distance (GSD) of 1.3 m. Both Pavia images are composed of 9 land cover classes (Figures 2.e, 2.f): Asphalt, Meadows, Gravel, Trees, Painted Metal Sheets, Bare Soil, Bitumen, Self-Blocking Bricks, Shadows for Pavia University and Water, Trees, Meadows, Self-Blocking Bricks, Bare soil, Asphalt, Bitumen roofing, Tiles roofing, Shadows for Pavia Center. For both datasets, MS and HS images were generated. MS images were generated using Pleiades satellites spectral configuration (limited to three RGB bands), with a GSD of 1.3 m, while HS images were resampled at a lower spatial resolution of 7.8 m and at the full original spectral range (i.e. 103 and 102 bands). To sum up, both datasets include a MS image (RGB with 1.3 m GSD), and a HS image (Full spectral range with 7.8 m GSD).

The last dataset called “Toulouse Center”, is a scene over Toulouse city in France. It has 405 spectral bands ranging from 400 to 2500 nm, and an initial GSD of 1.6m. Its associated land cover is composed of 15 classes (Figure 2.d): Slate roofing, Asphalt, Cement, Water, Pavements, Bare soil, Gravel roofing, Metal roofing1, Metal roofing 2, Tiles roofing, Grass, Trees, Railway tracks, Rubber roofing, Shadows. MS and HS images were created for the fusion purpose; a MS image using Pleiades satellite spectral configuration (four RVB - NIR bands), with a GSD of 1.6 m, and a HS image which is a resampled version of the original image at a lower spatial resolution of 8 m. To sum up, Toulouse dataset has a MS image (RGB with 1.6 m GSD), and a HS image (Full spectral range with 8 m GSD).

Figure 2. RGB Datasets and corresponding ground truth; (a),(d) Toulouse center, (b),(e) Pavia University, (c),(f) Pavia Center.

4.2 Results and discussion

4.2.1. Sources comparison: MS image is characterized by a high spatial resolution and a few bands, while the HS one has a low spatial resolution and a hundred(s) of bands. The SVM classifier was applied over these images leading to 1) a sharp objects delineation in the MS image due to its good spatial resolution, but also a lot of artefacts (Figures 3.b, 4.b, 5.b), and 2) a good discrimination of the different classes in the HS image but also a blurry objects delineation due to its low spatial resolution (Figures 3.a, 4.a, 5.a). The corresponding classification accuracies are listed in Table 2: the accuracies are better using HS image.

4.2.2. Decision fusion classification: 10 different decision fusion rules were first tested and compared over the 3 datasets, the quantitative results of Table 1 let us consider that the Compromise, Bayesian Product, Margin-Max and Dempster-Shafer rules are the most efficient. The comparison must also take into consideration the visual inspection of the results, as ground truth data remains very limited on these datasets. For Pavia University, four of the best accuracies were reached for Min, compromise, Bayesian Product, and Dempster-Shafer rules. In practice, the Min/Compromise rules, gives the best classification rendering, especially regarding class Self-Blocking Bricks which is a conflictual class (Figures 3.a, 3.b, Magenta color class), the two other rules seem to overestimate this class, and are more
considering the HS classification map in the fusion process, this explains their better accuracy (Table 1). The Min rule acts in cautious way when taking the best of the lowest memberships, the Compromise rule acts depending on the degree of conflict between sources. The Bayesian Product rule is a good and simple tradeoff if the initial classification maps are not highly conflictual, otherwise, the result will be degraded by wrong information. Concerning Pavia Center, all the rules seem accurate (Figure 4.c, example: Dempster-Shafer rule), with an overall accuracy superior to 98% (Table 1). When inspecting the classification images, all rules gave similar good results excepting Prior 1, showing a result guided by the HS classification map rather than the MS one.

Toulouse dataset, is the largest one, with up to 15 classes. Thus, the accuracies are lower. The best results were given by the Max, Prior 2, Bayesian Sum and Dempster-Shafer rules. In practice the Max, Prior and Sum rules seem to overestimate certain classes, especially tile roofing’s and vegetation, the best rendering in terms of classification fusion, is given by the Min, compromise and Dempster-Shafer rules. Despite a reasonable global accuracy, the AD rule presents many misclassifications regarding tiles roofing (i.e. underestimation), metal roofing 1 (i.e. overestimation), and a bad detection of the gravel roofing, this is mainly due to the global accuracy measure which is included in the rule and calculated thanks to the ground truth data.

As a conclusion, the quantitative accuracies doesn’t necessarily transcribe the real potential of the fusion rules, and the best ones from a quantitative and practical qualitative point of view are; the Compromise, the Bayesian Product and the Dempster-Shafer rules.

4.2.3 Global classification regularization: Once the decision fusion done, a global regularization procedure was applied to enhance the classification results and eliminate the artefacts. Table 2 presents the optimization results for the best rules per dataset, the optimization procedure permits to enhance further the classification. Quantitatively, global regularization process permits to slightly enhance the decision fusion classification (by 1-2%) but offers a better visual rendering with an elimination of the artefacts, a better decimation of the classes borders, and a regularization of the scattered pixels (Figures 3.d, 4.d, 5.d). These optimized maps seem better modeling the real scene. The optimization effect is more visible over Pavia University and Toulouse Center, concerning Pavia Center, the decision fusion gives already good results and thus, the optimized maps are only slightly improved (Table 2). Quantitative accuracy results obtained over the Pavia datasets are comparable to other studies (e.g. Fauvel et al, 2007). For Pavia University; the painted metal sheets are better recovered and no mismatches with the surrounding road are noticeable. The proposed method permits to extract some bitumen buildings that were difficult to differentiate from roads (i.e. up right and down right, Figure 2.b), the gravel buildings could be better refined. For Pavia center, the global rendering is enhanced with a minimization of the classification artefacts.

---

![Figure 3. Pavia University classification results using the best decision fusion rule; (a) SVM classification of HS image, (b) SVM classification of MS image, (c) Classification fusion by Min rule, (d) Global classification regularization.](image1)

![Figure 4. Pavia Center classification results using the best decision fusion rule; (a) SVM classification of HS image, (b) SVM classification of MS image, (c) Classification fusion by Dempster-Shafer rule, (d) Global classification regularization.](image2)
Figure 5. Toulouse Center classification results using the best decision fusion rule; (a) SVM classification of HS image, (b) SVM classification of MS image, (c) Classification fusion by Compromise rule, (d) Global classification regularization.

<table>
<thead>
<tr>
<th>Data</th>
<th>Accuracy</th>
<th>Image HS classification</th>
<th>Image MS classification</th>
<th>Decision fusion</th>
<th>After regularization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>Min</td>
<td>Compromise</td>
<td>Prior 1</td>
</tr>
<tr>
<td>Pavia University</td>
<td></td>
<td>OA (%)</td>
<td>92.8</td>
<td>96.1</td>
<td>96.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kappa (%)</td>
<td>90.7</td>
<td>94.9</td>
<td>95.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-score (%)</td>
<td>90.6</td>
<td>95.1</td>
<td>95.0</td>
</tr>
<tr>
<td>Pavia Center</td>
<td></td>
<td>OA (%)</td>
<td>98.5</td>
<td>98.6</td>
<td>98.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kappa (%)</td>
<td>97.8</td>
<td>98.0</td>
<td>98.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-score (%)</td>
<td>96.0</td>
<td>96.3</td>
<td>96.7</td>
</tr>
<tr>
<td>Toulouse Center</td>
<td></td>
<td>OA (%)</td>
<td>75.6</td>
<td>72.2</td>
<td>73.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kappa (%)</td>
<td>62.4</td>
<td>58.7</td>
<td>60.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>F-score (%)</td>
<td>69.8</td>
<td>65.8</td>
<td>68.0</td>
</tr>
</tbody>
</table>

Table 1. Classification accuracy after fusion procedure, 10 fusion rules at decision level are compared.

5. CONCLUSION

This paper proposes a two-step method dealing with multisource data fusion and global regularization. Several decision fusion methods were tested and compared, including fuzzy rules, Bayesian combinations, margin-based rule, and Dempster-Shafer based rule. Among the fuzzy rules, the Min and Compromise rules are the most efficient. The Max rule is often affected by misclassifications due to the fact it pays more confidence to the highest membership. The prioritized rules favor a source rather than the other; indeed, the reliability is not ensured as noticed for Prior 1, which gives confidence to the less reliable source. AD rule accuracy is dependent on the ground truth reliability; the rule gives encouraging results for Pavia datasets, the accuracy was not sufficient for Toulouse dataset. The Bayesian Sum and Product rules could be interesting in case of low conflict between sources, since they give acceptable results over Pavia Center and Toulouse. Concerning the proposed margin-based rule, it performs well over Pavia center, and correctly over Toulouse, but it is not enough sufficient over Pavia University. Finally the Dempster-Shafer rule gives homogeneous performance over the three datasets leading to interesting results.

Even if the decision fusion enables to increase the classification accuracy compared to the initial classification maps, the results are affected by classification artefacts, and unclear borders, the final maps are either guided by one of the initial maps or by both, the final result is, therefore, a better version of the initial maps. The final step is a global regularization optimization of the decision fusion results, for classification enhancement. The principle is based on regularization regarding each pixel membership and its neighborhood, and regarding an image contrast measure when comparing neighboring pixels. The optimization procedure gives encouraging results, with clear borders among the different classes, and artefacts elimination.

The method also has the possibility to integrate other decision rules in a fully tunable way, the optimization model is simple and flexible and could be tuned depending on the used dataset. Further work will investigate the explicit use of conflict measures from fusion step within the regularization framework. At the moment the optimization parameters selection is rather manual, some automation could be included, and other contrast measures could be tested to improve the accuracy. The proposed fusion scheme is quite generic and could also be applied to other similar configurations, such as low spatial resolution times series and VHR-MS data. For this study, the HS image was generated from the VHR-MS one leading thus to quite optimistic

This contribution has been peer-reviewed.
results. Real cases will be tested, including temporal differences between two images.

ACKNOWLEDGMENTS

This research was supported by the French National Research Agency (ANR) through the HYEP project on Hyperspectral imagery for Environmental urban Planning (ANR-14-CE22-0016).

REFERENCES


