Land cover Mapping using Triangular-Norms

SHAHEERA RASHWAN
Informatics Research Institute, MuCScat, Borg ElArab, Alex, EGYPT
M. A. ISMAIL, SOHEIR FOUAD
Computer Science Department, Faculty of Engineering, University of Alexandria, Alexandria, 21544, EGYPT

Abstract: - To acquire detection performance required for an operational system in the detection for satellite image for environmental changes, it is necessary to use multiple images over years to know the environmental changes over years. This paper describes a method for decision-level fusion technique where the fusion can compensate for correlation among images. The fusion is done using possibilistic combiners based on T-norms families that better represent the correlation of images. This technique was applied to satellite images for free vegetation of Africa (1998 April 01, 1998 September 01). The performance of this technique, compared to the other techniques such as naïve Bayes, Dempster-Shafer, voting, rule-based and linear discriminant, is evaluated by simple theoretical models.

Key-Words: - Decision Fusion, Correlation, naive Bayes, Dempster-Shafer, voting, linear discriminant, T-Norm, Satellite Imagery for environmental change.

1 Introduction

Multi image fusion has become an active field of research as more and more applications such as medical imaging, security, avionics, surveillance and night vision utilize multi sensor imaging arrays. Such arrays provide a wider spectral coverage and reliable information even in adverse environmental conditions at a price of a considerable increase in the amount of data. Image fusion deals with the data overload by combining visual information from multiple image signals into a single fused image with the direct aim of preserving the full content value of the multi sensor information.

The production of land cover / land use maps is a common application of multi spectral satellite images. There are numerous examples of land cover maps derived from multi spectral satellite imagery at global, regional and local level. The most widely used approach is to classify each image pixel as an independent observation, regardless of its spatial context. Recently, at local level, a number of data sources have been used to derive land cover products, including Land sat TM data for high resolution studies.

These studies have been carried out for a number of different applications, including estimation of biomass and vegetation mapping. It was evaluated the potential of ASTER VNIR (Visible and Near Infrared) and SWIR (Short Wave Infrared) sensors for land cover mapping. Information in the VNIR image contributed to the enhancement of vegetation and water classes. Rock and soil units were enhanced due to the contribution by the information in the SWIR images.

One of the main problems when generating land cover maps from digital images is the confusion of spectral responses from different features. Sometimes two or more different features with similar spectral behavior are grouped into the same class, which leads to errors in the final map. The accuracy of the map depends on the spatial and spectral resolution and the seasonal variability in vegetation cover types and soil moisture conditions. Attempts have been made to improve the accuracy of image classification.

Our aim in this paper is to develop a new approach for multi spectral image fusion or satellite imagery for environmental change based on possibility theory. This approach is expected to handle the problem of correlation which degrades the performance fusion.

This approach will then be used to satellite images for free vegetation of Africa (1998 April 01, 1998 September 01).

2 Fusion Techniques

Fusion techniques [3, 4, 5] can be seen as a *discriminant* function, $F(\vec{c})$ in satellite image confidence space defined in such a way that:

$F(\vec{c}) \geq t \quad assign \ \vec{c} \rightarrow$ Object of Interest

$F(\vec{c}) < t \quad assign \ \vec{c} \rightarrow$ Background

where $\vec{c} = (c_1, c_2, ..., c_R), c_i \in [0,1] \ \forall i \in [1, R]$ is an image output (confidence) vector with R the number of images and t the threshold. Image output vectors are defined only at locations where the images from co-registration and where image data is present.

The general layout of the image-fusion methods is shown in Figure 1. The input of each image-fusion method is a confidence level per grid cell.

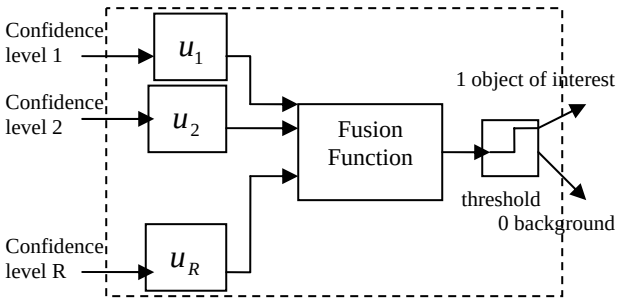

Figure 1. The generic decision-level image-fusion layout.

The output of the fusion process is one for detection and zero for no detection per grid cell. Each of the methods scales the influence of each of the images in a different way.

This mapping may remove the differences in definitions of the confidence levels. The mapped inputs are combined in a fusion function to acquire a single value per grid cell. The mapping functions and the fusion function are given in Table 1.

Method	Mapping Function	Fusion Function
Naive Bayes	linear scaling around 0.5	product
Linear discriminant	linear scaling	summation
Dempster-Shafer	uncertainty level	Dempster's rule of combination
Voting	threshold	Summation
Rule-Based	linear scaling	Disjunction of conjunction clauses

Table 1. The different functions for scaling the input and combining these into a single (fused)

2.1 Naive Bayes

The naive Bayes assumes that the confidence levels scale with the likelihood ratio per image. The likelihood ratio is the quotient of the conditional probability densities for both classes (mines and background). Based on this likelihood ratio the optimal Bayes decision can be made.

If these conditional probabilities are image independent, then the joint likelihood ratio is the product of these individual likelihood ratios. Since the confidence levels most likely do not linear scale with the likelihood ratios, a mapping factor is included for each image.

The naive Bayes fusion function $F(\vec{c})$ and the mapping are given by:

$$F(\vec{c}) = \prod_{i=1}^{R} ((1-u_i)c_i + \tfrac{1}{2}u_i) \quad (1)$$

with c_i the confidence levels and u_i the mapping parameters.

2.2 Linear Discriminant

The linear discriminant is another (naive) implementation of the Bayesian optimal decision. If confidence levels are interpreted as the logarithm of the likelihood ratio per sensor and these likelihood ratios are independent, then the summation gives the optimal decision boundary. Since the confidence levels are most likely not proportional to the likelihood ratio, a scaling factor is used. A linear classifier is also optimal when the images values have equal normal distributions.

The linear discriminant fusion function and the mapping are given by:

$$F(\vec{c}) = \sum_{i=1}^{R} u_i c_i \quad (2)$$

with c_i the confidence levels and u_i the mapping parameters.

2.3 Dempster-Shafer

For application of the Dempster-Shafer theory to image fusion, three inputs per image are needed: the probability mass assigned to an object of interest $m(M)$, the probability mass assigned to background $m(\overline{M})$, and the unassigned probability mass $m(M \cup \overline{M})$. The sum of these masses always equals one, so there are only two independent masses ($m(M)$ and $m(\overline{M})$). The mass $m(M)$ represents a belief in an object of interest, the mass $m(\overline{M})$ represents the belief in background, and the mass $m(M \cup \overline{M})$ reflects the uncertainty of the image. Each image produces one confidence level at each sample

location, which must be mapped onto the three required probability masses. This is done by using the uncertainty as an optimization parameter.
The confidence levels for image i are mapped onto probability masses, using:

$$m_i(M) = (1-u_i)c_i \quad (3)$$

$$m_i(M \cup \overline{M}) = u_i \quad (4)$$

with u_i the mapping parameter and c_i the confidence level for image i . The probability masses for image 1, 2,..,R are combined using Dempster's rule of combination:

$$m_{1,2,...,R}(M) = (m_{1,2,...,R-1}(M) + m_{1,2,...,R-1}(M \cup \overline{M})) * m_R(M) + m_R(M \cup \overline{M}) * m_{1,2,...,R-1}(M) \quad (5)$$

$$m_{1,2,..,R}(M \cup \overline{M}) = \prod_{i=1}^{R} m_i(M \cup \overline{M}) \quad (6)$$

with $m_{1,2,...,R}(M)$ the combined probability mass assigned to object of interest, and $m_{1,2,...,R}(M \cup \overline{M})$ the combined uncertainty mass. The output of the Dempster-Shafer fusion function is given by the three combined probability mass assigned to an object of interest plus half the uncertainty:

$$F(\vec{c}) = m_{1,2,...,R}(M) + \tfrac{1}{2} m_{1,2,...,R}(M \cup \overline{M}) \quad (7)$$

with c_i the confidence levels and u_i the mapping parameters.

2.4. Voting

Voting fusion is another decision-level fusion method. Our voting fusion is described by (R+1) thresholds: one for each image and one for the required number of votes. A vote is given by a satellite image if the confidence level of this image is larger than the threshold. The votes are summed and the final threshold select between one out of R ("or" voting), or R out of R ("and" voting) votes.
The voting fusion function and the mapping are given by:

$$F(\vec{c}) = T(\sum_{i=1}^{R} T(c_i, u_i), u_{R+1}) \quad (8)$$

with c_i the confidence levels and u_i the mapping parameters.
The threshold function is defined as:

$$T(c,u) = \begin{cases} 1, & if \quad c \geq u \\ 0, & otherwise \end{cases} \quad (9)$$

2.5. Rule-based fusion

Decision rules form an intuitive and flexible approach to satellite image fusion as it is very easy to incorporate any available a priori knowledge into the system.
The general form of a rule is the following:

$$c_1 > u_1 \wedge ... \wedge c_i > u_i \rightarrow \text{Object of Interest} \quad (10)$$

which consists of a conjunction ('and function') of clauses $c_i > u_i$. Each clause states that an image confidence level is higher than a certain threshold. Each clause states that a satellite image confidence level is higher than a certain threshold. A clause may be omitted by setting its threshold value to zero. Disjunction ('or function') of clauses is achieved by combining rules:

$$(c_1 > u_{1,1} \wedge ... \wedge c_i > u_{i,1}) \vee (c_1 > u_{1,2} \wedge ... \wedge c_i > u_{i,2}) \rightarrow \text{Object of Interest} \quad (11)$$

in which are thresholds. A rule set is derived by selecting a minimal (in the sense of number of rules) rule set for a training set which covers each sample in the training set. Each object is covered by at least one rule. This rule set can be simplified by removing rules already covered by the rest of the rules in the set.
Only the maximum number of rules and the quantization of the thresholds constrain the resolution of the discriminant and as such the approximations of the other discriminant functions. As such, the discriminant function of the rule-based system is the superset of all the discriminant functions of the other image-fusion methods.

3 T-Norm Fusion

We propose a general method for the fusion process, which can be used with satellite image outputs that may exhibit any kind of (positive, neutral, or negative) correlation with each other. Our method is based on the concept of Triangular Norms, a multi-valued logic generalization of the Boolean intersection operator. With the intersections of multiple decisions one needs to account for possible correlation among the sources, to avoid under- or over-estimates. Here we explicitly account for this by the proper selection of a T-norm operator. We combine the outputs of the images by the generalized intersection operator (T-norm) that better represents the possible correlation between the images.

3.1 The Triangular-Norm

A triangular norm (briefly t-norm) is a binary operation T on the unit interval [0, 1] as follows

$T : \forall (x, y) \in [0,1]^2, \max(0, x + y - 1) \le xy \le \min(x, y)$ The T-norm operation is commutative, associative, monotone and has 1 as neutral element, i.e., it is a function $T : [0,1]^2 \to [0,1]$ such that for all $x, y, z \in [0,1]$:

(T1) $T(x, y) = T(y, x)$,

(T2) $T(x, T(y, z)) = T(T(x, y), z)$,

(T3) $T(x, y) \le T(x, z)$, whenever $y \le z$,

(T4) $T(x,1) = x$,

In I. Bloch [10], T-norms were considered as fuzzy CICB operators which satisfy the requirements of the conjunction operators. There exist a lot of parameterized T-norm families of operators which range continuously from one operator to another depending on the value of the parameter. This parameter can be used to express correlation as explained in later sections**.**

3.2 Correlation of Image-Decision Fusion

Since correlation affects the performance analysis. The larger the correlation index, the larger the redundancy. In particular, the correlation index goes to zero if the individual incorrect answers are disjoint for all answers. In other words there is always at least one correct answer for any class. The ρ correlation coefficient [8] gets larger as the number of wrong answers is the same for many answers. Let N^f be the number of experiments where all tools had a wrong answer, N_i^c be the number of experiments with combinations of correct and incorrect answers; c is the combination of correct and incorrect answers; n is the number of tools. The correlation coefficient is then

$$\rho = \frac{nN^f}{\sum_{i=1}^{2^n-2} N_i^c + nN^f} \quad (12)$$

3.3 The T-Norm fusion technique for correlated images

In our work, we suggest a new decision-level fusion method based on possibilistic fusion for a better representation of the correlation among images.

From the associativity of the T-norms, we can derive the associativity of the fusion by:

$$F(\vec{c}) = TNorm(TNorm(c_1, c_2), c_3) = TNorm((c_1, TNorm(c_2, c_3)) \quad (13)$$

with c_1, c_2, c_3 the confidence levels for three images and this equation (13) can be computed recursively for R images.

For instance the operator $h(x, y, \alpha)$ is CIVB (Context Independent Variable Behavior) whose behavior depends on the value of α [10]. According to [11], Dempster-Shafer is a special case of possibilistic fusion where correlation = 0 and the function is equal to the product xy. From this, we can choose a suitable α to have a fusion technique sensitive to correlation.

- Schweizer-Sklar: $TNorm(x, y) = \frac{xy}{\max(x, y, \alpha)}$, which ranges from product xy for $\alpha = 1$ to $\min(x, y)$ for $\alpha = 0$ (this family is decreasing w.r.t. the parameter α)
  - We choose α such that $\alpha = 1 - (\rho / \infty)$
- Generalized:
  $TNorm(x, y) = \max[0, (x^\alpha + y^\alpha - 1)]^{1/\alpha}$, which ranges from $\max(0, x + y - 1)$ for $\alpha = 1$ to the product $\alpha = 0$ (this family is increasing w.r.t. the parameter α)
  - We choose α such that $\alpha = \rho$
- Family of Hamacher
  $TNorm(x, y) = \frac{xy}{\alpha + (1 - \alpha)(x + y - xy)}$, which ranges from $\max(0, x + y - 1)$ for $\alpha = +\infty$ to the product $\alpha = 1$ (this family increasing w.r.t. the parameter α)
  - We choose α such that $\alpha = 1/(1 - \rho)$
- Family of Frank:
  $TNorm\ (x, y) = \log_\alpha [1 + \frac{(\alpha^x - 1)(\alpha^y - 1)}{\alpha - 1}]$
  ,which is equal to the "min" for $\alpha = 0$, to the product for $\alpha = 1$ and to $\max(0, x + y - 1)$ for $\alpha = +\infty$ (this family is decreasing family w.r.t the parameter α)
  - We choose α such that $\alpha = 1/(1 - \rho)$

4 Performance Evaluations

In the performance evaluation table, the accuracy is compared to the correlation between different images taken by satellite NDVI images each 50 days. The images are acquired from http://free.vgt.vito.be/result.php/ for free vegetation products of Africa. The accuracy here is

defined by comparing the best image – the image which had the minimum error when applying the clustering techniques- with the fused image.
In order to create this comparison it is of extreme importance to have adequate simultaneous information on the detection rate over the entire diagram for the algorithm. Image 1 is Africa 1998 April 01 and the 2nd image is as in the following table

Image 2	**Duration (days)**	**Correlation**
Africa 98 May 21	50	0.9768
Africa 98 July 11	100	0.9641
Africa 98 September 01	150	0.9389

Table 2. The correlation between 1st and 2nd image

The table comparing accuracies is as follows:

Technique /Accuracy	Image A980521	Image A980711	Image A980901
Naïve Bayes	0.96081	0.93803	0.91357
Linear-Discriminant	0.9952	0.99546	0.99515
Dempster-Shafer	0.97138	0.95275	0.93418
Or Voting	0.79922	0.9718	0.97241
And Voting	0.96045	0.96001	0.9491
Rule-Based	0.9806	0.97193	0.9731
T-Norm(S-S)	0.97551	0.96284	0.94443
T-Norm(G)	0.78622	0.77815	0.75909
T-Norm(H)	0.97551	0.96284	0.94443
T-Norm(F)	0.78622	0.77815	0.75909

Table3. The performance fusion gain of different techniques

5 Conclusions and Future Work

We have proposed an approach based on possibility theory in this paper. The approach is based on calculating the correlation among different images taken at different times to study the change of the environment and use it as a parameter in four CIVB T-Norm techniques to handle the problem of high correlation. This approach performs better for correlated satellite images for environmental changes than the previous techniques. As shown in table 3, all the previous techniques are affected by the use of the mapping function which causes different output gains- and sometimes causes an output very similar to the best image. In this case, the performance gained by the fusion has no use. In T-norm techniques, the fusion plays its role in acquiring useful information given by more than one image.
For future work, we can use this idea in the landmine detection in the Alamen minefield, Egypt. Knowing the original map of the landmine field where mines were buried from years and calculating the correlation between satellite images of the area in this time considering the environmental change till now, we can deduce the recent map of the buried mines in the Alamen minefield.

6 Acknowledgements

The efforts of the correlated satellite imagery for environmental change project were supported and funded from Informatics Research Institute, Mubarak City for Scientific Research under the supervision of Dr. Muhammad Saleh Ibrahim.

References:

[1] M. Kam and Q. Zhu and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems", *IEEE Trans. Aerospace and Electronic Systems*, Vol. AES-28(3), pp.916-920, 1992
[2] Z. Chair and P. R. Varshney, "Optimal data fusion in multiple sensor detection systems", *IEEE Trans. Aerospace and Electronic Systems*, Vol. AES-22, pp.98-101, 1986
[3]Piet B.W. Schwering, Brian A. Baertlein, Sebastiaan P. van den Broek, Frank Cremer, "Evaluation methodologies for comparison of fusion algorithms in land mine detection", *Proc. SPIE Vol. 4742, Detection and Remediation Technologies for Mines and Minelike Targets VII , Orlando FL, USA*, Apr. 2002
[4] F.Cremer, K.Schutte, J.G.M. Schavemaker, E. den Breejen,"A comparison of decision-level sensor-fusion methods for anti-personnel landmine detection", *Information Fusion 2*, pp.187-208, 2001
[5] F.Cremer, K.Schutte, J.G.M. Schavemaker, E. den Breejen, "Towards an operational sensor-fusion system for anti-personnel landmine detection", *Proc. SPIE Vol. 4038, Detection and Remediation Technologies for Mines and Minelike Targets V , Orlando FL, USA*, Apr. 2000
[6] Hao Yin, ElSayed A. Orady, Yubao Chen, Chia-hao Chang, "A practical measure of the uncertainty level in a data set", *International Journal of Science & Technology*, Vol.14, 2003
[7] Kai Goebel, Weizhong Yan, "Choosing Classifiers for Decision Fusion", *Proceedings of the Seventh*

International Conference on Information Fusion, vol. I, pp. 563-568, 2004.

[8] K. Goebel, W. Yan, and W. Cheetham, "A method to calculate classifier correlation for Decision Fusion" *Proceedings of IDC 2002, Adelaide, 11-13 February, 2002*., pp. 135-140, 2002.

[9]K. Goebel, "Architecture and design of a diagnostic information fusion system", *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, Vol.15(4), pp.335-348, September 2001.

[10]Isabelle Bloch, "Information Combination Operators for data fusion: A comparative review with classification", *IEEE Trans. on Systems, Man, and Cybernetics-Part A: Systems and Humans*, Vol.26, NO. 1, pp.52-67, January 1996

[11] P. Bonissone, K. Goebel, W. Yan,"Classifier fusion using T-norms", *Lecture Notes in Computer Science: Proceeding of the Fifth International Workshop on Classifier Fusion*, pp. 154-163, 2004.