Multi-Sensor and Algorithm Fusion with the Choquet Integral: Applications to Landmine Detection

Paul Gader *Senior Member, IEEE*, Andres Mendez-Vasquez, Kenneth Chamberlin, Jeremy Bolton, Alina Zare

***Abstract*—We discuss the application of Choquet integrals to multi-algorithm and multi-sensor fusion in landmine detection. Choquet integrals are defined. Specific classes of measures, the full and Sugeno measures, are described. Full measures are optimized via quadratic programming. A steepest descent algorithm for optimizing Sugeno measures is derived by applying implicit differentiation. Multiple detection algorithms are applied to hyper-spectral and synthetic aperture radar imagery. In addition, a LWIR vegetation index is computed using statistics of apparent emissivity. The detection algorithms are combined using an OR operator and Choquet integrals with respect to full and Sugeno measures. The Choquet integral with respect to the full measure achieves lower false alarm rates.**

***Index Terms*— Choquet integral, Buried object detection, Multisensor systems, Vegetation mapping.**

I. INTRODUCTION

There are many sensor systems that provide value in landmine detection, but none that appear to perform well over all mine types and scenarios. These systems offer both complementary and reinforcing information.

In this paper, we discuss the application of Choquet integrals to multi-sensor fusion in landmine detection. We consider fusion operators defined by Choquet integrals with respect to full and Sugeno measures. We show how measures can be optimized for the fusion application. We show how the measure can be interpreted in an example of fusion of multiple detection algorithms applied to Synthetic Aperture Radar (SAR) imagery and Hyper-Spectral Images (HSI).

II. CHOQUET INTEGRALS AND FUZZY MEASURES

This section defines fuzzy measures and the Choquet integral, and derives the optimization methods.

A. Fuzzy Measures

Fuzzy measures are generalizations of classical measures [1,2]. A fuzzy measure, g, on a finite set X = $\{x_1, \ldots, x_N\}$ is a function on the power set of X, $g{:}2^X \rightarrow [0,1]$ which satisfies:

(i) $g(\varnothing) = 0, g(X) = 1$

(ii) $g(A) \le g(B) \quad if\ A \subseteq B, \quad where\ A, B \subseteq X$

Note that g is not necessarily additive but all additive measures are fuzzy measures.

The Sugeno measures are a class of fuzzy measures that have been used for fusion [3-6]. However, they have not been optimized using gradient descent in the past, although attempts have been made [7,8]. The Sugeno measure, g, is defined for sets $A, B \subseteq X$, with $A \cap B = \varnothing$, by

$$g(A \cup B) = g(A) + g(B) + \lambda g(A) g(B), \quad \text{with } \lambda > -1 \qquad (1)$$

When $\lambda = 0$, the Sugeno measure is a probability measure. Let $g_i = g(\{x_i\})$. The values $\{g_1, g_2, \ldots, g_n\}$ are called densities. Since g(X) = 1, finding λ is equivalent to solving

$$\lambda + 1 = \prod_{i=1}^{n} (1 + \lambda g_i), \text{ where } \lambda \ne 0 \qquad (2)$$

It has been shown that there exists a unique solution $\lambda \in (-1, \infty), \lambda \ne 0$ [3]. Thus, the n densities determine the 2^n values of a Sugeno measure.

In fusion, X represents a set of detectors. The measure of a set of detectors can be more, less, or the same as the sum of the measures of the detectors. The Choquet integral provides the mathematically based method by which confidence values from the detectors are combined with the measure of the detectors to produce a final confidence for fusion.

B. Choquet Integrals

A discrete Choquet integral is a nonlinear transformation that integrates a real function with respect to a fuzzy measure. Let h be a function h:X→[0,1] which represents the evidence provided by the detectors in X.

For the discrete case the values of h are sorted so that $0 = h(x_{(0)}) \le h(x_{(1)}) \le h(x_{(2)}) \le \cdots \le h(x_{(N)}) \le 1$. Let $A_{(i)} = \{x_{(i)}, \cdots, x_{(N)}\}$. The discrete Choquet fuzzy integral with respect to a fuzzy measure g is defined by

$$f=\sum_{i=1}^{n}h(x_{(i)})[g(A_{(i)})-g(A_{(i+1)})]$$
$$=\sum_{i=1}^{n}[h(x_{(i)})-h(x_{(i-1)})]g(A_{(i)}) \quad (4)$$

where $g(A_{(N+1)})=h(x_{(0)})=0$. The equations for f appear linear, but sorting makes them nonlinear.

C. Optimization

In this section, we describe the methods for optimizing full and Sugeno measures. The case of full measures is described by Grabisch [1]. It is necessary to select two labels for each class, α_1 and α_2, and use them to define the cost function

$$E(f)=\sum_{x\in Class1}(f(x)-\alpha_1)^2+\sum_{x\in Class2}(f(x)-\alpha_2)^2 \quad (5)$$

After some algebra, the cost function becomes

$$E(f)=u^tDu-2\Gamma^tu+\alpha \quad (6)$$

where

$$D=\sum_{x\in Class1,Class2}(c_1(x)\ldots c_n(x))^t\cdot\begin{pmatrix}c_1(x)\\ \vdots\\ c_n(x)\end{pmatrix},$$

$$\Gamma=2\alpha_1\cdot\sum_{x\in Class1}\begin{pmatrix}c_1(x)\\ \vdots\\ c_n(x)\end{pmatrix}+2\alpha_2\cdot\sum_{x\in Class2}\begin{pmatrix}c_1(x)\\ \vdots\\ c_n(x)\end{pmatrix},$$

$$\alpha=\sum_{x\in Class1}{\alpha_1}^2+\sum_{x\in Class2}{\alpha_2}^2,$$

and each $c_i(x)$ is equal to $h(A_{(i)})-h(A_{(i-1)})$, and u represent the full measure values in vector form. The requirement that g satisfy the axioms of a fuzzy measure can be represented by a set of linear inequalities which are constraints on the quadratic objective function given in (7). They can be expressed by an inequality of the form $Au\le b$. Thus, the problem of optimizing a full measure reduces to a quadratic programming problem, for which there exist standard solution software.

Steepest descent is used to optimize the densities of a Sugeno measure since no one has been able to transform the optimization problem into a standard form. This requires the partial derivatives of the Choquet integral with respect to the g_j. Implicit differentiation can be applied to find closed form solutions, something that has not been previously shown.

Let S be a set of training samples with desired outputs given by the vector D. The squared error for the k[th] sample is

$$E(k)=\frac{1}{2}(f(k)-D(k))^2 \quad (7)$$

The derivative of $E(k)$ with respect to g_j is used to determine the change in g_j. Let Δg_j denote the change of g_j.

$$\Delta g_j=\alpha\frac{\partial E}{\partial g_j}=\alpha(f(i)-D(i))\frac{\partial f}{\partial g_j} \quad (8)$$

The constant α represents the learning rate. We use (4) to compute the partial derivative of f with respect to g_j. Note that λ is not independent of g_j because, according to (2), λ changes whenever any density changes. Thus $\frac{\partial\lambda}{\partial g_j}\ne 0$. Also, since $g(A_{(i)})$ is dependent on λ, $g(A_{(i)})$ depends on g_j. The only $g(A_{(i)})$ value that may be independent of λ is $g(A_{(n)})=g(\{x_{(n)}\})=g_{(n)}$. Thus, the partial derivatives are

$$\frac{\partial f}{\partial g_j}=\sum_{i=1}^{n}h(x_{(i)})\left(\frac{\partial g(A_{(i)})}{\partial g_j}-\frac{\partial g(A_{(i+1)})}{\partial g_j}\right) \quad (9)$$

From $g(A_{(i)})=g_{(i)}+g(A_{(i+1)})+\lambda g_{(i)}g(A_{(i+1)})$, it follows that

$$\frac{\partial g(A_{(i)})}{\partial g_j}=$$
$$\begin{cases}1+\lambda g(A_{(i+1)})+g_{(i)}g(A_{(i+1)})\frac{\partial\lambda}{\partial g_j}+(1+\lambda g_{(i)})\frac{\partial g(A_{(i+1)})}{\partial g_j}, & \\ \quad if\ (i)\ne n,(i)=j & \\ (1+\lambda g_{(i)})\frac{\partial g(A_{(i+1)})}{\partial g_j}+g_{(i)}g(A_{(i+1)})\frac{\partial\lambda}{\partial g_j}, & if\ (i)\ne n,(i)\ne j\\ 1, & if\ (i)=n,\ j=n\\ 0, & if\ (i)=n,\ j\ne n\end{cases} \quad (10)$$

Eq. (2) implies that

$$\frac{\partial\lambda}{\partial g_j}=\frac{\lambda\prod_{i=1,i\ne j}^{n}(1+g_i\lambda)}{1-\sum_{i=1}^{n}g_i\prod_{k=1,k\ne i}^{n}(1+g_k\lambda)},\ \lambda\ne 0 \quad (11)$$

The fact that $\lambda+1=\prod_{i=1}^{n}(1+\lambda g_i)$ implies that $\frac{(1+\lambda)}{(1+g_j\lambda)}=\prod_{i=1,i\ne j}^{n}(1+g_i\lambda)$, so (11) can be simplified to

$$\frac{\partial \lambda}{\partial g_j} = \frac{\lambda^2 + \lambda}{(1+g_j\lambda)\left[(1-(\lambda+1)\sum_{i=1}^{n}\left(\frac{g_i}{1+g_i\lambda}\right)\right]}, \lambda \neq 0 \tag{12}$$

III. EXPERIMENTS

Landmine detection experiments were performed using airborne data collected over minefields at an arid test site. An imaging SAR and AHI were separately flown over the arid site. A variety of mines were emplaced on the surface and buried at the site. GPS information and image registration routines were used to assign GPS coordinates to the pixels of the images from each sensor [9]. Detectors produced locations of potential mines together with confidence values indicating the likelihoods that mines were present at the locations. A location together with confidence value is called an alarm.

Some of the detectors were developed by the authors and some were developed by other researchers who provided detection images and sometimes associated alarms available to the authors. A detection image is an image with pixel locations that are in one-to-one correspondence with an original image but whose intensity levels indicate the likelihood that a mine is present at the physical locations specified by the pixels in the image.

Three detectors were applied to the AHI images and one detector was applied to the SAR image. The three detectors applied to the AHI image are referred to as the RX, FCLS, and Reststrahlen detectors. The RX is an implementation of the well-known algorithm provided by Winter [10,11] who provided alarm files and detection images. The FCLS detector is an implementation of the algorithm in [12] by Meth et al. who provided detection images [13]. The Reststrahlen algorithm was developed by the authors. It makes use of the so-called reststrahlen effect to detect buried mines. In addition, a vegetation detection algorithm was developed for the AHI imagery by the authors. A vegetation index was developed for the Long Wave InfraRed (LWIR) AHI by statistically characterizing the vegetation. The blackbody-like characteristics of vegetation in these wavelengths suggest high mean and low standard deviation. In addition, skewness is useful. One detector was applied to the SAR data. The detector uses the stochastic hit-miss transform defined in [14].

AHI imagery is useful for finding buried mines and vegetation. SAR imagery is useful for finding surface mines. This suggests fusion an OR operator should find all mines detected by both sensors and that an AND operator would not work well. OR operators can lead to high false alarm rates. Four pairs of images were used for training and one pair for testing. Regions which the AHI and SAR sensors both imaged were computed using tools developed by Radzelovidge et al. [9] There are four types of mines in the overlap region, two plastic cased and two metal cased. The numbers of mines in the overlap region are shown in Table 1.

TABLE 1. MINE DISTRIBUTION

Mine Type	Burial Depth	# of Mines
Plastic Type 1	4 Inches	44
Metal Type 1	4 Inches	56
Metal Type 1	Top Flush with Surface	32
Metal Type 1	On the Surface	16
Metal Type 2	On the Surface	14
Plastic Type 2	On the Surface	5

Fusion was performed by an OR operator and Choquet integrals with respect to optimal full and Sugeno measures. Results on the test image are shown in Figure 1. The full measure significantly outperforms the other methods. Notice that from about 40-60% probability of detection, the false alarm rates are approximately halved by the full measure Choquet fusion compared to the OR.. Above 70% detection, the false alarm rates are essentially the same but the detections at that level are probably "lucky" detections as indicated by the flattening of the ROC curve.

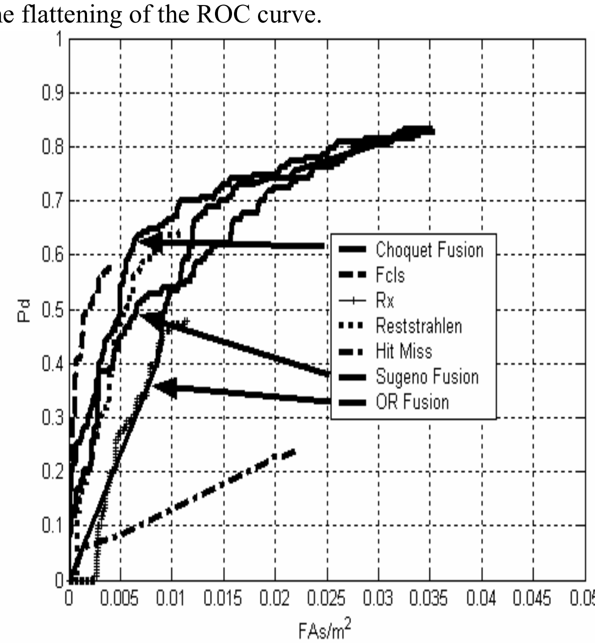

Figure 1. Probability of detection vs. False Alarm Rate for the four detectors on the SAR and AHI images, the OR fusion, and the Choquet integral fusion.

The physical area imaged by the sensors contains trees that obscure mines. Thus, the detection rate does not reach 100%. However, the false alarm rate is low and the detection rate should be good enough for *minefield* detection.

The densities for the full measure are in Table 2.

TABLE 2. DENSITIES ESTIMATED BY TRAINING

Algorithm	Density
RX	0.00
Reststrahlen	0.22
FCLS	0.09
Hit-Miss	0.18
Vegetation	0.17

The Shapley index [15,16] is a measure of importance of each detector. The Shapley indices for the full measure learned in this experiment are shown in Table 3. It is interesting to note that the densities and the Shapley indices do not always agree. For example, the RX algorithm has a density of 0 but a Shapley index of 0.19. The RX algorithm is an important algorithm component. We feel the Shapley index is a better indicator of the importance of a detector than the density value.

TABLE 3. SHAPLEY INDICES OF DETECTORS

Algorithm	Shapley Index
RX	0.19
Reststrahlen	0.20
FCLS	0.13
Hit-Miss	0.16
Vegetation	0.33

Consider the real alarm from the test image in Table 4. At a threshold that yields 60% detection, it is a false alarm using the OR (it is a false alarm at any threshold using the OR since the maximum detector output is 1) but is rejected using the Choquet integral with respect to the full measure.

TABLE 4. A FALSE ALARM BY THE OR REJECTED BY CHOQUET.

RX	Rest	FCLS	HM	Veg	Choquet	MAX
0.61	1.00	0.71	0.36	0.60	0.68	1

Even though there are detectors with high values, the output of the Choquet integral is low. The vegetation index, which is low when there is likely to be vegetation, contributes to this. If we change the vegetation index to 0.95 and leave the other detector values fixed, then the output of the Choquet integral changes to 0.84, which is much higher. This is why the vegetation index has a high Shapley index.

ACKNOWLEDGMENTS

Several individuals and institutions were involved in the collection and ground-truthing of the data and the development of the detectors used in this research. The authors would like to thank William Clark of ARO and Michael Cathcart of Georgia Tech for their support of this research, Miranda Miller of the U. S. Army NVESD for providing us with resources and direction, William Radzelovidge of Raytheon Inc. for providing us with ground-truthed data and software tools for computing overlap of multi-sensor imagery \, Ed Winter of TRA Associates for the RX outputs and also for ideas on LWIR vegetation indices, and Reuven Meth of the University of Maryland for the FCLS detector outputs. The idea to use implicit differentiation for optimizing Sugeno measures was suggested to the authors by James Keller of the University of Missouri – Columbia.

REFERENCES

M. Grabisch, H. Nguyen, and E. Walker, *Fundamentals of Uncertainty Calculi with Applications to Fuzzy Inference*, Kluwer Academic Publishers, 1995.

W. Rudin, *Real and Complex Analysis,* 2nd ed., McGraw-Hill Series in Higher Mathematics, 1974.

H. Tahani, J. M. Keller. "Information Fusion in Computer Vision Using the Fuzzy Integral," IEEE Trans. On Systems, Man and Cybernetics, Vol. 20, No. 3, May-June 1990, pp. 733 – 741.

P. D. Gader, M. Mohamed, and J. Keller, "Fusion of Handwritten Word Classifiers," *Pattern Recognition Letters*, Special Issue on Fuzzy Pattern Recognition, Vol. 17, No. 6, pp. 577-584, May 1996.

P. D. Gader, Bruce N. Nelson, A. Koksal Hocaoglu, Sansanee Auephanwiriyakul, Mohamed A. Khabou, "Neural versus Heuristic Development of Choquet Fuzzy Integral Fusion Algorithms for Land Mine Detection," in *Neuro-fuzzy Pattern Recognition* H. Bunke, A. Kandel (eds.): World Scientific Publ. Co., pp 205-226, 2000.

S. Beiraghi, M. Ahmadi, M. Sid Ahmed, M. Shridhar, "Application of Fuzzy Integrals in Fusion of Classifiers for Low Error Rate Handwritten Numerals Recognition", Proceedings of the International Conference on Pattern Recognition, Barcelona Spain, pp. 24872491, Sept 2000.

J. Chiang, "Choquet Fuzzy Integral-Based Hierarchical Networks for Decision Analysis", *IEEE Transactions Fuzzy Systems*, Vol. 7, No. 1, pp. 63-71, Feb. 1999.

Hocaoglu and Gader, "Comments on Choquet Fuzzy Integral-Based Hierarchical Networks for Decision Analysis", IEEE Transactions Fuzzy Systems, Vol. 7, No. 6, pp. 767-767, Dec. 1999.

William C Radzelovage, Personal Communication

X. Yu, I. S. Reed, and A. D. Stocker. "Comparative Performance Analysis of Adaptive Multispectral Detectors," *IEEE Trans. Signal Processing*, Vol. 41, No. 8, August 1993, pp. 2639-2656.

E. Winter, Personal Communication

D. Heinz and C. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery", IEEE Transactions Geoscience and Remote Sensing, Vol. 39, No. 3, March 2001, pp. 529-545.

J. Broadwater, R. Meth, R. Chellappa," A Hybrid Algorithm for Subpixel Detection in Hyperspectral Imagery" Proceedings IEEE Conference on Geoscience and Remote Sensing (IGARSS 2004), (to appear)

P. D. Gader, W. Lee, and A. Mendez-Vasquez, "Continuous Choquet Integrals with respect to random sets with applications to landmine detection", Proceedings IEEE Conference Fuzzy Systems, July 2004 (to appear).

L. S. Shapley, "A value for n-person games", in Contributions to Game Theory, Kuhn and Tucker (eds.), Princeton University Press, Princeton, NJ, 3-7-317, 1953.

M. Grabisch and M. Roubens, "An axiomatic approach to the concept of interaction among players in cooperative games", *International Journal of Game Theory*, Vol. 28, No. 4, pp. 547-565, Nov. 1999.