The Automatic Fusion of Classified Sidescan Sonar Mosaics Using CML-RTS and Markov Random Fields

S. Reed, I. Tena Ruiz , C. Capus , and Y. Petillot

Abstract— **This paper presents a framework for registering and fusing classified sidescan sonar data. It builds on recent advances in navigation and registration for improved mosaicing, applying novel fusion algorithms to integrate data from overlapping sidescan survey lines to produce large scale classified mosaics.**

While typical Mine-Counter-Measures (MCM) and Rapid Environmental Assessment (REA) missions provide various overlapping views of the same region of seafloor, research on sidescan image analysis has traditionally concentrated on the analysis of individual images. The available information from the other images, relating to the same region of seafloor, is generally not considered. The image registration and mosaicing process allows this complementary data to be fused, producing an improved final classification result.

The sidescan imagery is first pre-processed through the application of advanced radiosity correction algorithms. Following radiosity correction, texture segmentation for the data presented in this paper is achieved using features derived from the averaged normalised power spectral density.

The individual classification maps are geo-referenced and co-registered using a Concurrent Mapping and Localisation Rauch-Tung-Striebel (CML-RTS) procedure. This uses local landmarks within the individual images and the AUVs navigation data to generate a more accurate and smooth navigation trajectory. This trajectory is used to produce the registered classification mosaics.

The co-registered classification results are then fused to produce an improved class mosaic for the entire survey region. The fusion model uses a voting scheme to initialize the seafloor map after which a Markov Random Field (MRF) model is used to produce the final fused classification mosaic.

The entire process (Classification, Registration and Fusion) is demonstrated on real sidescan data taken at the Saclant Centre, La Spezia, Italy.

I. Introduction

The development of stable Autonomous Underwater Vehicle(AUV) platforms, fitted with high resolution sonars, has opened up the oceans to rapid and high resolution mapping. This paper tackles the difficult problem of generating accurate large scale maps of the seabed from sonar imagery obtained using these platforms.

Research into image based classification of sonar data has generally concentrated on analyzing single sonar images. The generation of large scale mosaics raises the problems of image registration, autonomous navigation in an environment deprived of GPS (Global Positioning Systems) and the automatic fusion of multiple, possibly contradictory, class maps. This paper presents a first attempt at producing large scale classified mosaics by integrating research conducted separately in the fields of navigation, classification and data fusion.

A. Image Formation Process and Pre-processing

The interpretation of sidescan imagery is a skilled procedure. There are many parameters of the image formation process contributing to intensity variations in recorded data which are quite separate from the influences of variations in seabed properties and textures. In the current work the sonar data are preprocessed to correct for the influences of the sonar beam pattern and time-varying gain (TVG) [1]. This enables the use of simpler and faster classification algorithms which is particularly beneficial for application to large area surveys.

B. Classification

Given the large numbers of images produced in a typical sidescan sonar survey, a fast classification algorithm is required to produce the seabed class maps. Many supervised techniques have been developed to tackle this problem. While Neural Networks and parametric statistical classifiers have dominated the scene [2], [3], other approaches such as fractal analysis [4], spatial point processes [5], grey level run-length measures [6], co-occurrence matrices [7] and fuzzy logic analysis [8] have also been investigated.

In this paper a variant of the power spectrum feature set [9] is used with a simple parametric classifier for rapid supervised classification of the individual images.

C. Registration and Mosaicing

To produce large scale mosaics it is necessary to register the individual sidescan sonar images. Given the position of the sensor in the world for each sidescan beam, the individual images can be mosaiced together [10]. The quality of the mosaics will depend on the precision of the position information of each sonar beam, i.e. the vehicle's navigation precision. Underwater, navigation is a difficult problem as conventional GPS receivers do not operate. Underwater platforms therefore usually rely on dead-reckoning sensors which drift over time. To fix this drift, we propose using a Concurrent Mapping and Localization (CML) technique based on the stochastic map developed for indoor robotics [11] and adapted to sonar imagery [12]. This method is based on the use of landmarks detected in the sonar imagery to help the navigation. It does not require the addition of additional sensors and does not interfere with the data acquisition constraints (stable platform, fixed altitude).

0-7803-9103-9/05/$20.00©2005 IEEE

D. Fusion

During a typical sidescan survey, multiple views of the same area are normally collected from different view points. The fusion of these views enables the generation of large scale classified mosaics with improved classification accuracy.

Data fusion is a well established research field and has used Bayesian Statistics [13], Fuzzy Logic [14] and Dempster-Shafer theory [15]. When little is known about the information sources, or when they produce information at a high level of abstraction, voting schemes have been successfully used [16]. This is particularly appealing for underwater imaging systems which currently favor 'black box' approaches to classification.

Fusion for classification within the image domain allows contextual information to be considered [17]. The fusion of multiple images is generally performed at the pixel level [18], [19], allowing information from surrounding pixels to be considered when classifying each pixel. An effective method for incorporating this spatial information is the use of Markov Random Fields(MRF) [20], [21].

The fusion model described in this paper uses a modified Majority Voting scheme to initialize the final mosaic. After, a MRF model is used to incorporate contextual information, smoothing the final result and 'inpainting' regions of pixels which are *unclassified* after the Voting process.

II. Side-scan Data Pre-processing

In low altitude surveys, small changes in vehicle altitude can affect the sonar image dramatically. Prior to classification the image data used has been preprocessed using an advanced radiosity correction algorithm [1]. This treats purely range-dependent artifacts, such as residual TVG effects separately to angular effects such as the influence of the sonar beam pattern. Separate correction factors are calculated for each. Whilst this gives better performance than standard radiosity correction algorithms in the presence of sensor altitude changes, platform stability is still assumed with respect to pitch and roll.

A sample raw image and the estimated beam pattern and residual TVG profile for these data are shown in Fig. 1. The complexity of the beam pattern is apparent with four significant lobes in the port channel and as many as six in the starboard channel.

The corrected image is shown in Fig. 2. In some places the beam pattern correction has failed, as indicated by the white arrow. This arises from the behavior of the vehicle, which rolls on turns.

A. Mosaicing

Each side of the sidescan sonar (port and starboard) insonifies a rectangular area on the sea floor. The length of the rectangle is determined by the slant range of the sonar (the maximum range of the sonar) and the height of the vehicle.

The seabed can be represented as a flat two dimensional grid of mosaic cells. Using simple geometry both channels can be superimposed if the orientation and position of the sonar is known. Each mosaic cell will take the value of the intensity cell on the beam closest to it. In this paper if a mosaic cell takes more than one value then, in the case of the pre-processed image mosaics, the data will be averaged or, in the case of the classified data, the cell will be left as unclassified.

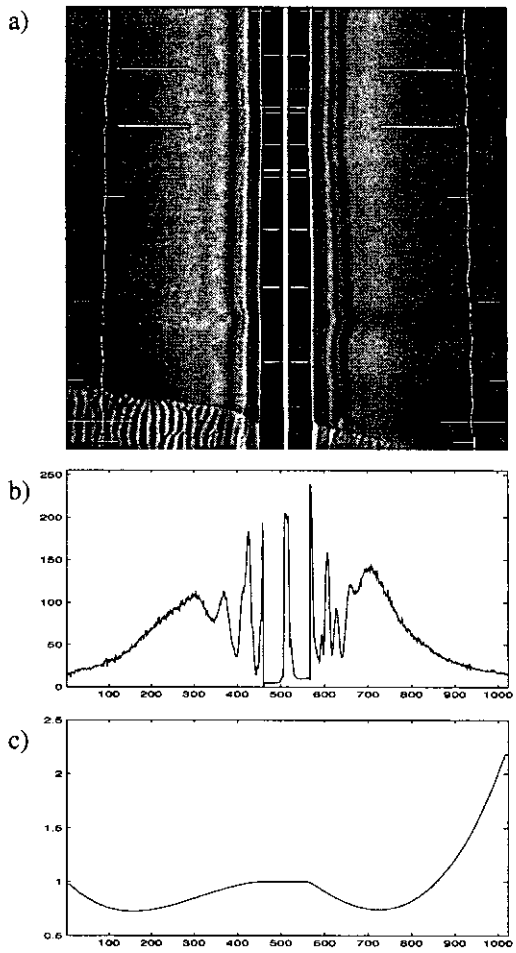

Fig. 1. **(a) sample raw image (1000 scan lines); (b) beam pattern estimate; (c) residual TVG estimate. Differences in ranges for y-axes derive from the methods of calculation. Values in (b) are referenced to a target grey level, hence range 0–255, values in (c) are referenced to a level of 1.0.**

III. Classification of Side-scan Data

Three seafloor textures have been identified for segmentation of the sidescan data, defining three classes: flat sediments, sand ripples, and complex regions. Suppression of the beam pattern effects and some of the residual TVG effects improves the images to the point where a fast supervised classification scheme can be combined with a relatively simple, easily generated feature set.

The features used are based on Pace and Gao's frequency based sediment classification scheme [9]. Overlapping 64-sample Gaussian windowed FFTs are used to generate the one-dimensional power spectra and this allows for identification of changes in texture across the sonar swath.

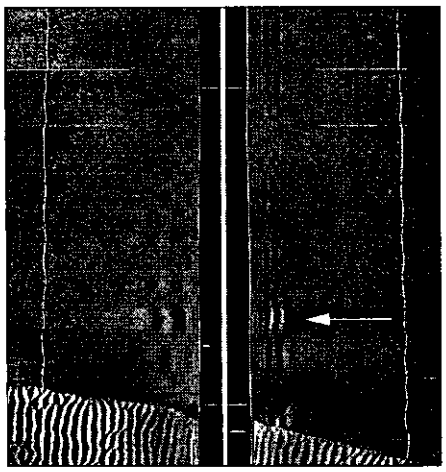

Fig. 2. **Corrected image. The correction algorithm cannot compensate for changes in sensor attitude, such as roll on trajectory corrections, see arrow above.**

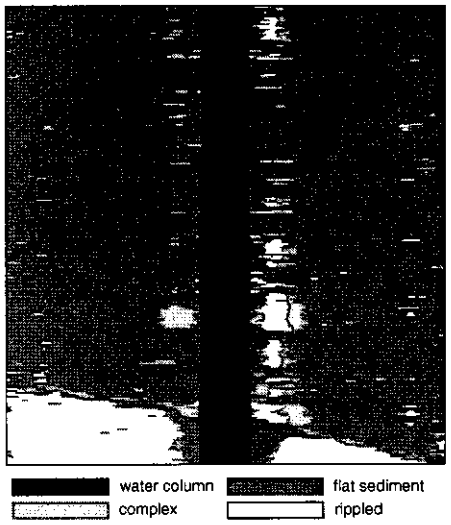

Fig. 3. **Classmap generated from the image introduced in Fig. 1. Water column – black; flat sediment – dark grey; complex – light grey; rippled – white. Errors are noted where the correction algorithm has failed due to vehicle attitude changes during course corrections and in the transition regions between textures.**

Spatial frequency bands within the normalized power spectra are identified which give a good separation between the classes. The training set used for these data comprised three small sections, one for each texture, of 200 × 200 pixels.

Three features are defined by the crossing points of the averaged normalized spectra derived from the training data. These give the proportion of the spectrum lying in sample bands 1-4, 4-12 and 16-32, corresponding to crossing points lying at $f_{\max}/8$, $3f_{\max}/8$ and $f_{\max}/2$.

$$\begin{aligned} D_{f1} &= \int_1^{f_{\max}/8} P_{iN}(f) \Big/ \int_1^{f_{\max}} P_{iN}(f) \\ D_{f2} &= \int_{f_{\max}/8}^{3f_{\max}/8} P_{iN}(f) \Big/ \int_1^{f_{\max}} P_{iN}(f) \\ D_{f3} &= \int_{f_{\max}/2}^{f_{\max}} P_{iN}(f) \Big/ \int_1^{f_{\max}} P_{iN}(f) \end{aligned} \quad (1)$$

In classifying a complete sonar image, the three features are generated from the averaged normalized spectral density formed from four successive lines of data. The same 64-sample sliding Gaussian windowed FFT is used and boundary problems between sonar channels are minimized by closing up the water column. This is done simply by shifting the scanlines on the assumption that there will generally be continuity in seabed textures between sonar channels.

Fig. 3 shows the initial classification result for the image introduced in Fig. 1 above. Misclassifications are greatest near the water column where the correction algorithm has failed. There are some boundary errors, with pixels classified as complex texture in the transition region between flat sediment and sand ripples. Misclassification further from the water column is due primarily to incomplete elimination of the influence of the surface return and crosstalk from other sensors. These effects are particularly prominent in this data set and add to the difficulty of the classification task. These misclassifications can frequently be rectified by the proposed fusion scheme.

IV. Concurrent Mapping and Localisation (CML)

In order to create an accurate mosaic of the classified maps, good navigation is crucial. The purpose of CML is to build a map of the environment and use that same map to localize the sonar [11]. Recently CML techniques have been developed to work with a side-scan sonar [10], [12]. This paper uses this method in order to geo-reference classified side-scan images.

The technique to fuse the navigation data uses a stochastic map smoothed using a Rauch-Tung-Striebel (RTS) fixed-interval smoother. It will be referred to as CML-RTS. The stochastic map keeps the estimates of the position and creates a map of landmarks to represent the environment. These landmarks are then used to aid localization of the vehicle. It is a CML method that works iteratively to provide an estimate of the position at the latest iteration. In order to improve the accuracy of the solution and to smooth it, post-processing is required. The next two sections provide a more detailed look at the algorithms.

A. The Stochastic Map

The stochastic map is an augmented state Extended Kalman Filter (EKF) [22]. It adds new states to the state vector to accommodate new landmarks as they are observed. A typical stochastic map state vector is of the form:

$$\mathbf{x} = [\mathbf{x_v} \mathbf{x_1} \ldots \mathbf{x_n}]' \quad (2)$$

where $\mathbf{x_v}$ holds the state of the side-scan sonar and $\mathbf{x_1}, \ldots, \mathbf{x_n}$ holds the state of the n landmarks in the map.

The stochastic map also stores and maintains all the covariances and correlations between the states. With fully correlated landmarks, an observation of any of the landmarks will help correct the whole map. It can also take advantage of the wealth of literature published on Kalman filters. The update equations of the stochastic map are the familiar EKF update equations. For more details on this implementation of the stochastic map, the interested reader should refer to [12].

B. CML-RTS

The Kalman filter and EKF use all measurements up to the last iteration to estimate the state at the last iteration. The RTS smoother uses all measurements *before* and *after* each iteration to estimate the state at each iteration. It is a post-processing filter that works on the stored outputs of a Kalman filter by re-processing it. The smoother works by combining a forward pass Kalman filter with a backward pass filter. It was originally designed to work with fixed size state vectors. However, the stochastic map adds new states to the state vector as it observes new landmarks. The CML-RTS algorithm adapts the RTS fixed-interval smoother to work with the stochastic map by fixing the size of the state vector to the size of the stochastic map on the last iteration. The CML-RTS algorithm ensures numerical stability in matrix operations by adjusting the estimates of the landmarks' states and covariances at all iterations before they have been observed to the values when they are first observed. The output of the CML-RTS has been shown to improve the accuracy of the stochastic map solution [12], as well as providing trajectories more suitable for creating and superimposing mosaics [10].

V. Fusion of Mosaiced Classification Data

This section presents the pixel level model for fusing the multiple classified sidescan sonar mosaics. The final fused mosaic is initialized using a voting scheme. The model is formulated within a Markovian framework to take advantage of contextual information and improve classification accuracy.

A. Initialising the Fusion Mosaic

The Fusion Mosaic is initialized using a voting scheme. The model assumes that each image j provides a classification result for each pixel label x_s^j. The fusion field W is initialized by using an adaptation of the Generalized Majority Voting [16]. In this model, a summed binary function $T_s(\omega_i)$ for pixel s, and each recognized seafloor classes $\omega_i, 1 \leq i \leq M$ is specified as

$$T_s(\omega_i) = \sum_{j=1}^{K} \delta(x_s^j, \omega_i) \text{ for } 1 \leq i \leq M \tag{3}$$

where the sum is over all the inputted class images and $\delta(.)$ is the Kronecker Delta Function. This function is not specified for the *unclassified* or *unmeasured* classes.

The initial fusion Field W is then specified as:

$$\begin{aligned} w_s &= \gamma \text{ if } \sum_{j=1}^{K} \delta(x_s^j, \gamma) = K \\ &= \tau \text{ if } T_s(\tau) = \max\ T_s(\omega_i) \geq \frac{2}{3} K^s \\ &= \alpha \text{ otherwise} \end{aligned}$$

where K^s is the number of images which do not provide an *unclassified* or *unmeasured* classification for pixel s.

Once the voting rule has been used to initialize Fusion Field W, a Markov model can be used to smooth the final fusion result and inpaint any *unclassified* pixels.

B. Markov Model for Image Fusion

Let us assume first that each of the input class maps is defined on a lattice S where label s specifies a specific pixel location. Two random fields $\mathbf{X}$ and $\mathbf{W}$ are defined. $\mathbf{X} = \{X_s, s \in S\}$ describes the classification field provided by each input map and $\mathbf{W} = \{W_s, s \in S\}$ describes the final fused classification map. For K input class maps, $X_s = (X_s^1, \ldots, X_s^K)$ takes its values from the finite set of classes $\Omega = \{\omega_1, \ldots, \omega_M, \alpha, \gamma\}$. The set Ω contains M recognized seafloor classes, the *unclassified* label α and the *unmeasured* label γ. Label α is allocated to $X_s^j, j \in \{1, \ldots, K\}$ when data is received regarding pixel s in image j but a classification based on the data provided is not possible. Label γ, *unmeasured*, is used when no data is received regarding pixel s, ensuring it is not possible to provide a classification X_s^j.

The fusion problem consists of estimating the true classified map $W = w$ from the individual classified maps $X = x$ where $x = x_s, s \in S$ are K classified maps of the same scene. The field $W = \{W_s, s \in S\}$ is said to be Markovian with respect to neighborhood $\eta = \{\eta_s, s \in S\}$ if its distribution can be written as

$$\begin{aligned} P_{W_s}(W_s = w_s | W_r = w_r, r \neq s) = \\ P(W_s = w_s | W_r = w_r, r \in \eta_s) \end{aligned} \tag{4}$$

For simplicity, the fusion model described in this paper assumes a second order isotropic neighborhood. Further reading regarding MRF models can be found in [20], [21].

The problem of maximizing probability $P_W(w)$ can be recast to the local problem of maximizing energy

$$U(w_s) = \sum_{t \in \eta_s} \beta \delta(w_s, w_t)[1 - \delta(w_t, \alpha)][1 - \delta(w_t, \gamma)] \tag{5}$$

for pixel s. In equation 5, $\delta(.)$ is the Kronecker Delta symbol and β controls the importance of the Markovian prior. For all cases in this paper, $\beta = 1.0$. The minimization of $P(W)$ is performed using the Iterated Conditional Modes method [21]. In this method, a raster scan is used to iteratively visit all the pixels in field W. If $w_s = \gamma$ *unmeasured*, the pixel is not considered further and the pixel remains *unmeasured*. Otherwise, w_s is allocated to the class which

locally maximizes $U(w_s)$. The ICM procedure is iterated until there are no pixel changes within a full image scan.

VI. Large Scale Mosaic Results

The following results were obtained by processing data gathered during the BP'02 experiments carried out by the SACLANT Undersea Research Centre in La Spezia, Italy. The side-scan data was gathered by a REMUS AUV [23]. The AUV mission lasted for 2 hours, 57 minutes and 8 seconds. It followed a set of parallel, regularly spaced and overlapping linear tracks (typical for rapid environmental assessment missions).

The data have been classified using the techniques outlined in section III.

Figs. 4 contains 4 mosaics of sector 1 (a pre-defined region) created by geo-referencing 17 linear tracks. All of the resulting mosaics created using the overlapping tracks are geo-referenced to the same reference frame (sector 1) and will constitute an input to the fusion algorithm.

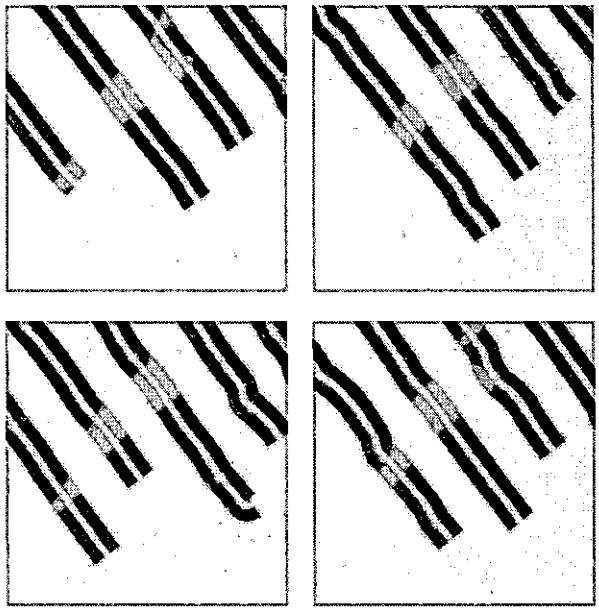

Fig. 4. **The 4 input mosaics of Sector 1 which are fused together to produce the final fused mosaic for sector 1.**

The mosaics in Fig. 4 contain a maximum of 5 classes. Each pixel is considered to belong to the sand, ripple, complex, *unmeasured* or *unclassified* class. The large light grey regions are *unmeasured* regions over which the AUV has not passed. The final fused result produced from the 4 input mosaics can be seen in Fig. 5.

The final fused result in Fig. 5 contain much more information than any of the mosaics considered in isolation. The regions of seafloor which are classified as *unmeasured* in all the input mosaics have also been left as *unmeasured*. The final fused map also contains no *unclassified* regions. These have been inpainted using the MRF model.

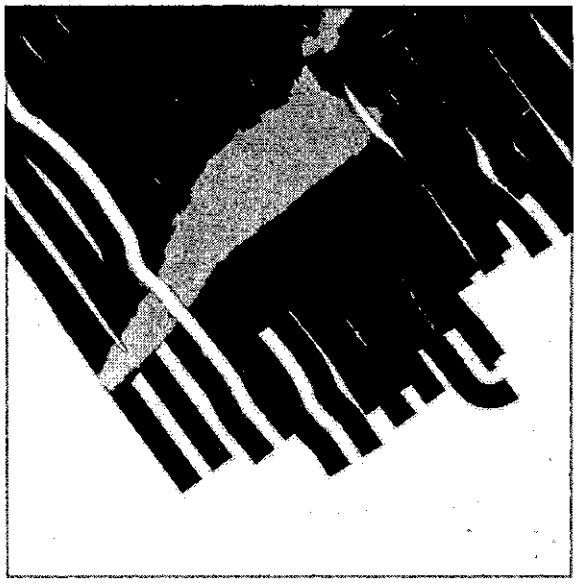

Fig. 5. **The Final Fused result obtained from the Voting/MRF fusion model for the sector 1 mosaics.**

Fig. 6 contains the 4 input mosaics for a second sector. Sector 2 has been created by geo-referencing 15 linear tracks.

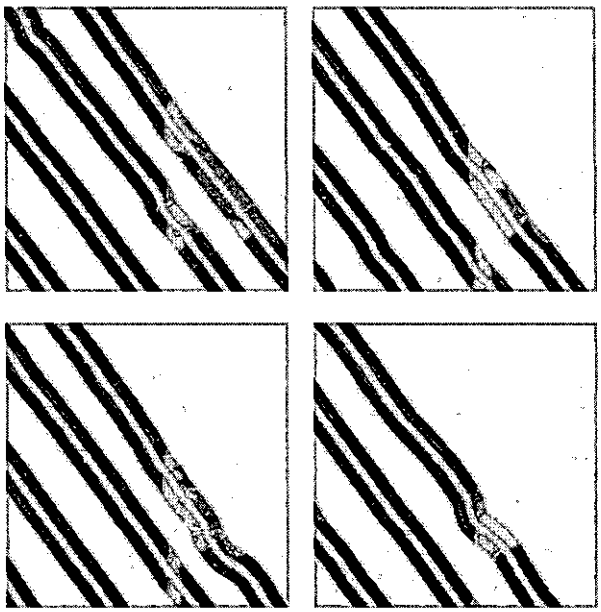

Fig. 6. **The 4 input mosaics of Sector 2 which are fused together to produce the final fused mosaic for sector 2.**

Again, there are 5 classes present within the input mosaics in Fig. 6. The *unmeasured* class is the dominant class and is represented by light grey. The *unclassified* class is represented by white. The sand, ripple and complex classes are also present. The final result from the fusion model can be seen in Fig. 7.

As Fig. 7 shows, the Fusion model has produced a more

Fig. 7. **The Final Fused result obtained from the Voting/MRF fusion model for the Sector 2 mosaics.**

complete and useful picture of the seafloor. The information from the individual mosaics has been fused to produce a smoothed map of the seafloor region. The benefits of the Fusion model can be clearly seen in both the examples shown in Figs. 5 and 7.

VII. Conclusions

This paper has presented a method for creating and fusing classified sidescan sonar mosaics of the seafloor. A normalization and classification model was then detailed which allowed inherent sidescan sonar problems such as the beampattern to be considered. This segmented the images into regions of flat seafloor, sand ripples and complex regions.

A mosaicing algorithm was then presented. This used CML techniques to produce high quality mosaics of the individual classification maps. The generation of these mosaics, where the images are geo-referenced in space, allows the possibility of multi-mosaic fusion.

A fusion model using simple voting techniques within a Markovian framework was then presented. Results were demonstrated on real classified sidescan mosaics. The mosaics were produced using the classification and mosaicing models presented in the first sections of this paper. Information from the individual mosaics was fused to produce a Fusion map of the entire region of seafloor surveyed by the AUV.

References

[1] C. Capus, I. Tena Ruiz, and Y. Petillot, "Compensation for changing beam pattern and residual tvg effects with sonar altitude variation for sidescan mosaicing and classification," in *Proc. 7th. European Conference on Underwater Acoustics*, Delft, The Netherlands, 2004.

[2] R. D. Muller, N. C. Overkov, J.-Y. Royer, A. Dutkiewics, and J. B. Keene, "Seabed classification of the south tasminrise from simrad em12 backscatter data using artificial neural networks," *Australian Journal of Earth Sciences*, pp. 689–700, 1997.

[3] B. Zerr, E. Maillard, and D. Gueriot, "Seafloor classification by neural hybrid system," in *Proc. IEEE Int. OCEANS Conf.*, 1994, pp. 239–243.

[4] D. R. Carmichael, L. M. Linnett, S. J. Clarke, and B. R. Calder, "Seabed classification through multifractal analysis of sidescan sonar imagery," *IEE Radar, Sonar and Navigation*, no. 3, pp. 140–148, 1996.

[5] L. M. Linnett, D. R. Carmichael, S. J. Clarke, and A. D. Tress, "Texture analysis of sidescan sonar data," in *Proc. IEE Conf. Texture Analysis in Radar and Sonar*, London, Nov. 1993.

[6] W. K. Stewart, M. Jiang, and M. Marra, "A neural network approach to classification of sidescan sonar imagery from a midocean ridge area," *IEEE J. Oceanic Eng.*, no. 2, pp. 214–224, Apr. 1994.

[7] S. Subramaniam, H. Barad, and A. B. Martinez, "Seafloor characterization using texture," in *Proc. IEEE South East Conf.*, New Orleans, USA, April 1993.

[8] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy, "Fuzzy logic modelling in sonar imagery: Application to the classification of underwater floor," *Computer Vision and Image Understanding*, pp. 4–24, 2000.

[9] N. G. Pace and H. Gao, "Swathe seabed classification," *IEEE J. Oceanic Eng.*, no. 2, pp. 83–90, 1988.

[10] I. Tena Ruiz, Y. Petillot, and D. M. Lane, "Improved AUV navigation using side-scan sonar," in *Proc. MTS/IEEE Internationnal Conference OCEANS'03*, 2003, pp. 1261–1268.

[11] J. J. Leonard, R. N. Carpenter, and H. J. S. Feder, "Stochastic mapping using forward look sonar," in *Proceedings of International Conference on Field and Service Robotics*, Pittsburgh, Pennsylvania, USA, August 1999, pp. 69–74.

[12] I. Tena Ruiz, S. de Raucourt, Y. Petillot, and D. M. Lane, "Concurrent mapping and localisation using side-scan sonar," *IEEE J. Oceanic Eng.*, vol. 29, no. 2, pp. 442–456, Apr. 2004.

[13] A. Solberg, A. Jain, and T. Taxt, "Multisource classification of remotely sensed data: Fusion of landsat TM and SAR images," *IEEE Trans. Geoscience and Remote Sensing*, no. 4, pp. 768–778, JULY 1994.

[14] I. Bloch, "Information combination operators for data fusion: A comparative review with classification," *IEEE Trans. Systems, Man and Cybernetics-Part A: Systems and Humans*, no. 1, pp. 52–67, JAN 1996.

[15] ——, "Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume into account," *Pattern Recognition Letters*, pp. 905–919, 1996.

[16] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. Systems,Man and Cybernetics*, no. 3, pp. 418–435, MAY/JUNE 1992.

[17] S. L. Hegarat-Mascle, I. Bloch, and D. Vidal-Madjar, "Application of dempster-shafer evidence theory to unsupervised classification in multisource remote sensing," *IEEE Trans. Geoscience and Remote Sensing*, no. 4, pp. 1018–1031, JULY 1997.

[18] A. Bendjebbour, Y. Delignon, L. Fouque, V. Samsom, and W. Pieczynski, "Multisensor image segmentation using Dempster-Shafer fusion in Markov fields context," *IEEE Trans. Geoscience and Remote Sensing*, no. 8, pp. 1789–1798, AUGUST 2001.

[19] S. L. Hegarat-Mascle, A. Quesney, D. Vidal-Madjar, and O. Taconet, "Land cover discrimination from multitemporal ERS images and multispectral landsat images: A study case in an argicultural area in france," *Int. Journal Remote Sensing*, no. 3, pp. 435–456, 2000.

[20] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary detection by constrained optimization," *IEEE Trans. Pattern Analysis and Machine Intelligience*, no. 7, pp. 609–628, JULY 1990.

[21] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc.*, no. 3, pp. 261–279, 1986.

[22] Y. Bar-Shalom and T. Fortmann, *Tracking and data association.*, ser. Mathematics in Science and Engineering. Academic Press, 1988, vol. 179.

[23] C. von Alt, B. Allen, T. Austin, N. Forrester, R. Goldsborough, M. Purcell, and R. Stokey, "Hunting for mines with REMUS: a high performance, affordable, free swimming underwater robot," in *Proc. MTS/IEEE Internationnal Conference OCEANS'01*, 2001, pp. 117–122.