Fusion of Multiple Side-scan Sonar Views

E. Coiras, I. Tena Ruiz, Y. Petillot, D. M. Lane

Ocean Systems Laboratory

School of Engineering and Physical Sciences

Heriot-Watt University, Edinburgh, EH14 4AS, Scotland, UK

http://www.ece.eps.hw.ac.uk/oceans/

Abstract

A new set of techniques for the construction of large-scale side-scan mosaics is presented in this paper. The complete procedure operates in two main stages. First an accurate registration of the source side-scan images is performed by filtering navigation data, where an augmented state *Rauch-Tung-Striebel (RTS)* filter is used to smooth the stochastic map obtained from a *Concurrent Mapping and Localization (CML)* algorithm– which stores landmarks as extra states in a Kalman filter. Then a fusion algorithm based on Gabor wavelets is used to combine the registered side-scan images and assemble the final mosaic. The algorithm can be tuned to either maximize the information content or to minimize signal noise in the mosaic.

The paper presents novel results created by fusing 14 views of a scene as observed by a side-scan sonar mounted on an *Autonomous Underwater Vehicle (AUV)*. The resulting mosaics are tailored to preserve the texture information of the individual views, accentuate the returns from important objects in the scene, and cover a wider area than any one single view.

1 Introduction

Imaging of the seafloor by AUVs has produced high resolution images of regions of interests [1]. AUVs offer a stable platform capable of operating at close distance from the seafloor. It is now possible

to covertly run *Rapid Environmental Assessment (REA)* shallow water missions, producing multiple views of the same regions. These type of missions are valuable for *Computer Aided Detection/Computer Aided Classification (CAD/CAC)* exercises and also for rapid characterization of the seafloor [2, 3, 4]. However, the accuracies of the navigation sensors used by AUVs are limited by the operating environment. Whereas in air DGPS or GPS aided INS systems can produce highly accurate solutions, electromagnetic signals attenuate quickly in water and satellite signals are limited to instances when the AUVs surface. Conventional navigation suites combine dead-reckoning sensors with either acoustic networks such as *Long Baseline (LBL)* networks, or acoustic trackers such as *Ultra Short Baseline (USBL)* systems [5, 6, 7]. The resulting navigation solutions are generally acceptable for most mission requirements, but do not meet the standards required by image fusion algorithms which are designed under the assumption that the images have been perfectly registered [8]. Distortions due to seabed topography cannot in general be corrected completely, for instance. Therefore, development of appropriate sonar registration tools is essential, as it is also the development of a fusion engine designed to work with non-perfectly registered input sources.

The *Ocean Systems Laboratory (OSL)* has recently developed a technique suitable for geo-referencing and co-registering multiple side-scan views of the same area [9]. Geo-referenced side-scan images are formed using the knowledge of the side-scan's position and orientation as the scene is observed. The appearance of the images is influenced by the relative orientation of the sonar and the seabed features being observed [10]. Thus image matching techniques to co-register side-scan data sets, as proposed in [11], will not be suitable under all circumstances. The OSL has instead produced a convenient system that allows a human operator to manually match the sonar returns, by extracting and matching landmarks. This system has also been extended to work with automatic landmark extraction techniques [12]. The system subsequently uses the information to correct the vehicle navigation. The chosen strategy is the stochastic map [13], an augmented state Kalman filter. This filter's state vector stores the sonar's states and adds new states to for each new observed landmark. The trajectory resulting from the stochastic map is not a smooth trajectory suitable for geo-referencing the side-scan data.

Each point in the trajectory is processed by using observations that occurred before that point. A smoother trajectory results by using measurements before and after that point. The OSL has adapted the RTS filter to work with the stochastic map. The new smooth trajectory creates accurate and smooth trajectories useful for geo-referencing the data.

Once the source images have been registered they must be properly assembled in order to construct the side-scan mosaic. The choice of the fusion algorithm [14] will determine how the information contained in the original images is managed and presented in the final result. An adequate selection of the fusion approach is therefore important to preserve the particular type of information in which the user is interested.

Thus, for the preservation of the visual information and image details a wavelet-based method should be preferred, while simpler blending models will reduce the influence of spurious image artifacts. To cover these two opposite goals, we have developed a tunable fusion algorithm based on Gabor wavelet decomposition [15, 16, 17] which can be adjusted to the particular requirements of the user. The method initially decomposes the registered images into their constituent spatial frequency components, which are blended separately and then recombined to form the final side-scan mosaic.

The rest of the paper is organized as follows:

In section 2 the CML-RTS strategy designed by the OSL is examined , and details on the functionality of the co-registration strategy are presented.

Section 3 describes the fusion algorithm designed to combine the registered sonar images into the final mosaics.

In section 4 novel registration and fusion results are presented by assembling 14 different side-scan views of the same seafloor region.

The last section summarizes the findings presented in this document and explores possibilities for future work.

2 Registration of Side-scan Sonar Using the Stochastic Map

The registration process uses the navigation information stored by the system and the side-scan data to register the data. The chosen tools are the stochastic map and the RTS filter. Both of these techniques are now re-visited.

2.1 The Stochastic Map

The stochastic map is obtained by augmenting the state vector of an *Extended Kalman Filter (EKF)* with new states representing newly observed landmarks. The EKF is a popular and well understood technique [18, 19] and the stochastic map can benefits from this fact. The stochastic map keeps the estimates of the vehicle-to-vehicle, landmark-to-vehicle and landmark-to-landmark correlations. Thus new observations of landmarks or vehicle states help correct the whole stochastic map. The importance of these correlations has been demonstrated through some telling research [20, 21].

The stochastic map's state vector $\mathbf{x}$ assumes the following form,

$$\mathbf{x} = \begin{bmatrix} \mathbf{x_v} \\ \\ \mathbf{x_l} \end{bmatrix} \tag{1}$$

where $\mathbf{x_v}$ holds the state of the vehicle and $\mathbf{x_l}$ holds the states of the landmarks. The estimated error covariance for this system,

$$\mathbf{P} = \begin{bmatrix} \mathbf{P_{vv}} & \mathbf{P_{vl}} \\ \\ \mathbf{P_{lv}} & \mathbf{P_{ll}} \end{bmatrix} \tag{2}$$

where the submatrices $\mathbf{P_{vv}}$, $\mathbf{P_{vl}}$ and $\mathbf{P_{ll}}$ are the vehicle-to-vehicle, vehicle-to-landmark and landmark-to-landmark covariances respectively.

The state and covariance are updated according to the EKF update equations [18, 19]. Generic incarnations of the stochastic map assumes that the landmarks are fixed in the world, although dynamic landmarks could also be modeled.

2.2 Smoothing the Stochastic Map

The *Rauch-Tung-Striebel (RTS)* backward filter is commonly used to smooth the output of a Kalman filter [22]. The OSL has successfully used this filter to smooth the output of the stochastic map [9, 12], the technique is referred to as CML-RTS. The smoothing process is a non-real-time data processing strategy that uses all measurements between times 0 and T to estimate the state at a certain time t, where $0 \leq t \leq T$. The RTS filter uses the stored predictions and corrections of a Kalman filter to obtain the smooth estimates.

The RTS was designed to operate with state vectors of a fixed size. Thus the OSL had to adapt the RTS input by increasing the size of the stored prediction and correction state vectors and covariances to match the final size of the stochastic map. In the adapted RTS the state vector at time 0 will be the same size as the stochastic map's state vector at time T. To do this the predicted and corrected landmark states must be defined for those instances before they were actually observed. Making these values zeros will cause a numerical instability. The strategy is to make them equivalent to the stochastic map corrected state after the landmarks were observed the first time. Thus you are no more, nor less, certain on the position of the landmarks before it was first observed. The correlation terms between the new landmarks and the other states in the map are set to zero, because only after it has been observed can the landmark be correlated to the rest of the stochastic map.

3 Gabor Fusion

After registration the side-scan images must be properly combined to build the final mosaic. Fusion of images based on wavelet techniques [14, 15] is generally preferable to simpler methods (such as blending, averaging, etc) because all the visual information contained in the original images is preserved in the fusion result. As a counterpart, these fusion techniques also tend to highlight all the noise and mis-registration errors present in the source images [23, 24], which can make the interpretation of the final mosaic difficult.

To overcome this limitation we have developed a tunable fusion algorithm based on Gabor wavelet decomposition [15, 16] which can be adjusted to the particularities of the source images. The method initially decomposes the registered images into their constituent spatial frequency components, which are blended separately and then recombined to form the final side-scan mosaic. The blending mode for each frequency component can be selected to either preserve all the initial image details or to reduce image noise.

Gabor wavelet functions have the following form [16, 25]:

$$g_{x_0,y_0,f_n,\theta_m} = g_{0,0,f_n,\theta_m}(x,y) * \delta(x-x_0, y-y_0) \tag{3}$$

where

$$g_{0,0,f_n,\theta_m}(x,y) = \exp(-\pi a^2((x\cos\theta_m + y\sin\theta_m)^2 + (x\cos\theta_m - y\sin\theta_m)^2)) \cdot \exp(\imath(2\pi f_n(x\cos\theta_m + y\sin\theta_m))) \tag{4}$$

and

$$a = \frac{1}{3}\sqrt{\frac{\pi}{\ln 2}} \cdot f_n \tag{5}$$

Parameters x_0, y_0 set the spatial location of the wavelet, f_n its frequency and θ_m its orientation. The value of a corresponds to the radial bandwidth of the Gabor function and is proportional to the radial frequency f_n the Gabor function is tuned to. In general these frequencies should be selected to fit the sizes of those elements of interest to the user– such as sand ripples, mounds, ground targets– in order to split the source image into meaningful spatial frequency bands. In this paper however, and as a proof of concept, empirically adjusted frequency values have been selected for the examples shown, which depend only on the dimensions in pixels of the source images. Also, for simplicity of implementation, wavelets corresponding to different orientations have been combined into a single integrated filter, as described in [17].

For the construction of the final mosaic, several additional considerations have to be addressed. Boundaries of the registered images, for instance, cause a powerful response in the high-frequency band (see Fig. 4), which means that they'll be clearly reproduced in the mosaic. To prevent it the boundaries

of the registered images were blurred to the size that cancels the response of each band filter, therefore completely removing their influence on the final result. More elaborate techniques [26] could also be used if the details contained in the border pixels must not be spared.

Image artifacts and noisy regions should also be eliminated from the final result whenever possible. To this end, each frequency band is processed separately in order to estimate the most probable pixel values at each position and scale. In the original fusion approach [15] the maximum value is always selected, which results in a mosaic that preserves all the image information but also all the noise and image boundaries (Fig. 4). In our case, a median filter is used to select the most probable pixel value at a given frequency band across the different images, which ensures that consistent features will be preserved and spurious defects removed. Fig. 5 shows the result of this fusion approach.

The algorithm weights and blending modes for the different spatial frequencies can be modified to enhance specific image features or reveal underlying patterns not clearly discernible by other means, as shown in the following section.

4 Results

These results were obtained by processing sidescan and navigation data recorded by the *Ocean Explorer (OEX)* AUV during the GOATS trials, organized by NATO SACLANT Undersea Research Centre [27]. The OEX observed a region of interest by running parallel transects over that area. These results were compiled by processing fourteen of those transects.

The CML-RTS solution is computed for all transects. To do this the operator extracts the landmarks manually from the data and matches them to the landmarks that already exist in the stochastic map. The landmarks and original trajectories, when plotted in a local East, North and Up (ENU) navigation frame, clearly show the error in the position estimates for the transects, see Fig. 1. The landmark observations display accross and along-track errors. The outcome of the CML-RTS system produces a set of smoothed transects and one estimate for each landmark position. This is a useful feature of the system that could be adapted for many different applications where landmarks (pipe

junctions, corals, underwater structures, wrecks, etc.) need to be inspected and mapped.

The trajectories were then used to create mosaics of transects. Fig. 2 shows the mosaic created by averaging the original stored data. The errors in the transect trajectories are apparent in the image. The resulting mosaic is not very informative, due to the errors in the co-registration of the data which result in a diluted image. Fig. 3 results from running the CML-RTS system and averaging the registered images. The figure clearly shows a much clearer result, associated to the more accurate geo-referenciation achieved by the registration method. Seafloor features and man made targets can be now clearly observed.

Minute details and dim areas are however blurred due to the blending by averaging image pixel values. Gabor fusion enhances every detail present in the original images, although that also includes image noise and distortion artifacts (see Fig. 4). The fusion approach presented in this paper is shown in Fig. 5, where the default algorithm settings– which give similar priorities to all frequency bands– were used. See how most image details that are consistent across the different images are preserved and most noise artifacts are removed.

The weights and blending modes for the different spatial frequencies can be modified to target specific image features. In Fig. 6 the algorithm has been tuned to reduce noise and image artifacts– note how most sensor noise has been removed while the details on dim areas and sand ripples are still present.

Appropriate tuning of the fusion algorithm can reveal underlying patterns not clearly discernible by other means. As an example see Fig. 7, where the settings have been set to enhance middle-sized features. Note for instance the rich seafloor structure in the central area of the mosaic, which is otherwise lost on the final mosaic.

5 Conclusions and Future Work

In this paper new registration and fusion techniques for the construction of mosaics of side-scan images have been presented. The CML-RTS registration method allows for accurate geo-referenciation of the

source images, which can then be combined to form the final mosaic. The proposed fusion algorithm, based on Gabor wavelet decomposition, has been designed to preserve all the information contained in the source images while simultaneously diminishing the influence of sensor noise and spurious geometrical distortions. The fusion algorithm can be tuned to enhance specific image features or reveal underlying structural configurations not observable by other means.

In the Ocean Systems Laboratory we are currently working on methods for the estimation of seafloor bathymetry by the rectification of side-scan images. Successful determination of seabed topography will permit more precise registration of the side-scan images, and will remove those mis-registration deformations that are associated to the simplified projection methods implicitly used in standard techniques for rendering side-scan data.

6 Acknowledgements

The authors would like to thank the SACLANT Undersea Research Centre, the US Office of Naval Research and Florida Atlantic University for allowing the inclusion of data from the GOATS 2000 experiment.

This research work has been partially supported by the 5th Framework Program of research of the European Community through the project AMASON (EVK3-CT-2001-00059).

References

[1] E. An, S. Smith, S. Dunn, T. Pantelakis, W. Whitley, J. Frankenfield, J. Kuntz, A. Burns, and R. Christensen. Ocean explorer AUV participation in the GOATS 2001 experiment. In E. Bovio, R. Tyce, and H. Schmidt, editors, *GOATS 2000 Conference. SACLANTCEN CP-46*, pages 61–78, La Spezia, Italy, August 2001.

[2] F. H. Vink, R. Hollett, and B. Zerr. Towards bottom classification with AUVs: segmentation of sidescan sonar imagery using an unsupervised artificial network. In E. Bovio, R. Tyce, and

H. Schmidt, editors, *GOATS 2000 Conference. SACLANTCEN CP-46*, pages 209–219, La Spezia, Italy, August 2001.

[3] 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*, pages 117–122, 2001.

[4] S. Reed, Y. Petillot, and J. Bell. An automated approach to the classification of mine-like objects in sidescan sonar using highlight and shadow informations. *IEE Proceedings Radar, Sonar and Navigation*, 2003. Accepted for publication.

[5] M. B. Larsen. *Autonomous Navigation of Underwater Vehicles*. PhD thesis, Technical University of Denmark, Department of Automation, Lyngby, Denmark, 2001.

[6] G. Grenon, P. E. An, S. M. Smith, and A. J. Healey. Enhancement of the inertial navigation system for the morpheus autonomous underwater vehicles. *IEEE Journal of Oceanic Engineering*, 26(4):548–560, October 2001.

[7] D. B. Marco and A. J. Healey. Command, control, and navigation experimental results with the nps aries auv. *IEEE Journal of Oceanic Engineering*, 26(4):466–476, October 2001.

[8] T. Macri-Pellizzeri, C. J. Oliver, and P. Lombardo. Segmentation-based joint classification of SAR and optical images. *IEE Proceedings Radar, Sonar and Navigation*, 149(6):281– 296, 2002.

[9] I. Tena Ruiz, S. de Racourt, Y. Petillot, and D. M. Lane. Concurrent mapping & localisation using side-scan sonar. *IEEE Journal of Oceanic Engineering*. submitted.

[10] J. M. Bell, M. J. Chantler, and T. Wittig. Sidescan sonar: a directional filter of seabed texture? *IEE Proceedings Radar, Sonar and Navigation*, 146(1):65–72, 1999.

[11] S. Daniel, F. Le Leannec, C. Roux, B. Solaiman, and E. P. Maillard. Side-scan sonar image matching. *IEEE Journal of Oceanic Engineering*, 23(3):245–259, July 1998.

[12] I. Tena Ruiz, S. Reed, Y. Petillot, J. Bell, and D. M. Lane. Concurrent mapping & localisation using side-scan sonar for autonomous navigation. In *Proceedings from the 13th International Symposium on Unmanned Untethered Submersible Technology*, Durham, New Hampshire, USA, August 2003.

[13] R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In I. Cox and G. Wilfong, editors, *Autonomous Robot Vehicles*. Springer-Verlag, 1990.

[14] Z. Zhang and R. S. Blum. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. In *Proceedings of the IEEE*, volume 87, pages 1315–1326, August 1999.

[15] J. Santamaria and M. T. Gomez. Visible-ir image fusion based on gabor wavelets decomposition. In *EOS Annual Meetings Digest*, volume 3, pages 97–98, 1993.

[16] O. Nestares, R. Navarro, J. Portilla, and A. Tabernero. Efficient spatial-domain implementation of a multiscale image representation based on gabor functions. *Journal of Electronic Imaging*, 7(1):166–173, 1998.

[17] R. Navarro and A. Tabernero. Gaussian wavelet transform: Two alternative fast implementations for images. *Multidimensional Systems and Signal Processing*, (2):421–436, 1991.

[18] P.S. Maybeck. *Stochastic models, estimation ans control. Volume 2*, volume 141-2 of *Mathematics in Science and Engineering*. Academic Press, 1982.

[19] Y. Bar-Shalom and T.E. Fortmann. *Tracking and data association.*, volume 179 of *Mathematics in Science and Engineering*. Academic Press, 1988.

[20] J. A. Castellanos, J. D. Tardos, and G. Schmidt. Building a global map of the environment of a mobile robot: the importance of correlations. In *IEEE International Conference on Robotics and Automation*, pages 1053–1059, Albuquerque, New Mexico, USA, April 1997.

[21] M. W. M. G. Dissanayake, H.F. Durrant-Whyte, S. Clark, and M. Csorba. A solution to the simultaneous localisation and map building (SLAM) problem. Technical Report ACFR-TR-01-99, the University of Sydney, Sydney, Australia, January 1999.

[22] A. Gelb, editor. *Applied Optimal Estimation.* The M.I.T. Press, Cambridge, MA, USA, 1974.

[23] D. A. Yocky. Artifacts in wavelet image merging. *Optical Engineering*, 35(7):2094–2101, July 1996.

[24] E. Coiras, J. Santamaría, and C. Miravet. Segment-based registration technique for visual-infrared images. *Optical Engineering*, 39(1):282–289, January 2000.

[25] B. Jähne. *Digital Image Processing, Concepts, Algorithms and Scientific Applications.* Springer-Verlag, 1991.

[26] C. Miravet, J. Santamaría, E. Coiras, J. Ureña, J.C. Escudero, and A. Sarasúa. Generación Semi-automática de Mosaicos. Aplicación de Técnicas de Fusión de Imágenes. *Revista de Teledetección*, (10):31–38, December 1998.

[27] E. Bovio, R. Tyce, and H. Schmidt, editors. *Autonomous Underwater Vehicle and Ocean Modelling Networks: GOATS 2000 Conference Proceedings.* SACLANTCEN Conference Proceedings CP-46. NATO SACLANT Undersea Research Centre, La Spezia, Italy, 2001.

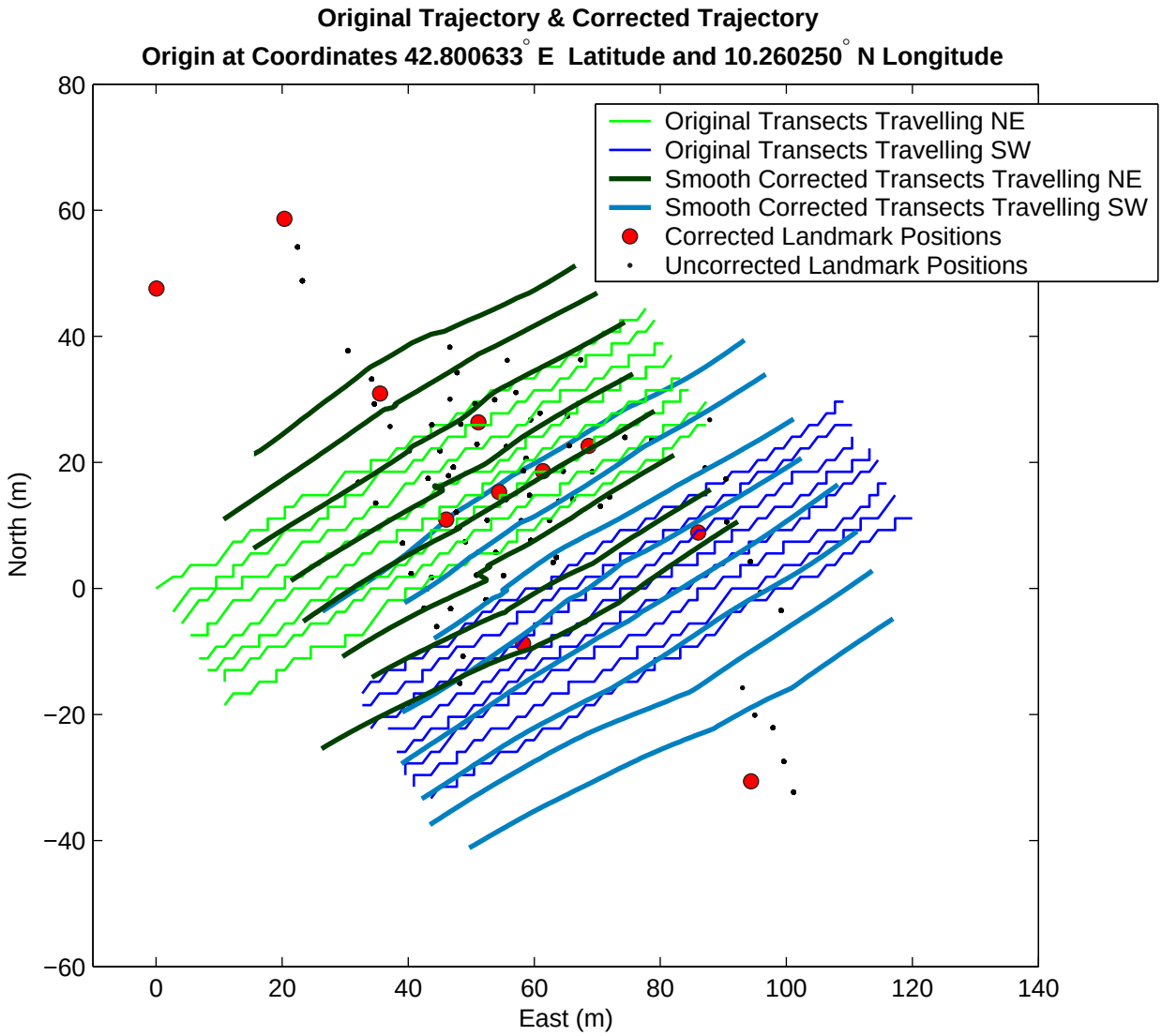

Figure 1: **Original Stored Trajectories and Smooth Corrected Trajectories.**

The image shows that the original absolute error was considerable. The stored vehicle trajectory had both along and across track errors. The CML-RTS strategy used the matched landmarks to correct the trajectories, the new landmark positions can also be seen in the image, the observations with respect to the original trajectories can also be seen.

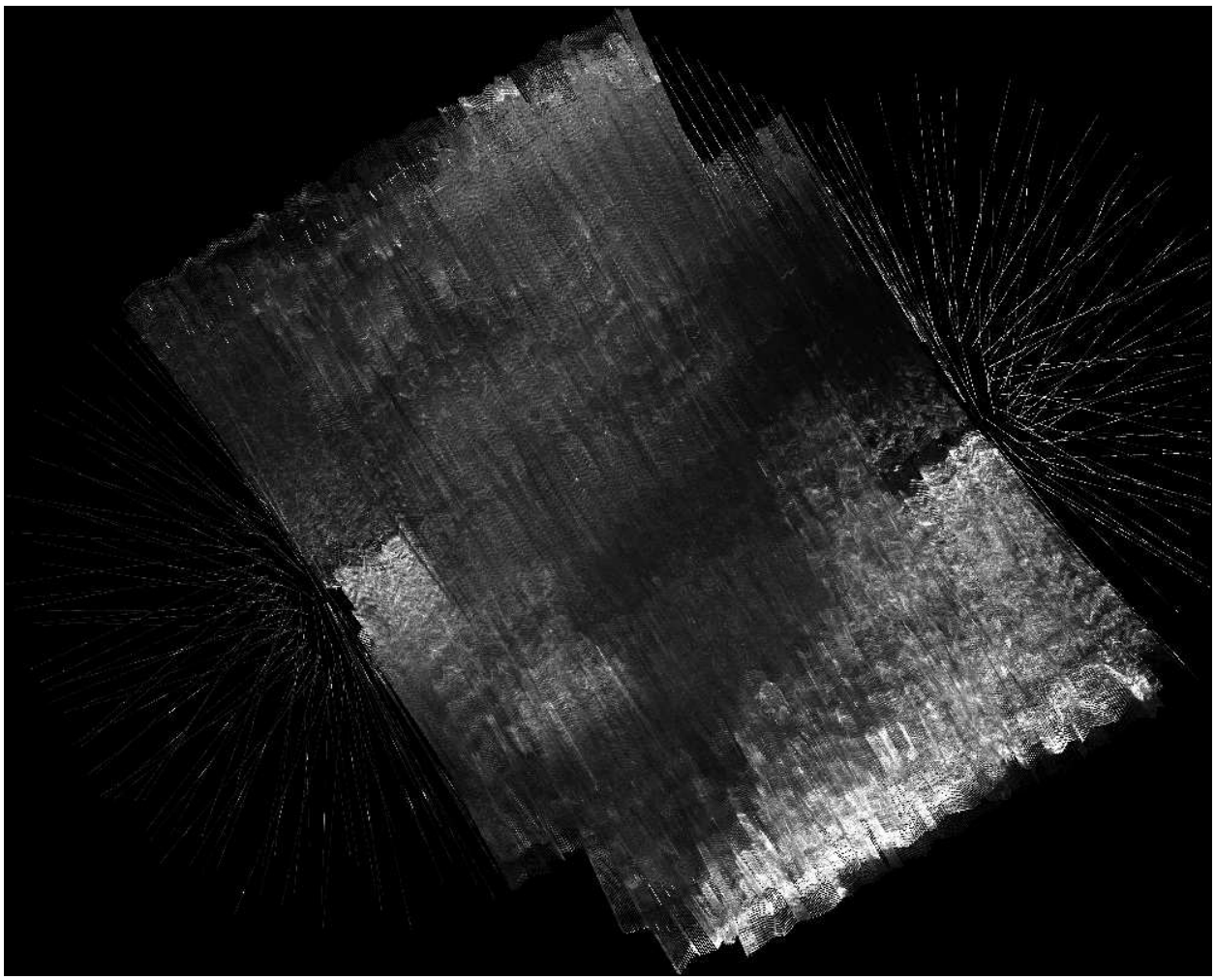

Figure 2: **Averaging of Transects Using Original Stored Trajectories.**

The mosaic has been created by geo-referencing each transect using the original stored trajectories. Collocated data has been averaged. Poor registration results in dilution and loss of data. The fan effects on the edges of the mosaic are the result of erroneous heading measurements.

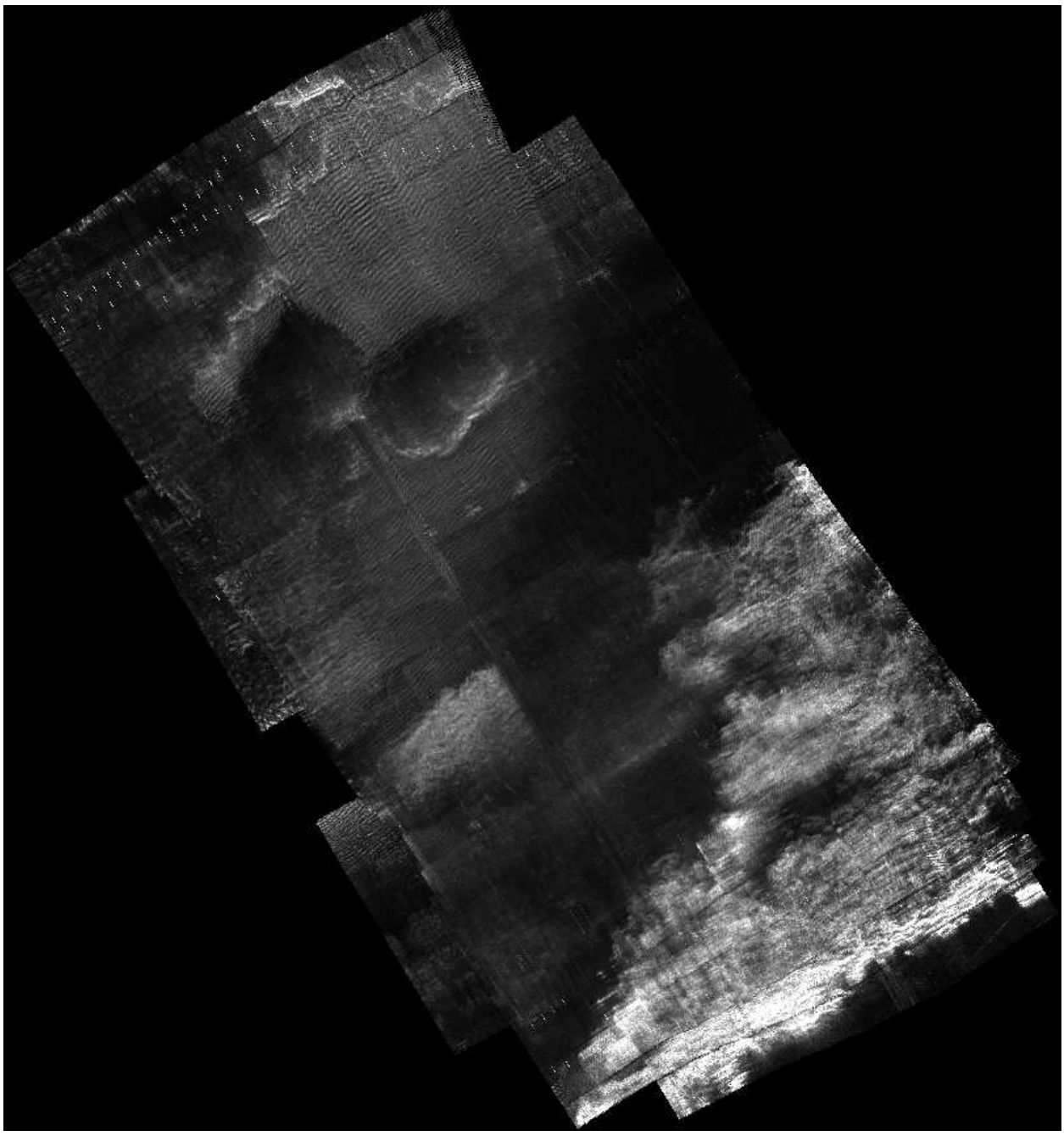

Figure 3: **Averaging of Transects Using Smooth Corrected Trajectories.**

The mosaic has been created by geo-referencing each transect using the corrected trajectories using the CML-RTS strategy. Collocated data has been averaged. Good registration produces a meaningful mosaic.

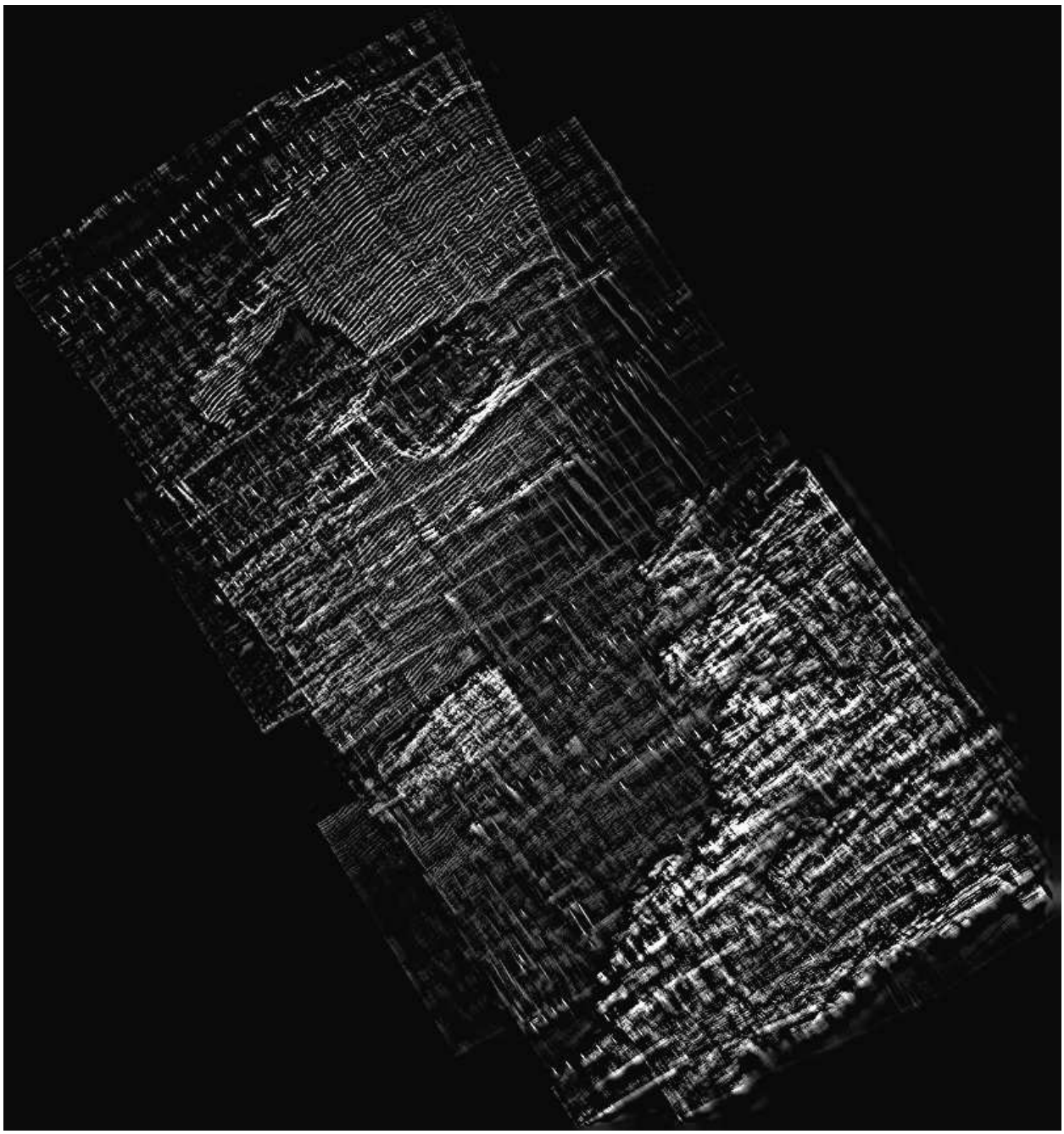

Figure 4: **Gabor fusion of registered transects.**

Result of the Gabor fusion of all the CML-RTS registered images. Every detail contained in the source images is preserved, but also are all noise and distortion artifacts.

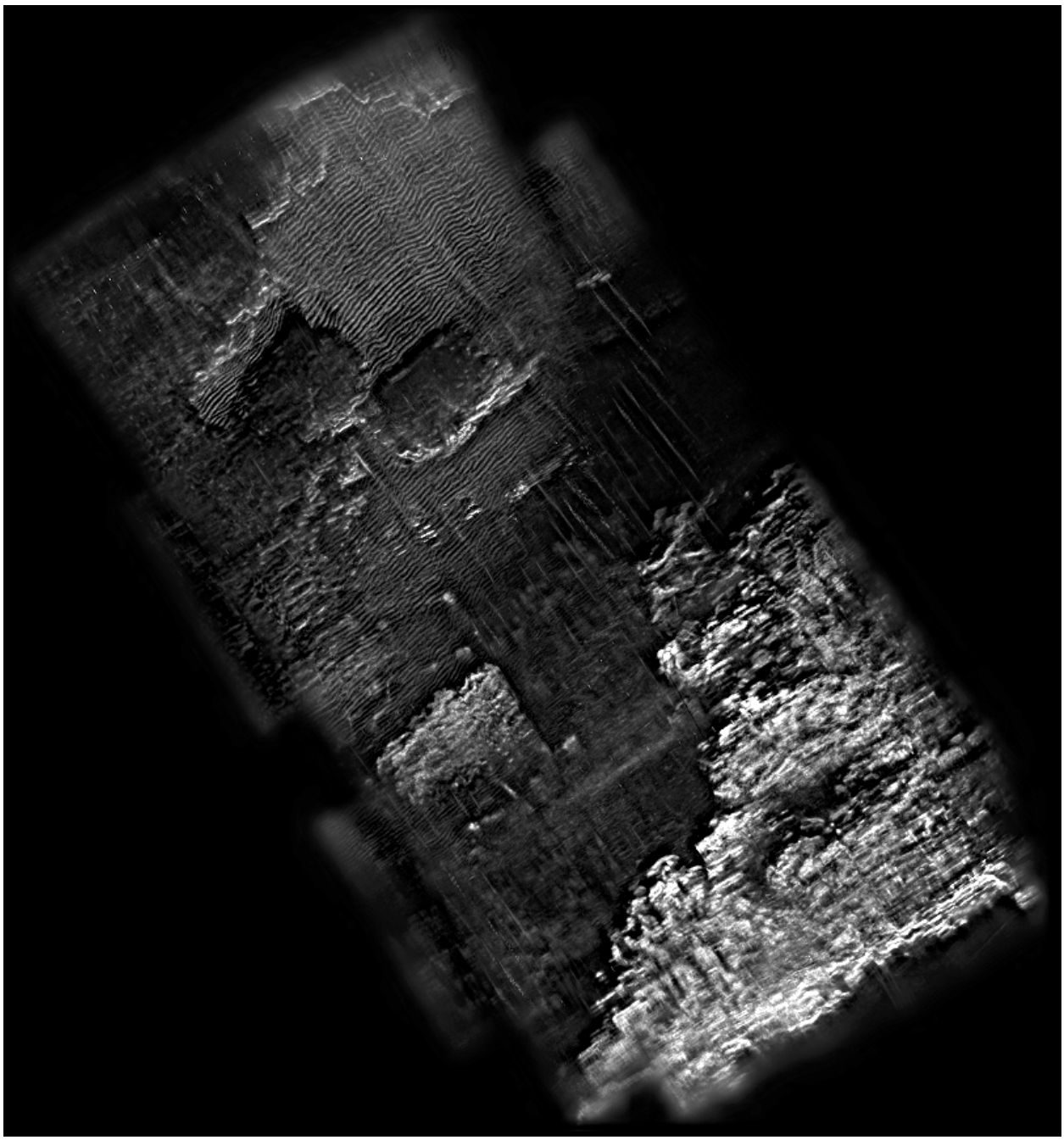

Figure 5: **Proposed fusion approach.**

The final mosaic is now constructed by Gabor decomposition of the geo-referenced images in their constituent spatial frequencies, which are then separately combined with equivalent priority given to all frequency bands. Note how all the original image details have been preserved while the presence of noise and distortion artifacts is diminished.

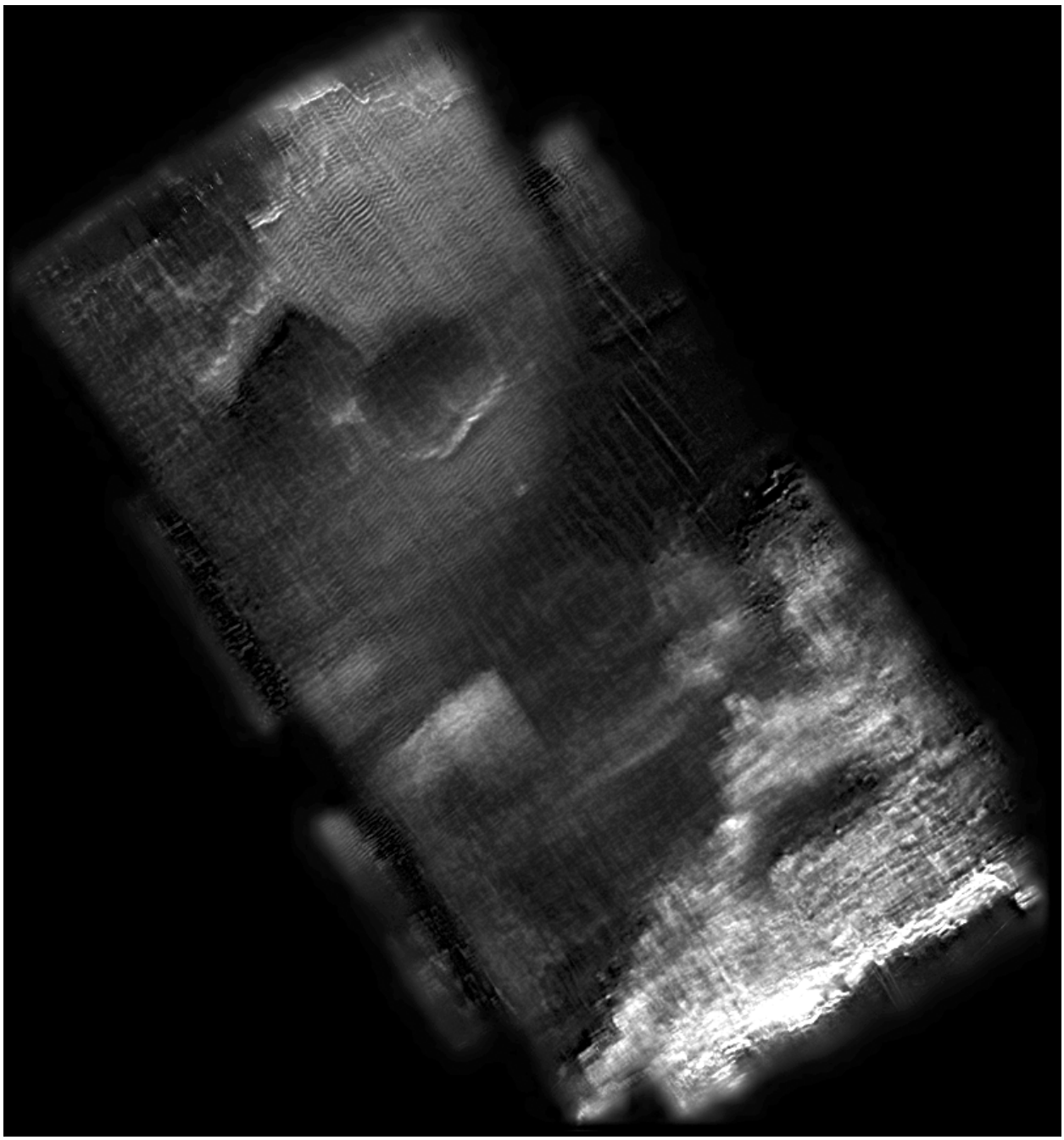

Figure 6: **Proposed fusion algorithm tuned for noise reduction.**

The proposed fusion algorithm is in this case adjusted to reduce the impact of sensor noise and image distortions in the final mosaic. Sand ripples and structures on dim areas are still present nevertheless.

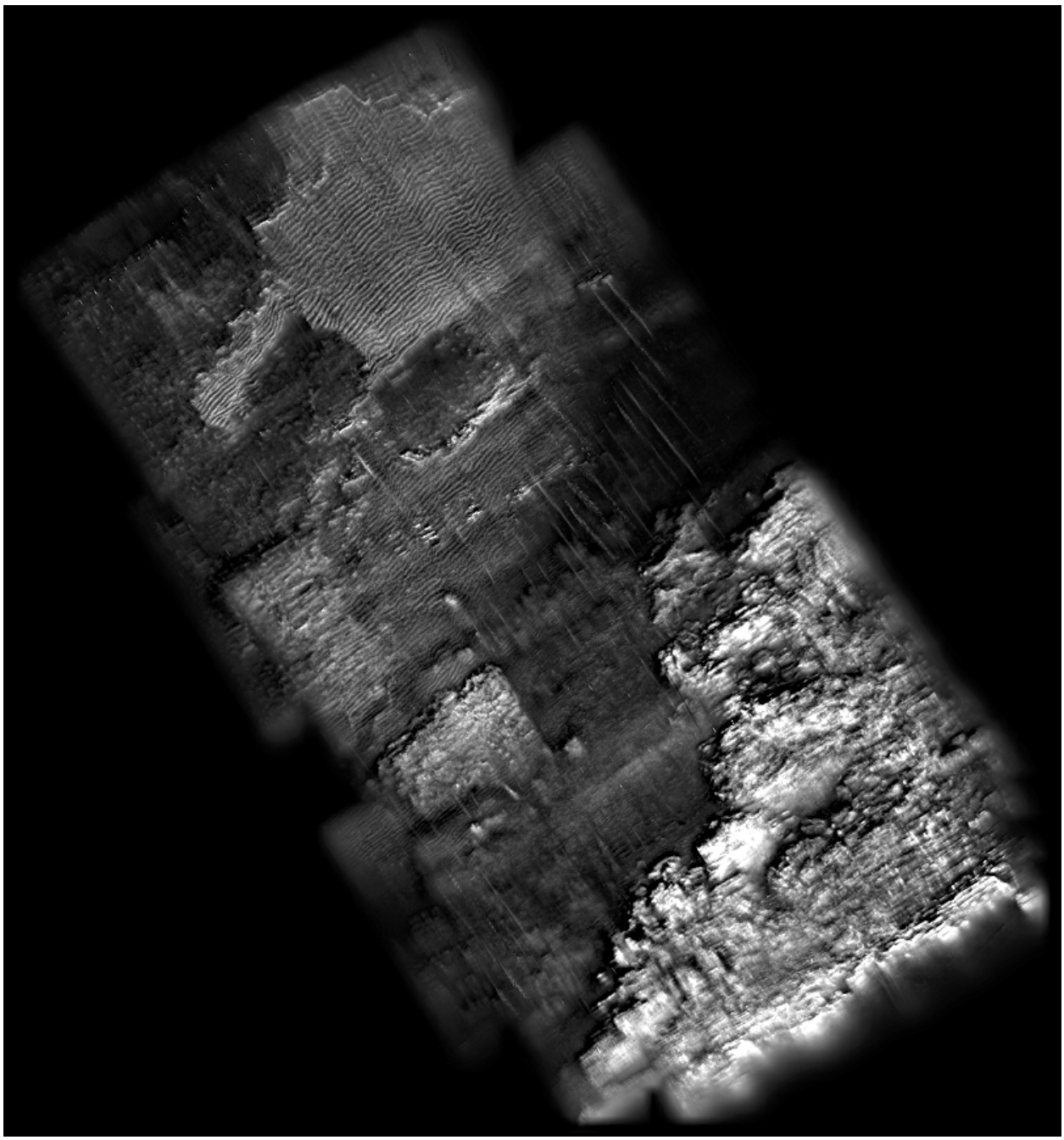

Figure 7: **Alternate settings for the fusion algorithm.**

The proposed fusion algorithm is now tuned to enhance middle-sized features. Note the rich structure of the seafloor in the central area of the mosaic, which is otherwise not preserved in the mosaicking result.