

Estimation of poly-metallic nodule coverage in benthic images

Timm Schoening, Biodata Mining Group, Faculty of Technology, Bielefeld University, Bielefeld, Germany

Thomas Kuhn, Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Hannover, Germany

Tim W. Nattkemper, Biodata Mining Group, Faculty of Technology, Bielefeld University, Bielefeld, Germany

ABSTRACT

Estimating the coverage of the seafloor by poly-metallic nodules from high-resolution photographs is a problem yet to be solved. Initial approaches to nodule density estimation have been done by sonar backscatter. This technique is capable of monitoring a large region but lacks a resolution sufficient to identify spots of high nodule density. Recently, image based monitoring has been proposed, which is able to display nodule density on much higher spatial resolution but introduces a bottleneck in data analysis since huge volumes of underwater video/image data have to be analyzed regarding the nodule coverage.

To move forward in the automated underwater image evaluation regarding a spatial quantification of nodule coverage, a pilot study was conducted on a set of benthic images taken in the Pacific Ocean.

Here we present a first prototype for a computational image analysis approach to quantify nodule coverage. The system first pre-processes the images to enhance color contrast and to reduce illumination artifacts. In a next step, a new artificial neural network approach is applied to map image sub-regions to coverage percentage values. A small subset of ten reference images is used to train the network and to tune the pre-processing parameters automatically.

We report a correlation of 0.95 between the expert's estimate and the automated approach for the training data and 0.99 for a small validation set.

INTRODUCTION

Vast regions of the Earth's oceans are yet unexplored. Vessels from several marine research institutions are currently exploring myriads of topics, including unknown species in the deep-sea, carbon sequestration and impacts of environmental change on habitats. Increasing scientific attentiveness also rises in the field of marine resources. While fishery has been around since millennia, oil and gas rigs since decades and even offshore wind parks are now operating, further unexplored resources exist, like methane hydrates and, as discussed in the context of this work, poly-metallic nodules.

Those nodules occur all around the globe, with the main, known hotspots on the abyssal plains in the Pacific and Indian Ocean. They lie on the sea floor and contain certain valuable metals like nickel, copper and manganese. It takes millions of years for the nodules to grow around a crystallization nucleus. Sources of the metals are the deep-sea sediment and the oceanic water column (Halbach et.al. 1988).

With increasing prices for several of the metals contained in the nodules, research in deep-sea mining of poly-metallic nodules has become more attractive for national institutes and consortia of private companies in the last years. Some countries have

thus obtained a license to explore a given region for a fixed amount of time from the International Seabed Authority. Germany is currently investigating the nodule occurrence in a region between the Clarion and Clipperton fracture zones in the eastern equatorial Pacific.

To assess the distribution of nodules, sonar backscatter has been applied on the whole German exploration area (Kuhn et.al. 2011). The results of this survey give an insight to the benthic conditions with a small resolution of 120 m per pixel. As an addition to this method, camera transects were captured with an Ocean Floor Observation System (OFOS). This technique has the advantage of a high resolution of about 0.5mm per Pixel, which is also its drawback since a full monitoring of the German exploration area would create about 800 Petabyte of data (150 million Gigabytes, more than five times the amount of data produced per year by the Large Hadron Collider). The imaging approach is thus used to take regional snapshots of nodule coverage to correlate the sonar backscatter data with high-resolution measures.

Measuring the coverage within the captured images has not yet been automated, similar to other fields of underwater imaging that have shown to be difficult to assess without human interaction.

MATERIALS AND METHODS

Images

In this study, two image transects from the eastern German license area are investigated, taken during an expedition with R/V Sonne in 2010 (Rühlemann et.al. 2010). The images were taken by the video sledge of an OFOS that was towed about 2m above the sea floor. The images have a footprint of 1.3 to 5.3 m² depending on the varying OFOS height and thus show different illumination settings from being dark blue when the OFOS is far away, to bright yellow when it is close to the sea floor. The footprint was computed by detecting the three laser points emitted from the OFOS and evaluating their positions relative to each other automatically.

Each image is 4224x2376 pixels in size and consists of three channels (Red, Green, Blue: RGB) with 8 Bit per channel. Due to strong illumination artifacts we limited the coverage estimation to a large central region of the images with 1688x792 pixel size in this initial study. Future work will aim at increasing the size of this region by improving the pre-processing and efforts towards stabilizing the camera-seafloor distance.

To develop and evaluate our algorithmic approach we selected ten example images (referred to as reference images) from the transects and gave those to one expert who quantified the coverage manually (see Figure 6). To this end, the images were fed into our online underwater image database system BIIGLE (Ontrup et.al. 2009). The expert inspected the images via an Internet browser and selected for each sub-region on the image a percentage value in steps of 10%.

Illumination correction

To enable an automated computational analysis of all images with the same setup (i.e. with one set of image processing parameters), we pre-processed the images to correct the varying illumination conditions. As all images feature a lightness falloff towards the corners, we subtracted a Gaussian filtered version of the image. The filter had a very large kernel of 701x701 pixels and the filtered image is thus only a

representation of the illumination cone. This procedure equalizes the illumination within one image.

To equalize the color contrast within all images of the transects, a histogram equalization was applied. Thereby the peak of the color histogram of each image was shifted to the center of the color scale. The complete method is explained in more detail in Schoening et.al. (2012).

Feature transformation

From the pre-processed images, we computed a color feature representation for each pixel. This feature representation consists of a color histogram. To compute the color features for one pixel p , a 7×7 pixel neighborhood $N(p)$ is considered. For each of the RGB channels, a 4 bit binning is applied, and the 3×16 bin counts are fused to a 48-dimensional feature vector $x(p)$ of the pixel p . The feature vectors of a large number of points (for instance from all pixels of the ten reference images) constitute a structure in a 48-dimensional space, which can be modeled using methods from machine learning and artificial neural networks regarding structural features (like clusters) and hidden regularities (e.g. cluster-specific features such as coverage percentage).

Machine Learning

For this study, a Hyperbolic Self-Organizing Map (HSOM) is applied which is an artificial neural network algorithm for data clustering and dimensionality reduction (Ontrup and Ritter 2001).

We randomly picked 20% of the feature vectors of five of the ten reference images for the training of the HSOM (which resembles less than 0.05% of all available data). The trained HSOM learned 161 clusters from this data, represented by prototypes $u^{(m)}$, ($m = 0, \dots, 160$) applying an Euclidean metric to compute distances in the feature space.

The trained HSOM was then used to identify the best-matching units of all other feature vectors from all images. This way each pixel p is assigned to one integer m by first mapping the pixel p to its feature vector $x(p)$ and mapping this feature vector to its best matching (i.e. most similar) cluster m with its prototype $u^{(m)}$. In a next step, the cluster indices m of all pixels in one grid cell C_j ($j = 1, \dots, G$; G : amount of grid cells in one image) are fused to one new feature vector $v(C_j)$, representing the prototype distribution within a grid cell. In other words, the HSOM clustering quantized 48-dimensional color features and projects them to a one-dimensional index, which is used to compute a new 161-bin histogram for this new index. Figure 1 shows two cells and their corresponding histograms.

As the HSOM features a hyperbolic geometry, its prototypes can be mapped onto the maximum-illumination disc of the Hue-Saturation-Value (HSV) color space. Thus each prototype is assigned a distinct color. Replacing each pixel's color values with the HSV color of its best-matching unit, the image can be transformed into a pseudo-color image that visualizes the clustering outcome.

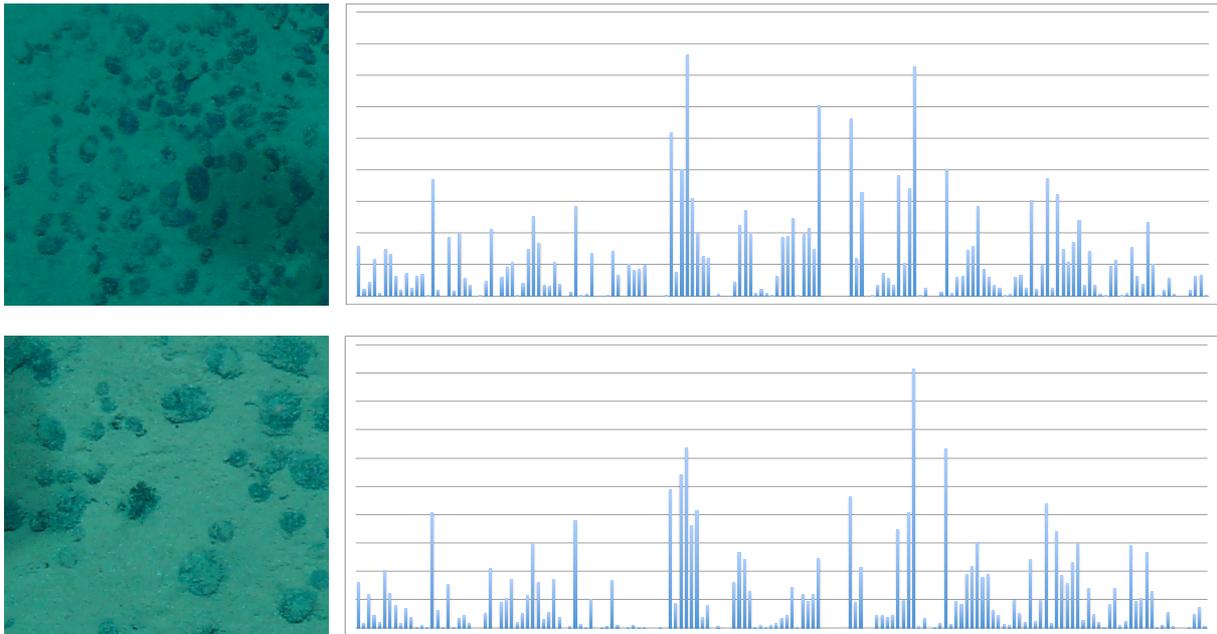


Figure 1: Two cells with different coverage and nodules of different size. Next to them are the histograms of the prototypes occurring within the cell. This histogram is a bag-of-prototypes and serves as a representation of the cell for further processing.

To map the grid cells feature vectors $v(C_j)$ to a coverage percentage value, a Principal Component Analysis (PCA) is conducted with the covariance matrix of all cells $\{v(C_j)\}$. The feature vectors are projected onto the first two principal components, reducing their dimensionality to 2.

Visualizing the resulting tuples of all cells in a scatterplot shows a continuous distribution of cells along an axis of increasing nodule coverage. This allows the relative classification of a cell according to cells with similar coverage.

Quantification

For a quantification of a cell's coverage, a manual annotation of some cells is necessary. As mentioned above, an expert of the field thus manually evaluated 80 cells, taken from the reference images using the BIIGLE system, assigning a coverage estimate of $i \times 10\%$ to each cell (i in $[0..10]$).

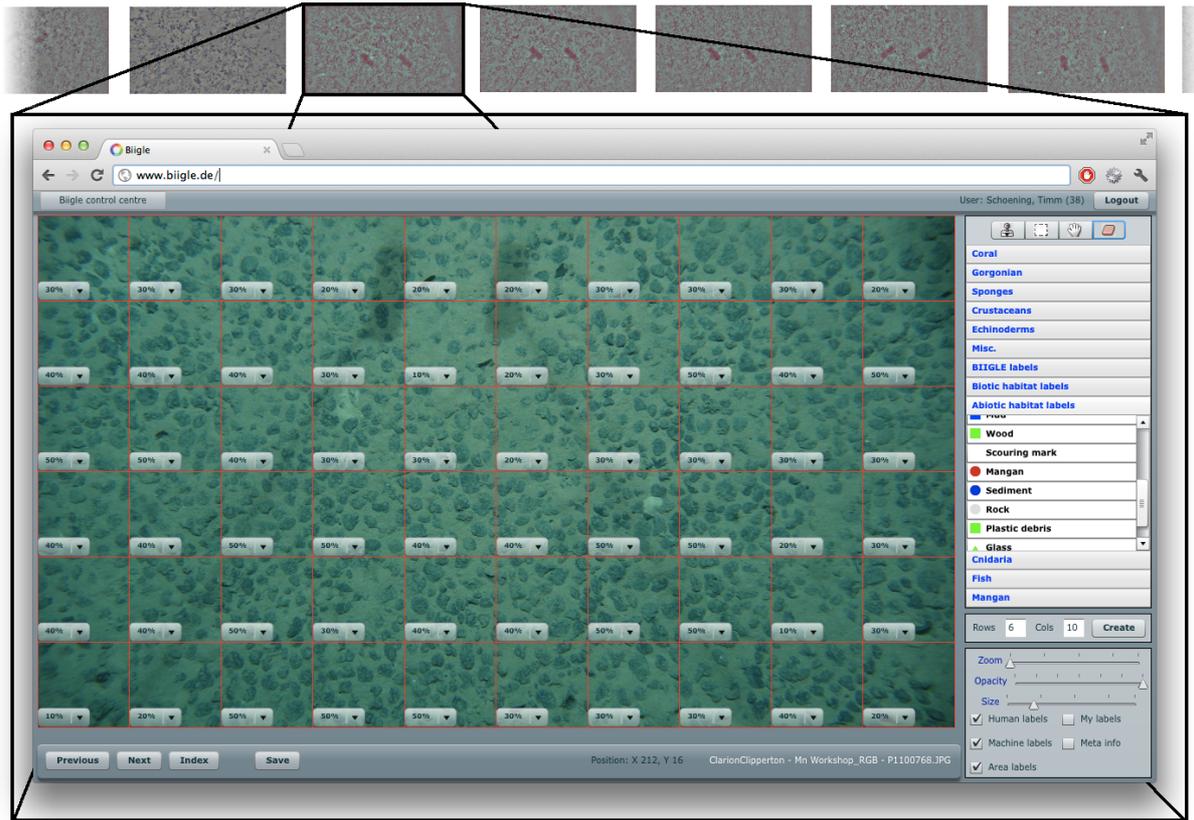


Figure 2: Only a small set of images has to be manually annotated for a qualitative coverage estimate (here: 10 images out of 2100). This can be done with the Grid annotation tool of BIIGLE in 1 to 5 minutes per image.

Image cells, which have not been quantified manually, i.e., the coverage of which is unknown, are processed accordingly. First, the prototype frequencies are counted to compute $v(C_i)$, which is projected onto the first two principal components. The projection result is assigned to that $i \times 10\%$ value of the closest annotated cell. This approach is also referred to as case-based reasoning.

RESULTS

The original images featured a wide variety of color spectra. One example is shown in Figure 3. The pre-processed images show increased brightness in the corners, more obvious laser points and higher color contrast between nodules and sediment as shown in Figure 4. Figure 5 shows the HSV pseudo color image of the clustering result.

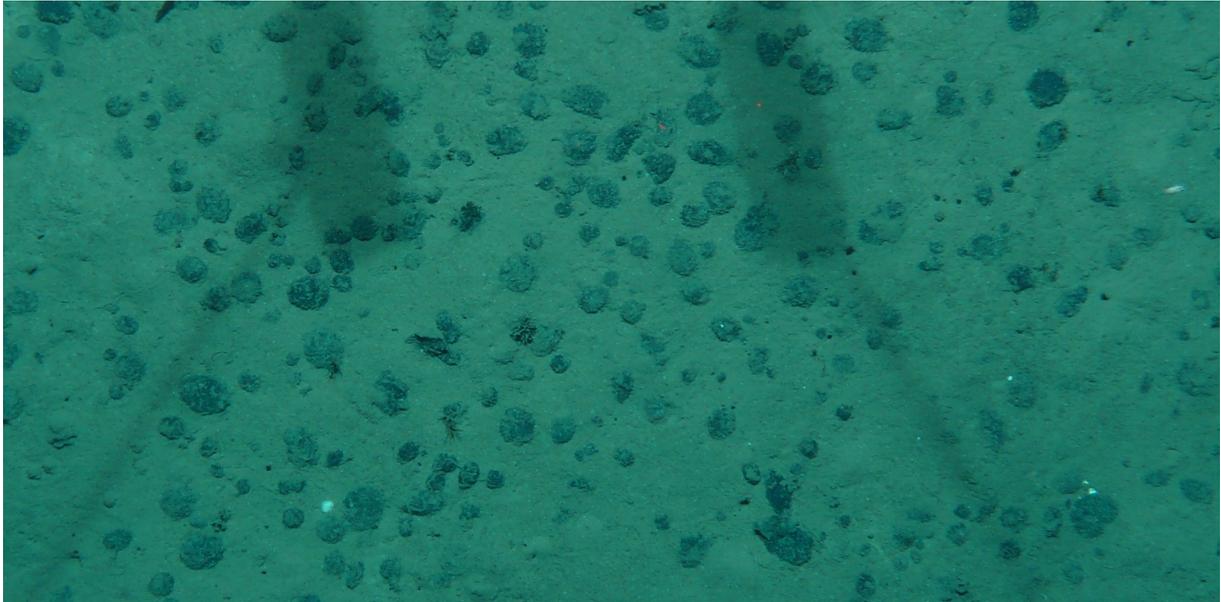


Figure 3: An example of a seafloor image with poly-metallic nodules. Shown is the center region of an image as it is used for further processing.

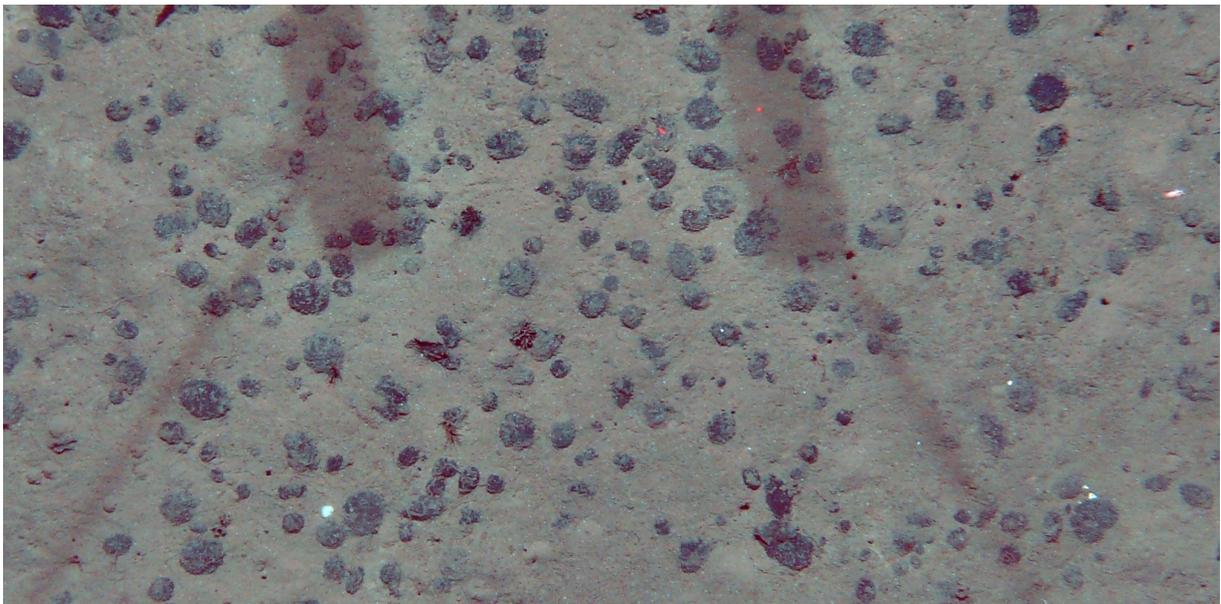


Figure 4: Pre-processed version of the image in Figure x. The nodules appear darker; the sediment was shifted towards higher values of red, making it appear brown/ochre. Also, the laser points are easier to perceive.

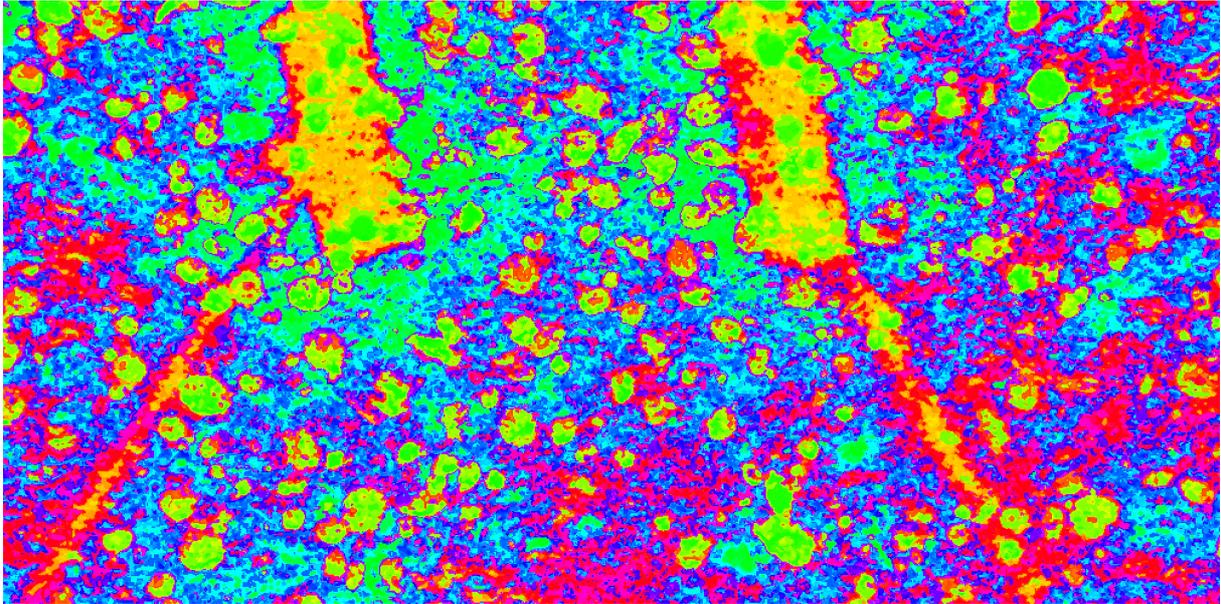


Figure 5: The clustering outcome of the HSOM as a pseudo color image. The input image was the one shown in Figure 3. Each pixel is colored according to the position of its best-matching unit on the HSV disc. The nodules appear primarily in green, the background (i.e. sediment and shadow of the OFOS pilot weight) appear in red, purple, blue and turquoise.

Figure 6 shows two sample annotations. Two benthic images are shown together with an overlay of the expert's coverage estimate. The coverage is encoded by gray value, with the lowest coverage shown in black and the highest coverage shown in white.

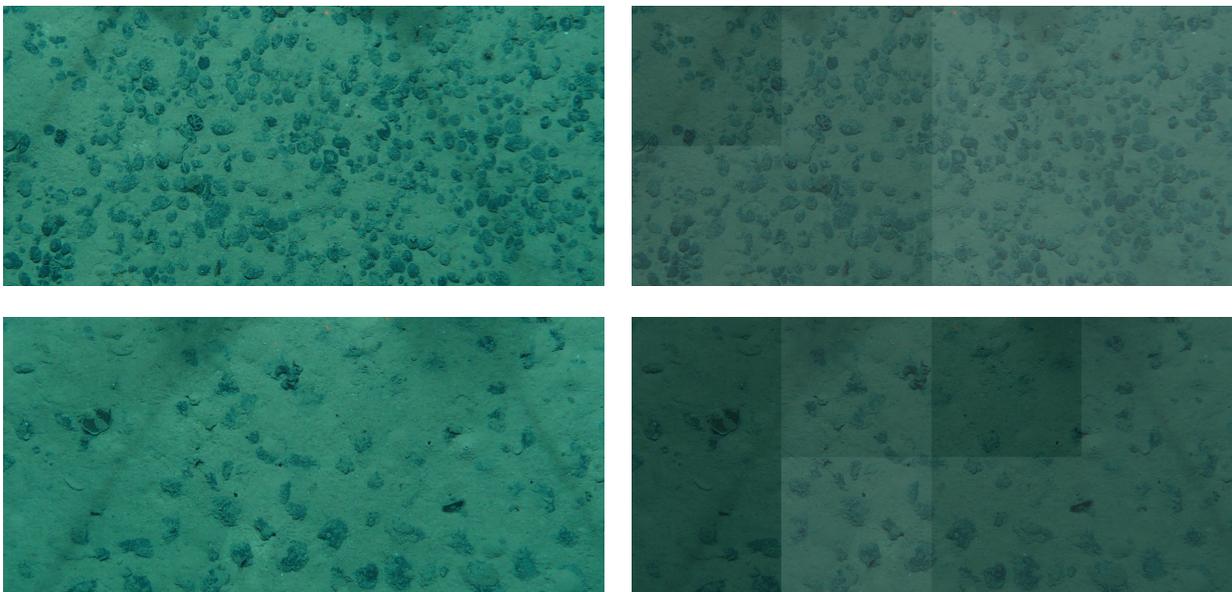


Figure 6: Two benthic images with different coverage and nodules of different shape and size are shown on the left. On the right are the same images with an overlay of the human experts coverage estimate. In the upper image, the coverage ranges from 20% (upper left cell) to 40% (right half of the image). In the lower image, the coverage ranges from 10% (two left-most cells) to 30% (lower row, second from left tile)

Figure 7 shows a scatterplot of the tuples of all cells, aligned by their increasing coverage along the x-axis. This axis depicts the second most important principal component.

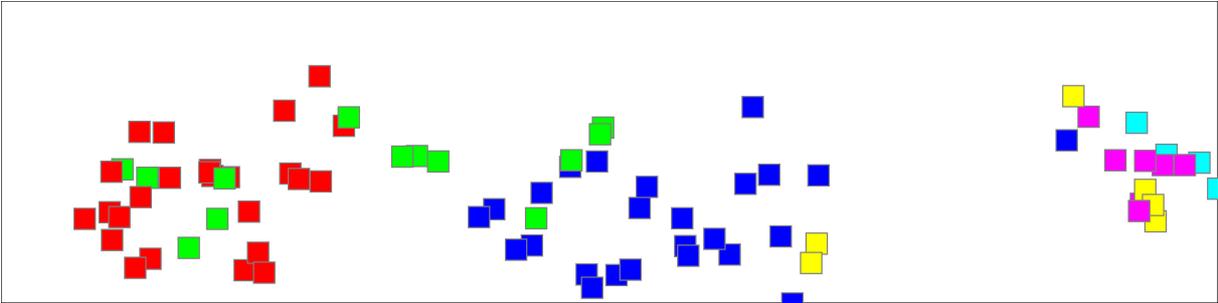


Figure 7: Scatterplot of all hand labeled cells. The colors encode the experts coverage estimate and are as follows: red 0%, green 10%, blue 20%, yellow 30%, purple 40%, turquoise 50%. The cells are aligned primarily along the x-axis. This axis represents the second most important principal component.

Re-evaluation of the training data showed a correlation of 0.96 between the human expert annotation and the bag-of-prototypes estimate. Figure 8 shows the amount of cells in each $i \times 10\%$ group for the expert as well as the automated estimate. There were no groups with a coverage of more than 50%, neither for the human expert nor for the automated results.

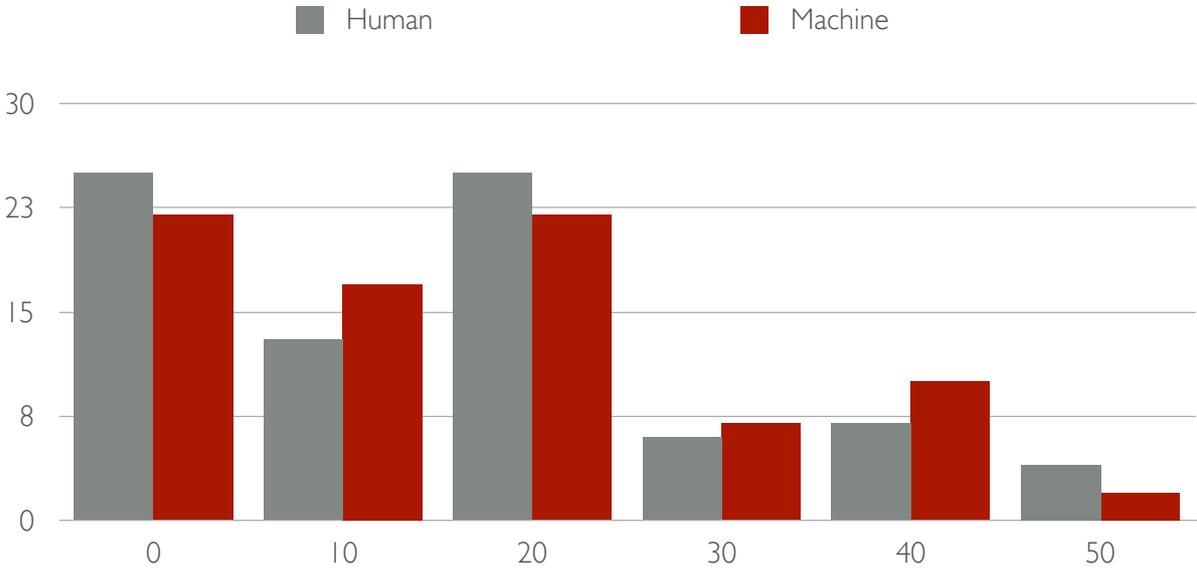


Figure 8: Histogram of the amount of cells of varying coverage for the training cells. The gray bars show the human experts estimate, the red bars show the automated estimate.

To validate the process, the coverage of 24 further cells was estimated with the bag-of-prototypes approach. Those cells have not been used for the training of the HSOM or the PCA and were also hand annotated by a human expert. The correlation between the human and machine estimates for the validation cells was 0.99. There were no cells with a coverage higher than 30%. The according group distribution is

shown in Figure 9, the scatterplot of the tuples, computed by the existing PCA, is shown in Figure 10.

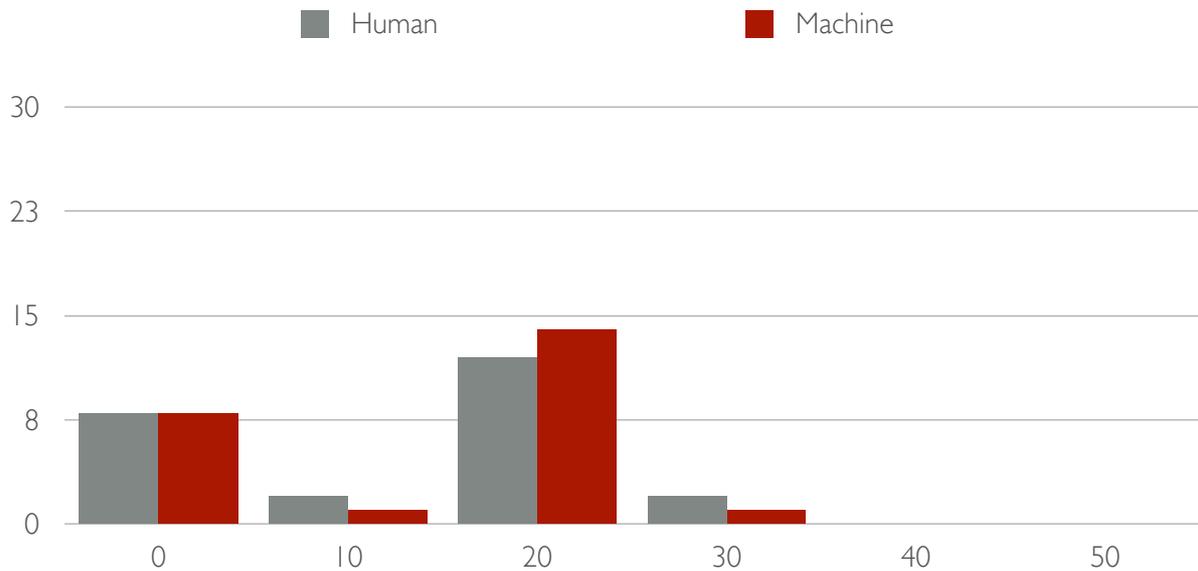


Figure 9: Histogram of the amount of cells of varying coverage for the validation cells. The gray bars show the human experts estimate, the red bars show the automated estimate.

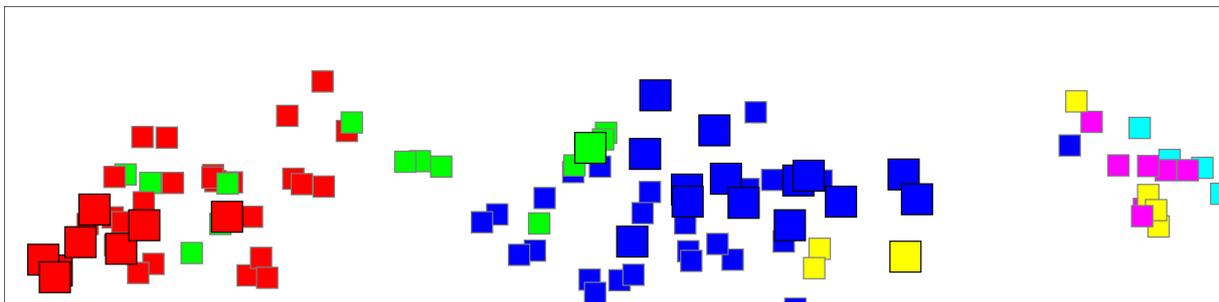


Figure 10: Scatterplot of all hand labeled training cells (small boxes) together with the validation cells (large boxes). The colors and axes are as in Figure 5.

Computation time

The pre-processing of an image takes about 3 minutes, since the computation of the Gaussian filtered image is time-consuming. The feature extraction takes less than one minute and expands the data per image from about 4.5MB to about 60MB. The training of the HSOM takes about 1 hour, but has to be done only once, as well as the PCA that takes less than a minute. The bagging of prototypes takes less than a minute per image. When a new image is recorded, the coverage estimate is computed within 3 to 5 minutes. All measures are averages for the execution times on a single-core machine.

DISCUSSION

This pilot study has shown, that the proposed approach of neural network-based classification of grid cell feature compositions is capable to qualitatively estimate the nodule coverage of a benthic image in steps of $i \times 10\%$. Nevertheless, the data basis

is still sparse, as only images from one region were evaluated. Also, the amount of validation cells is still too low to make a more significant estimate of the system performance on data from numerous large OFOS explorations. For a larger validation dataset, further expert annotations are required which are currently created for a follow up study.

The approach is currently dependent on the expert annotations by means of the case-based reasoning. A model of the point distribution of tuples within the PCA projection scatterplot could be used to derive a tessellation of the 2D plane for a more stable assignation of unseen cells to their coverage group. Also, a qualitative measure could be introduced according to the x-value of the tuple by which a continuous coverage value would be assigned to a cell instead of a coverage group. This would require more detailed expert annotations and could thus not yet be developed as the gold standard to test for is missing.

During the correlation of the human and machine results arose the problem of how trustworthy the expert results are. Within the scatterplot of the training data (Figure 5), the clusters of coverage groups overlap and visual inspection of those cells, lying considerably within a cluster dissimilar to their own type suggests, that they might have been erroneously annotated. The annotation by one or more further experts could give more detailed information about this issue.

In principal, the complete approach could be performed by the OFOS itself. Therefore, the first issue would be to speed-up the time-consuming Gaussian filtering of the illumination correction step considerably, e.g. by a hardware solution. This would allow the estimation of the nodule coverage without recording as much data as required for a digital image. An autonomous submersible could be launched, programmed to monitor a vast region of seafloor independently, recording only the nodule coverage. Thus the drawback of the high resolution would become an advantage.

CONCLUSION

We presented our initial efforts in quantifying the coverage of the deep seafloor with poly-metallic nodules. A pre-processing of the initial data has been proposed together with an unsupervised machine-learning procedure to gather a numeric estimate of the coverage.

The bagging of prototypes is a novel approach in data mining that allows the purely data-driven classification of sub-regions within images.

ACKNOWLEDGEMENTS

This work was supported by the Bundesanstalt für Geowissenschaften und Rohstoffe (BG), D-30655 Hannover, Germany. We thank Samy Slaih for the development of the grid cell annotation software and his help with the data management for BIIGLE.

REFERENCES

Halbach, P., Friedrich, G., von Stackelberg, U. 1988: *The Manganese Nodule Belt of the Pacific Ocean*. Enke-Verlag, Stuttgart, 284pp.

Kuhn, T., Rühlemann, C., Wiedicke-Hombach, M. 2011. *Development of Methods and Equipment for the Exploration of Manganese Nodules in the German License*

Area in the Central Equatorial Pacific, Proc. Ninth (2011) ISOPE Ocean Mining Symposium, 174-177.

Ontrup, J., Ehnert, N., Bergmann, M., Nattkemper, T.W. 2009 *Biigle - Web 2.0 enabled labelling and exploring of images from the Arctic deep-sea observatory HAUSGARTEN* IEEE Oceans 2009, pages 1-7

Ontrup, J., Ritter, H. 2001 *Hyperbolic self-organizing maps for semantic navigation* Advances in neural information processing systems volume 14, pages 1417-1424

Schoening, T., Bergmann, M., Ontrup, J., Taylor, J., Dannheim, J., Gutt, J., Purser, A., Nattkemper, T.W. *Semi-automated image analysis for the assessment of megafaunal densities at the Arctic deep-sea observatory HAUSGARTEN* (PLoS One, Accepted)

Rühlemann, C., Shipboard Scientific Party, 2010. *Cruise Report SO-205 MANGAN*. Bundesanstalt für Geowissenschaften und Rohstoffe, Hannover, 112 pp.