ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/278704818

Self-Organizing Artificial Neural Networks into Hydrographic Big Data Reduction Process

Conference Paper *in* Lecture Notes in Computer Science · July 2014

DOI: 10.1007/978-3-319-08729-0_34

CITATIONS
52

READS
662

2 authors:

Andrzej Stateczny
Gdansk University of Technology

99 PUBLICATIONS **1,211** CITATIONS

SEE PROFILE

Marta Wlodarczyk-Sielicka
Maritime University of Szczecin

31 PUBLICATIONS **339** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project SHREC - Automatic recognition and identification system for ships in video survelience areas View project

Project Sensors Special Issue - Security and Information Flow in Intelligent Systems for the Internet of Things View project

All content following this page was uploaded by Andrzej Stateczny on 10 August 2015.

The user has requested enhancement of the downloaded file.

Self-organizing Artificial Neural Networks into Hydrographic Big Data Reduction Process

Andrzej Stateczny[1] and Marta Wlodarczyk-Sielicka[2]

[1] Marine Technology Ltd., Szczecin, Poland
a.stateczny@marinetechnology.pl
[2] Maritime University, Szczecin, Poland
m.wlodarczyk@am.szczecin.pl

Abstract. The article presents the reduction problems of hydrographic big data for the needs of gathering sound information for Navigation Electronic Chart (ENC) production. For the article purposes, data was used from an interferometric sonar, which is a modification of a multi-beam sonar. Data reduction is a procedure meant to reduce the size of the data set, in order to make them easier and more effective for the purposes of the analysis. The authors` aim is to examine whether artificial neural networks can be used for clustering data in the resultant algorithm. Proposed solution based on Kohonen network is tested and described. Experimental results of investigation of optimal network configuration are presented.

Keywords: Kohonen network, big data, hydrography.

1 Introduction

The main component that contributes significantly to the safety of navigation is the information on the depth of the area. A navigational chart is the primary source of information for the navigator. The information on depth is the one of the most important layer of Electronic Navigational Chart (ENC) according to International Hydrographic Organization (IHO) standards. Although in Inland ENC there is no obligation to provide bathymetric data but it is often provided by electronic chart producers. Some aspects of electronic charts for inland and maritime navigation, chart production planning and navigational data evaluation were presented in [1,2,3,4,5].

Usually bathymetric data are gathered by multibeam echosounder (MBES), which uses acoustic waves is a device for bathymetric measurements and it measures the vertical distance between the head and the bottom or an object located at the bottom. For the purposes of the article, data was used from an interferometric sonar, which is a modification of a multi-beam echo sounder. During survey process large amount of data was collected. Data reduction is a procedure meant to reduce the size of the data set, in order to make them easier and more effective for the purposes of the analysis. There are several ways to perform data reduction. One of them is to transform a large quantity of variables into a single, common value. Another way to reduce data is to

M. Kryszkiewicz et al. (Eds.): RSEISP 2014, LNAI 8537, pp. 335–342, 2014.
© Springer International Publishing Switzerland 2014

use advanced statistical methods, that will make it possible to decrease the size of a data pack by breaking it down into basic factors, dimensions or concentrations, pinpointing the basic relations between the analyzed instances and variables. Another method is to deduct given number of instances from a large array, while maintaining its overall suitability for the analyzed population. The problem of bathymetric big data processing has been taken up by several authors [6,7,8,9,10,11,12].

Modern idea to solve computational problems with large amount of data processing is to use artificial intelligence methods, especially artificial neural networks (ANN). Some aspects of ANN solutions for navigation purpose are described in [13,14,15,16,17]. Very interesting idea is to process big data for compression or reduction purpose by using self-organized ANN. Especially useful to data reduction are self-organized Kohonen network [18,19,20,21,22,23].

Kohonen network idea was described in many articles and books [24,25]. Kohonen networks are an example of ANN taught through the self-organization. The idea of their functioning consists on self-adapting state of the network to present entrance data. In the process of the network's learning there is no association between input signals and output of the network. The learning consists in the specific concept of adapting of networks passed to images for their input. In case of Kohonen networks a competition between neurons is a philosophy for the value of its weights updating. During the competition between neurons the neuron closest to the input sources in the meaning of chosen distance method calculation is a winner for input data set. Euclidean distance is usually used. The degree of adjustment depends on the distance of the neuron from the input data.

In the article Kohonen networks are examined and offered as a step in the process of reduction of hydrographic big data.

2 Hydrographic Big Data Reduction Problem in Electronic Navigation Chart Production Process

One of the issues connected to bathymetric measurements is registering an large amount of the data, as well as various types of interference. The echo sounder is a device used to measure the vertical distance between the head of the equipment and the sea bottom, or an object located on the sea bed, using a acoustic wave. The establishment the depth of water is achieved by measuring the time in which the acoustic wave needs to reach the object, as well as to return to the receiving transducer as a reflected wave. The angle measurement is carried out in various manners, depending on the type of the echo sounder. The simplest example is the single beam sonar, which operates by sending a narrow acoustic signal beam vertically downwards. The MBES emits several signal beams in multiple directions, monitoring in all these directions. This solution allows for a much wider area of measurement, when compared to a single beam sonar, by increasing the width of the scanning zone. For collecting bathymetric data authors used the interferometric sonar system GeoSwath Plus 250kHz, which is able to gather high density data. It allows to simultaneously collect vertical data (like a standard multi-beam sonar), as well as horizontal data (like a side scan sonar). The depth data is received not only based on the measurement of time in which the acoustic wave reflected off the object returns to the receiving transducer, but also based on measuring the difference between phases of the wave reaching the piezoelectric sensors installed within the head. Bathymetric data processing is realized in several stages. First off, all values of corrections influencing the accuracy of the measurements which are taken into consideration, such as: water properties, head submersion, errors in the average speed of sound in water and erroneous offset inputs in the measurement devices. Subsequently, the system operator performs the initial rough filtration, using predefined data processing filters. The next step is to process data. The first stage of processing is to convert the raw data into 'swath' files. There are four basic types of filters used to process the data: the amplitude filter, the limit filter, the across track filter and the along track filter. If properly applied, the combination of these filters guarantees the optimal size and content of the files [26]. The authors` main purpose is to create a data reduction algorithm in aspects of production electronic navigational chart. The whole process is shown in the figure below.

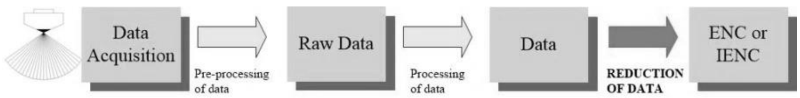

Fig. 1. Processing of bathymetric data

It is worth noting if the filtered samples are big amount of data sets. Data reduction is a procedure meant to reduce the size of the data set. Generally the hydrographic systems generate the grid with use the following methods: mean (select a mean depth value) or weighted mean (uses amplitude values to give higher weighting to data points which are higher in amplitude when calculating the mean depth value).

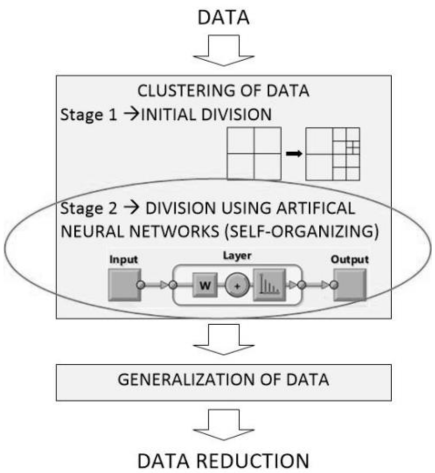

Fig. 2. Reduction algorithm of bathymetric data

The main objective of authors is compilation of reduction algorithm. The clustering of data is first part of search algorithm, as shown in figure above. The next step will be generalization of bathymetric data. The aim of article is to examine whether artificial neural networks can be used for clustering data in the resultant algorithm.

3 Experimental Results of Kohonen Network for Hydrographic Data Reduction

For researches Matlab software with the Neural Network toolbox was used. For the purpose of clustering problems, the self-organizing feature map (SOFM) was applied. SOFM has one layer with neurons organized in a grid. Self-organizing maps learn both the distribution and topology of the input vectors they are trained on. During the training, network applies the rule WTM (Winner Take Most) according to which the weight vector associated with each neuron moves to become the centre of a cluster of input vectors. In addition, neurons that are adjacent to each other in the topology are also moved close to each other in the input space.

The floating laboratory Hydrograf XXI, with the GeoSwath Plus 250 kHz sonar and supplementary equipment (GPS/RTK, satellite compass, motion sensor) installed was used . The measurement profiles were carried out due to maintain 100% coverage of the measured body of water. Test data was collected within the Szczecin Harbour, near the Babina Canal, on May 23, 2010.

Test area included very high density data , which is the main limitation of neural network usage. It is impossible to train the network from whole data set using standard computer. In order to solve this problem, authors divide original data points set into smaller subsets, which could be trained separately. Training data set is the square measuring 25 to 25 metres and it includes 28911 samples of 3 elements (X,Y,Z), as shown on figure 3.

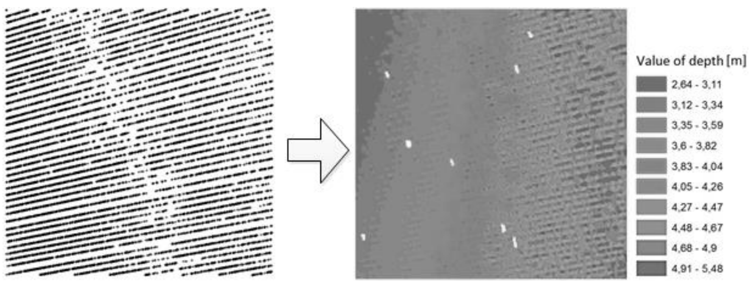

Fig. 3. Test data on the area 25x25 metres

While network was being creating, the numbers of rows and columns were specified in the grid. During tests, these values were set to 3x3, 7x7 and 10x10 which provided the number of neurons 9, 49 and 100, respectively. Authors decided to select the hexagonal topology of network, where each of the hexagons represents a neuron. Distances between neurons are calculated from their positions with a distance function. The most common link distances were selected in those researches: the number of links, or steps, that must be taken to get to the neuron under consideration. Initial neighbourhood size was set at 3 and ordered phase steps was set at 100. The training was running for the selected number of epochs: 10, 50, 100, 200, 500 and 1000.

There are many visualization tools that can be used to analyse the resulting cluster. One visualization for SOFM is the weight distance matrix presented on figure 4.

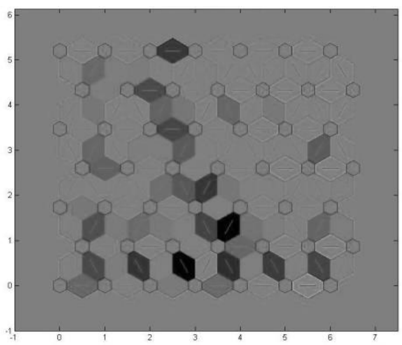

Fig. 4. Weight distance matrix for 49 neurons and 200 epochs

The grey hexagons illustrate the neurons and the red lines connect the neighbouring. The colours in the regions contain the red lines designate the distance between the received neurons: the lighter colour indicates smaller distance and darker colours indicates bigger distance.

Next type of presentation of the results is illustrating the locations of neurons in the topology and it indicates numbers of the bathymetric data associated with each of the cluster centres (Fig.5.).

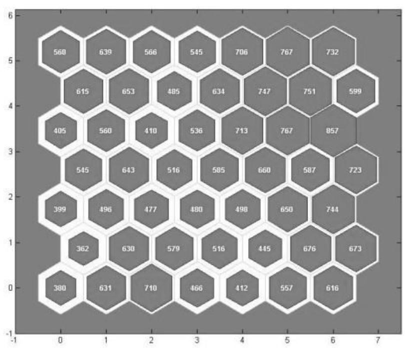

Fig. 5. Sample hits for 49 neurons and 200 epochs

Another useful visualization of the results is to show the locations of the data points and the weight vectors (Fig.6.).

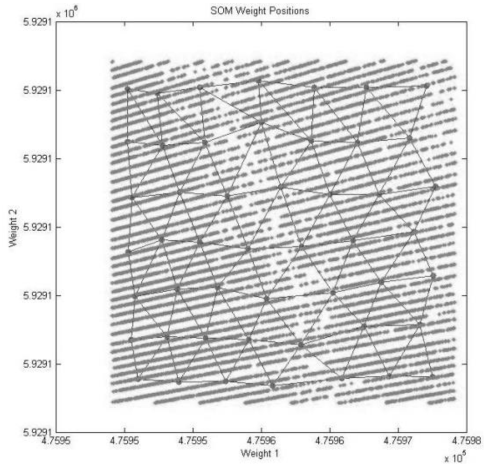

Fig. 6. Weight positions for 49 neurons and 200 epochs

It was found, that after only 200 epochs the map is well distributed through the input space. Results received at these settings are sufficient for clustering bathymetry data. The authors test sample data with the following settings: 49 neurons and 200 epochs one hundred times.

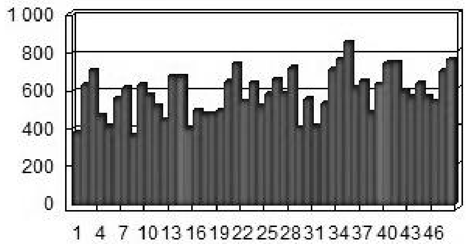

Fig. 7. Frequency distribution

Frequency distribution of mean outcomes is presented on figure above The horizontal axis represents number of neurons and the vertical axis shows bathymetric data points. The examples in a set of data are evenly distributed in each cluster.

4 Conclusion

Data reduction is a procedure meant to reduce the size of the data set. Generally the hydrographic systems generate the grid with use the following methods: mean or weighted mean. The main purpose of authors is elaboration of reduction algorithm of bathymetric data. The clustering of data is first part of search algorithm, as shown in figure above. The next step will be generalization of bathymetric data. The use of 200 epochs during training have the result in decomposition of the neurons – they are very evenly spaced. Number of neurons will be selected depend on the compilation scale and geographical coverage of electronic navigational chart. The test examples in a set of data are regularly distributed in each cluster. The main limitation of neural network usage is high density data. It is long-term to train the network from whole data set using standard computer. The results point that artificial neural networks are good methods of clustering large amount of data. Application of SOFM in Matlab involves grouping data by similarity and it can be used as one of the steps in reduction algorithm of bathymetric data.

References

1. Lubczonek, J.: Application of GIS Techniques in VTS Radar Stations Planning. In: Kawalec, A., Kaniewski, P. (eds.) 2008 International Radar Symposium, Wroclaw, pp. 277–280 (2008)
2. Weintrit, A., Kopacz, P.: Computational Algorithms Implemented in Marine Navigation Electronic Systems. In: Mikulski, J. (ed.) TST 2012. CCIS, vol. 329, pp. 148–158. Springer, Heidelberg (2012)
3. Lubczonek, J., Stateczny, A.: Aspects of spatial planning of radar sensor network for inland waterways surveillance. In: 6th European Radar Conference (EURAD 2009). European Radar Conference-EuRAD, Rome, pp. 501–504 (2009)
4. Przyborski, M., Pyrchla, J.: Reliability of the navigational data. In: Kłopotek, M.A., Wierzchoń, S.T., Trojanowski, K. (eds.) Intelligent Information Processing and Web Mining. ASC, vol. 22, pp. 541–545. Springer, Heidelberg (2003)
5. Stateczny, A.: Artificial neural networks for comparative navigation. In: Rutkowski, L., Siekmann, J.H., Tadeusiewicz, R., Zadeh, L.A. (eds.) ICAISC 2004. LNCS (LNAI), vol. 3070, pp. 1187–1192. Springer, Heidelberg (2004)
6. Przyborski, M.: Possible determinism and the real world data. Physica A-Statistical Mechanics and its Applications 309(3-4), 297–303 (2002)
7. Stateczny, A.: Methods of comparative plotting of the ship's position. In: Brebbia, C., Sciutto, G. (eds.) Maritime Engineering & Ports III, Rhodes. Water Studies Series, vol. 12, pp. 61–68 (2002)
8. Maleika, W., Palczynski, M., Frejlichowski, D.: Effect of Density of Measurement Points Collected from a Multibeam Echosounder on the Accuracy of a Digital Terrain Model. In: Pan, J.-S., Chen, S.-M., Nguyen, N.T. (eds.) ACIIDS 2012, Part III. LNCS (LNAI), vol. 7198, pp. 456–465. Springer, Heidelberg (2012)
9. Lubczonek, J., Stateczny, A.: Concept of neural model of the sea bottom surface. In: Rutkowski, L., Kacprzyk, J. (eds.) Neural Networks and Soft Computing. ASC, vol. 19, pp. 861–866. Physica, Heidelberg (2003)
10. Lubczonek, J.: Hybrid neural model of the sea bottom surface. In: Rutkowski, L., Siekmann, J.H., Tadeusiewicz, R., Zadeh, L.A. (eds.) ICAISC 2004. LNCS (LNAI), vol. 3070, pp. 1154–1160. Springer, Heidelberg (2004)
11. Stateczny, A.: The neural method of sea bottom shape modelling for the spatial maritime information system. In: Brebbia, C., Olivella, J. (eds.) Maritime Engineering and Ports II, Barcelona. Water Studies Series, vol. 9, pp. 251–259 (2000)
12. Maleika, W.: The influence of track configuration and multibeam echosounder parameters on the accuracy of seabed DTMs obtained in shallow water. Earth Science Informatics 6(2), 47–69 (2013)

13. Stateczny, A., Kazimierski, W.: Determining Manoeuvre Detection Threshold of GRNN Filter in the Process of Tracking in Marine Navigational Radars. In: Kawalec, A., Kaniewski, P. (eds.) 2008 Proceedings International Radar Symposium, Wroclaw, pp. 242–245 (2008)
14. Balicki, J., Kitowski, Z., Stateczny, A.: Extended Hopfield Model of Neural Networks for Combinatorial Multiobjective Optimization Problems. In: 2nd IEEE World Congress on Computational Intelligence, Anchorage, pp. 1646–1651 (1998)
15. Stateczny, A., Kazimierski, W.: A comparison of the target tracking in marine navigational radars by means of GRNN filter and numerical filter. In: 2008 IEEE Radar Conference, Rome, vol. 1-4, pp. 1994–1997 (2008)
16. Stateczny, A., Kazimierski, W.: Selection of GRNN network parameters for the needs of state vector estimation of manoeuvring target in ARPA devices. In: Romaniuk, R.S. (ed.) Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments IV, Wilga. Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 6159, pp. F1591–F1591(2006)
17. Stateczny, A.: Neural manoeuvre detection of the tracked target in ARPA systems. In: Katebi, R. (ed.) Control Applications in Marine Systems 2001 (CAMS 2001), Glasgow. IFAC Proceedings Series, pp. 209–214 (2002)
18. Chung, K., Huang, Y., Wang, J., et al.: Speedup of color palette indexing in self-organization of Kohonen feature map. Expert Systems with Applications 39(3), 2427–2432 (2012)
19. Ciampi, A., Lechevallier, Y.: Multi-level Data Sets: An Approach Based on Kohonen Self Organizing Maps. In: Zighed, D.A., Komorowski, J., Żytkow, J. (eds.) PKDD 2000. LNCS (LNAI), vol. 1910, pp. 353–358. Springer, Heidelberg (2000)
20. de Almeida, C., de Souza, R., Candelas, A.: Fuzzy Kohonen clustering networks for interval data. Neurocomputing 99, 65–75 (2013)
21. Du, Z., Yang, Y., Sun, Y., et al.: Map matching Using De-Noise Interpolation Kohonen Self-Organizing Maps. In: Conference: International Conference on Components, Packaging and Manufacturing Technology, Sanya. Key Engineering Materials, vol. 460-461, pp. 680–686 (2011)
22. Guerrero, V., Anegon, F.: Reduction of the dimension of a document space using the fuzzified output of a Kohonen network. Journal of the American Society for Information Science and Technology 52(14), 1234–1241 (2001)
23. Rasti, J., Monadjemi, A., Vafaei, A.: Color reduction using a multi-stage Kohonen Self-Organizing Map with redundant features. Expert Systems with Applications 38(10), 13188–13197 (2011)
24. Kohonen, T.: Self-Organized Formation of Topologically Correct Feature Maps. Biological Cybernetics 43(1), 59–69 (1982)
25. Kohonen, T.: The Self-Organizing Map. Proceedings of The IEEE 78(9), 1464–1480 (1990)
26. Wlodarczyk-Sielicka, M.: 3D Double Buffering method in the process of hydrographic chart production with geodata taken from interferometry multibeam echo sounder. Annals of Geomantic, issue X 7(57), 101–108 (2012)

View publication stats