Comparison of selected clustering algorithms of raw data obtained by interferometric methods using artificial neural networks

Conference Paper · May 2016

DOI: 10.1109/IRS.2016.7497290

3 authors:

Marta Wlodarczyk-Sielicka
Maritime University of Szczecin

Jacek Łubczonek
Maritime University of Szczecin

Andrzej Stateczny
Gdansk University of Technology

Comparison of Selected Clustering Algorithms of Raw Data Obtained by Interferometric Methods Using Artificial Neural Networks

Marta Wlodarczyk-Sielicka
Institute of Geoinformatics
Maritime University of Szczecin
Szczecin, Poland
m.wlodarczyk@am.szczecin.pl

Jacek Lubczonek
Institute of Geoinformatics
Maritime University of Szczecin
Szczecin, Poland
j.lubczonek@am.szczecin.pl

Andrzej Stateczny
Marine Technology Ltd.
Szczecin, Poland
a.stateczny@marinetechnology.pl

Abstract— **The article presents a particular comparison of selected clustering algorithms of data obtained by interferometric methods using artificial neural networks. For the purposes of the experiment original data from Szczecin Port have been tested. For collecting data authors used the interferometric sonar system GeoSwath Plus 250 kHz. GeoSwath Plus offers very efficient simultaneous swath bathymetry and side scan seabed mapping. During the use of Kohonen's algorithm, the network, during learning, use the Winner Take All rule and Winner Take Most rule. The parameters of the tested algorithms were maintained at the level of default. During the research several populations were generated with number of clusters equal 9 for data gathered from the area of 100m². In the subsequent step statistics were calculated and outcomes were shown as spatial visualization and in tabular form.**

Index Terms— **interferometric system; artificial neural network; clustering algorithm; bathymetry**

I. Introduction

The most important data for maritime and inland navigation are Electronic Navigational Charts (ENC). Some aspects of Electronic Navigational Charts using in navigation process was discussed in [1-5].

Data essential to chart production are gathering during hydrographical works. Usually, the most important are the bathymetric data. This data are gathered by synchronous registration of geographical coordinates (ϕ,λ) obtained most often by means of system GPS in RTK mode and hydrographic measurement of depths (h), converted to data sets of XYZ points. This data are a very big sets of bathymetric points. During post processing data should be reduced for presentation bathymetric data on charts. Clustering is the first step of the proposed reduction algorithm.

Spatial clustering is the task of grouping a set of points in such a way that points in the same cluster are more similar to each other than to those in other groups clusters [6].

Bathymetric data very often are collected used multibeam echosounders (MBES), among them interferometric systems are popular. For collecting data authors used the interferometric sonar system GeoSwath Plus 250 kHz. GeoSwath Plus offers very efficient simultaneous swath bathymetry and side scan seabed mapping with accuracies much better then specified in the IHO Standards for Hydrographic Surveys. The applied phase measuring bathymetric sonar technology provides data coverage of up to 12 times the water depth, giving unsurpassed survey efficiency in shallow water environments. The GeoSwath Plus turn-key solution comprises a dual transducer head with versatile mounting options as well as a deck unit containing the complete sonar electronics together with a high spec PC with hydrographic software. The software provides full acquisition, calibration and data processing capabilities for producing the final bathymetry map and side scan mosaic data products. All customary ancillary sensors can be directly interfaced. The measurement profiles maintaining 100% coverage of the measured body of water were realized. Some problem of multibeam echosounders data processing was described in [7-11]. Similar to MBES problems are with LIDAR data [12] as well other navigational data processing [13-16]. Another interesting problem is multisensory data fusion [17-21].

The main goal of this paper is to compare of selected clustering algorithms of raw data obtained by interferometric methods. Some aspects of data reductions was described in previous authors works [22-23].

II. Artificial Neural Networks

Artificial neural networks (ANN) can be treated as a certain kind of data structure, which changes in the course of

978-1-5090-2518-3/16/$31.00

the learning process adapting to the kind of problem to be solved. This structure is constituted by single neurons performing simple arithmetic functions bound into a network. The first and basic neuron model defined as early as in 1943 by McCulloch and Pitts is the nerve cell, the function of which consists in the weight sum of neuron entrances, and next subjecting the sum thus obtained to the action of non-linear activation function. ANN are very often used to solve navigational problems like sea bottom shape modeling [24-26] and others tasks [27-35].

For clustering problems solution especially useful are self-organizing ANN. Kohonen's networks are one of basic types of self-organizing neural networks. The ability to self-organize provides new possibilities - adaptation to formerly unknown input data. It seems to be the most natural way of learning, which is used in our brains, where no patterns are defined. Those patterns take shape during the learning process, which is combined with normal work.

Kohonen's networks are a synonym of whole group of networks which make use of self-organizing, competitive type learning method. At the beginning, signals on network's inputs were set up and then winning neuron is chosen, the one which corresponds with input vector in the best way. Precise scheme of competition and later modifications of synaptic wages may have various forms. There are many sub-types based on rivalry, which differ themselves by precise self-organizing algorithm.

During the use of Kohonen's algorithm, the network, during learning, use the Winner Takes All rule (further referred to as WTA) and Winner Takes Most rule (further referred to as WTM). In case of WTA rule neural adaptation relates only to the winner neuron. Neurons lose competition neurons do not modify their weights. While, WTM rule modifies not only the weight of the winner, but also its neighbors. The radius of the neighborhood decreases with learning time. In this case winner neuron and all neuron within a radius of its neighborhood subject to adaptation [36].

III. The specification of geodata reduction method

Spatial data obtained by interferometric methods is a large set of points. The essential purpose of the authors' research is the implementation of a new reduction algorithm for spatial data (XYZ points) to be used for the creation of bathymetric map. In short reduction of data is a procedure by which the number and hence size of a data set is reduced, in order to make the analysis easier and more efficient. In many cases, hydrographic systems generate a grid of bathymetric data by using means or weighted means. The authors aim to create a new reduction algorithm for bathymetric data using artificial neural networks. The clustering of data is the first part of the search algorithm and the next stage is the generalization of bathymetric data. Schema of the search algorithm is shown in Fig. 1.

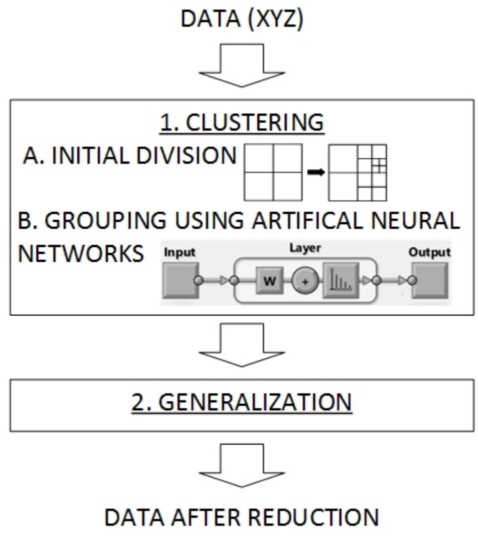

Figure 1. Schema of proposed reduction algorithm.

The goal of the authors is to classify a set of XYZ points into clusters and then represent each group by a single point with minimum depth depending on the compilation scale. It needs to be highlighted that, in this method the points of minimum depth will remain in their true position, and they will be visualized irrespective of the scale used on bathymetric map. For safety associated with navigation it is very important to retain points of minimum depth. The main objective of new reduction algorithm is that, the position of point and the depth at this point will not be an interpolated value.

IV. Experiment

For the purposes of the experiment original data from Szczecin Port have been tested. Artificial neural networks were used for data clustering. During the use of Kohonen's algorithm, the network, during learning, use the WTA rule and WTM rule. The parameters of the tested algorithms were maintained at the level of default. During the research several populations were generated with number of clusters equal 9 for data gathered from the area of 100m². In the subsequent step statistics were calculated and outcomes were shown as spatial visualization and in tabular form. The final step was their analysis. The test algorithms were implemented using Matlab software, developed by MathWorks.

A. Test area

During the bathymetric survey a large amount of data was gathered. When using a standard computer, very high-density data present the main operational limitation. For solve this problem, the authors separated the primary data point sets into smaller subsets. During the studies, test data gathered from the area of 100m² was used and this set contains 3 760 samples of XYZ elements. Test data was collected within Szczecin Port, near the Babina Canal. This area at the scale 1:25000 is presented in Fig. 2.

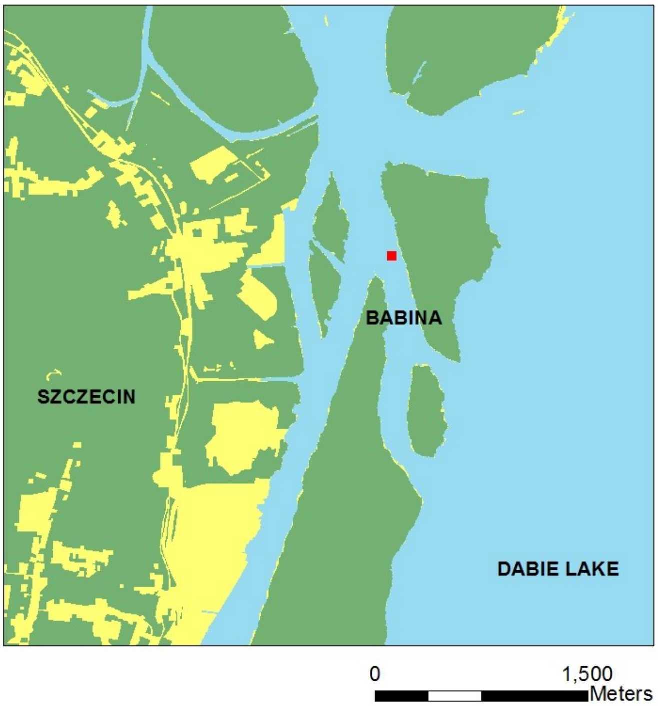

Figure 2. Partial view of Szczecin Port.

Several point has three attributes: latitude (X), longitude (Y) and depth at a given point (Z). The minimum depth within this area is 3.60 meters and the maximum depth is 5.23 meters.

B. Parameters of algorithms

For data gathered from the area of 100m², over the tests several populations were generated with number of clusters equal 9. For the purpose of clustering self-organizing map was applied. The authors selected the hexagonal network topology, where each of the hexagons represents a neuron. The numbers of rows and columns was set to 3×3, which provided 9 clusters. During each trainings the number of iteration was set at 1000. Distances are calculated from their positions by means of a link distance function, which is default function in software used. The link distance from one neuron is the number of links or steps that must be taken to get to the neuron under consideration. During the training the network applies the WTA rule and WTM rule. Consequently using the WTN rule the initial neighborhood size was set at 3 and the number of training steps for initial coverage of the input space was set at 100. During this phase, the neighborhood is gradually reduced from a maximum size of neighborhood down to 1, where it remains from then on.

C. Results

The results for 9 clusters are presented in this article. All sets of received clusters were analyzed. During the research the authors adopted the precision of two decimal places.

In this research the authors adopted the following evaluation criteria: time taken for calculations and distribution of data in each cluster. The authors focused on depth values, which are of significant importance for the safety of navigation. Tab. 1 introduces the results for 9 clusters.

TABLE I. COMPARISON OF STATISTICS FOR 9 CLUSTERS

		Clusters								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
WTA	Min[a]	4.09	4.16	3.93	3.66	3.96	4.00	3.60	3.67	4.08
	Max[b]	5.22	5.20	5.04	4.67	5.04	4.92	4.75	4.60	5.23
	Mean[c]	4.65	4.71	4.46	4.17	4.42	4.46	4.25	4.15	4.69
	SD[d]	0.22	0.20	0.19	0.18	0.17	0.17	0.19	0.17	0.22
	NoP[e]	355	368	348	467	484	518	416	478	326
WTM	Min	4.09	4.11	3.93	3.67	3.96	3.94	3.60	3.70	4.17
	Max	5.22	5.20	4.94	4.67	4.94	5.04	4.74	4.60	5.23
	Mean	4.64	4.69	4.48	4.17	4.39	4.47	4.21	4.17	4.72
	SD	0.22	0.20	0.17	0.18	0.17	0.19	0.19	0.17	0.22
	NoS	304	446	524	478	443	430	421	434	280

a. Minimum value of depth

b. Maximum value of depth

c. Mean value of depth

d. Standard deviation

e. Number of samples in each cluster

The results for different method in clusters are comparable. Minimum values of depth in each cluster are at a similar level. For four clusters they are the same. The major differences are in cluster designated as 9 and cluster marked 2. The differences range from 5 centimeters to 9 centimeters. The differences between minimum and maximum depth in each clusters for WTA range 0.92 meter to 1.15 meter. While, for WTM they range 0.90 meter to 1.14 meter. The mean values of depth are slightly different in each cluster, which is shown in Fig. 3. The biggest difference between the methods occurs for cluster designated as 7 and it is only 4 centimeters.

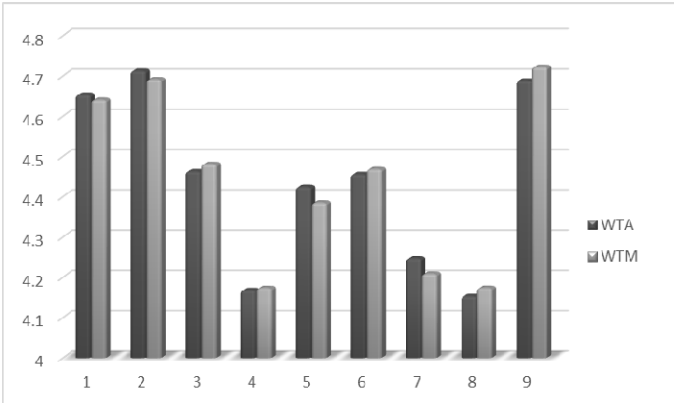

Figure 3. Comparison of the mean values of depth in each cluster.

A high standard deviation is in cluster marked 1 and 9 – at a level equal 0.22. It indicates that the data points are spread out over a wider range of values.

Fig. 4 presents spatial representation of the results for the for 9 clusters. It can be noticed, that results for WTA and WTM are close to each other.

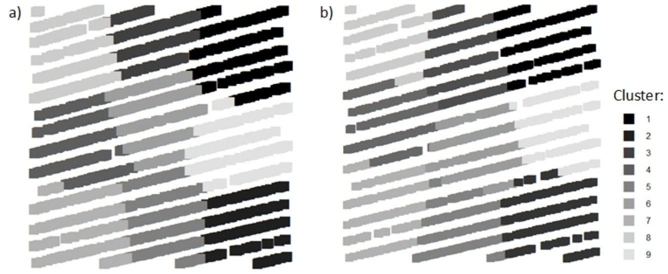

Figure 4. Spatial representation of results for a) WTA and b)WTM.

The final analyzed value is number of samples in each cluster. Fig. 5 presents distribution numbers of points in each clusters for tested methods. The axis X represents the number of clusters and the axis Y shows number of samples.

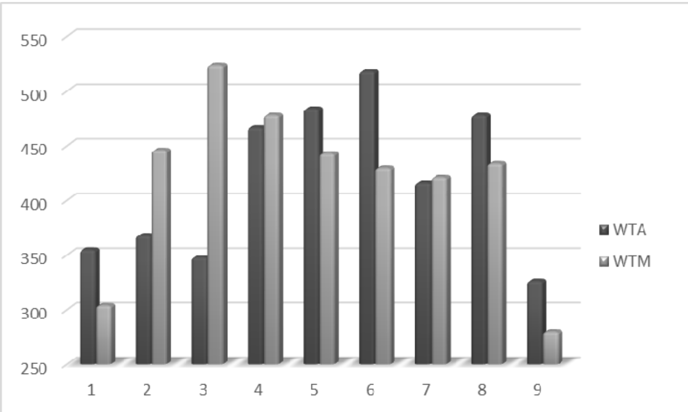

Figure 5. Distribution of the number of samples in each cluster.

The greatest difference can be seen for cluster designated as 3. However, it should be noted that for this cluster the minimum depth is the same and it assumes a value of 3.93 meters.
During tests the authors also paid attention on the time taken for calculations. It was shorter by about 5 seconds, when WTM rule was tested.

V. CONLUSION

Self-organizing networks have the ability to divide spatial data areas. They accumulate together data with similar values. The authors aim to create a new reduction algorithm for bathymetric data. The main criterion for evaluating each method for reduction of bathymetric maps is the legibility of the maps. After analysis of the above results it can be used both tested methods. The statistics related to depth values were taken into account and the results in particular clusters are comparable. Minimum values of depth in each cluster are very similar. However, it should be noted that the time taken for calculations was shorter when WTM rule was used. High-density data obtained by interferometric methods present the main operational limitation when using a standard computer. So, the original data point have to be divided into smaller subsets, which could be trained separately. In this case computing time is very important. Regular distribution of data in each cluster is important in case of a small slope bottom. In the next stages of the research, the authors will use the selected method over several different test areas. These areas will be characterized by varying the inclination and by a diverse distribution of samples.

REFERENCES

[1] A. Stateczny, I. Bodus-Olkowska, “Hierarchical Hydrographic Data Fusion for Precise Port Electronic Navigational Chart Production”. Mikulski J.(ed.) Telematics in the Transport Environment, Book Series: Communications in Computer and Information Science 471, pp. 359-368. Ustron, 2014.

[2] N. Wawrzyniak, T. Hyla, „Managing Depth Information Uncertainty in Inland Mobile Navigation Systems”. Proceedings of the Joint Rough Set Symposium, Granada and Madrid, Spain, Kryszkiewicz et al. (Eds), Lecture Notes in Artificial Intelligence, 8537, pp. 343–350, 2014.

[3] W. Kazimierski, A. Stateczny, “Radar and Automatic Identification System track fusion in an Electronic Chart Display and Information System”. The Journal of Navigation vol. 68, issue 6, pp 1141 - 1154 2015.

[4] W. Kazimierski, A. Stateczny, “Fusion of Data from AIS and Tracking Radar for the Needs of ECDIS”, IEEE Conference: Signal Processing Symposium (SPS), Jachranka 2013.

[5] M. Wlodarczyk-Sielicka, M., A. Stateczny, “Clustering Bathymetric Data for Electronic Navigational Charts”. The Journal of Navigation (in press), 2016

[6] Z. Li, “Algorithmic Foundation of Multi-scale Spatial Representation”, CRC Press, 2007

[7] W. Maleika, “The influence of track configuration and multibeam echosounder parameters on the accuracy of seabed DTMs obtained in shallow water”. Earth Science Informatics, vol. 6, issue 2, pp. 47-69, 2013.

[8] W. Maleika, “Development of a Method for the Estimation of Multibeam Echosounder Measurement Accuracy”. Przeglad Elektrotechniczny, 88 (10B), 205–208, 2012.

[9] W. Maleika, M. Palczynski, D. Frejlichowski, “Effect of Density of Measurement Points Collected from a Multibeam Echosounder on the Accuracy of a Digital Terrain Model”. 4th International Scientific Asian Conference on Intelligent Information and Database Systems (ACIIDS). Edited by: Pan, JS., Chen, SM., Nguyen, NT., Book Series: Lecture Notes in Artificial Intelligence, vol. 7198, pp. 456-465. Kaohsiung, Taiwan 2012.

[10] W. Maleika, “Moving Average Optimization in Digital Terrain Model Generation Based on Test Multibeam Echosounder Data”. Geo-Marine Letters, 35, 61–68, 2015.

[11] W. Maleika, “The Influence of the Grid Resolution on the Accuracy of the Digital Terrain Model Used in Seabed Modelling”. Marine Geophysical Research, 36, 35–44, 2015.

[12] P. Burdziakowski, A. Janowski, A Kholodkow et al. “Maritime Laser Scanning as the Source For Spatial Data” Polish Maritime Research, vol. 22, issue 4, pp.9-14, 2015.

[13] A. Stateczny, N. Wawrzyniak, “Method for determining stationary position of scanning sonar, involves determining head position based on sonar search for actual sonar image with set of synthetic images generated by performing ray tracing process based on model of bottom”. Patent Number: PL406523-A1. Patent Assignee: Marine Technology Sp. z o.o. 2015.

[14] A. Stateczny, “Methods of comparative plotting of the ship's position”. Book Editor(s): Brebbia, CA., Sciutto, G. Maritime Engineering & Ports III. Book Series: Water Studies Series vol. 12, pp. 61-68, Rhodes 2002.

[15] M. Przyborski, “Possible determinism and the real world data”. Physica A-Statistical Mechanics and its Applications, vol. 309, issue 3-4, pp. 297-303, 2002.

[16] M. Przyborski, J. Pyrchla, “Reliability of the navigational data”. International Intelligent In-formation Systems/Intelligent Information Processing and Web Mining Conference (IIS: IIPWM 03). Edited by: Klopotek, MA., Wierzchon, ST., Trojanowski, K., Book Series: Advances in Soft Computing, pp. 541-545. Zakopane, 2003.

[17] J. Łubczonek, M. Borawski, “Comparative analysis of radar image compression methods”, Proceedings of 16th International Radar Symposium (IRS), International Radar Symposium Proceedings, H. Rohling (Ed.), pp. 1117 - 1122, Dresden, Germany (2015).

[18] A. Stateczny, W. Kazimierski, “Sensor Data Fusion in Inland Navigation”. Book Editor(s): Rohling, H. 14th International Radar Symposium (IRS), vols. 1 and 2. Book Series: International Radar Symposium Proceedings pp. 264-269. Dresden 2013.

[19] W. Kazimierski, “Problems of Data Fusion of Tracking Radar and AIS for the Needs of Integrated Navigation Systems at Sea”. 14th International Radar Symposium (IRS), vol.1 and 2, Book Series: International Radar Symposium Proceedings, pp. 270-275, Dresden, 2013.

[20] T. Hyla, W. Kazimierski, N. Wawrzyniak, „Analysis of Radar Integration Possibilities in Inland Mobile Navigation. Proceedings of the 16th International Radar Symposium (IRS), International Radar Symposium Proceedings, Dresden, Germany, H. Rohling (Ed.), pp. 864-869, 2015.

[21] A. Stateczny, I. Bodus-Olkowska, “Sensor Data Fusion Techniques for Environment Modelling”. Proceedings of 16th International Radar Symposium (IRS), International Radar Symposium Proceedings, H. Rohling (Ed.), pp. 1123-1128. Dresden, 2015.

[22] A. Stateczny, M. Wlodarczyk-Sielicka, “Self-Organizing Artificial Neural Networks into Hydrographic Big Data Reduction Process”. 2014 Joint Rough Set Symposium, Granada-Madrit, Lecture Notes in Computer Science, vol. 8537, pp. 335-342, Granada-Madrit, 2014.

[23] M. Wlodarczyk-Sielicka, A. Stateczny, “Selection of SOM Parameters for the Needs of Clusterisation of Data Obtained by Interferometric Methods”. Proceedings of 16th International Radar Symposium (IRS), International Radar Symposium Proceedings, H. Rohling (Ed.), pp. 1129-1134. Dresden, 2015.

[24] J. Lubczonek, “Hybrid neural model of the sea bottom surface”. Edited by: Rutkowski, L., Siekmann, J., Tadeusiewicz, R. et al. 7th International Conference on Artificial Intelligence and Soft Computing - ICAISC 2004. Book Series: Lecture Notes in Artificial Intelligence, vol. 3070, pp.: 1154-1160. Zakopane, 2004.

[25] J. Lubczonek, A. Stateczny, “Concept of neural model of the sea bottom surface”. Book Editor(s): Rutkowski, L., Kacprzyk, J. Neural Networks and Soft Computing Book Series: Advances in Soft Computing, pp. 861-866. Zakopane 2003.

[26] A. Stateczny, “The neural method of sea bottom shape modelling for the spatial maritime information system”. Book Editor(s): Brebbia, CA., Olivella, J. Maritime Engineering and Ports II. Book Series: Water Studies Series vol. 9, pp. 251-259. Barcelona 2000.

[27] A. Stateczny, “Neural manoeuvre detection of the tracked target in ARPA systems”. Book Editor(s): Katebi, R. Control Applications in Marine Systems 2001 (CAMS 2001). Book Series: IFAC Proceedings Series, pp. 209-214, Glasgow 2002.

[28] A. Stateczny, W. Kazimierski, “Selection of GRNN network parameters for the needs of state vector estimation of manoeuvring target in ARPA devices”. Book Editor(s): Romaniuk, RS. Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments IV. Book Series: Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE) vol. 6159, pp. F1591-F1591. Wilga 2006.

[29] A. Stateczny, W. Kazimierski, “A comparison of the target tracking in marine navigational radars by means of GRNN filter and numerical filter”. 2008 IEEE Radar Conference, vols. 1-4. Book Series: IEEE Radar Conference, pp. 1994-1997. Rome 2008.

[30] A. Stateczny, W. Kazimierski, “Determining Manoeuvre Detection Threshold of GRNN Filter in the Process of Tracking in Marine Navigational Radars”. Book Editor(s): Kawalec, A., Kaniewski, P. 2008 Proceedings International Radar Symposium, pp. 242-245. Wroclaw 2008.

[31] A. Stateczny, “Artificial neural networks for comparative navigation”. Book Editor(s): Rutkowski, L., Siekmann, J., Tadeusiewicz, R., et al. Artificial Intelligence and Soft Computing - ICAISC 2004. Book Series: Lecture Notes in Artificial Intelligence vol. 3070, pp. 1187-1192, Zakopane 2004.

[32] J. Balicki, Z., Kitowski, A. Stateczny, “Extended Hopfield Model of Neural Networks for Combinatorial Multiobjective Optimization Problems”. 2th IEEE World Congress on Computational Intelligence, pp. 1646-1651. Anchorage 1998.

[33] A. Stateczny, W. Kazimierski, “Method and system of determining the vector of traced objects”. Patent Number: PL212560-B1. Patent Assignee: Akademia Morska w Szczecinie, 2012.

[34] W. Kazimierski, G. Zaniewicz, “Analysis of the Possibility of Using Radar Tracking Method Based on GRNN for Processing Sonar Spatial Data”, Proceedings of the Joint Rough Set Symposium, Spain, Kryszkiewicz et al. (Eds), Lecture Notes in Artificial Intelligence, 8537, pp. 319-326. Granada and Madrid, 2014.

[35] A. Stateczny, Method for determining control parameters of collision avoidance floating units of watercraft, involves detecting objects that threaten collision, cascading modules in appropriate manner, and determining parameters for optimal maneuver. Patent Number: PL408536-A1. Patent Assignee: Marine Technology Sp. z o.o. 2016.

[36] S. Osowski, “Sieci neuronowe do przetwarzania informacji”, Oficyna Wydawnicza Politechniki Warszawskiej, 2013.