Processing of Big Data for 3D Multibeam Sonar (Water) and LiDAR Point Cloud (Land) in Hydrography Field

Wioleta BŁASZCZAK-BĄK, Poland

Key words: big data, 3D multibeam sonar, LiDAR, OptD method

SUMMARY

This paper presents a new approach to the acquisition and simultaneous processing of two different datasets of measurements: (1) bathymetric observations, (2) laser scanning data. This integrated approach allows the processing of water and land areas. It is a new way of dealing with big datasets based on fragmentary data acquisition and fast reduction (performed by means of the Optimum Dataset method - OptD) within acquired measuring strips in almost real time.

The modification of OptD method was used to simultaneously acquire and reduce the dataset from multibeam echosounders (MBES) and Light Detection and Ranging (LiDAR).

All stages of the proposed approach are performed in a first phase of datasets acquisition, with the difference that during measurement not the whole dataset is processed, but a fragment of it. Data acquisition, data processing, reduction of the observations and 3D model generation in hydrography in almost real time is important research subject in context of comparative navigation and land and water management.

STRESZCZENIE

W artykule przedstawiono nowe podejście do pozyskiwania i jednoczesnego przetwarzania dwóch różnych zestawów danych pomiarowych: (1) obserwacji batymetrycznych, (2) danych ze skaningu laserowego. To zintegrowane podejście pozwala na przetwarzanie zbiorów danych reperzentujących dno zbiorników wodnych i obszarów lądowych.

Jest to nowy sposób przetwarzania dużych zbiorów danych, opraty na fragmentarycznym pozyskiwaniu danych i szybką redukcję (w oparciu o metodę Optimum Dataset - OptD) w obrębie pozyskanych pasków pomiarowych w czasie prawie rzeczywistym.

Modyfikacja metody OptD została wykorzystana do jednoczesnego pozyskiwania i zmniejszania zestawu danych z echosondy wielowiązkowej (MBES) oraz skaningu laseroweg (LiDAR).

Wszystkie etapy proponowanego podejścia są wykonywane w pierwszej fazie pozyskiwania zestawów danych, z tą różnicą, że podczas pomiaru nie przetwarzany jest cały zestaw danych, ale fragment zbioru danych. Akwizycja danych, przetwarzanie danych, redukcja ilości obserwacji i generowanie modelu 3D w hydrografii w prawie czasie rzeczywistym jest ważnym przedmiotem badań w kontekście nawigacji porównawczej oraz gospodarki lądowej i wodnej.

Processing of Big Data for 3D Multibeam Sonar (Water) and LiDAR Point Cloud (Land) in Hydrography Field

Wioleta BŁASZCZAK-BĄK, Poland

1. INTRODUCTION

Big data problems are closely related to the idea of single-beam echosounders measurements (Aykut et al., 2013; Moszynski et al., 2013) and LiDAR (Błaszczak-Bąk et al., 2018; Kogut et al., 2016). Big data cause problems with their rational use during processing. New measuring technologies give us the opportunity to obtain huge amounts of data, which very often are not necessary later in the process. Therefore, the different methods are being developed to reduce the number of observations in measuring datasets. However, it is important that these methods do not interfere with the nature of the object being tested.

In this paper a different approach is presented than existing in the literature on the subject (Calder and Mayer, 2003), (Yang et al., 2013), (Wlodarczyk-Sielicka, 2016). The approach includes fragmentary data acquisition (LiDAR and MBES data) with fast reduction (Optimum Dataset method (Błaszczak-Bąk, 2016)) within acquired measuring strips in almost real time. The OptD method was modified for this purpose and the details of this modification are presented in (Stateczny et al., 2019). This modification consists in the introduction of a loop (FOR instruction) for fragmentary data processing. Additionally, the approach considers obtaining the measuring strips without overlay between them. The results show that the proposed approach enables fast data acquisition and reduction in almost real time, and it can be used for example for comparative navigation.

The proposed data processing methodology can be used for integrated measuring systems. Such a system suggests the Marine Technology sp. z o.o. in Szczecin. The fully functional prototype of HydroDron-1 was made as a result of the project "Developing an autonomous/remote-controlled floating platform dedicated to hydrographic measurements in limited water bodies".

2. MATERIALS AND METHODS

2.1. Optimum Dataset method

Previous applications of the OptD method consisted of processing the entire dataset (Airborne Laser Scanning (Błaszczak-Bąk, 2016), Terrestrial Laser Scanning (Suchocki and Błaszczak-Bąk, 2019), Mobile Laser Scanning (Błaszczak-Bąk et al., 2018). In the case of MBES the strips with observations are reduced in almost real time and it happens in stages. The OptD method was modified for this purpose (Stateczny et al., 2019). This modification consists in closing the OptD method in a loop (FOR instruction) for fragmentary data processing. In this article additionally the process of simultaneous data processing from two sources: MBES data and LiDAR data was tested.
The observation datasets from these two sources were first integrated to combine them into one measurement dataset in one coordinate system.

2.2. Test area characteristics

The measurements were carried out by HydroDron-1 (Stateczny et al., 2018) on Klodno Lake, which is a gutter-type lake and located in the Kashubian Lake District in Chmielno Commune, in the Kartuzy administrative unit (Pomorskie Province), Poland. Data presented in the article were taken during a 9 day measurement campaign in April and May 2019. Materials and data used in the paper were made available by Marine Technology sp. z o.o. in Szczecin. The dataset LiDAR+MBES is presented in Figure 1.

Figure 1. Data acquisition

2.3. Processing with OptD method and results

The main aim of OptD method is the reduction of the set of measurement observations. The degree of reduction is determined by setting BEFORE reduction the optimization criterion (*c*). In this paper optimization criterion was the number of observations in dataset that user requires AFTER reduction and it was $c = 2\%$.

The strip's width can be determined in relation to the measuring speed. The first measurement strip is reduced while the next strip is acquired. The second strip is attached to the previous (already reduced) strip, and then the second is reduced while the third is obtained and so on, until the measurement is finished (Stateczny et al., 2019). Reduction conducted within each of the separated strips is based on the Douglas-Peucker cartographic generalization method (Douglas and Peucker, 1973). Figure 2 presents the steps of processing with OptD method in each strip.

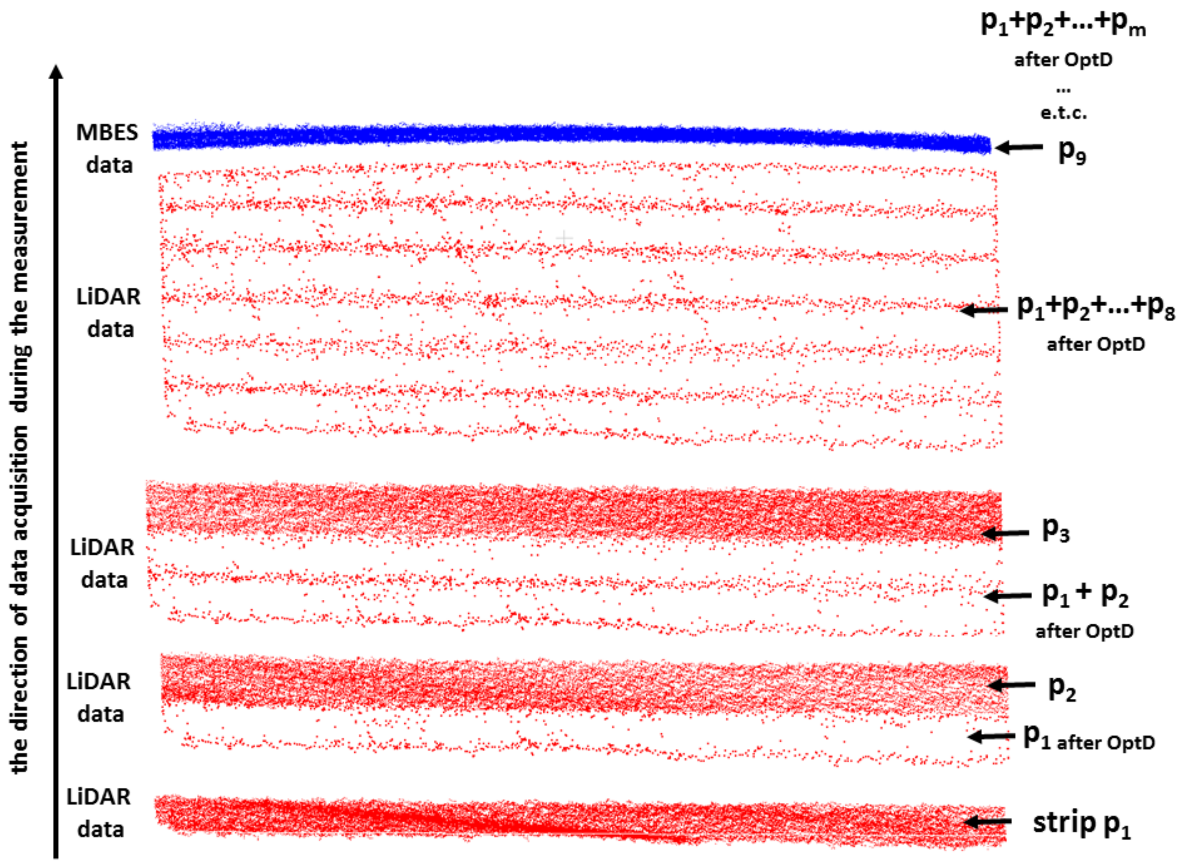

Figure 2. Processing based on the OptD method.

The Figure 2 shows successive processing of measuring strips. The test area is divided into strips p_i without overlay between them (p_i is a strip with observations, where i = 1, 2, 3 ..., m, and m is the number of strips)**.** Each dataset is reduced by the OptD method.
The whole dataset after reduction = p_1 after OptD + p_2 after OptD + ... + p_m after OptD.
The reduction within strips without overlay took 4–7 s. The data processing time was much shorter than for obtaining one strip.
Figure 3 presents the results of applying the OptD method in all strips (whole dataset after reduction).

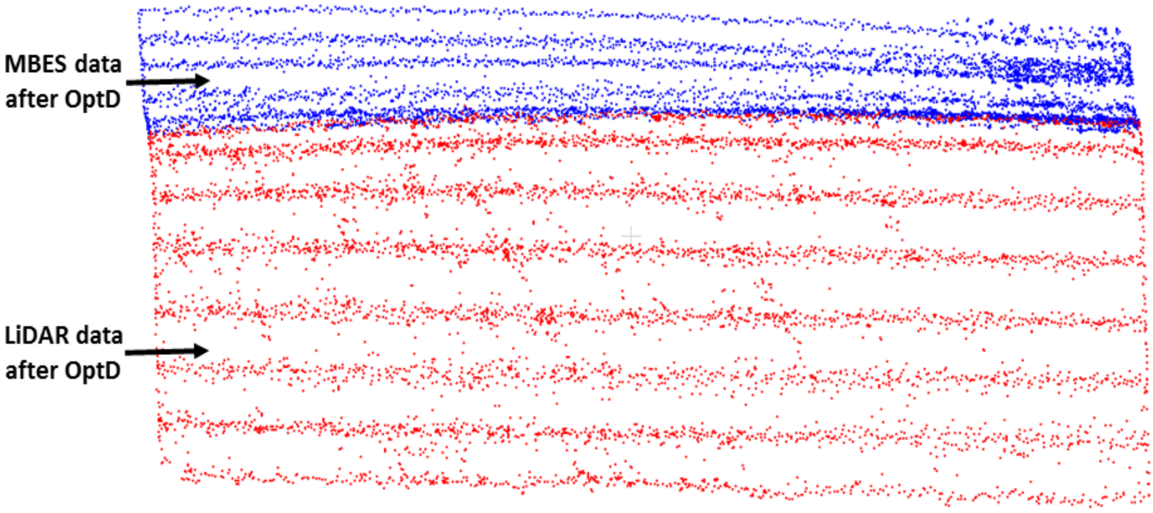

Figure 4. MBES data and LiDAR data after OptD method

The statistical characteristics of the original datasets and dataset obtained by OptD processing are shown in Table 1 and Table 2.

Table 1. Statistical characteristics of the original dataset
(where: H – height, R – range, STD – standard deviation)

		Number of points	H min. [m]	H max. [m]	|R| [m]	STD [m]
Whole dataset		669138	18.40	22.53	4.13	1.34
LiDAR data	p_1	34176	18.40	18.42	0.02	0.06
	p_2	35157	18.45	18.42	0.03	0.12
	p_3	54333	18.80	18.45	0.35	0.11
	p_4	43579	18.78	18.80	0.02	0.10
	p_5	52952	19.01	18.90	0.11	0.13
	p_6	55562	19.04	19.02	0.02	0.16
	p_7	69704	19.84	19.00	0.84	0.21
	p_8	84851	20.45	19.84	0.61	0.35
s d	po_9	79742	20.24	20.42	0.18	0.45

	po_{10}	48211	21.58	20.45	1.13	0.25
	po_{11}	54526	22.65	21.25	1.40	0.32
	po_{12}	56345	22.70	22.53	0.17	0.33

Table 2. Statistical characteristics dataset obtained by OptD processing (where: H – height, R – range, STD – standard deviation)

		Number of points	**H min. [m]**	**H max. [m]**	**|R| [m]**	**STD [m]**
2% of dataset		13383	18.40	22.53	4.13	3.23
LiDAR data	**p_1**	684	18.40	18.42	0.02	0.23
	p_2	703	18.45	18.42	0.03	0.45
	p_3	1087	18.80	18.45	0.35	0.65
	p_4	872	18.78	18.80	0.02	0.48
	p_5	1059	19.01	18.90	0.11	0.59
	p_6	1111	19.04	19.02	0.02	0.54
	p_7	1394	19.84	19.00	0.84	0.61
	p_8	1697	20.45	19.84	0.61	0.75
MBES data	**po_9**	1595	20.24	20.42	0.18	0.85
	po_{10}	1595	21.58	20.45	1.13	0.68
	po_{11}	1091	22.65	21.25	1.40	0.76
	po_{12}	1127	22.70	22.53	0.17	0.78

The presented results show that the OptD method gives good reduction effects. In this case, the dataset was reduced to 2% of the total dataset. It can be seen in the Table 1 and Table 2 that the minimum and maximum values in individual measuring strips have been maintained. Thus, the range R is the same for both the original data and after the reduction. The SD increased slightly, but in the case of terrain it is correct with such a small number of observations. SD is also calculated in relation to the average H value, which means that it presents the distribution of points in relation to the average H value.

3. CONCLUSION

New methodology for reduction of the dataset being acquired from MBES and LiDAR was proposed. It consists of the OptD method, that enables to conduct reduction of the number of observations in almost real time. The methodology is based on fragmentary processing of observations organized in strips without overlay. Our tests showed that using strips without overlay and performing reduction by using the OptD method is an efficient approach. It can

be fast way to obtain data for exapmle for 3D model generation. A major advantage of our method is that only points containing relevant information about terrain and depth differences are used for DTM construction and insignificant points representing flat areas are omitted.
The general conclusions can be formulated as follows:
1. New methodology is dedicated for LiDAR and MBES data.
2. New approach consists of steps: acquisition the fragment of data and reducing of data in almost real time.
3. In the same time the one fragment of data is processed with new methodology, the next fragment of data is acquired. This approach allows fast processing.

REFERENCES

Aykut, N.O., Akpınar, B., Aydın, Ö., 2013. Hydrographic data modeling methods for determining precise seafloor topography. Comput. Geosci. https://doi.org/10.1007/s10596-013-9347-1

Błaszczak-Bąk, W., 2016. New optimum dataset method in LiDAR processing. Acta Geodyn. Geomater. 13, 381–388. https://doi.org/10.13168/AGG.2016.0020

Błaszczak-Bąk, W., Koppanyi, Z., Toth, C., 2018. Reduction Method for Mobile Laser Scanning Data. ISPRS Int. J. Geo-Information 7, 1–13. https://doi.org/10.3390/ijgi7070285

Calder, B.R., Mayer, L.A., 2003. Automatic processing of high-rate, high-density multibeam echosounder data. Geochemistry, Geophys. Geosystems. https://doi.org/10.1029/2002GC000486

Douglas, D.H., Peucker, T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Can. Cartogr. 10(2), 112–122. https://doi.org/10.1002/9780470669488.ch2

Kogut, T., Niemeyer, J., Bujakiewicz, A., 2016. Neural networks for the generation of sea bed models using airborne lidar bathymetry data. Geod. Cartogr. https://doi.org/10.1515/geocart-2016-0007

Moszynski, M., Chybicki, A., Kulawiak, M., Lubniewski, Z., 2013. A novel method for archiving multibeam sonar data with emphasis on efficient record size reduction and storage. Polish Marit. Res. https://doi.org/10.2478/pomr-2013-0009

Stateczny, A., Błaszczak-Bak, W., Sobieraj-Złobińska, A., Motyl, W., Wisniewska, M., 2019. Methodology for processing of 3D multibeam sonar big data for comparative navigation. Remote Sens. https://doi.org/10.3390/rs11192245

Stateczny, A., Wlodarczyk-Sielicka, M., Gronska, D., Motyl, W., 2018. Multibeam Echosounder and LiDAR in Process of 360-Degree Numerical Map Production for Restricted Waters with HydroDron, in: Proceedings - 2018 Baltic Geodetic Congress, BGC-Geomatics 2018. https://doi.org/10.1109/BGC-Geomatics.2018.00061

Suchocki, C., Błaszczak-Bąk, W., 2019. Down-Sampling of Point Clouds for the Technical Diagnostics of Buildings and Structures. Geosciences. https://doi.org/10.3390/geosciences9020070

Wlodarczyk-Sielicka, M., 2016. Importance of neighborhood parameters during clustering of bathymetric data using neural network, in: Communications in Computer and Information Science. https://doi.org/10.1007/978-3-319-46254-7_35

Yang, F., Li, J., Han, L., Liu, Z., 2013. The filtering and compressing of outer beams to multibeam bathymetric data. Mar. Geophys. Res. https://doi.org/10.1007/s11001-012-9164-2

BIOGRAPHICAL NOTES

Wioleta Błaszczak-Bąk is a Professor at the Institute of Geodesy of the University of Warmia and Mazury in Olsztyn, Poland. She is conducting research on LiDAR point cloud processing. She is an author of papers on big data optimization.

CONTACTS

dr hab. inż. Wioleta Błaszczak-Bąk, prof uczelni
Institution: University of Warmia and Mazury in Olsztyn, Faculty of Geoengineering, Institute of Geodesy
Oczapowskiego street 1, 10-719 Olsztyn
Olsztyn
POLAND
Email: wioleta.blaszczak@uwm.edu.pl