Automation of contrast media boundary detection when processing GPR data

Viktor Pupatenko¹, Yuri Sukhobok¹, Leonid Verkhovtsev¹ and Gennady Stoyanovich¹
¹Far Eastern State Transport University, 47 Seryshev St, Khabarovsk, 680021, Russia
E-mail: khv_pvv@mail.ru

Abstract. The paper presents the field results of the inspection of transport facilities, in which the contrasting media in terms of radio wave propagation velocities are found. The possibility of high-precision determination of interface boundaries for such media at shallow depths (for example, the contact of asphalt concrete pavement and crushed stone and sand mixture), as well as at high depths (for example, the determination of the bottom of water bodies) is discussed. In addition, the possibility of using artificial neural networks to identify contrasting boundaries on the GPR data and the transition to less contrasting boundaries is justified.

1. Introduction
Non-destructive methods are widely used in engineering practice. One of them – ground penetrating radar (GPR) – appeared not so long ago, but it has already proven itself in solving the problems of engineering and geological surveys and monitoring the status of soil structures and infrastructure [1-4].

A wide range of tasks to be solved with GPR is due to a large selection of equipment. The main characteristic of GPR antennas – the central frequency – varies over a wide range. Low-frequency 50...25 MHz antennas are used for the deep inspection of subsoil bases, high embankments and slopes [5, 6]. In the frequency range of 100...150 MHz the problems of searching for local inhomogeneities in the bases are successfully solved [7]. The method allowed detecting lenses of ancient ice at the base of the railway embankment in the permafrost propagation zone. [8]. These antennas solve the problems of detecting the boundaries of different soil layers. For example, when one of the sea ports in Colombia was surveyed with a GPR unit LOZA, the thickness of weak silt soils at the base was determined [9]. Survey work was carried out both on the surface of the soil and on the bottom, during low tide. Often it is necessary to determine the soil layers that are under water. Such tasks are well done in winter. Areas of the Amur River in the Russian Far East were surveyed. The same tasks were solved in summer as well. The survey was carried out using a GPR unit mounted on a boat from the water surface.

If it is necessary to increase the accuracy of determination of borders up to the first centimeters, the work is performed by GPR units with high-frequency antennas. The depth of the survey in such cases is small, but the resolution of the set of equipment meets the high requirements for the accuracy of soil delineation. The central frequency of the antennas is in the range of 500 MHz ... 2 GHz. Tasks for such GPR units can be connected with determination of thickness of layers up to 5...10 cm, or, with
search of small inhomogeneities in the ground, for example, with burials, determination of a place of reinforcing elements position in the ground, concrete or asphalt.

For long transport facilities, such as road and railroad embankments, river and sea water areas, the processing of large amounts of GPR data is a significant problem. The large length of such profiles is only one side of the issue. In the case of a detailed examination, with small sizes of inhomogeneities, the sampling frequency increases. Even on small profiles with a survey step of up to 0.1...0.05 m the number of traces on the 1000 m GPR profile is already too high. The use of automated computer technology allows solving the problem. For processing lengthy or highly detailed GPR profiles, the need for using formal algorithms for processing GPR traces is obvious.

Each individual task of processing unique GPR profile is usually solved by a specialist of the appropriate qualification. However, in engineering practice, we can distinguish such tasks for which formal approaches have already been developed. Within these approaches, the work on processing a large number of GPR traces on long profiles should be performed by programs, and certain areas where the automated processing algorithm fails can be done by a highly qualified specialist.

One of such tasks for transport infrastructure objects is to identify the contrasting boundaries of soil layers, the characteristics of which are very different. These are the boundaries of "soft soil" – "strong soil", and the boundaries of soils in different conditions, for example, flooded soil – water repellent, frozen – thawed soil.

Today, in engineering practice, as a rule, semi-automatic algorithms of boundary search are used. The main idea of such algorithms is that the engineer sets the starting line and the starting point of the boundary on the trace. Then the program delineates a boundary between layers, and at any stage the engineer can interfere in process of construction and make corrections, and also change the parameters that control the operation of the algorithm [10, 11].

Algorithms based on the Hilbert transform (calculation of the signal envelope and instantaneous phase) [12], attribute and texture analysis [13, 14], cross-correlation function, and computer vision are usually used for semi-automatic boundary search.

The papers [15-17] show the application of deep learning technologies (convolution neural networks) to search for various local objects on radargrams. However, the problem of automatic search for boundaries without engineer intervention has not yet been solved.

Thus, the aim of this work is to develop, on the basis of deep learning methods (artificial neural networks), tools for processing large-scale GPR data obtained by examining media with contrasting boundaries.

2. Materials and Methods
Based on the experience of previously performed surveys, two groups of sites were adopted for the application of automatic GPR data processing.

Firstly, longitudinal GPR profiles obtained during the road survey. These profiles are characterized by a small number of boundaries with horizontal or subhorizontal position. When working with low-frequency antennas, the task is to determine the boundary of the embankment base and intermediate soil layers (Figure 1).

When working with high-frequency antennas (0.5..2 MHz) the task is different. Road pavement layers are detected at a depth of up to 0.5 m (Figure 2).

Figure 2, a shows the design of the top of the road pavement. The actual boundaries of the ground and structural layers are shown in Figure 2, b. The survey was carried out by the MALA GeoScience GPR unit (antenna central frequency is 1.7 GHz).
Figure 1. An example of a radargram and a geological section of an embankment and a bridge on the Yelizovo-Milkovo highway, Kamchatka Peninsula, Russia, 2015, the Loza-V GPR unit.

Secondly, GPR profiles obtained during the water area survey. The main task is to identify the water body bottom and one or two soil boundaries at the upper part of the section (Figure 3).
Figure 3. Scheme of the developed recurrent neural network.

The system of automatic search media boundaries based on a recurrent artificial neural network (on the data from GPR highways surveys) is developed. Each GPR profile is considered as a set of interconnected traces. Information about the position of boundaries on adjacent traces allows to specify the position of the boundary on the current trace. Sequences of measurements (traces) are fed into the network, at the output of the model, a sequence of the same length is expected with the selected boundaries.

3. Results
Processing of GPR profiles made by the OKO-2 unit with 450 MHz antenna is presented.

Two examples show that the results of automatic processing are close to the decoding performed manually. In the first GPR section (Figure 4, a) the boundary is contrasting and does not intersect with images from interference. The boundary line does not move to the lower layers. The result can be considered excellent. Deviations from the boundary delineated by the specialist can be considered not as a neural network error, but as a reason to check a radargram fragment again. Perhaps the error was made by the specialist who performed the processing manually.
The quality of a neural network can be evaluated formally, using a binary assessment – according to the percentage of coincidence of the results of automatic processing with the results of a “teacher”, i.e. a specialist. Only after transferring the time section to the depth section, it is possible to calculate the difference between the thicknesses of the two boundaries. Requirements for the accuracy of determining the thicknesses of soil boundaries in this case are determined by the layer velocities and equipment parameters.

The results obtained fully correspond to the accuracy of measurements specified during the survey.

4. Discussion

Despite the high quality of predictions, a number of drawbacks have been noted in the application of this approach. Firstly, the boundary between the layers of the pavement is usually at the top of the radargram, so the network is retrained for this location of classes and is unable to distinguish the boundaries of deeper layers. Secondly, considering the normalized initial signal of the OKO-2 GPR unit at a certain central frequency, it is impossible to work with other GPR units due to the specific distribution of amplitudes of the initial signal. Thirdly, the possibility of artificially expanding the source data by using augmentations is limited, except for reversing the sequence and adding noise to the data.

The described prevent both the introduction of models into GPR data processing practices and the rapid use of models for solving related problems (in particular, for automatic delineating of the bottom of water areas).

In the future, it is planned to modify the above approach and move from working with each trace to working with a GPR section as an image. This will open the possibility of using an image processing apparatus to solve the universal problem of finding boundaries in any conditions.

The following workflow is proposed:

![Figure 4. Examples of neural network boundary prediction results: solid line – manual delineating; dotted line – network prediction.](image)
1. Boundaries data is converted into layers data (transition from the classification task to the segmentation task); in this case, data augmentation can be used to artificially increase the amount of training data;
2. The data obtained are used to train a convolution neural network in two stages. At the first stage, only road data is used, after which at the second stage the network is retrained on new data (in particular, water area data)
3. The quality of network predictions is tested on a pre-delayed sample and, when satisfactory results are obtained, it is embedded in existing software for processing GPR data.

5. Conclusions
The quality of the result of radargram processing is determined by the engineer's competence, his experience, as well as the task properly assigned to him. The article proves the possibility of using neural networks in the processing of simple GPR profiles. In addition, the dependence of the result of automatic processing on the source samples is determined.

From a practical point of view, the use of such technologies is justified. It is necessary to control the results of computer calculations by a specialist. The advantage of automatic processing is a significant reduction in time spent on repetitive operations that the computer handles well.

The prospects for the development of automatic processing are related, on the one hand, to the improvement of the quality of the GPR profiles on which the neural network is learning, and, on the other hand, to the mathematically rigorous formulation of the signs of interference. It is planned to switch to neural network training on a small array of complex data after obtaining acceptable training results on a large number of simple GPR profiles.

References
[1] Lai W 2018 *NDT & E International* **96** pp 58 – 78 doi:10.1016/j.ndteint.2017.04.002
[2] Saarenketo T 2000 *Journal of Applied Geophysics* **43** pp 119 – 138 doi:10.1016/S0926-9851(99)00052-X
[3] Benedetto A 2017 *Signal Processing* **132** pp 201 – 209 doi:10.1016/j.sigpro.2016.05.016
[4] Benedetto A 2012 *Nondestructive Testing and Evaluation* **27** pp 253 – 262 doi:10.1080/10589759.2012.694884
[5] Pupatenko V 2017 *Procedia Engineering* **189** pp 643 – 649 doi:10.1016/j.proeng.2017.05.102
[6] Pupatenko V 2019 *MATEC* **265** 03003 doi:10.1051/matecconf/201926501001
[7] Sukhobok Y 2016 *Procedia Engineering* **143** pp 1236 – 1243 doi:10.1016/j.proeng.2016.06.110
[8] Stoyanovich G 2015 *Sciences in Cold and Arid Regions* **7** pp 0541 – 0546 doi:10.3724/SP.J.1226.2015.00541
[9] Sukhobok Y 2017 *Bearing Capacity of Roads, Railways and Airfields* **1** pp 1583 – 1587 doi:10.1201/9781315100333-227
[10] Liu H 2016 *Journal of Applied Remote Sensing* **43** pp 046023 – 046023 doi:10.1117/1.JRS.10.046023
[11] Varela-González M 2014 *Automation in Construction* **45** pp 42 – 49 doi:10.1016/j.autcon.2014.05.004
[12] Dossi M 2015 *Geophysics* **80** pp H23 – H35 doi:10.1190/IGCPR.2016.7572640
[13] Zhao W 2013 *Journal of Applied Geophysics* **97** pp 107 – 117 doi:10.1016/j.jappgeo.2013.04.010
[14] Zhao W 2016 *Pure and Applies Geophysics* **8** pp 2737 – 2751 doi:10.1007/s00024-016-1355-3
[15] Xu X 2018 *Scientific Programming* **2018** 4832972 doi:10.1155/2018/4832972
[16] Tong Z 2018 *Construction and Building Materials* **169** pp 69 – 82 doi:10.1016/j.conbuildmat.2018.02.081
[17] Tong Z 2017 *Construction and Building Materials* **146** pp 775–787 doi:10.1016/j.conbuildmat.2017.04.097