Reducing the Algorithmic Complexity of the Process of Determining the Reference Objects by Using Neural Networks

Boris Chernikov[1,2][0000-0002-6622-888X], Igor Gayduk[2][0000-0002-9759-4188] and Elena Chernikova[1][0000-0003-4453-6301]

[1] Plekhanov Russian University of Economics, Stremyanny lane 36, 117997 Moscow, Russia
[2] Moscow Institute of Electronic Technology, ld. 1, Shokin Square, Zelenograd, 124498 Moscow, Russia
bor-cher@yandex.ru

Abstract. One of the most important stages of the method of combining several point clouds into a single three-dimensional model - the stage of defining reference objects is considered in this paper. At this stage, the point cloud is processed in order to identify geometric primitives on it, the location data of which will allow for alignment. At the moment, such an analysis is performed in two stages: triangulation of the entire point cloud and search for primitives on the resulting surface. Since triangulation is an operation with quadratic algorithmic complexity, and the processed point clouds can contain millions and billions of points, the resource intensity of this stage becomes prohibitive. The use of the apparatus of artificial neural networks can eliminate the need to carry out such a resource-intensive operation, due to the ability to recognize patterns of primitives on raw data.

Keywords: triangulation, neural networks, 3D scanning, combining point clouds.

1   Introduction

Every year, 3D modeling technologies are increasing their presence in people's daily lives. From the class of entertainment technologies at the time of their inception, they have grown to an ordinary tool for researchers, engineers, designers and designers. The use of three-dimensional models allows you to reduce the cost of developing electronic circuits, cars, airplanes and many other industrial products, due to the early screening of bad samples without making physical models. 3D scanning technologies are inseparable from modeling tasks. In addition to their role as a tool to accelerate the creation of models, they are also an independent means of work and research. This technological innovation is being used to transfer architectural monuments to the digital space, which makes them immortal. In medicine, scanning a patient allows for high-precision prosthetics of damaged organs or body parts. With geodetic surveys, the time for obtaining data on the surface pattern is significantly reduced. In addition, thanks to the introduction of automated quality control systems for finished products, it is possible to improve the quality and volume of products on the LSI market.

In all areas, work is carried out with the final model of the object, but for its formation it is necessary to scan from a large number of angles, which leads to the problem of combining several point clouds. In work [1] the most universal and economical, from the point of view of resource intensity, the method of alignment by reference objects is recognized, but its use is not possible without detecting these objects.

2   Relevance and purpose of the study

In connection with the active increase in the presence of 3D scanning technologies in various areas of everyday life, there is a need to improve existing techniques in the areas of resource intensity, performance and versatility. A qualitative improvement of these parameters can be realized thanks to the development of tools, which include an improved algorithmic apparatus for key stages, including for combining several point clouds by determining reference objects.

The purpose of this article is to analyze the stage of defining reference objects in the existing method of combining several point clouds and identifying ways to improve it.

3   Determination of reference objects using triangulation

The classic way of defining reference objects is to use a mathematical apparatus to find characteristic primitives on a point cloud. This determination takes place in two steps.

The first step is to build a surface triangulation, since the raw point cloud resulting from the scan is not of sufficient practical value. After processing, one of the following possible results can be obtained [2]:
  · The same point clouds, but separated by morphological affiliation;
  · Digital elevation models;
  · Orthophotomosaics;
  · Recognized topographic objects;
  · Their models in three-dimensional representation.

The second step is to select fragments of planes that represent the contours of surfaces of characteristic objects. Most generally, this stage can be described as the selection of a certain number of adjacent triangles and an attempt to construct a plane from their vertices that approximates the surface under study. If, according to the results of checking by the least squares method, the total number of triangles falling on the plane does not exceed the threshold value, then this fragment is excluded from further consideration. The threshold value is selected depending on a number of parameters, such as point cloud density, scanning accuracy, and the like, and is usually 70% or more. Since this algorithm has a similar design both for planes and for sections of a more complex shape, consider the selection of objects such as a sphere, cylinder, cone, etc. on the triangulation model.

In general, the algorithm can be divided into the following stages [3]:

· A graph of topological connectivity of the faces is constructed. Two faces (vertices of the graph) have an edge if one face belongs to a triangle $T1$ adjacent to a triangle $T2$ belonging to another face;

A certain number of connected faces is selected, the number of which is sufficient to unambiguously determine the type of surface (the surface must be convex);

· For the resulting set of faces, a figure is constructed in the first approximation;

· By the points of the selected faces, a new approximation of the figure is iteratively constructed using the method of least squares;

· The obtained approximation is used to check whether other faces belong to this figure (Fig. 1). On the topological adjacency graph, when performing a breadth-first search, if a face with at least half of its points lies on a figure (with certain tolerances), then it is attached to this figure and is excluded from further consideration;

· If at the previous step enough faces were selected, then the shape is considered found, otherwise all changes are canceled.

Include a triangle in a shape with a positive point check
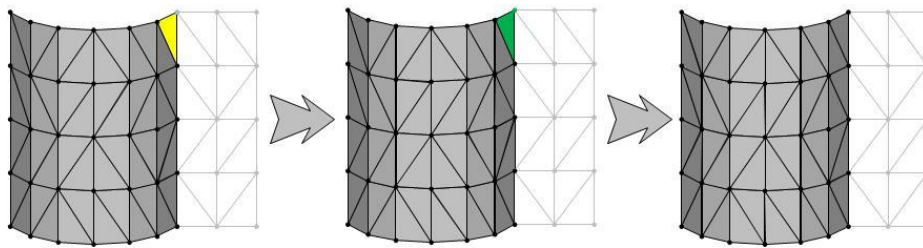
Ignoring triangle on negative point check
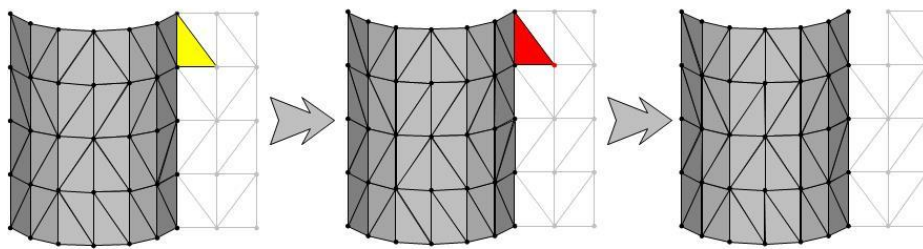
Fig. 1. Checking the belonging of the faces to the required figure

For each type of shape, further implementation is slightly different. To unambiguously select a sphere, you need four planes, for a cone and a cylinder - three, and for a cylinder, two of them must intersect with the third at right angles. Let's consider further actions for the sphere as for the most time consuming case.

Any sphere is uniquely defined through the four planes that touch it. To carry out the determination, it is necessary to solve a system of linear equations, each of which determines the distance from the plane to the center of the assumed sphere. The dis-

tance from the plane given by the equation $Ax + By + Cz + D = 0$ to the point $P_0 = (x_0, y_0, z_0)$ is calculated by formula (1).

$$d = \frac{Ax_0 + By_0 + Cz_0 + D}{\sqrt{A^2 + B^2 + C^2}} \qquad (1)$$

where d is the distance that takes into account not only the magnitude, but also the sign. It will be positive if the point P_0 and the origin lie on the same side of the plane. Since the coefficients of all four planes are known, then in formula (1) you can calculate the denominator and divide by it, which will give us r - the distance from the plane to the center of the sphere, without taking into account the sign.

$$r = |A`x_0 + B`y_0 + C`z_0 + D`|. \qquad (2)$$

To get rid of the modulus in formula (2), let us denote the sign in front of the radius r after the expansion of the modulus by $E`$ and obtain the following equation:

$$A`x_0 + B`y_0 + C`z_0 - E`r = -D`. \qquad (3)$$

Reducing all equations to the form (3), we obtain the system:

$$\begin{pmatrix} A`_1 & B`_1 & C`_1 & -E`_1 \\ A`_2 & B`_2 & C`_2 & -E`_2 \\ A`_3 & B`_3 & C`_3 & -E`_3 \\ A`_4 & B`_4 & C`_4 & -E`_4 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ r \end{pmatrix} = \begin{pmatrix} -D`_1 \\ -D`_2 \\ -D`_3 \\ -D`_4 \end{pmatrix}. \qquad (4)$$

The solution to system (4) will be the point $P_0 = (x_0, y_0, z_0)$ - the center of the sphere, which four planes touch, as well as its radius. Based on the data obtained, a final conclusion can be made about the correctness of the figure definition.

Thus, after defining the reference objects, the output contains a list of figures that have fallen into the scanning area and their coordinates. Based on this information, a characteristic map is compiled, which serves to compare two or more point clouds with each other.

4   Determination of reference objects using artificial neural networks mechanism

As an alternative to the classical method for finding reference objects, it is proposed to use the apparatus of neural network technologies. Recognition of reference objects by artificial neural networks, in essence, will be a multiclass classification, the result of which should be a verdict on the belonging of a fragment of the point cloud to the template of some geometric primitive.

This method can be roughly divided into three key stages.

1. Determination of areas of the original point cloud that potentially contain objects of interest to us. Determination of the area and its boundaries takes place in two steps:

· Taking several slices of the point cloud for analysis. The number of slices and their position relative to each other is dictated by the area of study;

· Detection by the neural network of the required images on the cloud of points and preservation of their position for further processing.

2. Triangulation of fragments of the cloud of points is carried out, on which the necessary images were recognized by the neural network apparatus.

3. This stage is no different from the second stage of the classical technique. Based on the available fragments, the figures are compared with the point cloud and the reliability of the constructed primitives is assessed.

The reason for working with slices of the point cloud, and not with the entire array of information, should be explained. The work [4] describes in detail the learning process of a neural network in terms of algorithmic complexity. Each iteration of training a classical backpropagation neural network contains a forward pass and a backward pass. Difficulty of direct passage $S_{d.p.}$ is calculated by the author of work [4] as follows:

$$S_{d.p.} \le n_0 n(\alpha + \beta) + n(\beta + \gamma) + (m - 2)[n^2(\alpha + \beta) + n(\beta + \gamma)] + nn_m(\alpha + \beta + nm\beta + \gamma, \tag{5}$$

where n_0 is the number of network inputs, n is the maximum number of neurons in all hidden layers of the network, n_m is the number of network outputs, m is the number of network layers, α are addition and subtraction operations, β is multiplication and division, γ is the calculation of the exponent.

For the complexity of backpropagation $S_{b.p.}$ the formula is valid:

$$S_{b.p.}$$
$$\le nn_m(5\alpha + 6\beta + 2\gamma) + n^2[3\alpha + 6\beta + 2\gamma + n_m(\alpha + \beta)]$$
$$+ (m - 3)n^2[3\alpha + +6\beta + 2\gamma + n(\alpha + \beta)]$$
$$+ nn_0[3\alpha + 6\beta + 2\gamma$$
$$+ n(\alpha + \beta)].$$

Since each iteration of the neural network training consists of a forward signal pass through all the connections, during which the result is calculated, and back propagation, during which the connection weights are adjusted, the total complexity of one training iteration will be equal to the sum of these two stages. Adding formulas (5) and (6) we get the following expression:

$$S_{p.} = S_{d.p.} + S_{b.p.} \sim O(n^3 m) + O(n^2 m) + O(n^2 n_m) + O(n^2 n_0) + (nn_m) + O(n) + + O(n_m), \tag{6}$$

where $S_{p.}$ is the algorithmic complexity of the training iteration of a classical neural network.

From expression (7), it can be seen that the number of neurons in hidden layers plays a key role in the complexity of training a neural network. The complexity of the problem solved by the neural network is also related to the number of these neurons.

Based on this, the conclusion suggests itself that a neural network apparatus is much more expensive than classical triangulation, but there are two points here:

a) cubic complexity concerns only the learning process, not the functioning of the neural network. If we take a closer look at formula (5), then

you can see that with forward signal propagation, the algorithmic complexity is at the quadratic level, which is a result close to triangulation;

b) training of the neural network is carried out at the stage of preparation for work. In the course of work, only direct passes are used, with the exception of reinforcement neural networks, but this is a separate topic.

Despite the points mentioned, the training of the neural network should still occur, and if a point cloud is fed to it in its original form, then any computing equipment will spend years learning on just one cloud. Because of this, it is logical to assume the use of slices, which will not only reduce the input volume of information, but also simplify the recognition task, transferring it from the category of three-dimensional to the category of analyzing the shape of a curve on a plane.

Determination of the area of triangulation becomes a completely trivial task when using neural networks. Vertically, the border can be traced by the presence of an object on the slices, and on the slices themselves, the window within which the network found the image will become frames.

Based on this, the question arises about the advisability of applying the new methodology, taking into account the fact that triangulation is used in it as well. For a correct assessment of this point, it should be borne in mind that triangulation in the improved technique is carried out on a set of small clouds, the total volume of which is several times less than the original cloud. In addition, do not forget that the total cost of triangulating one large cloud will be higher than for the same operation for many small ones, even if they are equal in terms of the final volume.

Let's note the main advantages and disadvantages of the proposed method using neural networks.

Disadvantages:

· It is inappropriate to apply the proposed methodology for one-time tasks, since it requires a resource-intensive learning process;

· It is necessary to have a labeled training sample.

Advantages:

· Faster performance in comparison with the classical method due to the smaller amount of input data;

· Lower computational complexity due to the definition of the type of object at the time of detection by the neural network, which allows triangulation of the object according to one template;

· High probability of successful matching due to the high probability of finding non-standard templates;

· The ability to work on a sparse point cloud.

Thus, the proposed technique using neural networks will reduce the resource consumption for determining the reference objects, but only if used in the long term. At this stage, it is impossible to name the exact figures of the performance gain, since in this issue it depends on the architecture of the neural network, the choice of which is the task of subsequent research.

5   Areas of further research

The sphere of applicability of three-dimensional modeling technologies is expanding every year. In addition to the listed opportunities for use in industry, various industries and medicine, there is a huge market for devices for obtaining information about the area, where a specialist poses a serious threat to life and health. This applies both to equipment scanning emergency monuments of architecture, and to serious automated devices that help rescuers in finding survivors in the aftermath of emergencies. Consequently, it is logical to expect that technologies for scanning space from many angles and obtaining a single three-dimensional model based on this data will progress and develop. To ensure the greatest user comfort, emphasis should be placed in the further development of technologies on the unification and reduction of the resource intensity of the process of obtaining the resulting model.

The purpose of further development of the process of determining reference objects on a set of raw data is to investigate various architectures of artificial neural networks to identify the most suitable for use within the framework of the accepted formulation of the problem. The most advantageous architecture will reduce resource costs for defining reference objects, as well as increase the accuracy of their recognition.

In order to achieve the set goals, it is necessary to solve the following tasks:

1. To develop and analyze the architectures of neural networks typical for image recognition tasks, since the task of finding reference objects on a raw data cloud is poorly formalized and will differ for different areas of research.

2. Form metrics for evaluating neural network architectures that take into account all key aspects of the subject area.

3. Justify and confirm the advantage of the chosen architecture in comparison with others, as well as with the classical method.

6   Conclusions

1. Based on the analysis of the methodology for determining the reference objects on the point cloud using preliminary triangulation of the entire cloud, the most resource-intensive stage of the selection of point cloud fragments was revealed, the improvement of which will allow to achieve a significant reduction in resource costs when compiling a single three-dimensional model in general and, in particular, when compiling a characteristic cards.

2. A new method for identifying reference objects is proposed, which allows, due to the elimination of the triangulation of the entire cloud and the compilation of a

characteristic map on the basis of raw data, to significantly reduce the time to obtain the result and reduce the algorithmic complexity of data processing from quadratic to logarithmic.

3. The goals and objectives for further research aimed at improving the methodology for obtaining a single 3D model based on the use of artificial neural networks have been formulated.

References

1. B.V. Chernikov, I.O. Gaiduk, E.A. Borisova The problem of creating a unified three-dimensional model of an object based on multi-angle scanning data. Modern science-intensive technologies. 2019; (10-1):83-91.
2. Danilin I.M., Medvedev E.M., Melnikov S.R. Laser location of land and forests. Geolidar, Geosmos, Moscow; Krasnoyarsk: Institute of Forest named after V.N. Sukachev SB RAS, 2007. 230 p.
3. Kostyuk Yu.L., Gulbin K.G., Peshekhonov S.V. Construction of surface triangulation and selection of spatial figures according to laser scanning data. Bulletin of the Tomsk State University. - Federal State Budgetary Educational Institution of Higher Professional Education "National Research Tomsk State University". 2006; (293):151-155.
4. Maksimushkin V.V., Arzamastsev A.A. Comparative assessment of the computational complexity of training an artificial neural network with a rigid core and a network with a classical structure. Bulletin of Russian universities. Maths. 2006; 11(2):190-197.