Research of Human Hand Movements Repeatability Using Robotic System

Paulius Sakalys
Vilniaus Kolegija University of Applied Sciences
Vilnius, Lithuania
p.sakalys@eif.viko.lt

Loreta Savulioniene
Vilniaus Kolegija University of Applied Sciences
Vilnius, Lithuania
l.savulioniene@eif.viko.lt

Dainius Savulionis
Vilniaus Kolegija University of Applied Sciences
Vilnius, Lithuania
d.savulionis@eif.viko.lt

Abstract - **The aim of the research is to determine and evaluate the repeatability of the robotic system by repeating the movements of the human hand, to identify the displacement using digital infrared projection equipment, skeletal methods and depth cameras. The article reviews and selects possible skeletal methods, motion recognition algorithms, reviews and substantiates the physical equipment selected for the technical stage of the experiment. The plan of experimental research stages, research stand, systematized research results, conclusions and usability suggestions are described.**

Keywords - robotic system, skeletal methods, IR projection, depth camera.

I. INTRODUCTION

There are many areas of activity where a sequence of interactions between a human and a robot could be implemented - a collaborative effort that requires training the robot or robotic system to observe, record, and accurately reproduce the movements of the human body.

The aim of this study is to investigate the accuracy of the robotic system control method based on skeletal methods and machine learning algorithms to determine the repeatability indices of the robotic system manipulator using digital optical means of human body motion recognition. The article reviews the most commonly used skeletal methods, the principles of operation of digital optical human body motion recognition devices, the methodology of their interconnection, and presents the results, conclusions and suggestions of the first tests of real robotic manipulator packaging.

II. LITERATURE REVIEW

Structured recognition of human movement is necessary for robotics movement control training. It is appropriate to use the skeletal method to identify the main human kinematic motion points. In order to select the most appropriate skeletal method, it is necessary to define the concept of skeleton and the issues related to the creation of a skeletal model.

Skeletal separation is important in a variety of fields: medical image processing, visualization, navigation, virtual endoscopy, computer graphics. Some authors often refer to the skeleton as the medial axis [4]. Set of maximum circles a maximum circle is one that is not completely drawn on any other circle within the surface o of the object in question [11] in three-dimensional space, instead of circles, maximum balls are drawn:

$$S_r = \{y \in R_3, d(x,y) < r\}, \qquad (1)$$

where r is the radius of the sphere (or circle in the case of two dimensions); x $O\in$ is the central point when 3 O R $\subset$; d (x, y) - distance between x and y points in three-dimensional space. Without any additional operations, the central axis becomes the central surface in three-dimensional space [1]. It can be argued that the skeleton is a subset of the central surface [9] [10]. By reviewing the relevant sources of information, two main types of skeletons can be distinguished:

Geometric skeleton - depicts the center of the object. The image does not always have to be one point wide (Fig. 1 a). Often the geometric skeleton is called the central axis [3], so the geometric skeleton does not necessarily have to be one pixel wide, but it always corresponds to the center of the object.

Skeleton of curves - a set of connected central lines is called a skeleton of curves (Fig. 1 b). The number of skeletal branches does not necessarily coincide with the branches of the skeletal object, so not every curved skeleton is at the same time a topological skeleton. The skeleton of the curves is one-dimensional everywhere except for the branching points, because the one-dimensional centerlines intersect at those points.
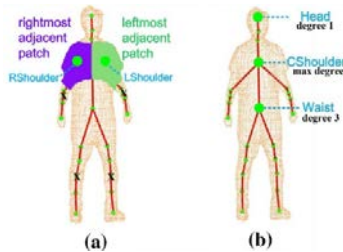

Fig. 1. The main types of skeletons [15]

Many skeletal extraction (skeletal) algorithms have been developed. Algorithms that implement skeletonization can be divided into the following groups [11]:

Iterative thinning. The main principle of this group of algorithms is the identification and destruction of simple points. Deleting a simple point does not affect the topology of the object. All iterative thinning algorithms operate in discrete space. Thinning iterations are performed until no simple points remain in the object. The main problem with this group of algorithms is limb clipping or uneven reconstruction, so additional conditions are introduced to preserve them. Also, the resulting skeleton is not always geometric, it is centered [8].

The distance field is formed when each point of the object is assigned a distance value to the nearest point of the surface. After calculating the distance field, an attempt is made to pave the way through the points where the distance to the boundaries of the object is the maximum. It is these points that correspond to the skeleton of the object [10]. The aggregation of points is performed using a gradient search or graph theory (searching for the shortest path between the selected maximum). The main advantage of these algorithms is that the distance field is quickly calculated.

Geometric methods are most commonly used in cases where the surface of an object is represented using a set of scattered points [2] or a grid of surfaces. These points are used to construct the Voronoi diagram, that is, to create the Voronoi cells. Each wall of a Voronean cell is equidistant from the vertices or points of the grid that constructs it (Fig. 2). The inner edges of Voronoi can be used to approximate the geometric skeleton. The main advantage of these algorithms is speed, because to determine a simple point it is enough to examine the position of its 3×3 adjacent points [9]. Additional transformation steps are required to isolate the one-dimensional skeleton.
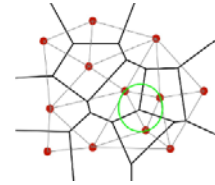

Fig. 2. Example of Voronoy's cell [7]

When creating a skeleton based on points in three-dimensional space, the change of human movement describing the coordinates using digitization means must take into account the calculation speed of the chosen skeletal method [13]. The fastest in terms of calculations is the latter described geometric method.

III. Methodology

The description of the research methodology is divided into four components: description of the research object (Part a), explanation of the principle of forming a digital skeleton of a moving object (human) (Part b), robotic system motion repetition training algorithm analyzes (Part c):

a. The object of the study is a robotic system (Fig. 3) - a robotic system printed by a 3D printer is identified with the human body, having two robotic arms with nine kinematic degrees of freedom - manipulators, and a head representing a two degrees of freedom robotic mechanism equipped with environmental video surveillance equipment. Servo drives are used for motion motors. The initial repeatability tests presented in the article will be performed with the right-hand movements of the model.

Fig. 3. Robotic system

b. The principle of digital skeleton formation of a moving object (human): The Kinect multi-infrared projection camera (Fig. 4 a) is used for the study. The device consists of a standard RGB VGA camera, a depth camera, an infrared projector, an array of microphones and a controller.

For the specified cameras and infrared projector, the device forms a three-dimensional image field in which it distinguishes moving points (Fig. 4 a):

1. The depth camera sensor illuminates the real space and determines the distances specified for each pixel of the image based on the reflections of the infrared point projections [14].

2. The software that combines the depth camera results and the RGB camera image creates a map of the depth pixel projections of each object in front of the camera, and creates RGB textures based on the location corresponding

to the RGB pixels. All information is generated in real time at a rate of 30 times per second.

3. The processed data is sent to the middleware on the receiving computer.
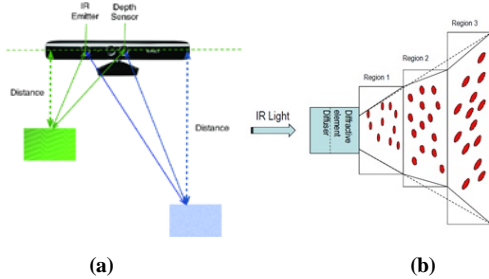

(a)   (b)

Fig. 4. Kinect Multi-Infrared Projection Camera and Formation of Three-Dimensional Image Field Points by Infrared (Stitilis, 2014)

The first version of the Kinect used for the study returns the joint positions with a three-dimensional vector. The OpenNI library (for access to the cameras) used on the receiving computer provides a function that returns the position of the joint in the orientation matrix (Fig. 5):

```
public float getJointOrientationSkeleton(int userId, int joint, PMatrix3D jointOrientation) {
    float[] mat = new float[9];

    float ret = getJointOrientationSkeleton(userId, joint, mat);

    jointOrientation.set(mat[0], mat[1], mat[2], 0,
                         mat[3], mat[4], mat[5], 0,
                         mat[6], mat[7], mat[8], 0,
                         0,      0,      0,      1);

    return ret;
}
```

Fig. 5. An OpenNI library function that outputs a position orientation matrix

After the calculations, the value of the returned matrix changes only in the first 3 rows. Therefore, we can reduce the size of this matrix to 3x3. Using Euler's characteristics for a given 3x3 matrix (Fig. 6):

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

Fig. 6. Transformed orientation coordinate matrix

The steering angles at the three coordinate points are calculated using formulas 1-3:

$$\theta_x = \text{atan2}(r_{32}, r_{33}), \tag{2}$$

$$\theta_y = \text{atan2}(-r_{31}, \sqrt{r_{32}^2 + r_{33}^2}), \tag{3}$$

$$\theta_z = \text{atan2}(r_{21}, r_{11}), \tag{4}$$

The coordinates of the obtained moving points are linked to each other to form a skeleton.

c. Robot system motion repetition training algorithm: positions of the obtained skeleton coordinate points - gestures are stored every 30 milliseconds in the control program, giving access rights to call them. Robot system training algorithm (Fig. 7).

Fig. 7. Robotic system training algorithm

The sequence of the motion capture function of the Kinect camera is shown in the activity diagram. When using this feature, the client makes a "Copy" request in the command input field or with voice commands. This request starts the repetition of the robot's movements with the Kinect camera. As already mentioned, the positions of the robot servo motors are saved in a new file called "positions" during the entire repetition of the movements. The customer can stop repeating the robot's movements at any time by sending the command "Off". When this is done, the robot repeats the movements again this time taking the positions of the servo drives from the generated file. The user is then asked if they want to save this gesture. For the gesture to be saved successfully, it needs to be given a name. The name must be unique because it is used in two places. The first place is to generate a Python script. This name is used to identify the script file. Also, in order to call the generated method, one of the AIML type files needs to be edited. The AIML interpreter is used to call this function during program operation.

AIML is a recursive language based on XML [6], which allows to parse natural language text input so that it

corresponds to the response that a robotic system can send (Fig. 8).

```
<?xml version="1.0" encoding="UTF-8"?>
<aiml>
    <category>
        <pattern>HELLO WORLD</pattern>
        <template>Hello to you too!</template>
    </category>
</aiml>
```

Fig. 8. Example of AIML code structure

The first "<aiml>" tag indicates the beginning and end of the AIML document. The <category> tag defines the unit of knowledge in the robot database. The number of these tags in AIML-type files is often large, so they simply separate each unique query content, of which we have only one in example 8. Thus the robotic system is trained to identify, record and reproduce the movement of the human body.

IV. RESEARCH RESULTS

The test procedure of the performed study is shown in Fig. 9. The training part of the robotic system is performed first: the fixed image is scleted, the motion is recognized, the point matrix is interpreted and obtained and converted into a coordinate motion point matrix of the kinematic connection servo drives, which is saved in the form of an AIML file and stored in a database. Part of the next study: the motion matrix is called from the database, the servo drives create a displacement of the manipulator axes (the manipulator has a built-in writing device with which it must reach the initial reference point memorized during the training part). When the system reaches a possible reference point, it returns to its original servo drive coordinates, the deviation data is recorded, the motion matrix is recalled from the database, and the test is repeated (20 times).
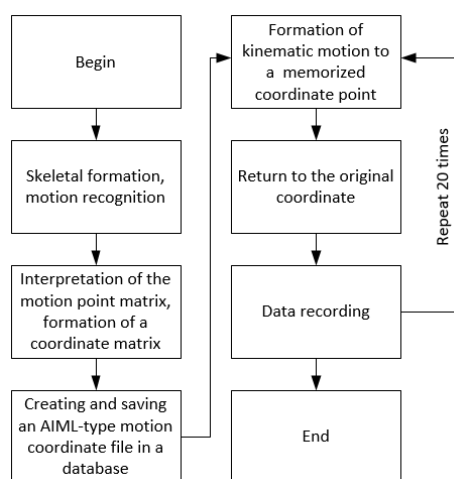
Fig. 9. Test Flow Chart

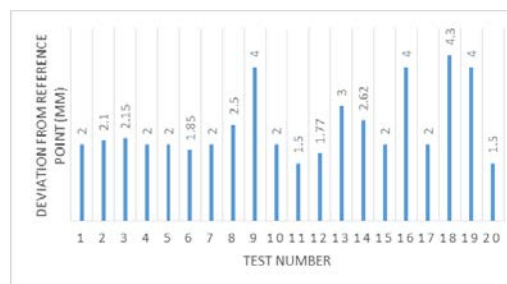Data for all twenty test deviations are shown in Fig. 10 of the diagram.

Fig. 10. Deviation from the Reference Point Diagram

V. CONCLUSIONS

The study developed a primary methodology that combines skeletal methods, machine learning algorithms and digital optical equipment to identify human body motion by creating a coordinate system and storing it in a robotic system database with the ability to retrieve servo drive motion displacement for manipulator freedom.

1. After reviewing the most suitable skeletalization methods available, the fastest geometric skeletalization method in terms of calculation was selected;

2. Kinect digital optical infrared projector used to isolate the human body from the environment;

3. OpenNI library was used for three-dimensional vector processing, algorithm for calculation, setup, recording and calling of motion angles was developed using AIML software interpreter;

4. Tests with a robotic right-hand manipulator (hand) containing a writing instrument to reach the reference point during the tests have a minimum deviation of 1.5 mm, a maximum of 4.3 mm and a mean deviation of 2.46 mm from the fixed starting point of the training point;

The description of the performed research methodology and the obtained research results can be used in areas where synergistic interaction of human and robotic system synchronous movements is possible: performing dynamic, time-varying functions that do not operate according to a predetermined sequence of actions. The results of the study also have a wide range of applications in the training process: researching the properties of machine training, the development of robotics, control systems for robotic systems and algorithms.

REFERENCES

[1] Carmelo, M., Acutis, A., Carrabba, M., Criscenti, G., Vozzi, G. (2016). Nanobiomaterials in Soft Tissue Engineering: Applications of Nanobiomaterials. Volume 5. Machine design for multimaterial processing, 111-140. DOI: https://doi.org/10.1016/B978-0-323-42865-1.00005-2

[2] Chaminade, T., & Cheng, G. (2019). Journal of Physiology-Paris. Volume 103. Social cognitive neuroscience and humanoid robotics, 286-295. DOI: https://doi.org/10.1016/j.jphysparis.2009.08.011

[3] Davies, E.R. (2017). Computer Vision. London: AcademicPress.

[4] Jin, D., Chen, C., Hoffman, E.A., Saha, P.K. (2017) Skeletonization: Theory, Methods and Applications. Curve skeletonization using minimum-cost path, 151-180. DOI: https://doi.org/10.1016/B978-0-08-101291-8.00007-9

[5] Kar, A. (2011). *Skeletal Tracking using Microsoft Kinect.* Kanpur

[6] Nava, A., Brena, R., Garrido, L. (2014). Recognizing Activities Using a Kinect Skeleton Tracking and Hidden Markov Models. *13th Mexican International Conference on Artificial Intelligence (MICAI),* DOI: https://ieeexplore.ieee.org/document/7222846.

[7] N.D. *Example of Voronoy's cell.* Retrieved from: https://stackoverflow.com/questions/42047077/voronoi-site-points-from-delaunay-triangulation

[8] Palionyte, A (2011). *Discretization of Continuum Structures via Image Algebra Methods.* Vilnius: Vilnius Gediminas Technical University.

[9] Paulinas, M., & Rokicki, J. (2008). Panasios formos pavirsiu, centrines asies radimo ir filtravimo metodika. *Biomedical Engineering,* 209-212.

[10] Rokicki, J. (2010). Skeletavimo metodu apzvalga. *Science - Future of Lithuania, Vol. 2 (No. 1),* 19-22.

[11] Saha, P.K., Borgefors, G., Baja, G.S. (2017). Skeletonization: Theory, Methods and Applications. *Skeletonization and its applications – a review*, 3-42. DOI: https://doi.org/10.1016/B978-0-08-101291-8.00002-X

[12] Stitilis, R. (2014). *Kinect Controlled Computer Game.* Siauliai: Siauliai University

[13] Vassilevski, Y., Olshanskii, M., Simakov, S., Kolobov A., Dailov, A. (2020) Personalized Computational Hemodynamics: Models, Methods, and Applications for Vascular Surgery and Antitumor Therapy. *Computational hemodynamics in vascular surgery, 189-215. DOI: https://doi.org/10.1016/B978-0-12-815653-7.00009-9*

[14] Palagyi, K. (n.d.) *Skeletonization.* Retrieved from: http://www.inf.u-szeged.hu/~palagyi/skel/skel.html

[15] Yong, Z., Fei, T., Shaofan, W., *3D human body skeleton extraction from consecutive surfaces using a spatial–temporal consistency model.* Retrieved from: https://link.springer.com/article/10.1007/s00371-020-01851-3