Gesture Recognition of Somatosensory Interactive Acupoint Massage Based on Image Feature Deep Learning Model

Yukun Jia[1], Rongtao Ding[2], Wei Ren[2], Jianfeng Shu[3], Aixiang Jin[4*]

[1] Special Equipment Institute, Hangzhou Vocational & Technical College, Hangzhou 310018, China
[2] School of E-commerce, Zhejiang Business College, Hangzhou 310053, China
[3] Department of Acupuncture and Massage, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310014, China
[4] Department of Head, Neck & Thyroid Surgery, Zhejiang Provincial People's Hospital, People's Hospital of Hangzhou Medical College, Hangzhou 310014, China

Corresponding Author Email: jinaixiang77@163.com

ABSTRACT

During rehabilitation, many postoperative patients need to perform autonomous massage on time and on demand. Thus, this paper develops an individualized, intelligent, and independent rehabilitation training system for based on image feature deep learning model acupoint massage that excludes human factors. The system, which innovatively integrates massage gesture recognition with human pose recognition. It relies on the binocular depth camera Kinect DK and Google MediaPipe Holistic pipeline to collect the real-time image feature data on joints and gestures of the patient in autonomous massage. Then the system calculates the coordinates of each finger joint, and computes the human poses with VGG-16, a convolutional neural network (CNN); the calculated results are translated, and presented in a virtual reality (VR) model based on Unity 3D, aiming to guide the patient actions in autonomous massage. This is because the image feature of the gesture recognition and pose recognition is hindered, when the hand or the human is occluded by the body or other things, owing to the limited recognition range of the hardware. The experimental results show that, the proposed system could correctly recognize up to 84% of non-occluded gestures, and up to 93% of non-occluded poses; the system also exhibited a good real-time performance, a high operability, and a low cost. Facing the lack of medical staff, our system can effectively improve the life quality of patients.

1. INTRODUCTION

The advancement of science and technology has promoted the application of artificial intelligence, and brought progress in human pose and gesture recognition, turning it into a novel way of somatosensory human-computer interaction [1-4]. In recent years, the technology of human pose and gesture recognition is widely applied in many fields, such as artificial intelligence [5], electronic games [6], smart home [7], sports training [8], teaching and training [9, 10], traffic guidance [11, 12], to name a few.

In addition, perception technologies like human pose and gesture recognition boast application prospects in life health and other medical fields [13]. For example, Liu et al. [14] used Leap Motion to capture the real-time gestures of patients during rehabilitation exercise, which realizes the somatosensory interaction between the upper limbs and the virtual environment, and stimulates their willingness for self-rehabilitation, and constructed a low-cost evaluation system for self-rehabilitation exercise of patients. With the pose estimation method of OpenPose, Hang et al. [15] processed the key action features in pose videos, extracted the joints from human skeleton, and proposed a human rehabilitation action recognition algorithm based on the gated recurrent unit (GRU) network, which improves the recognition accuracy of

rehabilitation actions to 98.14%. Chen [16] captured the real-time data on hand movement with Leap Motion somatosensory sensor, modeled the time-varying finger joint angle of grasping by analyzing the kinematics data on joints, designed a multi-joint virtual rehabilitation system for the upper limbs, and realized the interactive control of virtual joint movements, making the relevant rehabilitation training more immersive. Wang et al. [17] developed a virtual environment training software, which relies on Leap Motion somatosensory sensor to track the real-time position and gesture of the patient's hands; the patient is allowed to move the virtual hand naturally, and their system will give the patient visual and vibrotactile feedbacks, making his/her rehabilitation training more effective. Yang et al. [18] combined Kinect DK, an automatic human pose recognition technology, with medical rehabilitation, creating a medical rehabilitation action database, presented a pose judgement method based on the Euclidean distance and angle between joints, and applied the method to medical rehabilitation training. Qian et al. [19] adopted Kinect DK to extract the data on human poses, proposed an action model to extract human movement features from the joint angle series, and developed a Kinect DK-based system for habilitation trainers.

Human pose and gesture recognition is mainly realized based on various sensors or machine vision image analysis. (1)

Sensor-based human pose and gesture recognition: Zeng et al. [20] designed a portable, lightweight, and intelligent gesture recognition device based on elastic conductive coated yarns; the device adopts the gloves made of silver-plated nylon spandex with strain-resistance effect. The sensor-based recognition methods can achieve an accuracy as high as 96%, without being controlled by the ambient light. However, the relevant devices are inconvenient to use, because the user must wear a variety of sensors. (2) Machine vision-based human pose and gesture recognition: Ravinder Ahuja et al. [21] built a gesture dataset from OpenCV, and extended convolutional neural network (CNN) into a machine vision control system based on gesture actions. Despite its simplicity, the machine vision-based methods are less accurate than the sensor-based methods. Besides, this type of methods is susceptible to the ambient light; their recognition rate drops sharply when there is occlusion.

To sum up, the traditional gesture and posture recognition only improves the recognition rate from the algorithm, and does not integrate two single recognition methods into the same device and apply them. Taking advantage of somatosensory interaction, this paper develops an individualized, intelligent, and independent rehabilitation training system for acupoint massage that excludes human factors. In the system, the binocular depth camera Kinect DK is adopted to collect the data on joints and gestures of the patient in autonomous massage, and the coordinates of each finger joint are calculated; the poses of the patient are recognized by the CNN, and translated for display in a virtual reality (VR) model based on Unity 3D, aiming to guide the patient actions in autonomous massage.
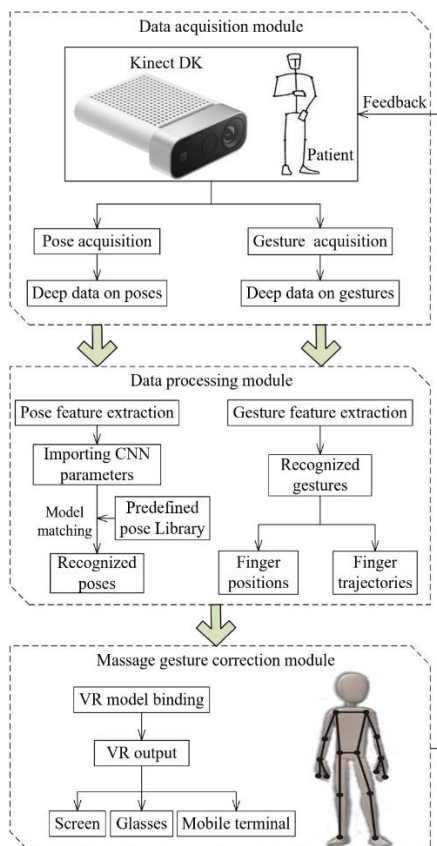
2. SYSTEM CONSTRUCTION

Figure 1. System modules

As shown in Figure 1, our system consists of three modules: a data acquisition module, a data processing module, and a massage gesture correction module.

(1) The data acquisition module collects the body and hand movements of the patient with the deep binocular camera Kinect DK. After removing the invalid frames, the deep data on poses and gestures are computed, and stored in the backend database, providing original data for subsequent modules.

(2) Based on deep learning, the data processing module extracts the pose features of the patient, imports the CNN parameters to match the pose features with the predefined pose action library, and outputs the recognized poses. In addition, the hand features of the patients are extracted, followed by the calculation of finger positions, finger joint positions, and optimal finger trajectory under the coordinate system of Kinect DK.

(3) According to the results on poses and gestures, the massage gesture correction module introduces the kinematics model parameters, binds the patient's poses and gestures to the VR model, and translates them for display in the VR model. The results, outputted via the screen, glasses, or terminal, guide and correct the patient's massage actions in real time, and ensure the quality and quantity of postoperative self-rehabilitation training.

3. DEEP DATA COLLECTION

The Azure Kinect DK, launched by Microsoft in 2019, was selected as the binocular depth camera. Using Kinect DK to collect the deep data on poses and gestures. It is a developer toolkit that integrates advanced artificial intelligence sensors. Kinect DK contains a 1 million pixel time of flight (TOF) depth camera, a 360° 7-microphone array, a 12 million pixel full high-definition camera, and a direction sensor. This small and portable binocular depth camera supports complex computer vision and voice models [22].

(1) Kinect DK coordinate system

Kinect DK has a three-dimensional (3D) coordinate system, with every sensor as the origin [0, 0, 0]. That is, the 3D coordinate system consists of multiple subsystems: the depth camera coordinate system, the color camera coordinate system, the gyro coordinate system, and the accelerometer coordinate system. Among them, the depth camera coordinate system is selected for this research. Every second, Kinect DK generates 30 frames of data per second, and captures the depth 3D coordinates of 25 key points in real time, providing the data support to pose and gesture recognition.

(2) Digital skeleton of key points

Using a ToF imaging chip, Kinect DK realizes a high modulation frequency and depth accuracy, and operates at a rate as fast as 30 frames per second (fps). The key points of the human body were computed from the depth image via ToF ranging: a light pulse is projected into the target space, and the depth data of the reflection points is derived from the phase difference of the reflected lights. Int this way, 32 joints are obtained and tacked. The position and direction of each joint form a coordinate system of that joint. The coordinate systems of all the joints make up a digital skeleton.

(3) Fingertip coordinates

Kinect DK provides 4 joint points for each hand, which is not enough for the recognition of massage gestures. Thus, Kinect DK is supplemented with Google MediaPipe Holistic pipeline to recognize the gestures. Then, the spatial

coordinates and fingertip direction of each finger are collected, and used to deduce the spatial coordinates of other joints of the finger.

4. ACUPOINT MASSAGE GESTURE RECOGNITION

In our system, the binocular depth camera Kinect DK is supplemented with Google MediaPipe Holistic pipeline to detect and track the real-time position and direction of each palm and finger. Each second, the system can capture 30 frames of data at the most. The working range is 0.5-5m. The following data can be tracked and collected in real time by our system: (1) The spatial coordinates of the five fingertips, mm; (2) The direction vectors of the five fingers, i.e., the direction from the root to the tip of each finger, when the hand is fully open. The spatial coordinates of fingertips and direction vectors are determined based on the 3D coordinate system K_D of the Kinect depth camera.

To recognize the gestures in acupoint massage, the primary task is to extract the coordinates of the finger positions of the self-rehabilitation patient during the massage, and then derive the coordinates of all joints in the hand, as well as the optimal trajectory of each fingertip. On this basis, the patient's massage actions will be guided and corrected in real time through the natural interaction in the virtual environment of the self-rehabilitation system, such that the postoperative patient can complete sufficient high-quality massage training and tasks autonomously during the self-rehabilitation.

4.1 Determining the coordinates of each finger joint

Take the tip of the middle finger as an example. Under ideal conditions, the middle finger is considered as a cylinder (Figure 2), with its center point as the origin O.
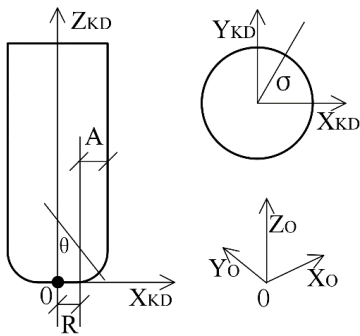

Figure 2. Coordinates of fingertip

Suppose the fingertip of the middle finger is an ideal partial spherical surface. Let A be the radius of the fingertip; R be the distance between the contact point and the origin on the centerline of the middle finger during massage. In the global coordinate system $(O-X_OY_OZ_O)$, any direction on the fingertip is selected as the X_{KD} axis, with the center point O of the middle finger as the origin, and the direction along the centerline of the finger as the Z_{KD} axis. Then, Y_{KD} is determined by the right-hand rule, producing the local coordinate system $(KD-X_{KD}Y_{KD}Z_{KD})$, where the origin is the center point O of the middle finger. During the massage, the contact point falls on the projection plane of X_{KD}–Y_{KD}. Let α σ be the angle between the normal vector of the contact point; θ be the angle between the normal vector of the spherical surface at the contact point and the Z_{KD} axis. Then, the parameter

equation of the contact surface in the local coordinates of the fingertip $(KD-X_{KD}Y_{KD}Z_{KD})$ can be established as:

$$\delta_{KD}(\varphi,\theta) = \begin{bmatrix} (R + A\sin\theta)\cos\theta \\ (R + A\sin\theta)\sin\theta \\ A - A\cos\theta \end{bmatrix} \sigma \in [0,2\pi], \theta \in \left[0,\frac{\pi}{2}\right] \tag{1}$$

In the above global coordinate system $(O-X_OY_OZ_O)$, it is assumed that the surface equation of the body part contacted by the tip of the middle finger is $S=S(\mu, v)$. Then, the arbitrary trajectory of the tip of the middle finger during the massage can be described as:

$$P_i = S(\mu(x), v(x))(i = 1,2,3,\cdots, n) \tag{2}$$

On the i-th trajectory, the contact point between the tip of the middle finger and the massaged body part can be expressed as:

$$D = D_{i,j} = S(\mu(x_j), v(x_j))(S = S(\mu, v) \quad j = 1,2,3,\cdots, n) \tag{3}$$

Taking the current middle finger tip, i.e., the contact point D of the massaged part, as the origin, the Z_D axis is established on the normal direction of the surface of the massaged part, the X_D axis is established with the tangent of the massage direction as the positive direction, and the Y_D axis is determined by the right-hand rule, creating a local coordinate system for the contact point of the middle finger tip on the massaged part: $(D-X_DY_DZ_D)$.

From the spatial geometry, it is learned that the spherical surface of the middle finger tip is tangent to the surface of the massaged part at D. In $(D-X_DY_DZ_D)$, the following parameters can be determined: the spatial position of middle finger tip (X_D, Y_D, Z_D), the inclination angle α of the middle finger relative to the massaged part, and the rotation angle β of the middle finger. Specifically, the inclination angle α refers to the angle between the middle finger and the normal surface of the massaged part, and the rotation angle β stands for the angle between the massage direction and the centerline of the middle finger. According to the relative positions, $(KD-X_{KD}Y_{KD}Z_{KD})$ can be transformed into $(D-X_DY_DZ_D)$ by:

$$[X_{DK}, Y_{DK}, Z_{DK}] = [X_D, Y_D, Z_D]M_a \tag{4}$$

where, M_a is a matrix:

$$M_a = \begin{bmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}$$
$$= \begin{bmatrix} \cos\beta\cos\alpha & -\sin\beta & \cos\beta\sin\alpha \\ \sin\beta\cos\alpha & \cos\beta & \sin\beta\sin\alpha \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix} \tag{5}$$

During the massage, the center point O of the middle finger moves to the middle finger tip, i.e., the contact point $D_{i,j}$ of the massaged part. At this moment, the Z_D axis of the two local coordinate systems coincide. Suppose the middle finger continues with the original trajectory. Let R be the movement distance of the middle finger along the X_D axis; α be the rotation angle of the middle finger about the axis Y_D, with the center point O as the origin; β be the rotation angle of the middle finger about the axis Z_D, with KD as the origin. Then,

the 3D spatial position of the origin of the middle finger can be described as:

$$O = D + AZ_D - RX_D \qquad (6)$$

From formulas (4) and (6), we have:

$$\begin{cases} O = D - R\cos\beta\cos\alpha X_{KD} - R\sin\beta\cos\alpha Z_{KD} + (A + R\sin\alpha)Y_{KD} \\ Z_{KD} = \cos\beta\sin\alpha X_{KD} + \sin\beta\sin\alpha Z_{KD} + \cos\alpha Y_{KD} \end{cases} \qquad (7)$$

According to formula (7), the spatial position of the middle finger can be determined, as long as the following three parameters are known: the contact point D of the massaged part, the inclination angle α of the middle finger relative to the massaged part, and the rotation angle β of the middle finger. Then, the spatial position of every angle on the middle finger can be derived by the plane quadrilateral model [23] and the law of cosines. By analogy, it is possible to obtain the spatial coordinates of all the fingers of the hand.

4.2 Optimizing fingertip trajectory

The massaged surfaces on human body differ greatly. Some body parts are convex, and some are concave. Let (x_i, y_i, z_i) be the coordinates of a random point selected from the massage trajectory on a concave surface. Then, a coordinate plane O-XZ is established at the point $(0, y_i, 0)$. This plane is perpendicular to the Y axis, and intersects the concave surface of the massaged part along the trajectory S. The trajectory is explained in Figure 3, where dm is the tiny increment of each finger movement.
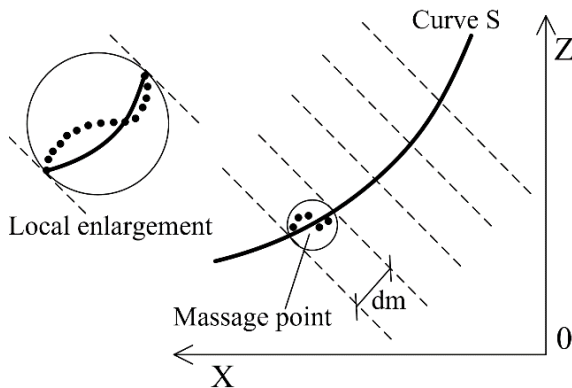

Figure 3. Fingertip trajectory during massage

Then, the function G=f(d, s) is set up for the message process, where parameter d is any point on the massage trajectory S. Let G be the shortest normal distance of point d along trajectory S, and dm be the tiny increment of each finger movement. Then, the t contact points within the dm-long trajectory, $G_1, G_2, G_3,\ldots, G_t$, can be solved. Since the t points are randomly arranged on the ideal trajectory, there must be an inflection point, that is, the crest point or the trough point.
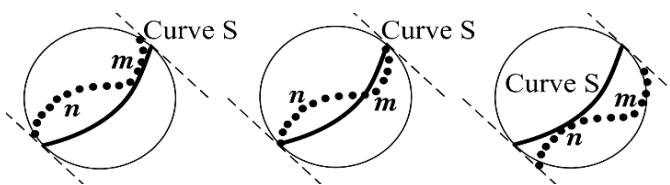

Figure 4. Distribution of contact points on massage trajectory S

Figure 4 shows the distribution of contact points on massage trajectory S. There are three possible distributions:

(1) If the contact points are inside trajectory S, then the function G=f(d, s) maximizes at point n and minimizes at point m. In this case, point m is the contact point for all the trajectories in the tiny increment, that is, the function value equals G_m.

(2) If the contact points are on both sides of trajectory S, then the function G=f(d, s) maximizes at points m and n, with point n on the outside. In this case, point n is the contact point for all the trajectories in the tiny increment, that is, the function value equals G_n.

(3) If the contact points are outside trajectory S, then the function G=f(d, s) maximizes at point m and minimizes at point n. In this case, point n is the contact point for all the trajectories in the tiny increment, that is, the function value equals G_n.

Depending on the situation of the above three cases, it is not difficult to fit the optimal trajectory for each fingertip.

5. DEEP LEARNING-BASED POSE RECOGNITION FOR ACUPOINT MASSAGE

For pose recognition, the binocular depth camera Kinect DK collects the depth images on the patient's poses during the autonomous massage. Then, the pose features are extracted from the images, and imported to the CNN for matching with the predefined pose library, aiming to recognize the patient's massage poses in self-rehabilitation.

The CNN is very suitable for image processing. With the speed improvement of computer hardware, the CNN has been widely adopted for face recognition [24, 25], gesture recognition [26, 27], human pose recognition [28], license plate recognition [29, 30], security, and other fields. In this paper, our CNN for pose recognition in acupoint massage includes four parts: convolutional layer, activation layer, pooling layer, and fully connected layer.

(1) The convolutional layer is the feature extraction layer. The convolution kernels and the activation function work together to extract specific data, and combine them into a feature map. The convolutional operation can be expressed as:

$$C(i, j) = \sum_{h=1}^{h} \sum_{w=1}^{w} x_{i+h, j+w} \times w_{h,w} + b \qquad (8)$$

where, C(i, j) is the convolution result; i and j are the coordinates of the kernel; h and w are the height and width of the kernel, respectively.

(2) The activation layer adopts the rectified linear unit (ReLU) function:

$$f(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases} \qquad (9)$$

(3) The fully connected layer connects the data of convolutional layers and those of pooling layers, and output the feature information. The output of the fully connected layer can be expressed as:

$$h_{w,b}(x) = \theta(w^T x + b) \qquad (10)$$

(4) VGG-16-based pose recognition

The depth images, which are in the size of 224×224, on massage poses are collected by Kinect DK. The deep learning algorithm is VGG-16, a CNN developed by Computer Vision group from the University of Oxford [31], in association with Google DeepMind. Our CNN contains 13 convolutional layers(3x3), 5 pooling layers(2x2), 3 fully connected layers(4096), and 1 output layer (softmax). The kernel size is set to 3×3.

The three 224×224 depth images collected by Kinect DK are imported to VGG-16. Then, the data flow is as follows: two convolutional layers, two convolutions + ReLU operation, max pooling layer, two convolutional layers, max pooling layer, three convolutional layers, max pooling layer, three convolutional layers, max pooling layer, three convolutional layers, max pooling layer, fully connected layer, outputting recognized poses and performing softmax classification. To ensure that the system can really recognize the patient's gestures, (Figure 5).
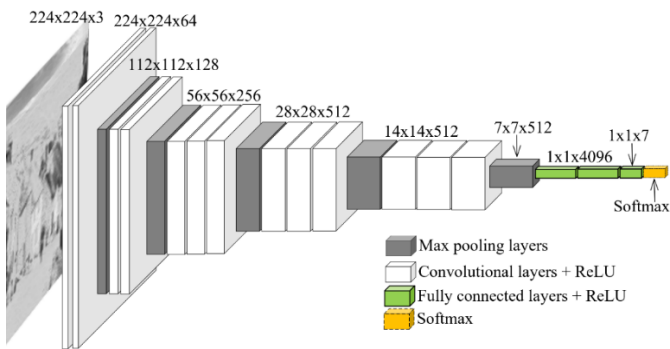

Figure 5. Pose recognition model

6. EXPERIMENTS AND RESULTS ANALYSIS

6.1 Software and hardware environments

The following software and hardware environments are established to verify the accuracy and effectiveness of our system.

Hardware environment: Intel Core I7-6850K@3.6GHz 6-core 12-thread CPU; DDR4-3200 32GB memory; Samsung 970evo plus 500G hard disk; NVIDIA GeFoce1080 GTX GPU with 8GB video memory; Kinect DK binocular depth camera.

Software environment: operation system: Windows10, 64-bit professional edition; key point detection and tracking tool: Google MediaPipe Holisticipeline; VR platform: Unity3D engine.

The test environment is well lighted with no occlusion between Kinect DK and testers.

6.2 Effectiveness of massage gesture recognition

The gesture data come from the fingertip data collected by Kinect DK and Google MediaPipe Holistic pipeline. The actual massage gestures of traditional Chinese medicine (TCM) are collected in real time, producing a series of gesture movements. The data are processed in the unit of frames by removing the invalid frames and retaining the valid ones. A total of five massage gestures commonly used in TCM massage are selected for the test: thumb massaging manipulation, pressing manipulation, kneading manipulation,

grabbing manipulation, and pulling manipulation [32]. Each gesture is tested for more than 8s.

To ensure the applicability of the system, the binocular deep camera Kinect DK is placed on a tripod. The left and right of the lens directly face the palm of the patient, while the vertical direction faces the patient with a certain angle. The gestures are made 1.0-2.5m away from the camera, and tested at 1.0m, 1.5m, 2.0m, and 2.5m, respectively, for more than 2s. By the method specified in Section 4, the coordinates of each joint in the hand are obtained, and rendered with the Unity3D engine. Then, the gestures are displayed on the screen in a virtual form (Figure 6), where the joints are presented as black dots and linked up with black lines.
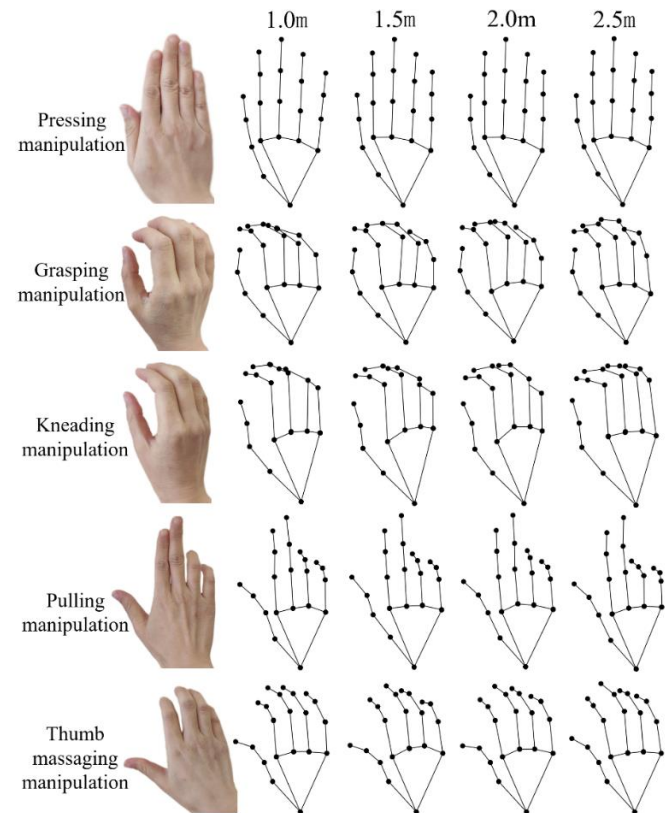

Figure 6. Massage gesture test

During the test, the palm faces the Kinect DK, but the back of the hand faces the camera. For convenience, the recognized gestures are rotated by 180°. As shown in Figure 6, some recognized gestures are slightly deformed, with an obvious jitter during screen playback; the best recognition effect is achieved at the distances of 1.5m and 2.0m; when the finger is bent too much, part of the hand is occluded, causing falling recognition rate and intensified jitter.

6.3 Autonomous rehabilitation system test

In the system test, 40 patients aged about 35-55 were selected, including 20 males and 20 females. During the test, the patient stands at 2m away from the camera of Kinect DK, with no occlusion over the distance. Kinect DK is mounted on a tripod, and the center point of the camera is 1.2m above the ground. The screen is placed at 0.5m behind the camera. The height of the screen is adjusted such that the patient can see the displayed contents clearly.
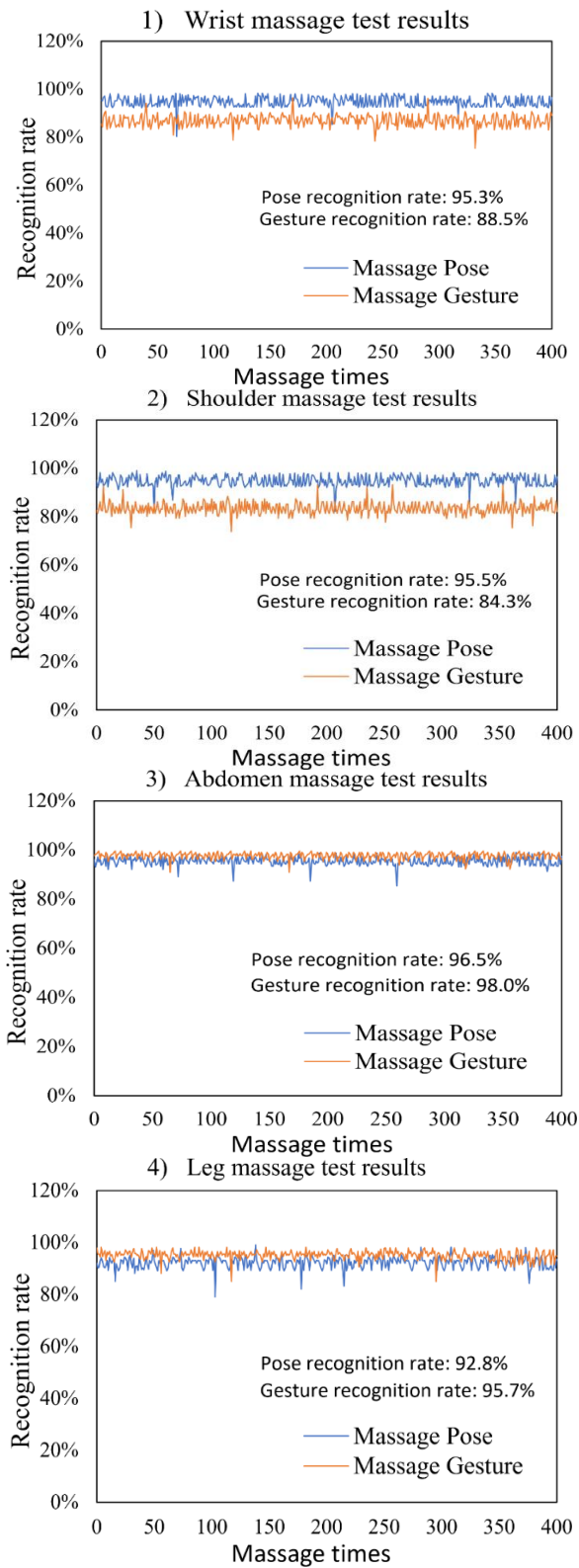
Figure 7. Mean recognition rates of gestures and poses for the four massage actions in self-rehabilitation

According to the sound prompt, each patient performs self-rehabilitation actions, including self-rehabilitation poses and autonomous massage gestures, on four different body parts. Each action is executed 10 times, that is, every patient performs 40 self-rehabilitation actions. In total, each action is executed 400 times, and all actions are implemented 1,600 times.

Four different massage actions of self-rehabilitation are tested:

(1) Wrist massage: the left hand makes a fist and raises to the height of the nose bridge, while the right hand kneads the left wrist: with the palm facing inward, the thumb is placed on the right of the wrist, and the other four fingers on the left.

(2) Shoulder massage: the left arm droops naturally, while the right arm raises to massage the left shoulder: the thumb is placed in front of the shoulder, and the other four finger behind the shoulder.

(3) Abdomen massage: the five fingers draw close to each other to massage the abdomen.

(4) Leg massage: the body slightly bends over, and the four fingers other than the thumb massage the front of the thigh.

To ensure the recognition accuracy and precision, the patient needs to enter the scanning range of the camera, and stand at the test point during the test. After the end of the test, the patient must leave the scanning range. At least five seconds is provided to recognize each self-rehabilitation action. Since Kinect DK generates 30 frames of data per second, a total of 150 frames of data can be obtained for each action. After removing invalid frames, the valid frames are imported to our algorithm to obtain the mean recognition rates of massage gestures and poses (Figure 7).

As shown in Figure 7, over 93% massage poses in the four actions are correctly recognized, and the recognition rate, with only a slight fluctuation, fully meets the system needs. During leg massage, the recognition rate is slightly dropped due to the occlusion effect of the body on the arm. The recognition rate of the massage gestures surpasses 84%, but with large fluctuations. This is because the gesture recognition is hindered, when the hand is occluded by the body or other things, owing to the limited recognition range of the hardware. The recognition rate can be improved by adopted two Kinect DKs [33].

7. CONCLUSIONS

The rehabilitation system based on somatosensory interaction and autonomous massage is an important research direction of life health. It offers the patient an immersive treatment experience that unknowingly reduces the pain of disease. In this paper, massage gesture recognition is innovatively integrated with human pose recognition, and fused with somatosensory interaction to develop an individualized, intelligent, and independent rehabilitation training system for acupoint massage that excludes human factors.

In the system, the binocular depth camera Kinect DK is adopted to collect the data on joints and gestures of the patient in autonomous massage, and the coordinates of each finger joint are calculated; the poses of the patient are recognized by VGG-16, and translated for display in a VR model based on Unity 3D, aiming to guide the patient actions in autonomous massage.

The experimental results show that, the proposed system could correctly recognize up to 84% of non-occluded gestures, and up to 93% of non-occluded poses. However, obvious jitter is observed on screen replay, when the patient is too close to or too far from Kinect DK, causing slight deformation of the recognized gestures; when the hand is occluded, the recognition rate plunges. To improve the recognition accuracy, two Kinect DKs could be installed to detect the patient's poses and gestures from different angles. The data of the two devices can complete each other, resulting in better recognition effect.

The proposed system boasts a good real-time performance, a high practicality, a strong operability, and a low cost. Facing the lack of medical staff amidst the ongoing Coronavirus (COVID-19) pandemic, our system can enable postoperative patients to perform autonomous massage on time and on demand, and improve their quality of life.

ACKNOWLEDGMENT

REFERENCES

[1] Rautaray, S.S., Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: A survey. Artificial Intelligence Review, 43(1): 1-54. https://doi.org/10.1007/s10462-012-9356-9

[2] Gao, Y.Q., Lu, X., Sun, J.B., Tao, X.L., Huang, X.M., Yan, Y.X., Liu, J. (2020). Vision-based hand gesture recognition for human-computer interaction-a survey. Wuhan University Journal of Natural Sciences, 25(2): 169-184. https://doi.org/10.19823/j.cnki.1007-1202.2020.0020

[3] Lahiani, H., Neji, M. (2020). A survey on hand gesture recognition for mobile devices. International Journal of Intelligent Systems Technologies and Applications, 19(5): 458-485. https://doi.org/10.1504/IJISTA.2020.111065

[4] Jaramillo-Yánez, A., Benalcázar, M.E., Mena-Maldonado, E. (2020). Real-time hand gesture recognition using surface electromyography and machine learning: A systematic literature review. Sensors, 20(9): 2467. https://doi.org/10.3390/s20092467

[5] Mo, T., Sun, P. (2019). Research on key issues of gesture recognition for artificial intelligence. Soft Computing, 24: 5795-5803. https://doi.org/10.1007/s00500-019-04342-3

[6] Morelli, T., Folmer, E. (2014). Real-time sensory substitution to enable players who are blind to play video games using whole body gestures. Entertainment Computing, 5(1): 83-90. https://doi.org/10.1016/j.entcom.2013.08.003

[7] Peng, Y., Peng, J., Li, J., Yao, C., Shi, X. (2019). Smart Home based on Kinect Gesture Recognition Technology. International Journal of Performability Engineering, 15(1): 261. https://doi.org/10.23940/ijpe.19.01.p26.261269

[8] Park, K.S. (2016). Development of Kinect-based pose recognition model for exercise game. KIPS Transactions on Computer and Communication Systems, 5(10): 303-310. https://doi.org/10.3745/KTCCS.2016.5.10.303

[9] Wang, J., Liu, T., Wang, X. (2020). Human hand gesture recognition with convolutional neural networks for K-12 double-teachers instruction mode classroom. Infrared Physics & Technology, 111: 103464. https://doi.org/10.1016/j.infrared.2020.103464

[10] Wang, J., Liu, T., Wang, X. (2020). Human hand gesture recognition with convolutional neural networks for K-12 double-teachers instruction mode classroom. Infrared Physics & Technology, 111: 103464. https://doi.org/10.1016/j.infrared.2020.103464

[11] You, Z., Liu, J., Hou, W., Wang, X., Liu, W., Song, W. (2017). A wearable system designed for Chinese traffic police based on gesture recognition. in transdisciplinary engineering: A paradigm shift. Proceedings of the 24th ISPE Inc. International Conference on Transdisciplinary Engineering, 5: 385-393. https://doi.org/10.3233/978-1-61499-779-5-385

[12] Ma, C., Zhang, Y., Wang, A., Wang, Y., Chen, G. (2018). Traffic command gesture recognition for virtual urban scenes based on a spatiotemporal convolution neural network. ISPRS International Journal of Geo-Information, 7(1): 37. https://doi.org/10.3390/ijgi7010037

[13] Hammal, Z., Huang, D., Bailly, K., Chen, L., Daoudi, M. (2020). Face and gesture analysis for health informatics. In Proceedings of the 2020 International Conference on Multimodal Interaction, 874-875. https://doi.org/10.1145/3382507.3419747

[14] Liu, Z.H., Mo, W.P., Tang, Z., Sun, Y., Wang, S.Z. (2016). An upper-limb active exercise rehabiligtation system for stroke patients based on leap motion. Journal of Donghua university (Natural science), 42(4): 572-575.

[15] Hang, Y., Chen, G., Tong, Y., Ji, B., Hu, B.C. (2021). Human rehabilitation action recognition based on pose estimation and GRU network. Computer Engineering, 47(1): 12-20. https://doi.org/10.19678/j.issn.1000-3428.0058201

[16] Chen, Y. (2018). The Design of Virtual Upper Limb and Its Movement Control for Rehabilitation Training. Chongqing University.

[17] Wang, Z.R., Wang P., Xing, L., Mei, L.P., Zhao, J., Zhang T. (2017). Leap Motion-based virtual reality training for improving motor functional recovery of upper limbs and neural reorganization in subacute stroke patients. Neural Regeneration Research, 12(11): 1823-1831. https://doi.org/10.4103/1673-5374.219043

[18] Yang, H.Q., Qian, T. (2020). Application of Kinect-based dynamic posture recognition method in medical rehabilitation. Computer Technology and Its Applications, 46(12): 94-96, 102. https://doi.org/10.16157/j.issn.0258-7998.200147

[19] Qian, C.H., Zhang, X.H., Tao, J., Liu, J.L. (2020). Design and research of rehabilitation training system based on Kinect. Journal of Jilin University (Information Science Edition), 38(1): 92-98. https://doi.org/10.3969/j.issn.1671-5896.2020.01.013

[20] Zeng, W.C., Liu, Q. (2021). Design of gesture recognition system based on elastic conductive coated yarn. Cotton Textile Technology, 49(594): 21-26.

[21] Ahuja, R., Jain, D., Sachdeva, D., Grag, A. Rajput, C. (2019). Convolutional neural network based American sign language static hand gesture recognition. International Journal of Ambient Computing and Intelligence, 10(3): 60-73. https://doi.org/10.4018/IJACI.2019070104

[22] Xavier-Rocha, T.B., Carneiro, L., Martins, G.C., Vilela-Junior, G.D.B., Passos, R.P., Pupe, C.C.B., Monteiro-Junior, R.S. (2020). The Xbox/Kinect use in poststroke rehabilitation settings: A systematic review. Arquivos de Neuro-Psiquiatria, 78(6): 361-369. https://doi.org/10.1590/0004-282X20200012

[23] Hong, H., Chao, J., Jin, Y., Zhao, Z., Lin, W. (2015). Key

point model for hand pose estimation based on leap motion. Journal of Computer Aided Design & Computer Graphics, 7: 1211-1216.

[24] AlBdairi, A.J.A., Xiao, Z., Alghaili, M. (2020). Identifying ethnics of people through face recognition: A deep CNN approach. Scientific Programming, 2020: Article ID 6385281. https://doi.org/10.1155/2020/6385281

[25] Xie, Z., Niu, J., Yi, L., Lu, G. (2021). Regularization and attention feature distillation base on light CNN for hyperspectral face recognition. Multimedia Tools and Applications, 1-17. https://doi.org/10.1007/s11042-021-10537-4

[26] Saqib, S., Ditta, A., Khan, M.A., Kazmi, S.A.R., Alquhayz, H. (2021). Intelligent dynamic gesture recognition using CNN empowered by edit distance. CMC-Computers Materials & Continua, 66(2): 2061-2076. https://doi.org/10.32604/cmc.2020.013905

[27] Cardenas, E.J.E., Chavez, G.C. (2020). Multimodal hand gesture recognition combining temporal and pose information based on CNN descriptors and histogram of cumulative magnitudes. Journal of Visual Communication and Image Representation, 71: 102772. https://doi.org/10.1016/j.jvcir.2020.102772

[28] Wang, H., He, P., Li, N., Cao, J. (2020). Pose Recognition of 3D human shapes via multi-view CNN with ordered view feature fusion. Electronics, 9(9): 1368. https://doi.org/10.3390/electronics9091368

[29] Shivakumara, P., Tang, D., Asadzadehkaljahi, M., Lu, T., Pal, U., Anisi, M.H. (2018). CNN-RNN based method for license plate recognition. CAAI Transactions on Intelligence Technology, 3(3): 169-175.

[30] Kim, H.H., Park, J.K., Oh, J.H., Kang, D.J. (2017). Multi-task convolutional neural network system for license plate recognition. International Journal of Control, Automation and Systems, 15(6) 2942-2949. https://doi.org/10.1007/s12555-016-0332-z

[31] Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition.

[32] Zhang, B.M., Wu, H.G., Shen, J. (2004). Tuina therapy for prolapsed lumbar intervertebral disc. Journal of Acupuncture and Tuina Science, 2(2): 58-60. https://doi.org/10.1007/BF02877117

[33] Sun, G.T., Huang, P.Y., Liu, Y.P., Liang, R.H. (2018). Interactive 3D visualization with dual leap motions. Journal of Computer-Aided Design & Computer Graphics, 30(7): 1268-1275. https://doi.org/10.3724/SP.J.1089.2018.16680