SCIENCE CHINA
Information Sciences

• **MOOP** •

Identifying Gait Abnormality with a Single Click

Jin HUANG, Shuai MA, Feng TIAN*, Xiang LI, Jie LIU & Hongan WANG

Beijing Key Laboratory of Human-Computer Interaction, Institute of Software Chinese Academy of Sciences, Beijing 100190*, China;*
State Key Laboratory of Computer Science, Institute of Software Chinese Academy of Sciences, Beijing 100190*, China;*
University of Chinese Academy of Sciences, Beijing 100049*, China*

Citation Jin HUANG, Shuai MA, Feng TIAN, et al. Identifying Gait Abnormality with a Single Click. Sci China Inf Sci, for review

Gait abnormality is one of the major symptoms of nervous system diseases such as Parkinsons disease. Clinically, gait abnormality can be diagnosed by monitoring patients abnormal behaviors via neurological assessment tools such as Wisconison Gait Scale (WGS) Gait Abnormality Rating Scale (GARS). However, these assessments tools usually requires patients to complete a long and tedious testing process under the supervision of a doctor, which is a tremendous pressure for both patients and hospitals.

Clinical gait analysis systems have many applications in diagnosis, monitoring, treatment and rehabilitation [1]. It has been proven that Kinect based approach can efficiently and accurately extract comprehensive gait information from all parts of the body [2] and, in the meantime, the Kinect solution also has the characteristics of low-cost and easy-deployment [3]. We consider that the Kinect based systems can provide accurate and inexpensive gait analysis solutions. However, there are still challenges to actually deploying these systems in the field. On one hand, besides the target subject, doctors or other unrelated personnel may also enter the visual field of the camera in the consulting room, this makes it difficult for us to extract the behavior of target subject. On the other hand, in a practical clinical examination, a subject may perform several testing activities such as sitting or standing other than simply walking. We must intercept specific segments from the video stream to extract gait features, otherwise it will seriously affect the accuracy of diagnosis.

We propose a novel Kinect based system, which integrates identity recognition, behavior recognition, and a built-in gait detection model to accelerate the clinical diagnosis. The system use the identity information provided by the color image and the behavior information embodied in the skeleton sequences to extract the walking segments of the target subject, and finally use these segments to diagnose gait abnormality. In an experiment involving 68 subjects (35 patients and 33 age-matched healthy subjects), we found that our system achieved 92.96% accuracy in automatically selecting walking streams of target subject, and 80.89% accuracy for gait abnormality detection. With our system, it is possible that in the future clinic, doctors can diagnose abnormal gait of patients with nervous system diseases automatically by simply clicking a ”start” button.

System Overview. The system architecture is shown in Figure 1. Firstly, a Kinect is used to collect RGB video stream and 3D skeleton sequences separately. Then, in an Identity Recognition module, the 3D skeleton sequences of the target patient are extracted from the whole data stream. Secondly, the selected skeleton sequences are put in to the Behavior Recognition module, and the segments that the target patient is in the walking state are obtained. In other words, segments that the target patient is not in walking state were filtered

* Corresponding author (email: tianfeng@iscas.ac.cn)
The authors declare that they have no conflict of interest.

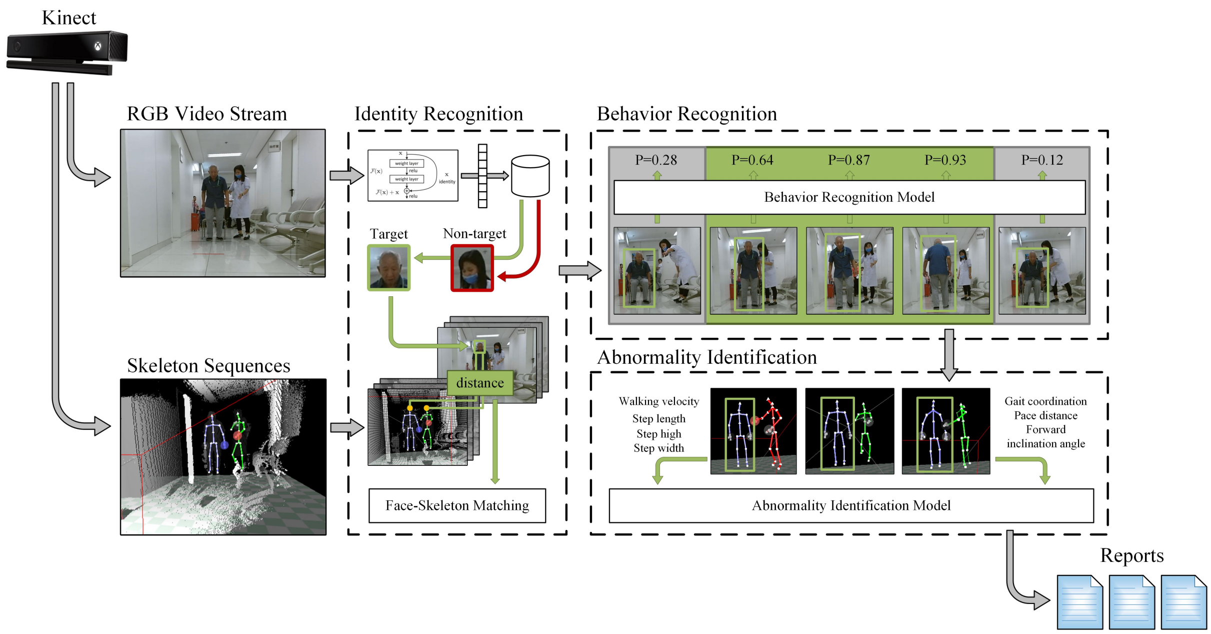

Figure 1 Caption 1. System Architecture for Gait Abnormality Identification

out to eliminate their impact on the subsequent process. Thirdly, in an Abnormality Identification module, gait features of the target patient are calculated using the walking segments obtained from the previous step, and a pre-trained Abnormality Identification model is used to generate the final gait function reports of the target patient.

Identity Recognition. A Residual Network (ResNet) with 29 convolutional layers was used to recognize identity for each person in our RGB video stream. The network is provided by dlib [1)], which is a modified version of the Kaiming Hes ResNet-34 [4] with a few layers removed and the number of filters per layer reduced by half. For each face in each frame of the stream, it extracts a 128-dimensional feature vector and determines the target by searching the minimum Euclidean distance of the vector between each face in the frame and the target.

While ResNet provided the identifications for each person in the stream, the Kinect API provided up to 6 human skeletons and their coordinate information in the stream. In the stream, the skeletons are organized in form of tracking unit, in which, each recognized skeletons would maintain a unique track ID from entering to leaving the camera. To map the skeletons to a recognized person, we transferred the head joint position of the skeleton of a track ID in to 2D image coordinate, and then marked it with the nearest human face by frame. The identity of most marked face would be finally set as the identity of the track ID. In practice, for the sake of efficiency, we used 50 frames evenly selected from the skeleton sequence of a track ID.

Behavior Recognition. A Long Short-Term Memory (LSTM) based classifier was used to identify whether the target is walking or performing other activities in the steam. LSTM [5] is an artificial recurrent neural network (RNN) architecture. Unlike standard feedforward neural networks, LSTM has feedback connections, which can not only process single data points (a skeleton frame in our case), but also entire sequences of data (a series of skeletons). The behavior recognition contains three steps. Firstly, the skeleton sequence data was divided into a series of overlapped 90-frames (about 3 seconds) windows, each window advanced forward for 30 frames. The duration of the window was determined based on a pilot analysis on all our data, which is approximate 2 times of mean length of a gait cycle for all subjects. Secondly, time-series feature including spatial coordinates and spatial directions of the 25 skeleton joints and length of 24 skeletons were extracted for each window. Finally, a pre-trained LSTM-based classifier was used to identify whether the target person is walking in the current time window. The LSTM-based classifier was trained by a labeled human behavior data set that contained 1000 samples. All samples in the data set were generated by randomly clipping 90-frames fragments from the

1) Davis King, Dlib C++ library, http://dlib.net/.

stream data collected in this study.

Abnormality Identification. In this module, 23 features was calculated for each walking segment to evaluate the gait functionality of the subject. The selected features mainly include mean and variance of walking speed, step height, step length, step width, step distance and body tilt angle, which were confirmed to be related to the cerebral injury, motor function and age of subjects [6]. All the features were calculated from the data of skeleton sequences of the target (see Appendix for a complete list of features and their calculation methods). In practice, the target patient may perform attempt multiple walking activities. If the system collected multiple walking segments from the target patient, a linear combination of the feature vectors weighted by the length of the segments were used. Finally, with the extracted 23 features, A linear-kernel support vector machine (SVM) was used to distinguish subjects with gait abnormality from health ones. The SVM classifier was trained and tested by a data set with pure walking segments which is collected and manually processed. A 10-folds cross validation in the data set achieved an 82.91% classification accuracy with 0.85 precision, 0.80 recall for gait abnormality class.

Experiment. We validated the proposed system in a clinical trial involving 68 subjects (35 gait abnormality subjects and 33 age-matched healthy subjects). All subjects in this study were recruited in hospital clinic of Peking Union Medical College Hospital (PUMCH), and the gait abnormality subjects were clinically diagnosed by neurologists in PUMCH.

In a trial, subjects were indicated to perform three sub tests including an upper limb function test in sitting posture, a 3-meters walking test and a balance ability test in standing posture. The first and third tests were routine neurological examinations designated by the hospital. The walking test included two rounds of 3-meters back-and-forth walking in front of the Kinect camera. In each round, the subject first sat in a chair three meters away from the camera, then stood up and walked to the camera at the usual walking speed. After he arrived at the camera, the subject was asked to turn around, walk to the chair and sit down. All the three sub tests were recorded by the camera, and at least two people (usually the subject and the supervisor) were videotaped.

To evaluate the identity recognition and behavior recognition modules, all recorded skeletons were labeled with classes *target*/*non-target* and *walking*/*non-walking* and compared with predicted skeleton state by frame. Results shown that the identity recognition module achieved high precision (0.94) and recall (0.95) for *target* class, and the classification accuracy reached 95.57%. The behavior recognition module also obtained high precision (0.96) and recall (0.91) for *walking* class, and the classification accuracy was 97.39%. By combined the two modules together, the system discriminate target subject's walking frames form others well with a 92.96% accuracy. To test wether the abnormality identification module still work fine with the automatic screening data, the skeleton data selected by the first two modules in the last module to identify target subjects' gait function. Results shown that the module maintained a good enough performance with 0.84 precision, 0.80 recall for gait abnormality class and 80.89% classification accuracy. The system takes about 30 seconds on average to process a data, which is much faster than manual filtering of data. Therefore, we can believe that this system can largely reduce the time consumption of gait examination and provide relatively accurate gait detection recommendations for future clinical practice.

Acknowledgements This work was supported by National Key R&D Program of China (Grant No. 2016YFB1001405), and National Natural Science Foundation of China (Grant No. 61802379).

Supporting information Videos and other supplemental documents. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 T. A. Wren, G. E. Gorton III, S. Ounpuu, et al. Efficacy of clinical gait analysis: A systematic review. Gait & Posture, 2011, 34(2): 149C153
2 Gabel, M., Gilad-Bachrach, R., Renshaw, E., et al. Full body gait analysis with Kinect. In Proceedings of Engineering in Medicine and Biology Society (EMBC), San Diego, California, USA, 2012. 1964-1967
3 Z. Zhang. Microsoft Kinect Sensor and Its Effect. In IEEE MultiMedia, 2012, 19(2): pp. 4-10
4 He K., Sun J., and Tang X.. Single image haze removal using dark channel prior. In Proceedings of Computer Vision and Pattern Recognition, Miami, Florida, USA, 2009. 1956-1963
5 Sepp Hochreiter and Jrgen Schmidhuber. Long short-term memory. Neural Computation, 1997, 9(8): 1735C1780
6 Su N., Liang X., Zhai F., et al. The consequence of cerebral small vessel disease: Linking brain atrophy to motor impairment in the elderly. Human Brain Mapping, 2018, 39(11): 4452-4461
7 Su N., Zhai F., Zhou L., et al. Cerebral Small Vessel Disease Burden Is Associated with Motor Performance of Lower and Upper Extremities in Community-Dwelling Populations. Frontiers in Aging Neuroscience, 2017, 9: Article 313