Abstract

Joint kinematics has been identified as a critical point of physical rehabilitation assessment. By tracking the gait patterns (i.e., stride length, step frequency, speed, knee flexion angle, and ankle flexion angle), walking abnormalities could easily be detected. Despite the substantial role of gait analysis, existing approaches are generally assisted by sophisticated equipment, making the popularization of gait analysis tough.

In this project, an offline surveillance-camera-based motion tracking system is introduced. Joint points projected on video frames would be detected with the implementation of a deep-learning-oriented open source library named OpenPose. After the pre-processing of regression, an image distance detection method on account of trigonometry is implemented and calibrated. Fast Fourier Transform is then applied to extract the periodicity of the gait patterns. Besides, the system compares the calculated gait identities with healthy gait trends using time series similarity measurements.

The project reduces the complexity of conducting gait analysis trails by limiting the required equipment to a single surveillance camera. By using a device as common as a smartphone, modern people could keep an eye on their walking patterns and follow routine physical rehabilitations easily.

Acknowledgments

I would like to give my sincere gratitude to Dr. Rosa Chan, my Final Year Project supervisor, for her consistent guidance, enthusiastic encouragement, valuable suggestions of the project, as well as the drone from her Laboratory. I would like to thank Yuzhou Shi. Her intelligent work is indispensable and inspired me a lot. I would like to express my thanks to Dr. Roy Cheung, for using his Gait & Motion Analysis Lab at PolyU and his professional knowledge and assistance in gait rehabilitation. My heartful appreciation also goes to Dr. Kelvin Yuen, my Final Year Project assessor, for his constructive advice to the project. I would like to thank Pak Kwan Chan, Chao Ying Chen, Xueru Ding, and Rui Zhong, for their assistance in the progress of data collection. I want to acknowledge Computer Terminal Laboratory of EE Department, CityU, for borrowing the facilities. My grateful thanks are also extended to my friends, who offered me endless support whenever I need it. They are Jingzhi Xu, Yijie Jin, and Qiqi Ouyang.

Contents

List of Figures

List of Tables

Introduction

This report provides a comprehensive description of the project that focuses on analyzing human gait with a single surveillance camera.

In the background section, the necessity of analyzing human gait gestures, especially using tools as simple as a single camera, is stated. A survey regarding the attitude of target users is elaborated.

As for methodology, this report follows the logic of project code to illustrate the implementation of system functionalities. Assessments for selecting the most appropriate approaches are also attached.

After the theoretical contents, detailed results (input and output of the system, calculation results of different data sets, and accuracy measurements) are provided.

Besides, further discussions concerning the shortage of the current system and the work to be implemented in the future exists in the discussion section.

Background

1. Gait Analysis

Gait analysis tracks the movement of human to interpret the kinematic activities like walking or running. Gait analysis is of high importance in understanding the correctness of body postures and assisting rehabilitation therapy. Traditional gait analysis emphasizes the importance of observation, with the limitation that the changes of joints in the left-side limbs and the right-side limbs are not synchronized. Gait quantification is introduced to overcome this obstacle. Tools like the electronic protractor, infrared markers, photoelectric markers, and motion sensors are widely applied to precisely obtain numbers to describe the gait patterns. However, the involvement of these tools inevitably increases the cost and difficulty in conducting gait analysis.

2. Objectives

This project aims to find a cheap, accessible, and flexible method to meet the same results that can be generated using the tools mentioned above. Previous work[1] in camera-based gait analysis achieves part of the objective. The method requires to know the pitch angle (hard to measure) and field of view (not provided by all of the devices) of the camera, preventing the application from the promotion through the market. As a result, this project intends to realize an application with no specific limitations (for example, indoor or outdoor video shooting, device with or without detailed parameters, maximum or minimum height of the environment).

Besides, although previous work could calculate some kinematics data (such as the speed of running, frequency, and stride length), these parameters are more entertaining than rehabilitative. Many of the trajectory features are more closely associated with pathological diagnosis and treatment. For example, flexion contracture of the knee requires extra muscular activity to stabilize the flexion-bearing knee joints. Insufficient knee flexion reflects quadriceps weakness. Incorrect spontaneous rest posture (15° to

[1] Details in Shi, Y. (2019). Gait analysis with Drone. Retrieved from http://dspace.cityu.edu.hk/handle/2031/9164

60° for knee flexion angle of normal samples) renders deformities, one of the major physiological responses to pain. Thus, apart from the cadence of gait, this project focuses more on the detecting and analyzing of joint trajectories.

The target customer of this project's application is not limited to professional rehabilitation institutions; people who are just interested to know more about themselves are also welcomed. Consequently, to comprehensively understand the market of gait analysis, I conducted a survey in which people in all age groups (divided as 1-20, 21-40, 41-60, >60) are involved. The primary purpose of the survey is to comprehend people's concerns about gait abnormalities (Necessity), curiosity in their gait patterns (Interest), and their intention in using the application developed in this project (Practicability).

In total, 174 valid feedbacks are collected.

Age Groups	Number of Questionnaires	Percentage
1-20	10	5.75%
21-40	53	30.46%
41-60	100	57.47%
>60	11	6.32%
Sum	174	100.00%

Table 1: Survey Feedbacks

3. Necessity

a) Problematic Gait Postures

A stereotype of self-consciousness in problematic gait posture is that older adults will pay more attention to their gait postures than young people and be more likely to regard their postures as questionable. According to the survey results, however, young people are more sensitive to posture abnormalities, while elderlies are more confident in their gait postures.

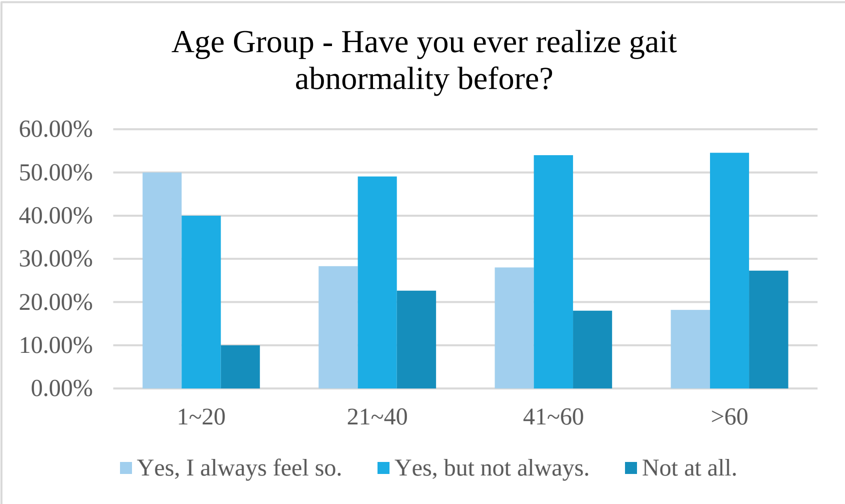


Figure 1: Age Group - Problematic Gait

Interestingly, although seniors are more likely to be optimistic for themselves, they intend to regard that more people that are under the same age group have the problem of gait abnormality. On the opposite, young people would expect their peers less likely to feel uneasy during walking/ running.

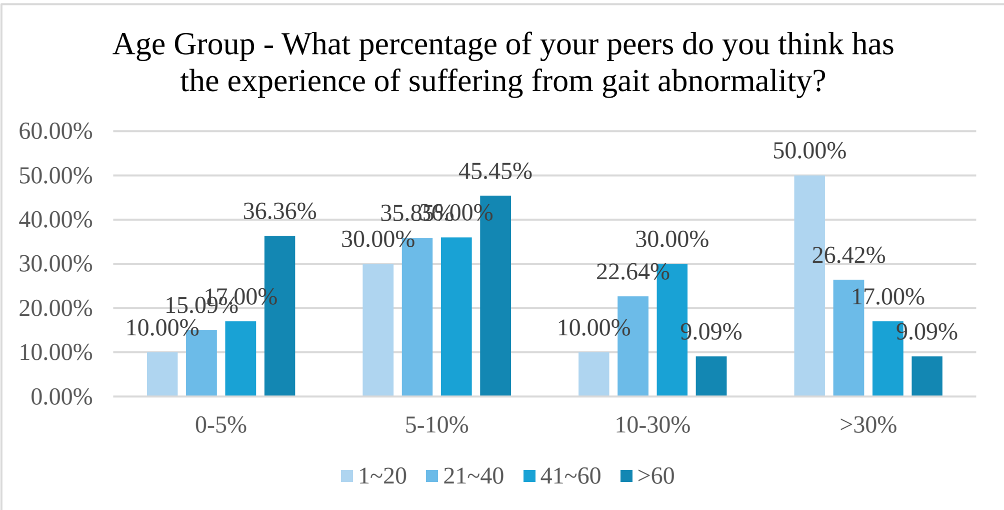


Figure 2: Age Group - Expectation of Peers

In general, around 80% of people aged under 60 (the main target customers of Internet products) had the experience of realizing their walking/ running postures problematic. Even though elderlies are more confident in their gait patterns, 70% of them have had the problem of gait abnormality.

4. Interest

a) Walking Speed

Around 80% of people indicate that they are curious about their walking speed. Given nowadays a variety of wearable devices provide the function of monitoring the walking speed, the application developed here may not be the first nor the only one that could meet people's demands. However, according to the survey, over half of the people who are curious about their standard walking speed haven't tried to measure it. Shortly, the market space for the application of the project exists.

b) Gait Properties

People's expectations about "gait properties" include muscle stress, joint stress, the scope of postures, the association between posture abnormalities and diseases, and personal identities of postures.

The expectations of different age groups variate slightly. Young people focus more on the posture identities, and seniors concern more about the stress placed on joints. The attention in joint stress also increases with age. The result reminds us that although not all the properties of gait the application could measure are essential for rehabilitation purposes, but many of them are good attraction points for the promotion of the application.

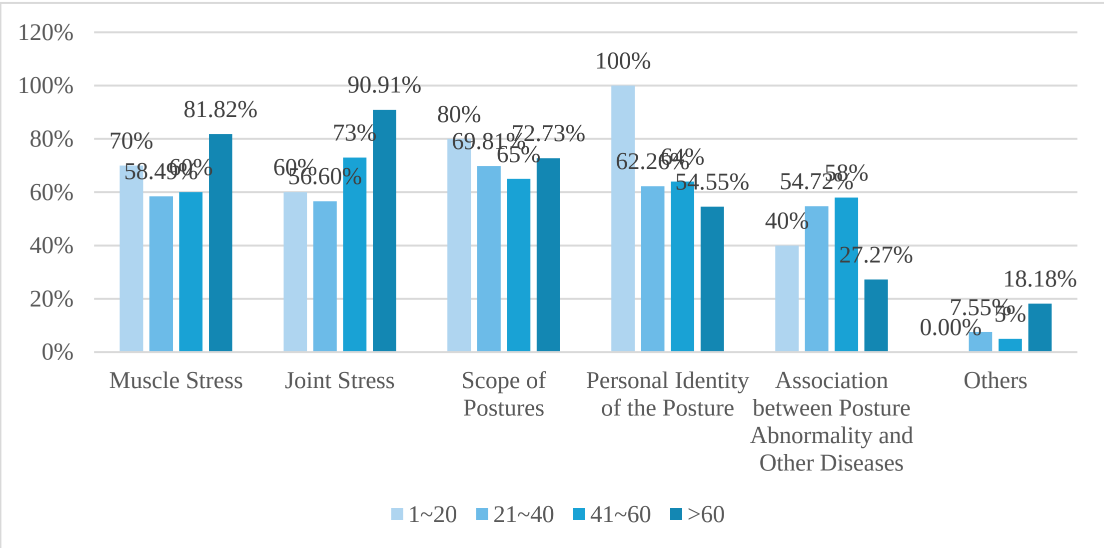


Figure 3: Age Group - Gait Properties

5. Practicability

a) Gait Analysis Application's Assistance

Around 80% of people are positive in seeking help if they face problems in gait. Over 80% of them are willing to use gait analysis applications to monitor their gait conditions. Around 55% will only refer to the suggestion given by the gait analysis application while not conducting further checking with the help of professional hospitals or rehabilitation institutions.

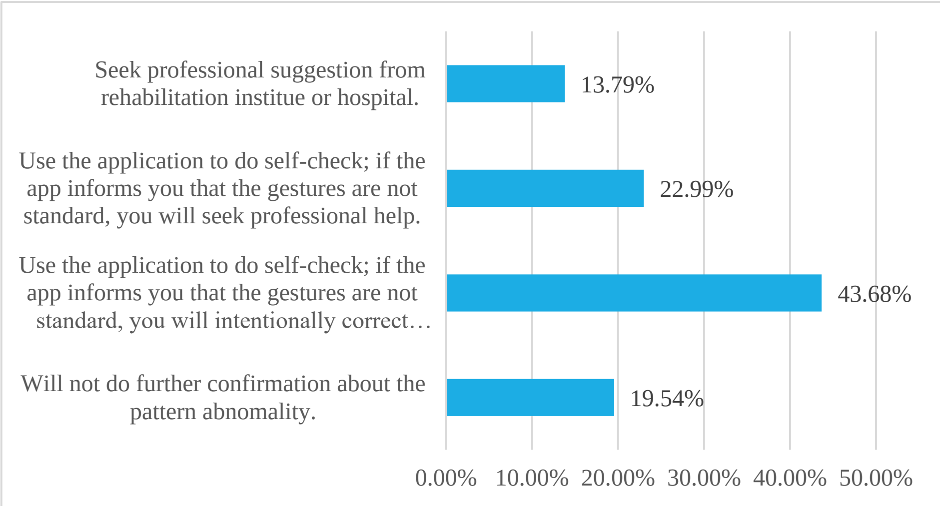


Figure 4: People's Intention When Facing Gait Abnormality

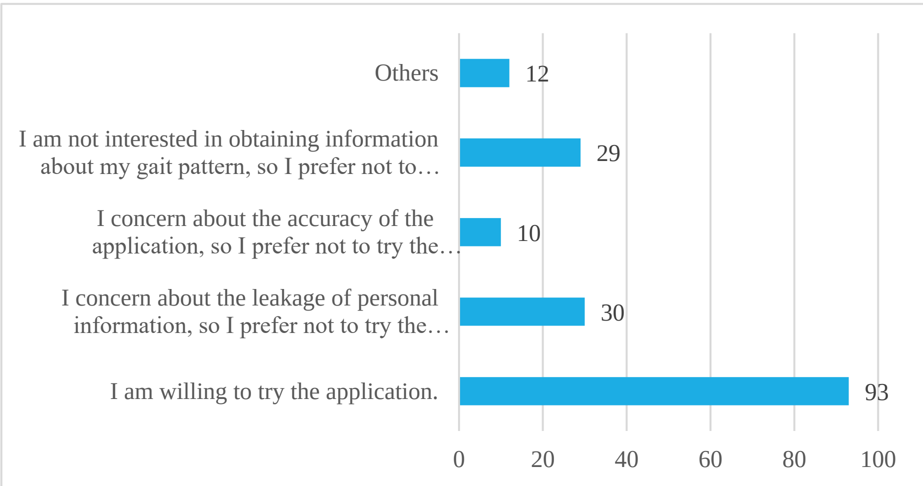


Figure 5: Intentions and Concerns

Apart from lacking interest in gait postures, the concerns in personal information leakage (since the input of the application are videos that capture the person's body and their postures) is the main point that drives people's negative attitudes in using the application. If the application is to be promoted to the market, functionalities that could improve the security condition will be highly valuable.

b) Real-time Application

Around 70% of people will be optimistic about using the gait analysis application if it could conduct calculations in real-time. 40% of people who are not interested in using the gait analysis application declares that they are willing to try if it could generate results in real-time. For the people who are pleased to use the current gait analysis application, over 90% of them would like a premium version that could process analysis during the shooting progress.

To sum up, a high proportion of people has the experience of suffering from gait abnormality and are more willing to use gait analysis applications to monitor their gait behaviors instead of seeking professional help from rehabilitation institutions. The development of this application not only providing an option for people to pay closer attention to physical health, but also perfectly matches the demands of the market by delivering kinematic information to people who are interested to know more.

Methodology

1. OpenPose Key Point Detection Library

OpenPose is a deep-learning-based human pose estimation model. Layers of the neural network will pack feature maps and input the batch to a stage, during which the prediction generated by the current feature map will be combined with the one rendered by the previous feature map. After iterations of the co-refinement, the forecast is more likely to be accurate. This action is implemented for both the detection of confidence maps and the production of Part Affinity Fields (PAFs) stages.

(a) (b)

Figure 6: The Principle of OpenPose. Confidence Map (a) and Part Affinity Fields (b). Retrieved from Cao, Z., Hidalgo, G., Simon, T., & Wei, Y. (2019, May 30). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields.

Confidence map shows the level of confidence that the modal “believes” a pixel is in the position of a joint. Referring to the confidence map, points will be weighted with an accuracy probability, and the pixel with the highest confidence rate will be marked as the projected joint position.

PAF is used to associate discrete joint points detected by the confidence map in the unit of person. By applying PAF to store a 2D vector for each joint, the orientation of limbs will be considered besides the position, and the representation of location is no longer a single point. PAF is of vital importance in the cases that more than one person exists.

2. Pre-Processing

a) Interpolating

Interpolating is generally used for filling up missing points. A pre-processing of interpolation is applied in the previous version of the gait analysis application. However, even though there are data points missing in the output of OpenPose, most of the missing data distribute in face/ upper body/ arms. Since only the data related to legs and feet are essential in this project, interpolation is not the best fit that could help to improve accuracy.

b) Smoothing

i. Bezier Curve

Bezier Curves are parametric curves using the starting and the ending points of motion trajectory to estimate the intermediate data positions. Computing Bezier Curve helps to reduce the input noise of trajectory data.

Given n+1 points with known positions named as q_k ($k \in [0, n]$), the path approximated by Bezier Curve can be represented as:

$$Q(i) = \sum_{k=0}^{n} q_k \binom{n}{k} i^k (1-i)^{n-k},$$

where Q(i) are the position vectors containing the position in both the X-coordinate and the Y-coordinate.

Extreme points do not affect the trend of the approximated path severely, but the result generated by Bezier Curve is slightly different from the original input data for many frames. This happens because Bezier Curve is not used for interpolation but estimation. By applying Bezier Curve, not only the noise will be re-printed, but the accurate points will also be modified to fit the whole trend. However, for this project, correct data-position values are of vital importance since a small variation in positions may cause a huge difference when calculating angles. Thus, Bezier Curve is considered not suitable for this project.

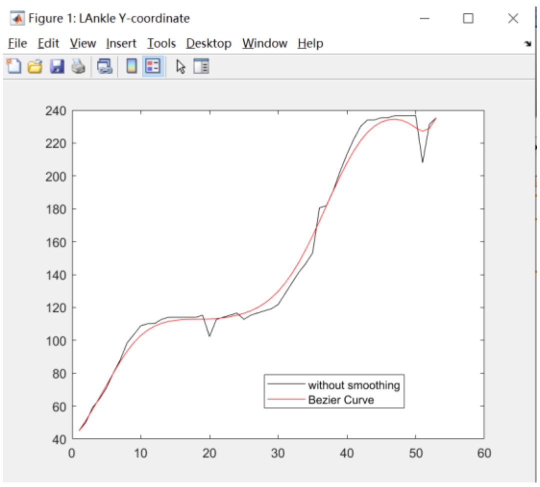


Figure 7: Pre-Processing using Bezier Curve

ii. Robust Locally Weighted Regression

1. Method

LOWESS (Locally Weighted Regression) estimates a first-degree polynomial model using weighted linear least squares. The polynomial is fitted with a subset of data points at the beginning. Explanatory variable values are assigned to points with different weights—locations that are closer to the point that is being estimated will have a larger influence than farther points. LOWESS is a model-less method, which means that we do not need to estimate a model from the data before the regression applied. This simplicity makes it fitful for this project.

The principle of RLOWESS (Robust Locally Weighted Regression) is similar to LOWESS, while "robust" is implemented by assigning zero weight to data points that are typically far (outside 6 mean absolute deviations) from the estimation point. This helps to reduce the influence that extreme noise points could bring to the regression results.

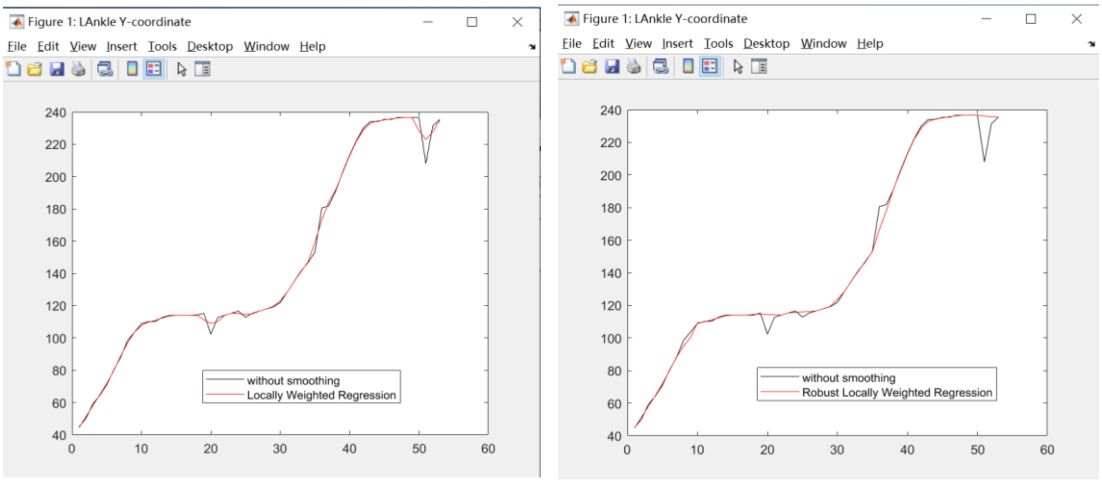


Figure 8: Data Pre-processed by Locally Weighted Regression (left) and Robust Locally Weighted Regression (right). Pre-processed result of LOWESS is more sensitive to extreme values.

When people walk at high speeds, OpenPose may not correctly mark the joint points for each frame. While for the calculation of feet length, flexion angles, and many other parameters, the precision in the input is of high importance since extreme points will cause problematic effects. Thus, RLOWESS method is applied in the project as pre-processing.

2. Span

“Span” stands for the window size of data points that are used to calculate the smoothed value. To make the value suitable for each data set, the window size in this project is set as 5/N, where N represents the number of frames contained by the video. Since N can be automatically generated when data are read into the MATLAB program, this transformation helps to determine a suitable span value intelligently for each case.

3. Image Depth Detection

a) Trigonometry-Based Image Depth Detection

The method works for cases that the surveillance camera with known pitch and field of view (FOV) angles. The camera should be placed high enough to shoot downward.

The vertical angle α and the horizontal angle β that the point is projected at the particular pixel (i, j) can be calculated by:

$$\alpha = \theta + \left(\frac{H}{2} - j\right) \cdot \frac{FOV_v}{H},$$

$$\beta = \left(\frac{W}{2} - i\right) \cdot \frac{FOV_h}{W}.$$

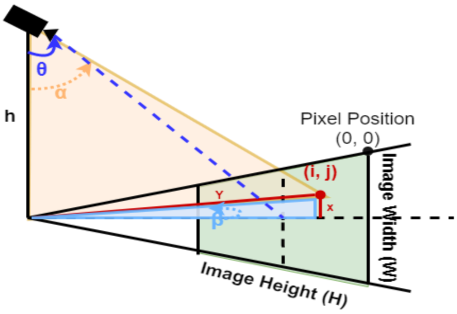


Figure 10: Image Depth Detection with Triangulation

Thus, the position of the point with respect to the origin set as the vertical projection of the camera could be represented as:

$$y = h \cdot \tan(\alpha),$$

$$x = h \cdot \tan(\beta).$$

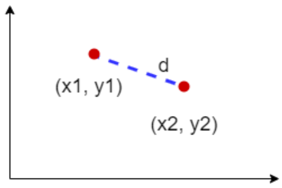


Figure 11: Absolute Distance

Although the point position (x, y) is relative to the origin point, when calculating the length of human feet or stride length, the absolute distance (d) is good enough.

$$d = \sqrt{(x1 - x2)^2 + (y1 - y2)^2}.$$

b) Ratio-Proportion

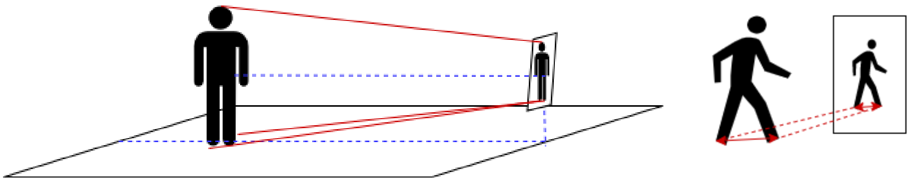


Figure 9: Ratio-Proportion

Ratio-Proportion is used to broaden the executable situation of the application. For cases that user cannot obtain the number of FOV, pitch angle, or a height over 2.2 meter to ensure the surveillance camera to shoot downward, ratio-proportion is a compensate method to determine the step length. The ratio is extracted by measuring the feet length to compare with the projected feet length:

$$\frac{projected\ left\ foot\ length + projected\ right\ foot\ length}{left\ foot\ length + right\ foot\ length} = \frac{projected\ step\ length}{step\ length}.$$

Since the trace of walking may not always be parallel to the FOV of the camera, the ratio may continuously change. By using the above equation, the proportion-ratio is supplemented dynamically, which guarantees that the step length is determined with a timely-calculated ratio.

c) Calibration

It is not always easy to get all the data required for the calculation of the trigonometry-based image depth detection. In general, the height of the camera can be clearly measured since the camera is placed by the user. However, for the FOV angle and the pitch angle, the properties may not always be available.

For the cases that the surveillance camera's FOV angle is unknown, this calibration is of high importance. With the real length of the user's feet known, the bisection method is used to find the FOV angle reversely.

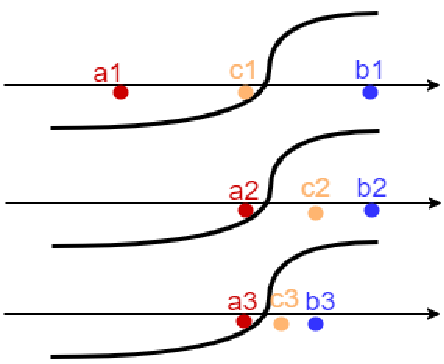


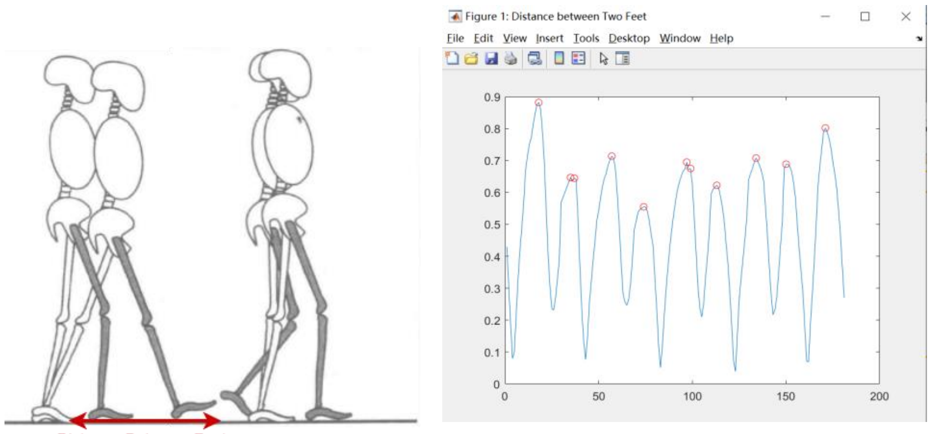
Figure 12: Bisection Method

Given a function F and we know that there will be one and only one x for which F(x) = 0 in the [a, b] interval, we can shrink the range of [a, b] by checking the sign of value of F(c) where c = (a + b) / 2. If $F(a) \times F(c) < 0$, there will be an x between [a, c] that F(x) = 0. Thus, in the next round, b substitutes the position of c in this round. The same rule applied between F(b) and F(c). For this project, the inline function F is the trigonometry-based image depth detection method, a, b, and c are used to represent the guessing range of FOV angle, which will always be between [0°, 90°].

In this project's data collection process, Nokia phone, iPhone, and DJI drone are involved. For the drone, the height and the pitch angle could be obtained during the flight. The FOV angle is fixed and could easily be found online[2]. The FOV angles for different iPhone types are also provided[3]. The height can be easily measured. Pitch angles could be measured while the value may not be entirely accurate. In these cases, the above-mentioned calibration method is a good fit. For the Nokia phone, however, the FOV angle is not provided and the pitch-angle measurement is not accurate. Here, using the ratio-proportion method is the best supplement.

4. Gait Cycle Determination

Intuitively, we can determine a gait cycle by observing the distance between two feet: for each step, there will be a time point that the length in-between becomes the largest. Thus, time can be sliced into the unit of steps by marking the frames that the distance between feet is the largest.



(a) (b)

Figure 13: Distance between Two Feet. The difference varies in a periodic format and a period is recognized as a gait cycle (a). The peaks in (b) can be used to separate gait cycles. Figure (a) retrieved from Perry, J., & Burnfield, J. (2010). Gait analysis: Normal and pathological function (2nd ed.).

The repeat-marking problem that can be observed from Figure 13(b) could be solved when introducing min-peak-distance: at least a number of frames must be between two peaks found. Nevertheless, the system needs to be responsive to the variance of walking speeds of users and the difference of the video shooting rate in

2 Information could be found at https://www.dji.com/hk/mavic-air/info
3 Information could be found at https://developer.apple.com/library/archive/documentation/DeviceInformation/Reference/iOSDeviceCompatibility/Cameras/Cameras.html

order to hard-code the min-peak-distance.

A better way is to use Fast Fourier Transform. A signal continuing in the time domain can be represented in the frequency domain as well. By transferring between time domain and frequency domain, trivial information will be dropped, and significant features will be preserved. This helps to reduce the effect of repeat-marking peaks without complicated modifications. A concomitant benefit is the reduce of computing time. Fourier coefficients are distributed symmetrically, and thus calculating half of the values achieves the demand of analyzing data.

5. Trajectory

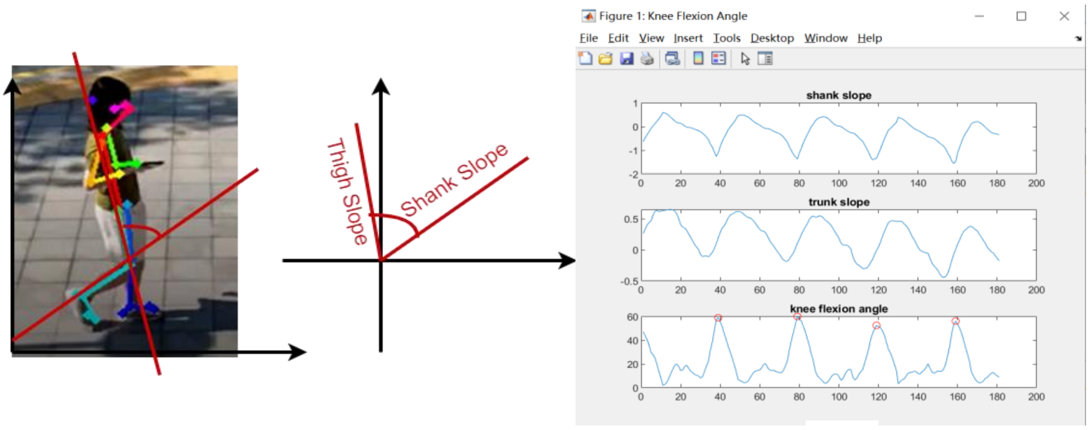
a) Synchronization of the Ground Truth Data

Noraxon myoMOTION, a system of sensors that can be attached to the human body to trace timely movements and gestures, measures the ground truth of the gait postures. Trajectory data—the flexion angles of different joints—are captured accurately in high frequency (100 fps).

To synchronize with the video used as the input of this project, the sensor system will start when a light bulb under the surveillance camera's shooting field is toggled to on. Thus, the data captured by the sensor system before lighting up will be regarded as redundancy.

Videos shot by different devices have different frequencies, but the rate will be around 30 fps in general. This is much lower than the rate of the ground truth sensor system. Before any further comparison between the calculated result and the ground truth could be applied, the time unit of the ground truth will be re-scaled to be the same as the video.

b) Knee Flexion Angle



(a) (b)

Figure 14: Knee Flexion Angle. The flexion angle can be regarded as the intersection angles of thigh and shank (a). The slope of thigh and shank is calculated with respect to the pixel positions of the joints (b).

Given the pixel positions of the hip (i), knee (j), and ankle (k), the flexion angle of the knee can be derived as:

$$Knee\ Flexion\ Angle = \arctan\left(\frac{x_i - x_j}{y_i - y_j}\right) - \arctan\left(\frac{x_j - x_k}{y_j - y_k}\right).$$

The accuracy of knee flexion angle calculation will be measured by comparing the dissimilarity between the sensor-captured ground truth and the calculated results. For the application, the gait trajectory data will be calculated and compared to the healthy trend of human gait. If the dissimilarities are under the threshold, the gait will be considered as healthy.

The synchronization of the general trend and the input trajectory is done by matching the peak. Firstly, the input trajectory will be linearly interpolated to be of the same scale of the general trend (the general trend is representing in 100fps, while the input trajectory in high probability is around 30fps). Then, the program will find the peaks of the two trends separately and "pin" them together.

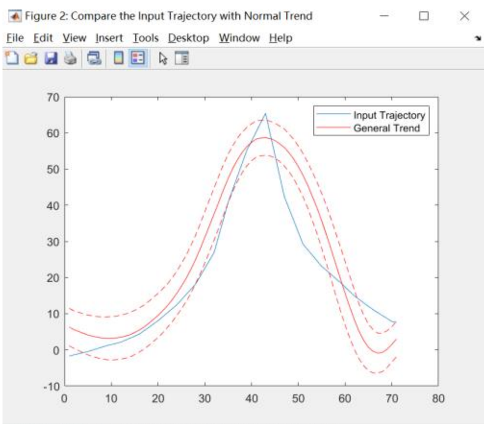


Figure 15: Gait Abnormality Detection by Comparing Input Trajectory with General Trend.

c) Accuracy Evaluation Methods

Different accuracy evaluation methods could be applied to measure the performance of the system quantitively. The evaluation of flexion angle detection accuracy is converted into the comparison of the similarity between two time series: the calculated results and the ground truth.

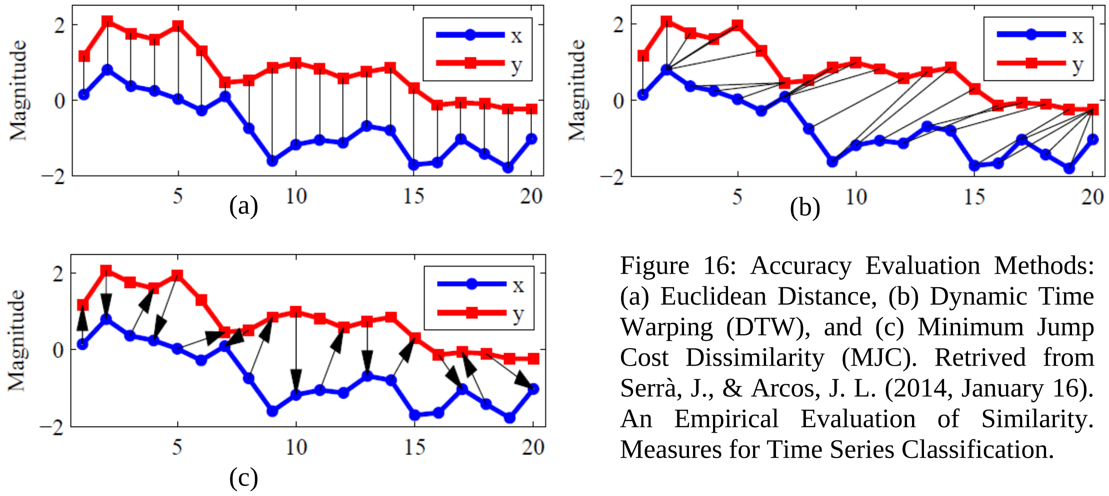


Figure 16: Accuracy Evaluation Methods: (a) Euclidean Distance, (b) Dynamic Time Warping (DTW), and (c) Minimum Jump Cost Dissimilarity (MJC). Retrived from Serrà, J., & Arcos, J. L. (2014, January 16). An Empirical Evaluation of Similarity. Measures for Time Series Classification.

i. Euclidean Distance

The simplest way to measure the dissimilarity between two time series is to calculate the Euclidean distance as:

$$\mathrm{d}(\mathrm{x},\mathrm{y}) = \sqrt{\sum_{i=1}^{M}(x_i - y_i)^2},$$

where x and y are the two-time series, M is the length of the series, x_i and y_i are the i-th element of the series.

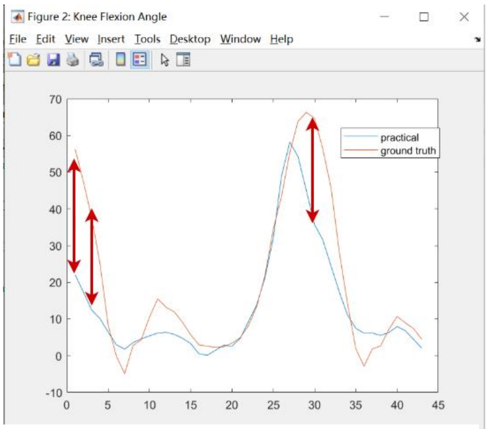


Figure 17: Performance Evaluation using Euclidean Distance

Euclidean distance is widely used, given its simplicity and indexing capability. However, the shortage of using Euclidean distance also exists. Even though the trends are similar, if the peaks are not hit at the same frame, the gap will be marked as exceptionally far. This property confuses the interpretation of results since users will not know the far distance is caused by the peak/ valley shift or the difference in trends.

ii. Fourier Coefficient

An extension of Euclidean Distance is to calculate the similarity between the Fourier Coefficients of two time series instead of comparing the raw time series. The results in the frequency domain is symmetric, and the sum will only be performed over half of the coefficients. Using the Fourier Coefficient helps to reduce the computation complexity by half. Moreover, the high frequencies close to M/2 can be filtered out to avoid rapid fluctuation of the signal.

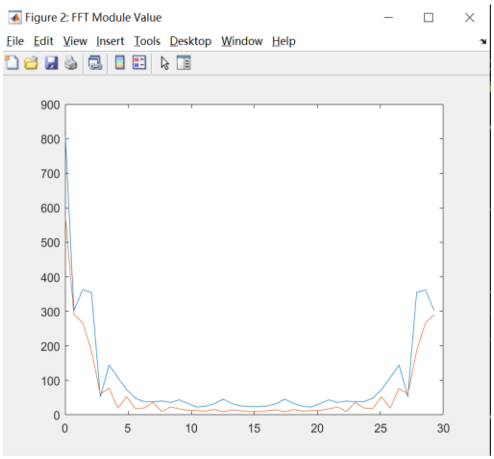


Figure 18: Performance Evaluation using Fourier Coefficient

The distance calculated with Fourier coefficients is:

$$\mathrm{d(x,y)} = \sqrt{\sum_{i=1}^{\frac{M}{2}} (\hat{x}_i - \hat{y}_i)^2},$$

where $\hat{x}_i$ and $\hat{y}_i$ are the complex value pairs representing the i-th Fourier Coefficient.

iii. Dynamic Time Warping (DTW)

Dynamic Time Warping works by minimize the accumulated cost of aligning (warping) the time series in the temporal domain. The cost is recursively calculated by:

$$D_{i,j} = f(x_i, y_i) + \min\{D_{i,j-1}, D_{i-1,j}, D_{i-1,j-1}\},$$

where $i \in [0, M], j \in [0, N]$. The initial value of $D_{0,0} = 0$, and $D_{i,j}$ where $i \in [1, M], j \in [1, N]$ is initialized as infinitely large.

The local cost function $f(x_i, y_i)$, called sample dissimilarity function, is Euclidean Distance:

$$f(x_i, y_i) = (x_i - y_i)^2.$$

The final dissimilarity is the accumulated cost, i.e., $d_{DTW}(x, y) = D_{M,N}$.

iv. Minimum Jump Costs (MJC)

The principle of MJC is that, if two time series resembles each other, the cumulative "jumping" cost between samples should be small.

Given that the previous jump is from the time step $t_y - 1$ in the series y, the cost of jumping from time step t_x in the series x is $c_{t_x}^{t_y+\Delta}$, where $t_y \leq t_y + \Delta \leq N$. The definition of the cost is:

$$c_{t_x}^{t_y+\Delta} = (\emptyset\Delta)^2 + f(x_{t_y}, y_{t_y+\Delta}),$$

where $\emptyset$ represents the cost of advancing in time as:

$$\emptyset = \frac{4\beta\sigma}{N},$$

where σ is the standard deviation of time series y. $\beta \in [0, \infty)$ controls the weight of the penalty assigned to advance in time.

The local cost function is represented as:

$$f\left(x_{t_y}, y_{t_y+\Delta}\right) = \left(x_{t_y} - y_{t_y+\Delta_i}\right)^2,$$

The minimal cost will be chosen for each jump:

$$c_{min}^{(i)} = \min\left\{c_{t_x}^{t_y}, c_{t_x}^{t_y+1}, c_{t_x}^{t_y+2}, \ldots\right\}.$$

The accumulate cost of a "jumping" process starting from x is:

$$d_{XY} = \sum c_{min},$$

A similar principle applies to y. The final distance $d_{MJC} = \min\{d_{XY}, d_{YX}\}$, which means to select the smaller cost between "jumping from y to x" and "jumping from x to y".

Results

1. The input of the Gait Analysis System

The data that should be input by users include:

a) A video that is shot from the right side of the user during his/ her walking process. Face and the upper part of the body may not show since this information will not be used. This also prevents confidential information leakage since it will not lead to the disclosure of sensitive personal information like the human face.

b) The resolution, the number of frames, and the frequency of the video. These parameters could be directly extracted by the system but input for confirmation purposes. The parameters could be omitted to develop a user-friendly application.

c) The height of the surveillance camera. If the height is higher than a particular threshold (2.2 meters), the following calculation will base on the trigonometry-oriented image depth detection, or else will use the ratio-proportion supplementary method.

   i. For the cases shot higher than 2.2 meters, a user needs to provide the Field of View (FOV) of the camera in degree. If the pitch angle during shooting is available, the user could input it to check the accuracy of the system; if not, the system could also calibrate with the length of human feet.

   ii. For the cases shot lower than 2.2 meters, a user only needs to provide human feet length for calibration.

2. Data Sets

a) Case 8, 9: Videos shot with a DJI drone at the height of 2.7 meters.

b) Case 12-21: Videos shot by mobile phone at the height of 1.95 meters.

3. Performance

a) Kinematics

i. Results

Case	Frequency		Stride Length (m)		Speed (m/s)	
	Calculated	Ground Truth	Calculated	Ground Truth	Calculated	Ground Truth
8	1.0015	1.0000	1.3581	1.2400	0.6949	0.6200
9	1.0176	1.0000	0.7980	0.6200	0.4068	0.3100
12	2.4300	2.5316	1.3228	1.4967	1.5230	1.9101
13	2.4975	2.2222	1.5446	1.4066	1.8795	1.5394
14	2.2477	2.1978	1.4145	1.5168	1.5418	1.6781
15	2.1407	2.0408	1.6183	1.6684	1.6699	1.7048
16	2.3054	2.2222	1.4614	1.5572	1.6091	1.6875
17	2.6061	2.2388	1.1031	1.7176	1.4646	1.9227
18	2.3054	2.2727	1.4252	1.6402	1.7159	1.8528
19	1.729	1.8405	1.5069	1.4363	1.3241	1.3154
20	1.9980	2.1429	1.5898	1.5460	1.5896	1.6327
21	2.0434	2.2222	1.2866	1.4368	1.2869	1.5756

Table 2: Kinematic Data Calculation Result

ii. Accuracy Evaluation

Case	Frequency	Step length	Stride length	Speed
8	99.85%	88.08%	90.48%	87.92%
9	98.24%	71.03%	71.29%	68.71%
12	95.99%	83.07%	88.38%	79.74%
13	87.61%	91.37%	90.19%	77.91%
14	97.73%	89.83%	93.25%	91.88%
15	95.11%	93.39%	97.00%	97.96%
16	96.26%	91.91%	93.85%	95.35%
17	83.59%	65.44%	64.22%	76.17%
18	98.56%	91.30%	86.89%	92.61%
19	93.94%	92.85%	95.08%	99.34%
20	93.24%	95.58%	97.17%	97.36%
21	91.95%	88.82%	89.55%	81.67%
Average	**94.34%**	**86.89%**	**88.11%**	**87.22%**

Table 3: Kinematic Data Accuracy Evaluation

iii. Assessment

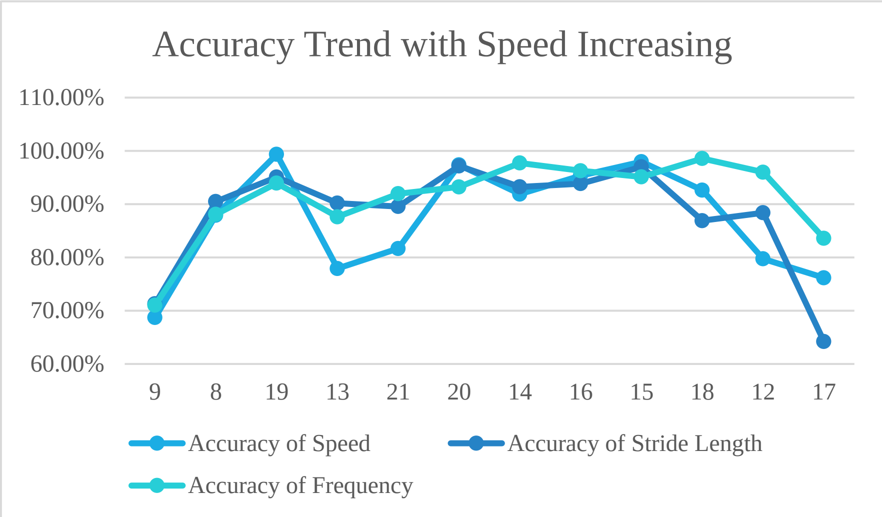


Figure 19: Accuracy Trend with Speed Increasing

Case	Speed (m/s)	Accuracy of Speed	Accuracy of Stride Length	Accuracy of Frequency
9	0.31	68.71%	71.29%	71.03%
8	0.62	87.92%	90.48%	88.08%
19	1.3154	99.34%	95.08%	93.94%
13	1.5394	77.91%	90.19%	87.61%
21	1.5756	81.67%	89.55%	91.95%
20	1.6327	97.36%	97.17%	93.24%
14	1.6781	91.88%	93.25%	97.73%
16	1.6875	95.35%	93.85%	96.26%
15	1.7048	97.96%	97.00%	95.11%
18	1.8528	92.61%	86.89%	98.56%
12	1.9101	79.74%	88.38%	95.99%
17	1.9227	76.17%	64.22%	83.59%

Table 4: The Accuracy (Speed, Stride Length, Frequency) with the Increasing of Speed

When the protagonist walks faster, OpenPose is more likely to mark joint points mistakenly and causes problems in the following calculation. When the speed is slow, however, a small difference in numbers will result in a large difference when compared with the ground truth since it is in the format of percentage. Above all, the accuracy is higher when people using a moderate speed walking through the camera.

4. Trajectory

a) Method Selection

Euclidean Distance, Fourier Coefficient, Dynamic Time Warping, and Minimum Jump Costs are all applied to evaluate the knee flexion angle dissimilarity between the calculated result and the ground truth captured by sensors.

Dissimilarity	Euclidean Distance	Fourier Coefficient	Dynamic Time Warping	Minimum Jump Costs
Case 12	2.1728	6.4694	17.105	0.005217
Case 13	1.6228	6.3242	9.5708	0.00796
Case 14	1.2511	4.715	10.593	0.001197
Case 15	1.8473	5.6681	16.027	0.000496
Case 16	1.828	3.9837	13.2	0.15657
Case 17	2.5124	6.7	30.766	0.002621
Case 18	2.4575	5.7472	13.486	0.000884
Case 19	2.7328	7.0306	24.337	0.000926
Case 20	2.9312	5.3518	20.419	0.001049
Case 21	2.0845	3.5632	17.855	0.000233
Average	2.14404	5.55532	17.33588	0.017715
Variance	0.246155	1.214909	37.86154	0.002148

Table 5: Knee Flexion Angle Dissimilarity Measurement

Since the distance calculation criteria of different methods vary, it will be inconclusive to compare the distance between the methods. However, the comparison for each technique among all the data sets could help us to choose the one best fit the abnormality detection function and to determine the benchmarks of performance evaluation functionality.

According to Table 5 and Figure 20, the evaluation of DTW is most suitable: The distance rendered by DTW follows the general dissimilarity trend. Also, the variation between cases is apparent by using DTW, leaving sufficient space for application administration to set the dissimilarity benchmark.

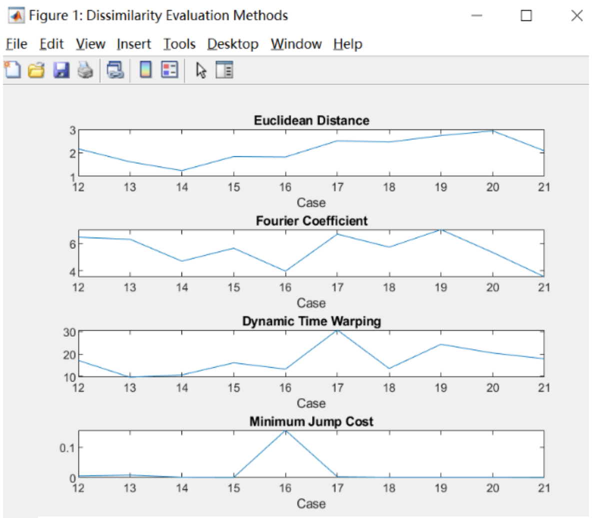


Figure 20: Dissimilarity Trends Comparison

In this project, I set the benchmark of DTW evaluation as 30. If the DTW distance calculation result of input is different from the healty trend by over 30, the user who is captured in the video is considered as performing gait abnormality.

b) Abnormality Detection

Abnormality Detection	DTW (Normalized)
Case 12	3.6665
Case 13	6.3654
Case 14	4.2178
Case 15	3.366
Case 16	3.2644
Case 17	NaN
Case 18	2.3707
Case 19	2.7381
Case 20	2.6914
Case 21	2.7936

Table 6: Comparison Results between the General Trend[4] and the User Input Trajectory

The peak in the input trajectory of Case 17 is not detected, which renders the DTW evaluation of abnormality failed. This may be caused by inaccurate detection since the accuracy of walking speed of this case is low and the DTW dissimilarity between the calculated results and the ground truth is high.

The DTW distances in the table are similarly low. The result indicates that the person is not suffering from any gait abnormality. This conclusion also matches with reality.

[4] The data of general trend is given by Perry, J., & Burnfield, J. (2010). Gait analysis: Normal and pathological function (2nd ed.). Thorofare, NJ: SLACK.

Discussion

1. Trajectory Performance Evaluation

Currently, the evaluation method (DTW) is chosen empirically. In this project, only ten sets of data are used to determine the method, while all of the ten sets capture the same person's trajectory information. However, to determine the way that is the best fit for the application may need more theoretical elaborations and empirical tests.

For further improvements, developers are highly recommended to capture videos for both people with gait abnormalities and people that walk in a healthy pattern. Besides, rehabilitation suggestions given by professional instructors will be powerfully helpful for the classification of "normal" and "abnormal".

2. Ankle Flexion Angle

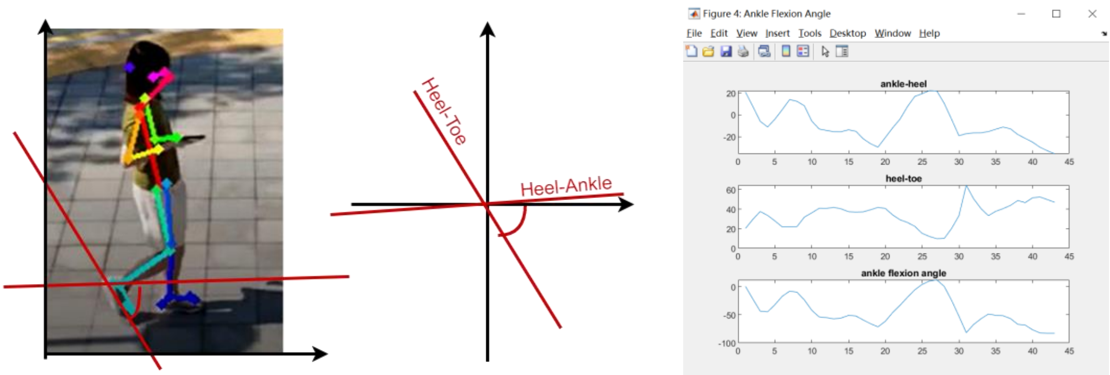


Figure 21: Ankle Flexion Angle

Similar to the knee flexion angle, the ankle flexion angle is of vital importance in pathological diagnosis. For example, contracture up to 30° reflects the flaccid paralysis of the tibial anterior muscle, and premature heel lift caused by excessive plantar flexion of the ankle reduces the walking speed to 70% of the normal.

The ankle flexion angle is calculated with the projected position of heel, ankle, and toe. However, the performance of extracting the ankle flexion angle variation trend is not as good as the knee flexion angle.

In Figure 22, extreme value point exists. This is not caused by noise since this abnormality is not distinct for a particular case. Thus, this peak can be regarded as happening in every gait cycle. Smoothing functions like regression or Bezier Curve should not be applied to remove the identity of the slope.

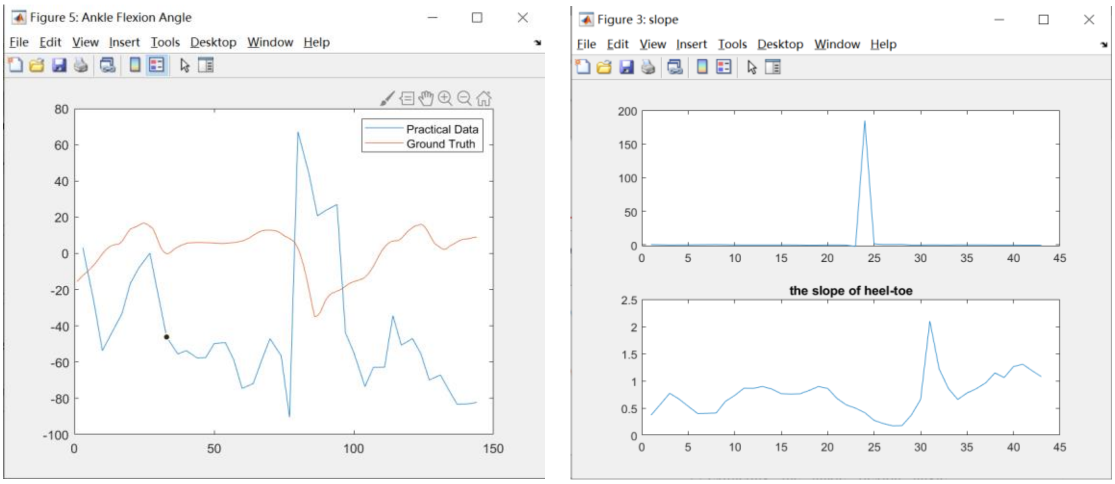


Figure 22: Ankle Flexion Angle and the Slope of Ankle-Heel and Heel-Toe without Smoothing. It could be recognize that their will be a sudden change in ankle-heel slope. This peak value renders extreme values in the ankle flexion angle. However, this is not unique for one case but applies to all the case, so it should not be marked as a noise.

The comparison of the calculated ankle flexion angle and the ground truth is inconclusive. This may be the results of:

- Single camera's limitation. Figure 23 shows the difference between the walking orientation and the ankle-toe direction. In the calculation, the mismatching is ignored since there is only one camera shooting the video. Thus, the angle between "absolute-front" and "ankle-toe" cannot be directly obtained.

Figure 23: Orientation of Walking and the Ankle-Toe Direction. Retrieved from Perry, J., & Burnfield, J. (2010). Gait analysis: Normal and pathological function (2nd ed.).

- OpenPose's inaccurate marking of joint points. Heel, ankle, and toe are close to each other, especially in frames with low resolution. A distance as short as one pixel could still generate a vast difference when measuring the angle in degree.

- Surveillance camera's low shooting frequency. Use equipment with a frequency higher than 60fps is a must to capture all the identity during a gait cycle. For mobile phones, however, the general rate is around 30fps. The ground truth used in this project is taken in 100fps, by which the sensitivity is guaranteed.

3. Real-time Gait Analysis System

It could be concluded from the survey that “real-time” is a function strongly preferred by intentional users. This project does not contain the feature yet, while this functionality is recognized as highly valuable.

Current obstacles for implementing real-time analysis include:

- the computing speed of OpenPose. For videos with a resolution 960×544, OpenPose could generate JSON files at a rate of 4fps (using Intel i5-8265U CPU only, with the processing progress shown frame-by-frame). This is much lower than the speed of video shooting (around or larger than 30 fps). The GPU version of OpenPose and high-performance computers should be used in order to keep up with the pace of surveillance cameras’ shooting rates.
- the computation capability of mobile phones. First and foremost, the tool OpenPose has not yet been developed to run on mobile devices. Moreover, the performance of mobile phone CPU may not be at the same level as laptops, let along with high-performance computers. If the application of the project is launched on mobile phones, it may need to work by communicating with remote servers (such as cloud service). The application on mobile phones will merely be used as a view to collect user input and display the calculated results and rehabilitation suggestions.

Conclusion

In this project, an offline motion-tracking system is developed to analyze human gaits. The project uses OpenPose, an open-source deep-learning library, to extract projected joint points in videos. After the implementation of Robust Locally Weighted Regression, trigonometry-based image depth detection, ratio-proportion, and calibration, the system applies Fast Fourier Transform to determine periodic features (i.e., stride length, frequency, walking speed). The accuracy of frequency reaches an average of 94.34%, and all kinds of kinematics data obtain over 85% of accuracy. The system is approved to be more sensitive to moderate-speed walking. Besides, trajectory information (knee flexion angle) is extracted from videos. The user’s walking pattern can be compared with the general trend of walking and evaluated with Dynamic Time Warping, which is useful in discovering gait abnormality and stimulating further rehabilitation suggestions from professional institutions.

Compared with previous work, this project unfreezes the limitation of surveillance camera placing height and devices by using ratio-proportion and calibration. The regression method also improves the accuracy of the results. The most constructive progress is the implementation of trajectory information detection, by applying which more rehabilitation possibilities and market attractiveness are granted.

Appendix

Survey Results:
https://bitbucket.org/tianxigao2/gaitanalysissurvey/src/master/Survey%20Results%20and%20Analysis.xlsx

Reference

[1]. Cao, Z., Hidalgo, G., Simon, T., & Wei, Y. (2019, May 30). OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. Retrieved from https://arxiv.org/abs/1812.08008

[2]. CMU-Perceptual-Computing-Lab. (2020, April 5). OpenPose GitHub Page. Retrieved from https://github.com/CMU-Perceptual-Computing-Lab/openpose

[3]. Perry, J., & Burnfield, J. (2010). Gait analysis: Normal and pathological function (2nd ed.). Thorofare, NJ: SLACK.

[4]. Qiao, S., Wang, Y., & Li, J. (2018, February 27). Real-time human gesture grading based on OpenPose. Retrieved from https://ieeexplore.ieee.org/document/8301910

[5]. Salih, Y., & Malik, A. S. (2012, August 16). Depth and Geometry from a Single 2D Image Using Triangulation. Retrieved from https://ieeexplore.ieee.org/document/6266436

[6]. Serrà, J., & Arcos, J. L. (2014, January 16). An Empirical Evaluation of Similarity Measures for Time Series Classification. Retrieved from https://arxiv.org/abs/1401.3973v1

[7]. Shi, Y. (2019). Gait analysis with Drone. Retrieved from http://dspace.cityu.edu.hk/handle/2031/9164