Human Gait Analysis

Rasmus Ramsbøl Jensen
Informatics and Mathematical Modelling
Technical University of Denmark
Email: s011908@student.dtu.dk

Introduction

The goal is to do a gait analysis based on a simple video sequence of the patient. The analysis should be done with as little interaction by the user as possible.

Software exist that does analysis on gait but requires markings of heel and toe touch down etc.

The wanted tracking algorithm should be marker less and demand no special requirements such as background screen or special clothing.

An energy minimization function is used based on likelihood, a smoothness prior, a contrast term and a shape prior.

The data

The input data are simple sequences from a stock video camera placed on a tripod (stationary). There are no requirements such as blue screen background and the background is completely random. The patients wear different clothing as well as bare skin.

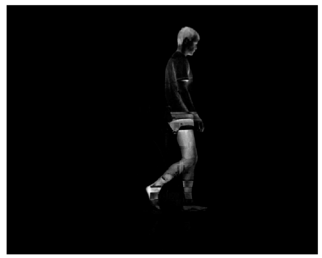
Background and object statistics

Because the camera is stationary a statistical model of the background is easily created as the mean and standard deviation of the whole sequence. This is done for every pixel ignoring the fact that some of the values are skewed when pixels take on values from the object. This will however only be briefly relative to the duration of the sequence. To get an estimate of the object statistics, a frame is found that differs greatly from the background and an area around the maximum difference in that frame gives basis for the object statistics.

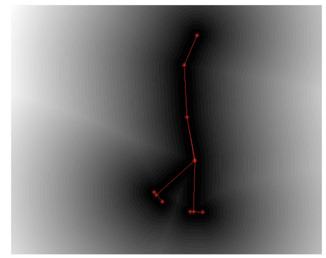
While the estimated background statistics are fairly good the estimate for the object is very rough.

Figure A.1 shows a likelihood function for the background. It is clearly a very god model of what is background as every pixel in the background has very low energy.

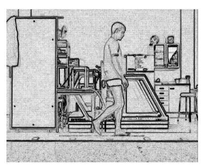
A.1 Background likelihood



C.2 Distance map



B.1 Contrast term

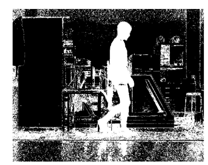


Abstract

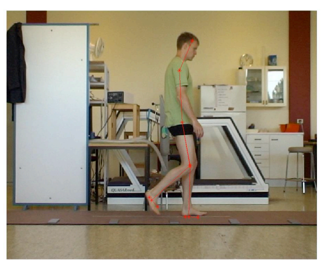
The motivation was to create a cheap alternative to expensive gait analysis using high-speed camera and marker suits. This alternative could help the average patient based on an algorithm that uses simple stationary video sequences.

Results: The algorithm is able to track the patient. Giving basis for outputs such as stride length, speed, joint angles etc.

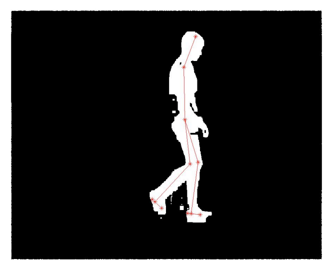
D.1 Initial segmentation



C.1 Initialization by user



D.2 Segmentation and shape optimization



Contrast term

In segmenting between background and object a penalty is given to a given pixel if the label of the surrounding pixels have different labels but the values differ very little. Said in another way, pixels with the same label are expected to have same values as their neighbours while pixels with different label as their neighbours should have a big difference in value. Figure B.1 shows the penalty given to pixels with different label than their neighbour.

Shape prior

Because the object to be found is a person, a human stickman model is used. As an initialization of the algorithm a stickman is placed in a given frame by the user to avoid a brute force search on a very big problem to find an initial pose and position of the shape prior. Figure C.1 shows the initialization of the stickman. By rasterizing the stickman and using the SEDT (Signed Euclidean Distance Transform) a distance map is calculated as shown in Figure C.2. Pixels labeled object are given high energy if far from the shape while pixels labeled background get high energy if close to the shape.

Energy minimization

Figure D.1 shows labeling based solely on the likelihood functions, while D.2 shows the segmentation based on likelihood, contras term, shape prior and finally a smoothness prior (Potts model). When the segmentation is done the stickman is optimized to minimize energy. Because the changes from frame to frame are relatively small, the stickman is used as an initialization for the next frame and the whole process is repeated and a new optimal pose and position of the stickman is found using a local search algorithm. This is repeated in a forward and backward manner from the initialization frame until the pose of every frame is found.

Conclusion

The algorithm is able to track the patient through the frame. It does however have some shortcomings such as very extensive running times. Detection when knees are close together is somewhat erroneous. This could be resolved with a prediction model based on the output.