

Automated Human Gait Recognition

Ifueko Igbinedion, Ysis Tarter
CS 229, Fall 2013

1. Task Definition

Motivation

A reliable way of human identification is in high demand in today's society. There has been a lot of work done identifying humans by faces or fingerprints, but this evidence is not always available. For example, in serious crime, established forms of human identification are purposely distorted, hidden, or obscured. Nevertheless, everyone must walk so gait is usually evident. In a more domestic setting, human gait recognition has many practical applications. You may be able to find someone in a crowd based on walking patterns observed on security cameras or identify which user is moving in a video game more easily. We aim to design a system of gait recognition in which a profile may be built by training a profile on one gait sample.

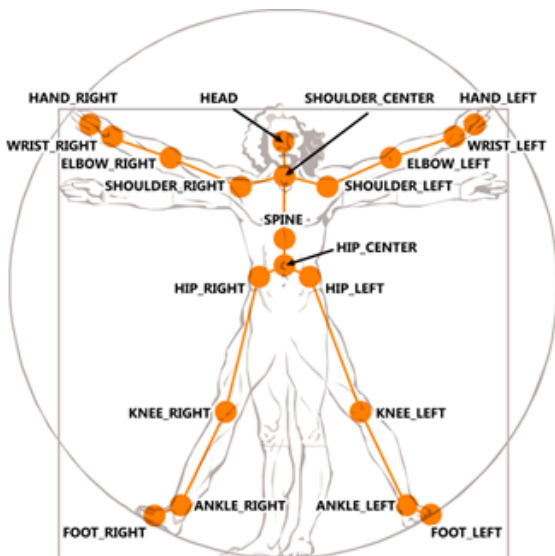


Figure 1: The kinect representation of the human body. Data collection was performed using a kinect program that creates CSV files of the x,y, and z coordinates of each of the joints pictured above.

Background and Significance

Established technology exists in which video can be converted into a three-dimensional joint

mapping [5]. The specific technology that we chose to use for this project is the X-box kinect system's sensor, which has the capability of extracting three-dimensional skeletal joint positions (Figure 1), but does not attempt to recognize the user. In the current field of automated gait recognition, phases of the gait cycle must be extracted manually and, for this reason, current gait recognition algorithms tend to identify features based on continuous image frames. Furthermore, most established methods of gait detection rely on measuring a subject at multiple paces to build a profile. We aim to design a system of gait recognition in which a profile may be created autonomously using a state-based model for feature extraction and identified accurately invariant to speed and sample size (number of cycles recorded). It is worth noting that the kinect system, as well as other gaming platforms, offer a user recognition system based on a range of biometric data, including facial features and body shape. Our system is unique in that it is designed to require only one sample of a subject walking to build a profile, and this is more generalizable to situations in which the physical condition (or even cooperation) of the observed subject is uncertain. In addition, our approach to this obstacle is novel in that we model the motion of the joints as a Markov Decision Process (MDP) in order to learn distinct profile features before identifying the observed subject based on proposed distributions of the existing profiles using particle filtering. We recognize that without time to explore optimal features, more precise recording tools, and much larger data set, our accuracy may not be optimal. Nevertheless, this project served as a valid proof of concept (and offers profound novelty) in the implementation of gait analysis **using a state-based model** and we are confident in its generalizability for biometric recognition systems.

Problem Definition

Input: three-dimensional joint tracking data for an unknown subject and three-dimensional joint tracking data of known subjects

Output: a match of the unknown subject to the most likely known subject

Assumption: the unknown subject is one of the known subject

This task can be broken down into two parts:

1. parsing the tracking data into different phases of the gait cycle autonomously and creating subject profiles
2. matching the unknown subject to the most likely profile

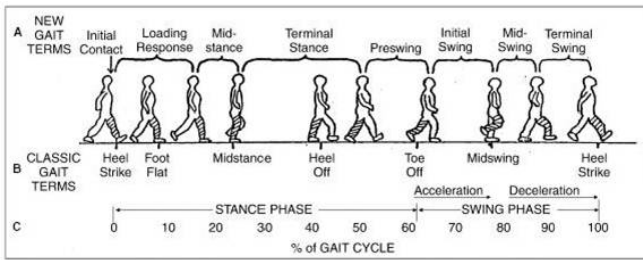


Figure 2: The above figure shows the body's positioning at each phase in the gait cycle. We use the new gate terms rather than the classic gate terms for our state-based model.

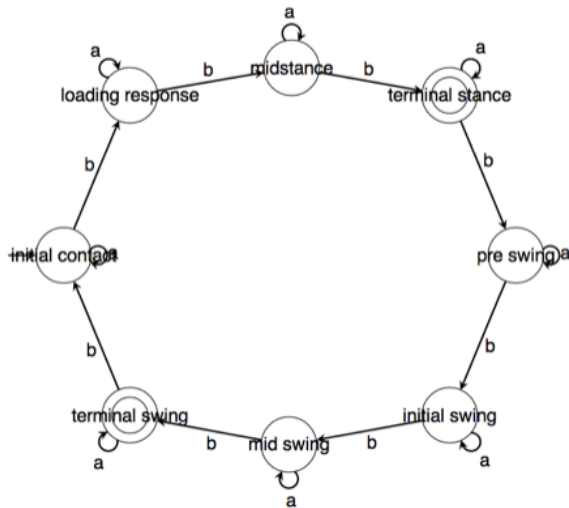


Figure 3: The figure above shows the state based model of the gait cycle. At each state the model can either choose to classify a data point as a point in the current state (action a) or classify the point as one in the next state (action b).

2. Approach

Identifying gait cycle phases and forming subject profiles

Challenge: Every person is unique, with their own range of movement and time distribution in individual gait periods

Baseline and previous approaches: Current gait recognition algorithms process continuous frames to extract features, such as joint angles and vertex points.

Gait phases are identified using peaks and troughs of joint locations and the phases themselves are only significant in that they serve as a marker for cycle length (a critical variable for certain normalizations), and even this was accomplished after manual preprocessing. Therefore the *baseline* we considered when testing the accuracy of our gait phase identification algorithm was our manual identification of the time-stamps for each phase.

Advanced Approach: In our approach, the gait cycle is considered in the form of a state-based cycle, so rather than extract features linearly

To accurately identify individual gate periods, we began with a generative learning algorithm in which we defined binary features for each phase based on known joint patterns [9] We initially began

segmenting the points using a support vector machine model, but realized that the gait cycle is considered in the form of a state-based cycle, so we modeled our algorithm as a **Markov Decision Process** in which:

States: Every phase of the gait cycle is a state and *starting state* is initial contact

Actions: at each time step (round), you may either remain in your state or proceed to the next state in the cycle

Transition probabilities: based on the average amount of time spent in each phase; updated at the end of each cycle

Rewards: the squared difference between features scores of a time step in the states resulting from staying in a phase or moving on

IsEnd: terminal stance and terminal swing
At the end of each cycle, we performed a **policy iteration** and updated accordingly

Matching the unknown subject to the most likely profile

Challenge: A human itself does not walk consistently, pace and rhythm are subject to change. Formally, this means we must predict an entire set based on a subset of information; therefore our profile features must be specific enough to accurately classify the subject, but generalizable enough to recognize the subject's gait in a range of conditions (speed variance).

Baseline and previous approaches: The most current techniques [11] typically proceed with a hybridization of support vector machines and two- or three-dimensional silhouette projection.

Advanced Approach:
To accurately identify the subject, we first had to extract features that would be adequately specific to each profile; to do so, at each state, we recorded joint angle progressions and normalized time, in addition to

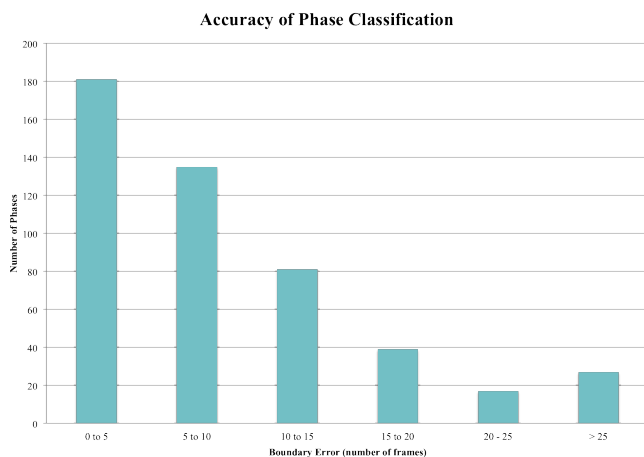
recording features such as hip rotation patterns at every cycle. With these features extracted, we created a **kernelized support vector machine**, which we trained on the joint tracking data of known subjects before performing empirical risk minimization on the hypothesis class.

We then treated the unknown joint tracking data as a **Markov network** and used **particle filtering** to use the profiles to predict the unknown subject's next state transition. Potentials (and proposed distributions) are based on the kernels in our svm and the state-based model used on the training examples and proposals are updated based on demonstrated variation in speed (angles, normalized times, and ranges are approximated are reweighted). The unknown subject is assigned to the profile that is the most accurate overall.

3. Results and Analysis

Evaluation Metric 1: Accuracy of Phase classification

The above figure shows the accuracy of our algorithm for accurately classifying phases of the training set, the golden standard being manual identification of the starting and ending frame of each phase. The training set consists of six subjects walking for ten gait cycles, 480 phases in total. Boundary error is the total number of frames the algorithm's starting and ending frames are off by for each frame. The figure indicates that the algorithm is able to accurately predict phases within 1 second (10 frames) 66% of the time.



Analysis:

As seen in evaluation metric one, our algorithm was able to accurately predict phases within one second (ten frames) 66% of the time. Upon evaluation of this less than ideal metric, we found that the boundary errors within the range of 10 to 20 frames were primarily due to the “snowball” effect: that is, if the previous frame was off by two and the current frame's boundary error is three, the observed boundary error

will be five. When re-parameterizing our figure to discount previous error in evaluation, our accuracy improved to correctly predict phases within one second 89% of the time. There were a few key feature incorporations we explored that were key in achieving this accuracy. The first features we added outside of strict joint biometric data were rhythmic sensibilities: that is, for consistency, the coordinates of the foot, ankle, and knee should be as close as possible at the end of each cycle. By incorporating this into our reward function, we were able to improve accuracy by 8%, as well as the accuracy of the profile recognition. After looking more closely, we believe this is due to the fact that regulating cycle length reduced the range of the distribution we used in particle filtering. In addition, feature selection was also a huge factor; we explored over fifty unique features before finding which held the most significant weight consistently across test-subjects. The most significant are listed below.

Figure 4: The chart below shows the most significant features across all subjects.

* fl: foot left; fr: foot right; al : ankle left, ar: ankle right, vi : speed at time step i; hl: hip left; hr:hip right
 ** These are all indicator functions.

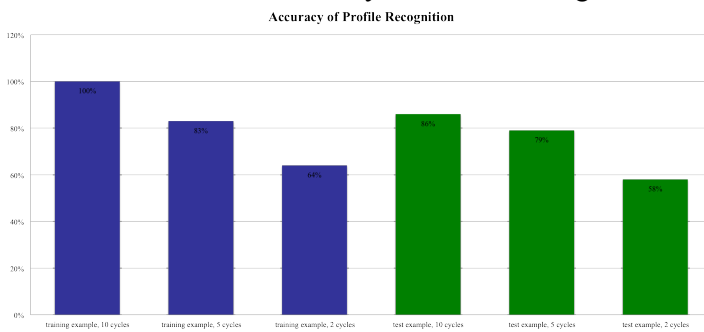
***The last letter is the x-y-z point we consider.

1. Left foot over right	$I(\text{fly} > \text{fry})I(\text{fly} > \text{fry})$
2. Flat front foot	$I(\text{flz} = \text{alz})$ if left is forward; otherwise $I(\text{frz} = \text{arz})$
3. Right foot over left	$I(\text{fly} < \text{fry})$
4. Parallel feet	$I(\text{fly} = \text{fry})$
5. Hip flexion right	$I(\text{hlx} \ll \text{hrx})$
6. Hip flexion left	$I(\text{hlx} \ll \text{hrx})$
7. Parallel hips	$I(\text{hlx} = \text{hrx})$
8. Left foot off ground	$I(\text{flz} \gg \text{frz})$
9. Right foot off ground	$I(\text{flz} \ll \text{frz})$
10. Both feet on ground	$I(\text{flz} = \text{frz})$

11. Accelerating	$I(v_f - v_i > 0)$ between timesteps i and f
12. Decelerating	$I(v_f - v_i < 0)$ between timesteps i and f
13. Back ankle flexed	$I(flz \ll alz)$ if left is forward; otherwise $I(frz \ll arz)$
14. Front ankle flexed	$I(flz \gg alz)$ if left is forward; otherwise $I(frz \gg arz)$

non-specific speeds that ranged from very similar to very divergent from the tempos in the training data. When testing our algorithm against speed variance by simply scaling our results (proportional to the time length of the cycle), our results had a maximum accuracy of 76%. We were able to drastically improve accuracy by modifying the feature vectors extracted in the personal profile and proposed particle sampling; that is, we added additional parameters that recalculated expected angles, time spent in cycles, and their weights, depending on how divergent the test example was from the training example. One of the biggest contributing factors to the accuracy of our test examples were the time steps we predicted in particle filtering. Particle filtering requires proposal, re-weighting and resampling, and as opposed to doing it at every time step that was labor intensive, we used a time step of 5 iterations (every $\frac{1}{2}$ second). This preserved accuracy and reduces the runtime by 500%. Over the increased number of cycles, there was only an observed 3% increase in profiling accuracy (to 89%), which was not very significant.

Evaluation Metric 2: Accuracy of Profile Recognition



The figure above shows the accuracy of our algorithm for classifying an example against the training set using leave-one-out cross validation (LOOCV, blue). In addition, in order to test how generalizable the recognition algorithm is, we also tested it against unseen joint tracking data (green) from each of the six subjects. The figure indicates that our algorithm performs optimally with at least ten gait cycles, but still performs well above random with half of that amount. Test example results indicate that our algorithm was fairly accurate in generalization.

Analysis:

As mentioned above, our algorithm performed best when trained on ten c We performed our algorithm for these two sets when training with ten, five, and two cycles. This means that in “10 cycles,” the algorithm was trained on samples of ten continuous gait cycles per person (all six). The consistent, seemingly linear decrease in accuracy as the number of cycles decreases was due to the fluctuation of feature values; increasing the number of observed cycles helped in both weighting and averaging values, (and it should be expected that a policy improve over iterations). In recording our test examples, we asked the subjects to walk at varying,

Evaluation Metric 3: Efficiency of profiling with SVM’s vs. HMM’s

Model	% Accuracy	# Phases	Runtime (seconds)
HMM	91.2	140	2970
	92	150	2992
	92.2	155	3004
	95	160	3015
	97	165	3005
	99	170	3028
	99	180	3056
SVM	90	140	7854
	92.4	150	11013
	93	155	15370
	93	160	17120
	94	165	20103
	94	170	23470

	94.2	180	255632
--	------	-----	--------

The figure above compares the runtime to the number of training cycles of the SVM and the HMM to achieve >90% accuracy. Runtime was calculated using the deterministic profiling capabilities of python modules profile and cprofile. When the number of phases considered was increased slightly, using an HMM proved to be significantly more efficient than an SVM to achieve the same threshold of accuracy, furthering our claim that HMMs provide a better model for our problem.

Analysis

From this data we can draw the conclusion that the HMM is more efficient in terms of data storage and accessing data. The model handles data more robustly than the SVM and is more computationally efficient in its development. This can be attributed to the fact that the state based model does not need information about all of its previous states to make decisions.

Future Directions

While this model was able to achieve high accuracy over our dataset, assumptions were made that may lead to high generalization error across larger datasets. One important assumption of the model was that the subject to classify actually existed in the original dataset. A more robust model would not assume that the subject exists in the training set, and would create a profile for each user before attempting to classify. Restructuring the model in this way would preserve accuracy over our dataset and lower generalization error.

Additionally, more can be done to improve the features used in this model. One approach would be to attempt to eliminate noise in feature vectors by eliminating irrelevant features to create a larger margin. This can be done manually by performing a qualitative analysis on the relevance of features and ranking features this way, or by running LOOCV with different combinations of features.

4. References

[1] Chacon-Murguía, Mario I., Rafael Sandoval-Rodríguez, and Omar Arias-Enriquez. "Human Gait Feature Extraction Including a Kinematic Analysis

Toward Robotic Power Assistance." *International Journal of Advanced Robotic Systems* 9 (2012).

[2] Cunado, David, Mark S. Nixon, and John N. Carter. "Automatic extraction and description of human gait models for recognition purposes." *Computer Vision and Image Understanding* 90.1 (2003): 1-41.

[3] Kale, Amit, et al. "Gait-based recognition of humans using continuous HMMs." *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on.* IEEE, 2002.

[4] Kale, Amit, et al. "Identification of humans using gait." *Image Processing, IEEE Transactions on* 13.9 (2004): 1163-1173.

[5] Kar, Abhishek. "Skeletal tracking using microsoft kinect." *Methodology* 1 (2010): 1-11.

[6] Lawson, Wallace, and Zoran Duric. "Analyzing Human Gait Using Patterns of Translation and Rotation." *Image Analysis and Recognition.* Springer Berlin Heidelberg, 2009. 408-417.

[7] Nauman, Eric A., Charles M. Krousgrill, and Gordon R. Pennock. "Human Motion Kinetics."

[8] Nixon, Mark. "Automated markerless analysis of human gait motion for recognition and classification." *ETRI Journal* 33.3 (2011): 259-266.

[9] Pietraszewski, Bogdan., Slawomir Winiarski, and Sebastian. Jaroszczuk. "Three-dimensional human gait pattern-reference data for normal men." *Acta of bioengineering and biomechanics/Wrocław University of Technology* 14.3 (2012): 9.

[10] Rose, Jessica. "Introduction to Human Walking & Clinical Gait Analysis." Stanford University, Stanford. Lecture.

[11] Sudha, L. R., and R. Bhavani. "Performance comparison of SVM and kNN in automatic classification of human gait patterns." *Int. J. Comput* 6.1 (2012): 19-28.

[12] Taylor, Graham W., Geoffrey E. Hinton, and Sam T. Roweis. "Modeling human motion using binary latent variables." *Advances in neural information processing systems.* 2006.