Baseball Batting Stance as a Proof of Concept for Human Identification via Unique Pose

Paul Perez, David Castello
Syracuse University
College of Engineering and Computer Science
223 Link Hall, Syracuse, NY 13244
pperez01@syr.edu
dcastel@syr.edu

Abstract

When considered over a span of time, the characteristic features of an individual's walking pattern are an established behavior biometric for use in identification or authentication tasks. Model-free approaches commonly aggregate silhouettes or pixel intensities across a recorded walk, while model-based techniques track the changes in key body coordinates derived from a constructed skeletal model. In both cases, a single frame or set of keypoints is deemed insufficient to characterize the behavior of an individual well enough for recognition. In this work, we propose that a person's pose or posture, captured while that individual performs a learned task, carries a wealth of information that is well suited for identification at a single moment in time. To prove this concept, we select the game of baseball as a domain that is well understood mechanically, routinely recorded at high resolution from a multitude of angles, and is noted for the distinctive stances of professional batters. From fixed-view images of Major League hitters at bat, we construct a dataset of skeletal feature vectors consisting of center body-relative coordinates and calculated angles and distances informed by domain knowledge. Experimental results yield promising classification accuracy, demonstrating the feasibility of human identification via pose and inviting future application to less controlled settings.

1. Introduction

Still images of an individual, captured at enough distance to include an entire profile, have several desirable characteristics for use in biometric applications. Features derived from a posture are universal, easily collectable, and, upon maturity, stable over time. Conventionally, however, a person's posture is not sufficiently unique for purposes of verification or even identification. Gait recognition overcomes this challenge by introducing a behavioral element, observing distinct traits of a subject's walking pattern over a time series. While gait patterns are demonstrably unique, they are inherently more difficult to collect than a snapshot and invite further complications arising from the viewpoint and configuration of the sensors used for data capture.

Figure 1. Annotated images of one left-handed and one right-handed player from novel dataset of baseball hitters at bat.

1.1. Problem Statement

Single frames taken from a walking pattern are likely to be highly similar across subjects, as the uniqueness of gait emerges over time. The dynamic shifts when considering posture, pose, or stance during a more complex activity a person has learned. In a sense, static characteristics become behavioral in this context: features like the placement of a boxer's feet or the grip of a tennis player on their racket are distinct and distinguishable from those of others engaged in the same activity. In this work, we seek to prove the feasibility of this concept for use in a human identification task. We select the field of baseball for initial study due to the ready available of high quality labeled footage and the surfeit of mechanical analysis that has been applied to the sport in prior research. By constructing a novel dataset of professional hitters at bat and performing feature creation and classification against this data, we aim to demonstrate the potential to identify an individual from a single image in an operational setting.

1.2. Significance

Although some recent biometrics research has touched on the idea of pose identification, the concept has yet to be thoroughly explored, even in feasibility study. Leveraging behavior from a still image has the potential to improve collectability and therefore user uptake when compared to gait-based biometrics, while requiring less processing power and sensor capture quality than technologies like face identification. In a defense application, a deployed pose identification system might be able to differentiate members of a squad by their shooting stance. Such a system could function at night or in adverse environments via night-vision or thermal imaging sensors, operating conditions that would preclude the use of many traditional biometric technologies.

The following sections will briefly review the state of the art and related work, discuss the approach taken and experimental setup, present preliminary results and analysis, and conclude with future research directions.

2. Related Work

Gait authentication can be broadly categorized into model-free and model-based approaches [7]. Model-free approaches analyze aggregate values from silhouettes or profiles over the duration of a walk, rendering them less applicable to this study. Model-based gait biometrics build a mathematical model of the walking subject's skeletal keypoints at each frame, where the model itself is either constructed top down (human subject identified prior to skeleton generation) or bottom up (keypoints identified first and then assembled into skeletons). OpenPose [2] is a popular bottom-up model widely used in biometrics and biometrics-adjacent work, including applications to baseball as a coaching aid [4].

The location and angle of the camera or sensor relative to the subject introduces a so-called viewpoint issue common among gait analysis techniques, especially those that do not impose rigid constraints on the walking patterns under observation (free-style walks). [3] overcomes this challenge by introducing a "Center-of-Body (COB)" relative coordinate system in three dimensions, using three hip joints facing forward as a new origin point. [1] contributes joint-relative angle and distance features, which [5] combines with COB-relative coordinates to introduce "Joint Replacement Coordinates" (JRC). [6] discusses strategies to account for occlusion of body keypoints, an increasingly important problem under realistic conditions. Although only [5] mentions the potential for pose to be used as a biometric, the coordinate scheme and features of [3] and [1] have inspired the approach discussed below.

3. Approach

Given the consistent angle of footage across Major League Baseball broadcasts, we were able to leverage a fixed-angle view of the subject hitter for input to the OpenPose bottom-up keypoint detection library. Following generation of hitter skeleton keypoints, we translate all two-dimensional coordinates from a top-left origin to a new frame of reference with origin at the mid-hip keypoint, similar to the approach taken in [3]. We then augment these translated coordinate values with complex angles and distances informed by domain knowledge of batting stances. Distance metrics are normalized to the height (difference between top-most keypoint and bottom-most) of each subject. The full steps for feature creation are as follows:

1. Capture image of baseball player at bat from center field camera angle
2. Input image to OpenPose for keypoint generation
3. Translate all coordinates using BODY_25 MidHip keypoint [2] as new origin
4. Augment coordinate values with a set of complex features informed by knowledge of the domain. For baseball, we create the following 14 angle and distance features:
    - back angle: angle from knee to hip to upper back
    - left/right shoulder angles: angle from hip to shoulder to elbow
    - left/right elbow angle: angle from shoulder to elbow to wrist
    - left/right knee angle: angle from hip to knee to ankle
    - stance width angle: left ankle to center hip to right ankle
    - left/right ankle angle: knee to ankle to toes
    - foot distance: distance between feet
    - center to bat distance: distance from center hip to wrists

- nose to bat distance: distance from nose to wrists

Although elements of this approach are specific to baseball, we anticipate that the core concepts will generalize to other unique poses. The model used to generate keypoints can be exchanged to better fit new domains, especially in cases where view-dependence is impossible or 3D coordinates would better suit complex feature generation. Likewise, the set of complex features calculated will change based on the nuances of the pose, posture, or stance under consideration.

4. Experimental Setup

Working from a set of 25 left-handed and 25 right-handed Major League batters with a large number of plate appearances in the 2021 season, we capture 5 images of each player from published footage of games in which they appeared. From this set of 250 images, we apply the OpenPose keypoint detection library to create an initial dataset of 250 skeleton vectors in JSON format.

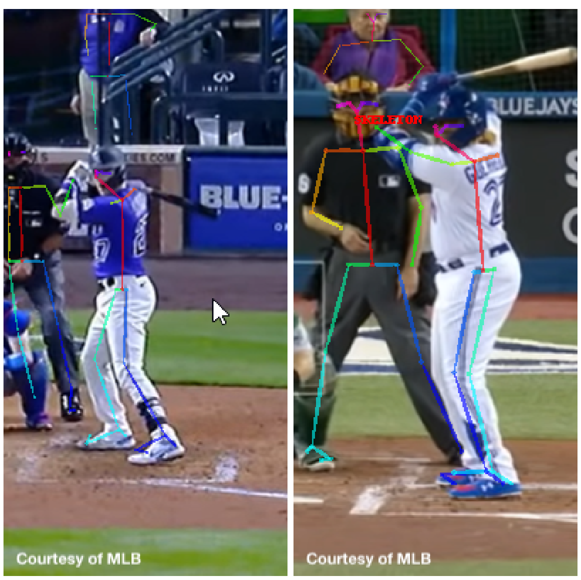

Figure 2. Common errors in initial keypoint dataset: batter's left forearm assigned to incorrect skeleton (left), wrong skeleton selected (right).

Several processing steps were needed to correct for common errors in the raw set of data. Keypoints were frequently detected for spectators in the background as well as the umpire. Due to the bottom-up approach of OpenPose keypoint detection, mistakes in skeletal generation and merge errors were common as well. See Figure 2 for examples of two common errors.

The following pre-processing steps were applied to correct errors where possible and prune bad data where not:

- For each image, select the most complete (largest number of keypoints) skeleton detected. Label this skeleton on the rendered image

- Manually review these labels for correctness. In some cases where the wrong skeleton is selected, the JSON could be updated to substitute the correct individual

- Where mistakes in keypoint detection were beyond repair (such as in Figure 2 left, where the player's arm has been mis-assigned), the keypoints were dropped from the dataset

Following these steps, the final dataset contained 232 skeletal feature vectors containing 75 values: 25 keypoint coordinates (X and Y) and a confidence score for keypoint correctness assigned by OpenPose for each. As described in Section 3, each coordinate was translated around middle hip as a new origin point. 14 complex angle and distance features were calculated for each skeleton. In cases where these features refer to derived distances, such as the set of both wrists, the midpoint was taken. As the new origin point, the middle hip keypoint was dropped from each feature vector, leaving a total of 62 features for each skeleton.

Even after correcting some generation errors and dropping malformed skeletons, keypoints where still missing in many feature vectors. In order to accommodate classification algorithms that do not accept missing values, these keypoints were imputed using the mean values from complete skeletons of other left or right-handed players, as appropriate. Of the maximum of five keypoint sets for each player, three were used for training and up to two were reserved for test. Three players had only three valid feature vectors remaining after pruning and were omitted from experimentation.

Classification Algorithms
Linear SVC (all features)
Linear SVC (20 best features)
SVC (RBF kernel)
Gaussian Naive Bayes
K Nearest Neighbors (k=3)
Decision Tree

Table 1. Classification Algorithms selected.

5. Results and Discussion

Table 1 lists the classifiers applied to the dataset described above. For the Support Vector Classification run on 20 features, a correlation-based feature subset algorithm was run to select the 20 most highly correlated features.

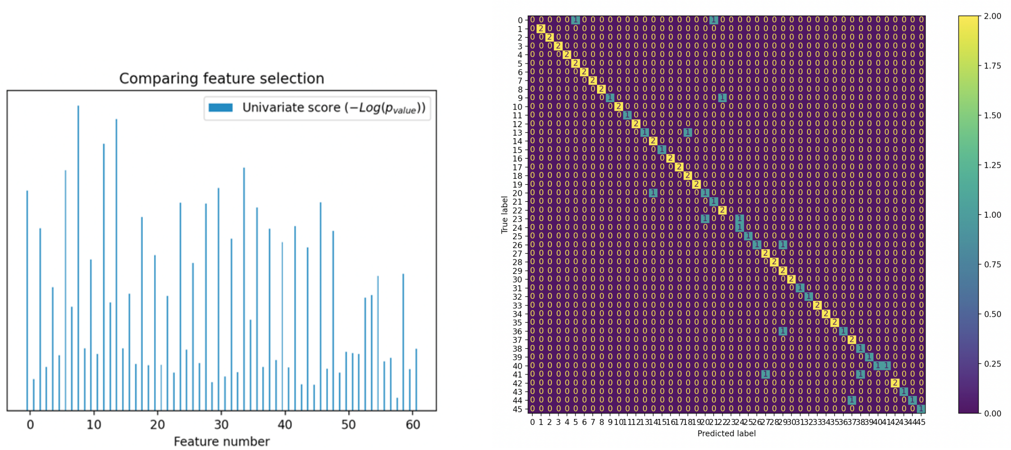

Figure 3. Results of ANOVA statistical filter feature selection (left) and confusion matrix resulting from best performing classifier (right).

Algorithm	OpenPose Confidence $\geq .1$	OpenPose Confidence $\geq .2$	OpenPose Confidence $\geq .3$
Linear SVC (all features)	**79.0%**	**82.7%**	**84.0%**
Linear SVC (20 best)	63.0%	63.0%	63.0%
SVC (RBF kernel)	74.1%	76.5%	76.5%
Gaussian Naive Bayes	33.3%	28.4%	32.1%
K Nearest Neighbors (k=3)	60.5%	64.2%	61.7%
Decision Tree	46.9%	54.3%	51.9%

Table 2. Preliminary classification results.

The results of this process are shown in Figure 3. OpenPose outputs a confidence value for each keypoint generated; thresholding this confidence value influences the accuracy of some classification algorithms. The experimental accuracy for each classification accuracy at three increasing thresholds is shown in Table 2.

5.1. Discussion

The statistical feature analysis shown in Figure 3 exhibit an alternating pattern for the first 48 features, which correspond to the translated X and Y coordinate values for the skeletal keypoints. This reflects the high correlation between X coordinates and the left or right-handedness of each player, and also serves to explain the weak performance of the support vector classifier run on the 20 most correlated features.

We hypothesize that Naive Bayes performs poorly due to the high dependence among many of the features, and K-nearest neighbors and decision tree classifiers lack sufficient data for strong performance. The confusion matrix from the best performing classifier (LinearSVC with OpenPose confidence threshold 0.3) highlights that left-handed players are never confused with right-handed players and vice versa, suggesting that an ensemble of classifiers, first to determine dominant hand and subsequently to identify the individual, could possibly yield better performance. Figure 4 shows perfect accuracy resulting from a LinearSVC differentiating between left and right-handed players in our dataset.

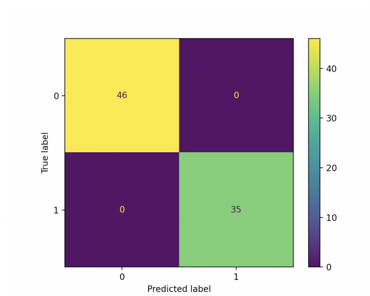

Figure 4. Confusion matrix from LinearSVC predicting left or right-handedness.

6. Conclusion

While gait recognition is gaining traction in the field of biometrics, pose analysis has not been viewed as a reliable source of identification. This may be true with common poses such as the walking or sitting pose, but we believe we have demonstrated that if the pose is unique enough, identification can be achieved. Through the creation of the skeleton and using domain knowledge of the pose, it was possible to construct a complex dataset that contains features that can be used to accurately identify. In this instance the baseball batter's stance was used as a proof of concept, however, there are other unique poses that may apply; and with more research and data collection they can be explored.

6.1. Technical Challenges

Missing values resulting from OpenPose keypoint detection was a technical limitations to our work, with the need to impute scores from the mean of other players likely lowering classification performance. Similarly, incorrect keypoints, some erroneously merged from adjacent skeletons, were largely impossible to solve without discarding whole individuals. In rare cases, JSON values could be repaired. Moving to a top-down model has the potential to improve skeleton detection, both in terms of missing values and the accuracy of located keypoints. Likewise, applying strategies from [6] to infer occluded keypoints could yield similar improvements.

6.2. Future Research Directions

While we believe we have succeeded in proving the concept of human identification from a unique pose, we suspect that classification approaches more finely tailored to the dataset created would yield better accuracy. For instance, an ensemble of classifiers as discussed in Section 5.1 might enable a deep learning approach to learn the nuances of each stance within the subset of left or right-handed individuals. An artificial neural network might also be applied to the task of predicting the next action taken by the subject; can it be predicted if the batter swing at the current pitch?

The implementation of a top-down model for keypoint detection would not only improve detection accuracy, but could also output three dimensional coordinates. We expect that this additional channel of information would produce better classification accuracy, especially when applied to less controlled circumstances.

Given that this work represents a feasibility study, the most significant future work lies in the application of the concept to an environment where the identification of an individual would be of more value. One key example is identifying a person holding a weapon by their shooting stance, which shares many characteristics with a batting stance (see Figure 5). As long as skeletal keypoints can be generated, such an approach could operate over nightvision or infrared sensors, conditions that would make more traditional identification techniques difficult or impossible to apply.

6.3. Acknowledgements

All baseball player images have been used with express written permission by Major League Baseball. All images are sole property of Major League Baseball. Thank you to Kevin Carey, Global Media Programming & Licensing Sr. Coordinator at Major League Baseball, for enabling this research.

Figure 5. Paper coauthor demonstrating keypoint detection of a shooting stance.

References

[1] F. Ahmed, P. P. Paul, and M. Gavrilova. Kinect-based gait recognition using sequence of the most relevant joint relative angles. *Journal of WSCG*, 23:147–156, 07 2015.

[2] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[3] N. Khamsemanan, C. Nattee, and N. Jianwattanapaisarn. Human identification from freestyle walks using posture-based gait feature. *IEEE Transactions on Information Forensics and Security*, 13(1):119–128, 2018.

[4] Y.-C. Li, C.-T. Chang, C.-C. Cheng, and Y.-L. Huang. Baseball swing pose estimation using openpose. In *2021 IEEE International Conference on Robotics, Automation and Artificial Intelligence (RAAI)*, pages 6–9, 2021.

[5] P. Limcharoen, N. Khamsemanan, and C. Nattee. View-independent gait recognition using joint replacement coordinates (jrcs) and convolutional neural network. *IEEE Transactions on Information Forensics and Security*, 15:3430–3442, 2020.

[6] Y. Ma, C. Wei, and H. Long. A hybrid approach using human posture and contour for gait recognition under body occlusion. In *2020 7th International Conference on Information Science and Control Engineering (ICISCE)*, pages 23–26, 2020.

[7] C. Prakash, R. Kumar, and N. Mittal. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review*, 49(1):1–40, 2018.