

# Optimized Pose-Based Gait Analysis for Surveillance

Apoorva Parashar<sup>1</sup>, Anubha Parashar<sup>2</sup>, Vidyadhar Aski<sup>3</sup>

<sup>1</sup> Maharshi Dayanand University, Rohtak, India

<sup>2,3</sup> Manipal University Jaipur, India

anubhaparashar1025@gmail.com

**Abstract.** In the most recent decade, researchers have lavished attention to human activity analysis, owing to its potential use in various fields. We suggest an alternative method for dealing with individual human joints (vertex positions) based on a model-based approach. The elliptic Fourier descriptors parameterize the movement formats that reflect the joint movement as inferred from walk inquiry. The heel strike data is manipulated to reduce the dimension of the parametric models. Spatial and temporal features characterize an Individual's walk, and we are capturing both. For indoor and outdoor information, the lower leg, knee, and hip joints are effectively and precisely extracted. Various experiments are done to make a marker-free inquiry, and various zones were created. The exploratory results confirmed the proposed technique's efficacy in perceiving walking objects with a 92% accuracy in characterization rate. Our suggested methodology applies to various surveillance scenarios, including pedestrian, subway, bus stop, airport, and bank surveillance.

**Keywords:** Model-Based, Gait, Pose based, Biometrics, Surveillance

## 1 Introduction

Most computer vision research is integrated into the study of explained structures and, more specifically, the examination of human expression. This inquiry is fueled by the breadth of applications for human movement analysis, including virtual reality, savvy reconnaissance, human-computer interfaces, and athletic execution. The three key phases in a dream-based paradigm for examining human movement are exploration, following, and discernment. The final stage delivers a high-level representation based on the highlights extracted from the worldly video stream during the preceding phases. Indeed, mental considerations showed that the action of human joints provides sufficient evidence to visualize human movement.

Now, most systems for movement analysis are marker-based and economically viable. The arrangement of markers is mostly determined by the markers or sensors attached to critical areas of the human body. Regardless, most applications, such as visual perception, necessitate the arrangement of a robotized marker less vision system to distinguish the directions of the joints. Conversely, automated joint location extraction is a tremendously difficult task since non-inflexible human activity encompasses a broad range of possible movement modifications due to its extraordinarily

adaptable form and susceptibility to self-impediment. Dress style, division faults, and multiple views all pose significant challenges to precise joint containment.

Given the abundance of vision approaches aimed at separating appendages and the absence of approaches specifically designed to determine vertices, we suggest another method for separating human joints with greater precision at certain point masses by consolidating priori knowledge to optimize exactness. Our novel methodology employs a model-based approach for illustrating human stride activity by using elliptic Fourier descriptors. The walk architecture is combined to create a model that can be used to monitor and highlight correspondence. The suggested scheme can accurately exclude moving joints of the human body for both indoor and outdoor data collected under unconstrained environments. Additionally, an acknowledgement ability survey was performed using the walk to demonstrate the efficacy of the approach described.

## 2 Literature Review

RGB-D devices such as the Microsoft Kinect or other depth sensors will be used to capture the video footage. The use of these instruments is growing due to their low cost and low computing criteria in contrast to other 3D body pose estimation approaches such as marker-based motion capture solutions, as well as their less invasive nature in comparison to wearable sensors.

Fall detection is a widespread application in this area. Fall detection is modelled as two events in [7], using the given depth details. After determining that the individual is on the deck, an ensemble of decision trees is used to determine if a fall preceded this condition. Experiments are conducted over many time-series records, outperforming other state-of-the-art approaches. [8] describes a procedure for gait control that is built on a robotic device. An autonomous robot maintains a safe distance from the individual at home. It tracks three separate metrics to execute the prediction based on skeletal data: stride length, stride duration, and center of mass motion. The device is compared extensively to a consumer marker-based motion capture system [9], and many benefits are highlighted, including low expense, simplicity of usage, and limitless recording volume.

Regarding human gait measurement and tracking Stone et al. assess the precision of calculating walking pace, stride time, and duration using depth data from the Kinect, achieving positive preliminary findings [10]. Similarly, [11] extracts lower and upper body gait parameters from skeletal details, including stride detail and arm kinematics. A machine learning algorithm is used to model gyroscope-derived function signals. The mean disparity between the assessed and wearable sensors was less than 1%. [12] is devoted to detecting abnormal gaits by online analysis of skeleton results. To reduce the dimensionality of the noisy skeleton results, a nonlinear manifold learning technique is used. The normal gait samples are used to construct a mathematical model, and identification is made by comparing the latest findings to the model using Markov assumptions. This technique is then extended to subjects ascending a flight of stairs to spot simulated gait abnormalities. There are findings with both event-based

and frame-based classification, with the event-based results being significantly higher due to a refinement stage that identifies pattern shifts. Our approach is like this work, and since their dataset is freely accessible, the present study includes a comparison of findings.

### 3 Methodology

Figure 1 shows proposed methodology, model-based approach is proposed for determining the mutual paths of people who are strolling. About the fact that the Fourier structure is the most precise method for illustrating stage movement, most previous approaches relied on simple models [1] to distinguish step precise movement by using an evidence social case with a few boundaries. This is primarily due to unpredictability and high computing costs. Grant et al. [2] adopted another temporary evidence social case approach to extricating arbitrary shifting forms.

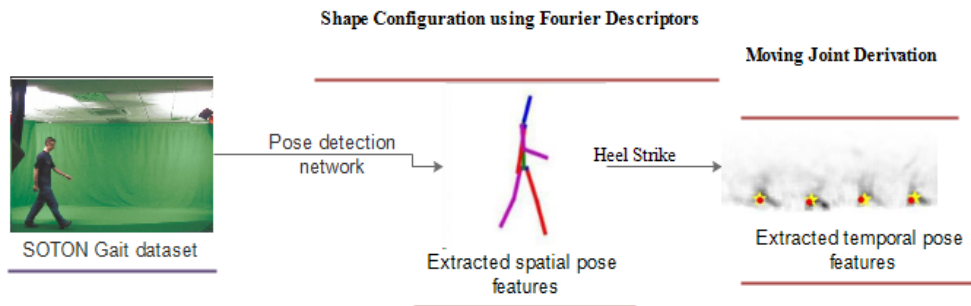


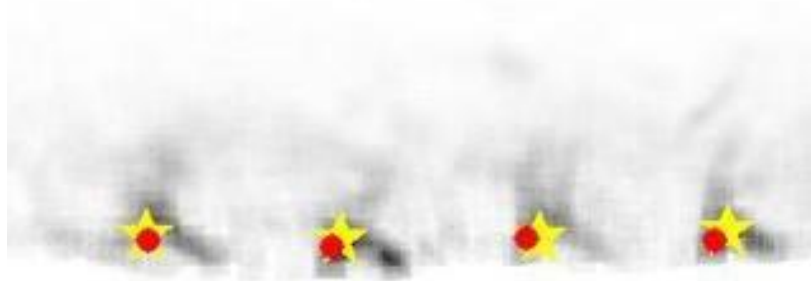
Fig. 1. Proposed methodology to collect spatial and temporal features.

#### 3.1 Shape Configuration using Fourier Descriptors

The layouts of moving shapes in a coherent framework are parameterized using Fourier descriptors. For a long time, the Fourier theory has been applied to the study of form bends and boundaries. The Fourier analysis allows you to extract highlights or descriptors from images that are essential for picture comprehension. Rising the parametric depiction of a bend in Fourier arrangement characterizes these descriptors. However, this will necessitate a massive amount of computing capacity, which would be impractical. On the other hand, the heel strike detail could be consolidated to reduce the multifaceted nature of the boundary space and, as a result, dramatically reduce the computational cost, as done here.

### 3.2 Heel Strike

The determination of the human phase time frame will provide useful information for determining the locations of human joints. Cutler et al. [3] suggested an ongoing method for estimating periodicity for an intermittent movement that is body-based. Rather, much like the walk move and step duration, the subject's heel strikes will have a specific measure of stride periodicity. In addition, the extraction of heel strikes may be used as a reliable cue to distinguish strolling people from other moving objects in the scene as shown in figure 2.



**Fig. 2.** Heel Strike Extraction.

During the strike process, the striking leg's foot stays in a single location for most of the walk period, whilst the rest of the human body moves forward. If we use a low-level element extraction technique (edges or corners), a dense region would be aggregated at the heel strike areas. Since the main goal of this investigation is to see human activity, it was agreed to use corners rather than edges because corners hold enough data to see human movement. In contrast, edges can create ambiguity in the extraction process due to the amount of data.

### 3.3 Moving Joint Derivation

The model-based technique is used to restrict the joints for walking persons to view them by their stride. The evidence social event mechanism is the most used method for extracting enunciated items. It is usually used in conjunction with the Hough Transform, which has two stages: I) global pattern extraction and ii) local feature extraction. Global extraction aims to find the best movement design based on a predefined standard, such as previous knowledge from the entire video sequence. The elliptic Fourier Descriptors are used to describe the movement model in a parametric structure. The Hough Change is used to determine the ideal boundaries by arranging highlight focuses (corners) over the entire sequence to the parametric potential. The increment often casts a vote in the aggregator room. The joint's exact condi-

tion is removed in the second step of the evidence get-together, an edge-by-outline quest. The illustration omitted at the principal stage guides the neighborhood extraction.

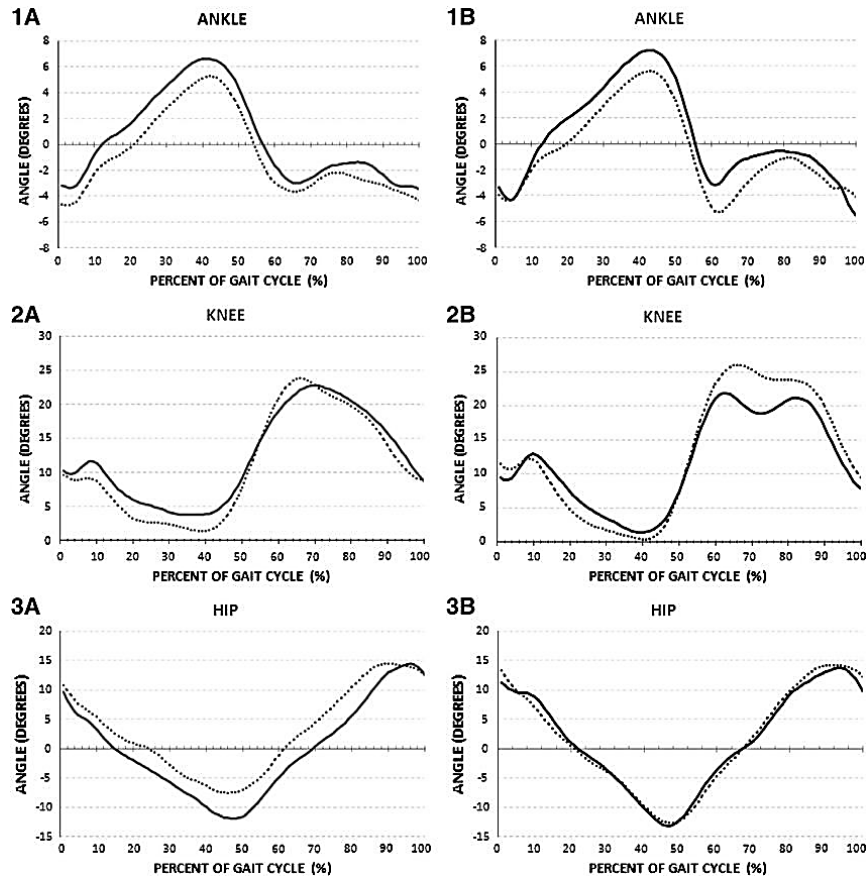
## 4 Results and Discussion

A calculation on a set of 120 unique subjects from the SOTON database [5] was used to demonstrate the feasibility of this technique. In an indoor domain with controlled conditions, 100 research database participants were photographed. The last ten subjects were photographed outside in natural light. The subjects walked from the review plane to the left, correct, then right.

Investigating visually marked details from 30 video arrangements, model formats representing joint movement were calculated as the mean qualities. Figure (2) depicts the joint mobility paths for the lower leg, knee, and hip between two progressive attacks on a similar leg. Since the elliptic Fourier descriptors can only speak to closed and ceaseless shapes, the model format was modified to a shut cyclic shape by circling back down the map. On corners pictures obtained using the Harris corner locator on each casing, the democratic method for the evidence social occasion technique is used. The investigation territory is constrained by anatomical details [4] of the body segment assets. Table 1 gives a summary and comparison of accuracy of various papers.

**Table 1.** Comparison of various accuracy.

Paper	Dataset	Accuracy
[13]	SOTON	55
[14]	CASIA B	63
[15]	SOTON	92
[16]	CASIA A	70
Ours	SOTON	99.2



**Fig. 3.** Motion Analysis of joints (1) Ankle (2) Knee (3) Hip.  
(A) Normal walking speed (B) Fast Walking speed

The proposed strategy effectively extracted 99.2 percent of the strikes from a total of 514 strikes for the position of the heel strike focuses as shown in figure 3. The average error for 65 strikes calculated by the equation vs strikes visually labelled is 0.52 percent of the person height. Since this is the most common explanation for standardization, the error is measured using Euclidean separation between the two strikes standardized to a measure of the individual's stature. The mean error for the locations of the extricated joints according to manual information of 10 visually named subjects is 1.36 percent of the subject's height. The measurement is performed on a subject who is dressed in Indian clothing that secures the legs. The locations of the joints are easily extricated, revealing the possibilities of this method of dealing with impediment. As a result, we may assume that the results obtained using this approach are predictable while using biomechanical data from Winter [6].

## 5 Conclusions

A model-based technique to separate moving joints by means of proof social event strategy is proposed. Spatial model formats for human movement are derived from the examination of walk information gathered from manual naming. Model formats are spoken to in a parametric structure dependent on elliptic Fourier descriptors. Walk information is exploited by means of heel strike extraction to decrease the boundary space dimensionality and therefore along these lines decrease the computational heap of the Hough Transform being utilized in the extraction procedure. The portrayed technique is demonstrated to work for both indoor and open-air situations with potential to limit joint positions with better exactness. Moreover, the acknowledgment intensity is assessed utilizing the information separated by means of the depicted model-based technique. The proposed arrangement has accomplished a characterization pace of 92% for individuals' acknowledgment. The model-based is fit to increasingly summed up deployment and this will be the core interest for future work.

## References

1. Cunado, D., Nixon, M., Carter, J.: Automatic Extraction and Description of Human Gait Models for Recognition Purposes. *Computer Vision and Image Understanding*90(1) (2003)
2. Grant, M.G., Nixon, M.S., Lewis, P.H.: Extracting moving shapes by evidence gathering. *Pattern Recognition* 35(5) (2002)
3. Cutler, R., Davis, L.S.: Robust real-time periodic motion detection, analysis, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2003)
4. Pheasant, S.T.: *Body sapce: Anthropometry, Ergonomics and Design*. Taylor & Francis (1988)
5. Shutler, J.D., Grant, M.G., Nixon, M.S., Carter, J.N.: On a large sequence-based human gait database. In: *Proceedings of Recent Advances in Soft Computing*, Nottingham, UK (2002)
6. Winter, D.A.: *The Biomechanics and Motor Control of Human Movement*. Second edn. John Wiley & Sons.
7. E. Stone and M. Skubic, "Fall detection in homes of older adults using the Microsoft Kinect," *Biomedical and Health Informatics, IEEE Journal of*, vol. 19, pp. 290–301, Jan 2015.
8. A. Staranowicz, G. R. Brown, and G.-L. Mariottini, "Evaluating the accuracy of a mobile Kinect-based gait-monitoring system for fall prediction," in *Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments, PETRA '13*, (New York, NY, USA), pp. 57:1–57:4, ACM, 2013.
9. "Vicon Motion capture system." <http://www.vicon.com/>.
10. E. Stone and M. Skubic, "Evaluation of an inexpensive depth camera for in-home gait assessment," *Journal of Ambient Intelligence and Smart Environments*, vol. 3, no. 4, pp. 349–361, 2011.
11. M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster, "Full body gait analysis with Kinect," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 1964–1967, Aug 2012.

12. "Microsoft Kinect SDK 2.0, Microsoft, Inc.," 2014.
13. H. L. Tavares, J. B. C. Neto, J. P. Papa, D. Colombo, and A. N. Marana, "Tracking and Re-identification of People Using Soft-Biometrics," in 2019 XV Workshop de Visão Computacional (WVC), Sep. 2019, no. March 2020, pp. 78–83, doi: 10.1109/WVC.2019.8876921.
14. J. Luo and T. Tjahjadi, "Gait Recognition and Understanding Based on Hierarchical Temporal Memory Using 3D Gait Semantic Folding," *Sensors*, vol. 20, no. 6, p. 1646, Mar. 2020, doi: 10.3390/s20061646.
15. C. Technology, "Multi-Level Feature Fusion for Robust Pose-Based Gait Recognition using RNN," *Int. J. Comput. Sci. Inf. Secur.*, vol. 18, no. 1, pp. 20–31, 2020.
16. W. Sheng and X. Li, "Siamese denoising autoencoders for joints trajectories reconstruction and robust gait recognition," *Neurocomputing*, vol. 395, pp. 86–94, Jun. 2020, doi: 10.1016/j.neucom.2020.01.098.