Grenze International Journal of Engineering and Technology, June Issue

AI-Driven Body Posture Recognition and Predictive Analysis of Long-term Health Risk

Kusum Hari[1], Malay Lokhande[2], Yash Marsattiwar[3], Khushi Lakhe[4], Riddhi Dongarwar[5] and Ujwalla Gawande[6]
[1-6]Yeshwantrao Chavan College of Engineering, Nagpur, India
Email: kusumhari7@gmail.com, malaylokhande007@gmail.com, yashmarsattiwar03@gmail.com, khushilakhe@gmail.com, ujwallgawande@yahoo.co.in, bhawalkrarita1991@gmail.com

***Abstract*— This research uses a model-fitting process guided by anatomical structures to estimate human body posture and skeletal system in three-dimensional space using just one RGB (Red Green Blue - refers to color) image. A three-dimensional mesh of a human body is reconstructed by generating the medial axis of an SMPLest-X body model with the help of Voronoi diagrams. The resulting representation of the body provides information about internal support structures without the need for multiple images of the same person from different points of view. When potential medial axis points are determined, a 17 joint, articulated skeleton that resembles the human body's structure is produced. The locations of the joints can then be used to extract key components of the individual's postural position including neck bending, upper body tilting, shoulder symmetry and alignment of the pelvis. In addition to determining these characteristics, this framework can also be used to evaluate whether the individual has a posture that is ergonomically safe (or poses risk for long-term health problems). From a cost-effective and accessible viewpoint, this model has been demonstrated to provide accurate body posture reconstruction and reliable posture-based health assessment, thus providing a viable means for real-time ergonomic monitoring.**

***Index Terms*— 3D Human Posture Estimation, SMPLest-x Model, Medial Axis Extraction , Anatomically Accurate Skeleton Reconstruction, Single Image 3D Modeling, Voronoi Skeletonization, Posture Feature Analysis Ergonomics Posture Assessment.**

I. INTRODUCTION

Across biomechanics, ergonomics, and the study of human movement, it is very important to accurately recreate how people stand and move. Traditional motion capture systems are very accurate but because they require cameras that capture multiple views and use markers, they cannot be used in less expensive, less constrained settings [1]. Geometric modeling began with the help of the Medial Axis Transform (MAT) algorithm. Internal symmetry was first introduced as a structural concept by Blum [2] and was later expanded to 3D solid objects by Choi and Min [3]. The applications of MAT methods to more complicated mesh geometries were further developed by Liu et al. [4]. While these techniques have been used to provide skeletal representations that keep the basic topology of the original surfaces, they are still limited by needing to have large amounts of surface information and mul tiple camera inputs.
Parametric models based on deep learning such as, SMPL [6] and SMPL-X [7] are able to produce 3D reconstructions of humans using only one image. The use of certain techniques, such as mesh recovery and scene awareness [8]-[10], produces even better results.

Grenze ID: 01.GIJET.12.2.298_22

However, they only produce surface meshes and do not create anatomically accurate skeletal models. The system that we propose uses the SMPLest-X system for 3D reconstruction and then uses the Voronoi/MAT techniques to extract the medial axis and then refine the joints using the EM algorithm [5]. The result of using this method is a well-defined 17 joint skeleton that can be used to help assess postures accurately and to assess the risk of injuries to the musculoskeletal system by using only a single RGB image [11], [12].

II. Literature Review

The Medial Axis Transform (MAT) was used in the early stages of the geometric research on recovering skeletal representations from 3D shapes; using a 3D shape represented as a model, Blum first explained how to describe a shape based on its medial axis from a 2D shape, which was then extended into complex 3D solids (in general) and later adapted to be used with more general mesh models (3D objects made up of points). Although Menier et al. were able to apply medial geometry for human pose recovery and demonstrated that skeleton estimation was improved with respect to internal symmetry, their approach relied on multiple views and introduced noise to the estimation process.

The emergence of a new class of deep learning-based models of the human body has revolutionized monocular 3D reconstruction. Advanced parametric models, including SMPL and SMPL-X, allow for highly expressive body modeling from a single RGB image, while later work improved the reconstruction of body meshes by using scene context and applying refinement techniques. While these advances were notable, they continue to miss out on the ability to provide anatomically consistent internal skeletons to use when estimating body posture.

REBA [11] has been applied as a method of determining the risk from poor posture in the area of Ergonomics. Recent studies using automation have utilised computer vision techniques for classifying posture (Parakkat et al., 2017) [12]. The main limitation of most of the current tools for analysing posture is that they primarily rely on either 2D Key-Point data or wearable sensor systems, which results in a decrease in spatial fidelity and therefore necessitates the use of a MAT algorithm based Skeleton Extraction and Reconstruction of the parametric body mesh for the purpose of implementing more comprehensive and potentially more accurate 3D Posture analysis based on anatomical structure.

III. Proposed Methodology

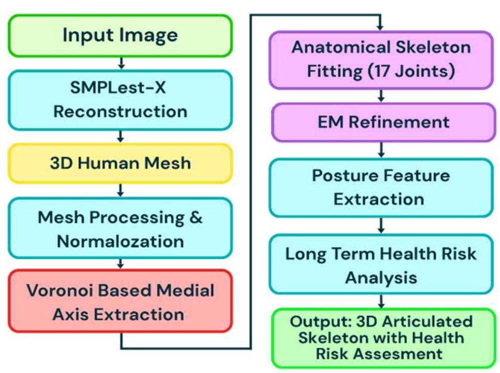

Figure 1. Workflow of the proposed method for 3D skeletal reconstruction and posture health evaluation.

A. Input RGB Image

Our pipeline's input is a standalone RGB image. In contrast with typical pose estimation systems which generally require multiple camera setups or motion capture markers [1],[14], the system can obtain a 3D Equivalent Representation from an RGB image without the use of multiple cameras and motion capture markers.

B. 3D Reconstruction using the SMPLest-X Model

Using our adjusted SMPL-X model, which is described in references [8] and [9], the 2D image is now processed into a mesh that can be analyzed as a three-dimensional, watertight surface. The primary purpose of this reconstruction is to create an accurate three-dimensional representation of the person depicted in the original 2D image based on their body shape and pose parameters determined by using a convolutional neural network trained via a supervised learning approach. Moreover, the resulting mesh will have realistic proportions, which are preserved via the parametric construction of the mesh.

C. Generation of the 3D Human Body Mesh

The process of converting a two-dimensional image into three-dimensional representations is comprised of multiple elements: the direct creation of three-dimensional meshes and the storage of three-dimensional meshes as representational data. This transformation between two dimensions and three dimensions is accomplished by reconstructing a person's three-dimensional form through a vertex-based disambiguation method. Vertex representations maintain all of the geometric detail of the original source image (i.e., depth and curvature) while allowing us to analyze all types of physical attributes of the person's body. These geometric attributes are the basis for all other processing steps, such as skeleton extraction and evaluating body posting.

D. Mesh Pre-processing and Geometric Normalization

The Delaunay tetrahedralization from uniformly sampled surface point data from the normalized mesh was used to extract the medial axes of the normalized mesh. The circum centres of the Delaunay tetrahedralization provides the locations of the Voronoi points used to define the medial axes [3],[4]. All the Voronoi vertices represent the location of the internal skeleton for the individual. Furthermore, Voronoi points that are located in close proximity to the surface or have a small radius are removed to eliminate noise [2],[5],[7]. Finally, the remaining interior Voronoi points can provide a smooth medial axis representation of an individual which can be used to determine the internal body structure of an individual.

E. Extract Medial Axis (Voronoi Method)

The Delaunay tetrahedralization from uniformly sampled surface point data from the normalized mesh was used to extract the medial axes of the normalized mesh. The circum centres of the Delaunay tetrahedralization provides the locations of the Voronoi points used to define the medial axes [3],[4]. All the Voronoi vertices represent the location of the internal skeleton for the individual. Furthermore, Voronoi points that are located in close proximity to the surface or have a small radius are removed to eliminate noise [2],[5],[7]. Finally, the remaining interior Voronoi points can provide a smooth medial axis representation of an individual which can be used to determine the internal body structure of an individual.

F. Fit Anatomical Skeleton Model

Anatomical skeletal structures are obtained by combining the extraction of a medial axis via a region-based method in conjunction with Principal Component Analysis to create a biologically-accurate articulated skeleton framework containing realistic joints, which will allow the body to move in a manner closely resembling an actual human being. In addition to using PCA, the skeleton is constructed using common human proportions provided that the human body is partitioned into major anatomical regions (i.e., hips, legs, arms, shoulders, and spine) based on the properties of their symmetry and the clustering of their parts spatially [1], [3], [13] (see Fig. 1). Additionally, there are 17 separate joint points located on each segment of the biological human skeleton to allow proper movement in a logical way in a biologically correct fashion.

G. EM Refinement

To improve the accuracy of the determined joint locations on the skeletal structure, the Expectation–Maximization (EM) algorithm [5] is applied to calculate the expected probability that a given medial point is associated with the skeletal segment nearest to it (E-step) and to modify the location of those joint points so that the distance (weighted by the number of points) between the medial and skeletal segment is minimized (M-step). This process continues iteratively until the joint points stabilize and converge upon one another. The resulting joint points are then combined into an articulated anatomical structure which smoothly follows the registered 3D (three-dimensional) mesh of points, reducing error in the placement of joints and solving small inconsistencies in the model.

H. Compute Posture Features

After the fitting of the skeleton, the system calculates the critical postural attributes of the body with respect to the place of the individual and Balance based on 3D coordinates of the Joints. Examples of the identified postural attributes include:

- Angle of Inclination in Your Spine: The angle created as the spine deviates from its vertical axis.
- Tilt of Your Neck: Between the head and neck joints relates to the angle at which the head has moved compared to the rest of the neck.
- Symmetry of the Shoulders: Left and right shoulder elevation can be evaluated based on their difference from an ideal shoulder position.

- Pelvis Position in Relation to Hips: The rotation of the left and right hip joints helps to find any symmetry about the pelvis.
- Leg Placement and Weight Distribution: Leg posture is used to identify postural conditions such as crossed or unnaturally positioned legs.

These extracted posture attributes are compared against known reference values for ergonomics [17], [18], to categorize a specific posture as upright, dowager's hum, leaning backward, or toward the side. These categories are used as a basis for generating kinematic indicators which can later be used to assess and analyze the health risks associated with body posture.

I. Analyze Long-Term Risk to Health

This component of the Health Risk Analysis module provides analysis of postural states using geometric measurements and angular limits derived from the 3-D skeleton. Every frame is assessed through a rule-based classification system that examines four key posture dimensions:

- Classification of Body Position: Determines the type of position (upright, slouched or backwards leaning) using the angle between the torso and head.
- Lateral Body Tilt Analysis: Measuring body position deviation from horizontal.
- Head Orientation Assessment: Determining whether the head is forward or backward from neutral, indicating possible neck strain or shoulder stress.
- Lower Limb Position Analysis: Detecting whether the legs are crossed or not equal leg positions that may lead to pelvic imbalance and reduced blood circulation.

The cumulative Posture is made up of the deviation noted by the system and its frequency and degree of non-neutral posture reported over time. In addition to generating the cumulative Posture, the system produces a textual report of the User's health assessment including summarised information about the health issues detected (e.g., slouched posture, head tilt, Lateral Body Tilt, crossed legs), and associated health effects (e.g., stress on the spine, neck discomfort, or muscle fatigue). The system converts raw geometric data into meaningful ergonomic information to help identify and prevent the risk of poor posture from occurring early in the period it is detected.

J. Output: 3D Articulated Skeleton with Health Assessment

A 3D body mesh, which shows the skeleton model of the person with the mesh is constructed with Open3D to visualize the person's anatomical locations when properly aligned to show the individual's posture. A complete health analysis report will also be provided, detailing the classification of posture, angles and the overall evaluation of posture at the joints. The use of graphical visualisation in combination with numerical analysis will allow the user to see and document the characteristics of their posture easily. The combination of both of these technologies will enable the user to monitor and evaluate their posture and will provide application benefits in healthcare, on-site ergonomics and rehabilitation.

IV. FITTING MODEL EM ALGORITHM

A joint fitting algorithm uses EM fitting techniques create joints at logical points to produce flowing transitions from medial axis to their skeletal model. The fitting algorithm utilizes a probabilistic optimization method which allows for the iterative optimization of the position of each joint and maximizes the likelihood of obtaining medial axis points on any of the bones. The expected values of medial axis points along their respective lateral lengths will provide a criteria against which to develop each joint during the iterative optimization of the joint position along the respective lateral length. The generative formulation of the model is expressed as:

$$P(S|\{Xi\}) \propto P(S) \prod i \sum a_i \ P(a_i|S)P(Xi| a_i,S)$$

where:

$p(a_i = j \mid S) \propto \text{segment_length}_j$represents the prior probability of a medial point $X_i$being associated with segment s_j.

$P(X_i \mid a_i = j, S) = \mathcal{N}(d(X_i, s_j), \sigma_j^2)$is the likelihood modeled as a Gaussian distribution of distance $d(X_i, s_j)$from segment s_j.

P(S)denotes a uniform prior over possible skeletal configurations.

The EM algorithm alternates between two main steps:

- E-step: Compute the posterior probabilities $P(a_i = j \mid X_i, S)$, is the probability that a medial point belongs to a specific segment of the skeleton.
- M-step: Maximize the expectation function:

$$F(S)=\sum i\sum j\ E(a_i = j)\log P(Xi|a_i = j,S)$$

The skeleton parameters (S) will be updated with respect to these medial points using an iterative refinement process based on the Levenberg Marquardt optimization technique.

This optimization method is advantageous for solving means and angle pairs through iterative convergence; thus producing a more anatomically accurate 17-joint framework which minimizes distances between medial points and the appropriate skeletal segments. The EM-based refinement process further guarantees smooth structural connections and proportionality within limbs, while providing stable connections between joints, enabling future extraction of posture features and evaluation of individual health.

V. Results and Discussion

This section provides an experimental evaluation of the "AI-Powered Body Posture Recognition And Predictive Health Risk Assessment" system described herein. In this phase of the project, the focus is to evaluate how well the entire pipeline works (reconstruction of human bodies in 3D using SMPLest X), derivation of bone structure/3D skeletons using Voronoi/MAT features, and creation of rules for identifying body posture and health indicators. Experimental evaluations were used to evaluate if the system could accurately and realistically generate a 3D mesh representation based on just a single RGB image, to create accurate 3D skeleton representations with the correct joint connections, and to provide accurate posture classification for all five identified postures (i.e., upright posture, slouching posture, backward leaning posture, side-leaning posture, and crossed-leg posture) and to develop understandable feedback regarding user's potential health risks associated with each of those five postures. Therefore, the results presented here provide evidence for both the accuracy of the methods proposed that are 3D geometric reconstruction as well as the accuracy of that proposed methods ability to assess how well users maintain appropriate body postures for optimal long-term health.

Skeleton Extraction Using the Medial Axis Transform (MAT)

After reconstructing 3D models of human bodies using the SMPLest-X model, we proceeded to conduct evaluations to extract a 3D representation of a medial skeleton (the internal structure) from these meshes, using Medial Axis Transform (MAT). Our goal for conducting these evaluations, was to validate the ability of MAT to produce a single-dimensional skeletal form, which retains the anatomy as well as the internal shape/growth of the human body. Through the evaluations conducted, we found that the MAT provided very accurate representations of the cross-section volume (volume of the mesh) of the SMPLest-X models through the extraction of the 'baseline' curves, hence providing the simplest and easiest representation of the 3D meshing forms; i.e., appropriate to fit skeletons, or find where joints should be placed.

A. MAT Processing Observations

To further evaluate how the MAT technique operates and performs when the technique is used on human body meshes, we examined the intermediate steps in the process of finding the medial skeleton structures.

Voxelization Output

The reconstructed three-dimensional mesh was converted to a three-dimensional voxel volumetric grid (or volumetric representation) of the human body in a uniform volumetric way, thereby providing a morphological representation of the human body that could be used to perform the morphological processing needed to find the human body's medial axis.

Medial Axis Derivation

In the thinning step of the MAT, the peripheral voxels located outside the medial axis were progressively removed until a single stable (or stable) medial curve structure remained. Comparison of the resultant skeletons with that of the SMPLest-X mesh indicates that the medial axes for both produced, in addition, the anatomical areas of major bone structures were aligned throughout the head, shoulder, torso, thigh, and arm areas.

Filtering and Noise Suppression:

Finally, through the use of noise reduction techniques, small clusters of detached voxels and redundant branches were removed from the skeletons generated. The resultant processed medial skeletons continued to remain smooth and anatomically significant as recorded across all subjects from this experiment.

EM-based Skeleton Fitting

Findings from the EM fitting process indicate that the EM based fitting process makes use of medial points derived from the medial curves to adjust the positions of the 17 joints in the articulated skeletal representation through a series of iteration-based optimizations that provide the minimum distance between the fitted skeleton (to the medial structure) after the EM fitting process has been completed. The qualitative results indicate that the fitted skeleton is fitted closely to the anatomical medial axis and accurately captures the anatomical representation of the human body, including the spine, shoulder girdle, pelvis, upper and lower extremities. Based on these results, it is clear that MAT-generated medial geometry provides an excellent reference point for fitting a model-based skeleton representation.

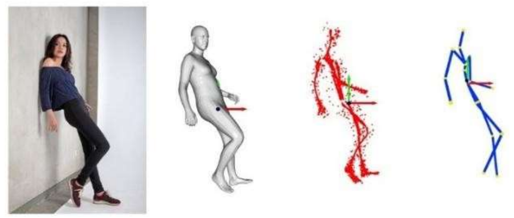

Figure 2. Workflow from input image to skeleton generation

The key steps in the entire proposed pipeline are depicted in Figure 2 with respect to creating a model of a fully articulated skeleton using only a single RGB image as input. After creating a full 3-D mesh from the RGB image, the next step is to compute the Voronoi-based medial axis of the skeleton for the purpose of determining the internal structure/geometry of the skeleton. Once the medial points are determined, an anatomical skeleton consisting of 17 joints is aligned with these points and subsequently refined using the EM algorithm. Through this step-by-step process, the system is capable of accurately reconstructing anatomically-generated, three-dimensional representations of a fully articulated skeleton based solely upon an individual's single two-dimensional image.

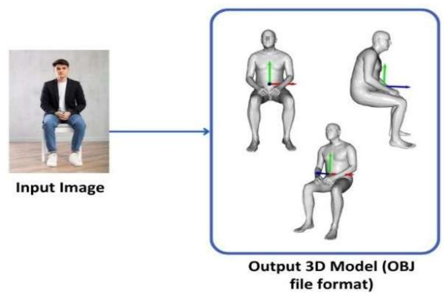

Figure 3. 3D human model reconstruction from an input image using SMPLest-X

The ability of the system to convert a one-dimensional RGB image into a three-dimensional human model can be seen in Figure 3. The left-half of the figure shows the physical body of the seated subject and in right side shows the generated 3D mesh model of the person using a combination of different viewing angles of the person created using SMPLest-X. This shows that the reconstructed model accurately represents how the human body should look based on certain input parameters such as how the subject sits and how long they are (height), along with how their arms and legs are positioned based on other anatomical reference points (vertical, lateral, forward) shown by the annotations on the figure. Based on these results, the reconstruction module can reliably retain important physical aspects of posture that will be used in further analysis to derive skeletons.

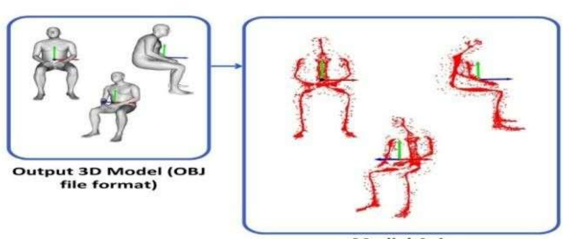

Figure 4. Medial axis point extraction from the 3D human model

Figure 4 shows how the reconstructed three-dimensional (3D) mesh can be turned into a two-dimensional (2D) representation of the mesh through its medial axis (Mat). The SMPLest-X mesh is shown on the left, and the second panel shows the clustering of dense medial points generated by the Voronoi-based Medial Axis Transform (MAT). The points represent the internal symmetry of the body, and correspond to anatomically important locations in the human body, including the spinal column, trunk, and arms & legs. Because the medial representation preserves the topology of the original mesh, this will enable a high-quality fit of the articulated skeletal model in the next stages of processing.

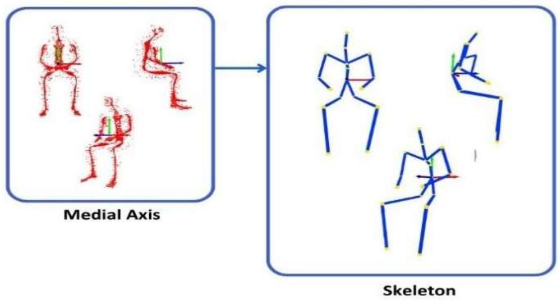

Figure 5. Skeleton generation from medial axis points

The transformation process from the medial axis representation to the 17 joint articulated final skeleton illustrated in the figure 5 is described. The left half of the figure is the medial axis points that show the internal geometric structure of the body while the right half of the figure is the anatomical skeleton from PCA-based segmentations and EM based refinement fitted to those medial points. The resultant skeleton retains the limb connectivity and joint hierarchy as they would normally occur in a body with the same posture as the reconstructed skeleton, thus indicating that the system can reasonably take a large number of points from a dense medial geometry and reconstruct them into an anatomically valid structure.

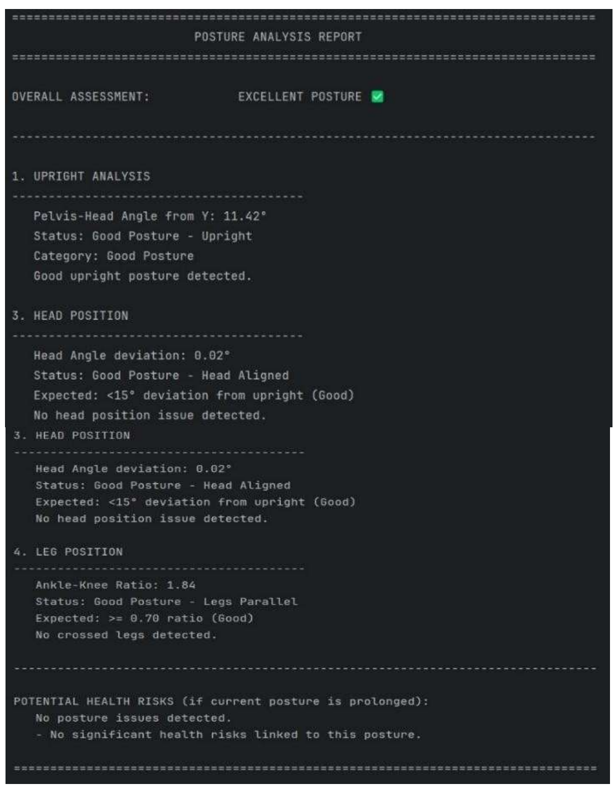

Figure 6. Posture Health Evaluation

In addition, figure 6 further shows the results from our rule based posture evaluation module, which analyses head orientation, upright verticality, lateral body incline, and proper leg position by utilising the extracted three-dimensional (3D) joint coordinates from the extraction phase. In this particular example, the results indicate that the system finds excellent posture (the angle of inclination between the pelvis and head is 11.42°, indicating proper vertical alignment), there was little to no head deviation from a centered position (0.02°), and the ratio of the legs was 1.84, indicating a parallel and not crossed position. There were no indications of slouching, side leaning, or crossed legs reported by the system, therefore the system does not identify any health risks associated with these results. These results indicate that the model can not only accurately reconstruct a 3-dimensional skeleton structure but can also provide accurate assessments of posture quality and ergonomic assessments that are meaningful to the user of the data.

VI. Conclusion

The proposed framework illustrates a comprehensive and effective path to creating anatomically precise 3D skeletal models that can be utilized for a health-related posture assessment based on the input of one RGB image. In integrating both 3D reconstruction from the SMPLest-X technique and skeleton extraction via Voronoi/MAT with EM refinement, this framework provides a means to reliably recreate the body shape and the body's interior skeletal structural organisation. Confirming that the rebuilt meshes, medial axes, and articulated skeletal configurations have preserved the correct anatomical arrangement, therefore providing an accurate identification of joint locations and major positioning inclinations throughout multiple types of sitting and standing configurations.

While the reconstructed model retains the structural consistency of the original model, the rule-based posture analysis component provides useful ergonomic information by detecting slouching; leaning to one side; leaning backward; misaligned heads; and crossed legs.

The resulting outputs produced by the system (eg. 3D skeletal visualisations and analysis reports) will offer potential real world utilities for the fields of ergonomics, workplace evaluations and real time health tracking applications.

On balance, this study presents a foundational framework for expansion with enhancements, including developing a health risk prediction system based on deep learning, temporal analysis via multi-frame recording, and connecting this framework to consumer-based wellness products.

Acknowledgment

The authors would like to express our gratitude to Dr. Ujwalla Gawande for her direction, support, and guidance throughout this project.

Her suggestions and insights contributed significantly to our project's success. Thanks to all the team members for their dedication, commitment, and cooperation in achieving the project objectives.

References

[1] sC. Menier, E. Boyer, and B. Raffin, "3D Skeleton-Based Body Pose Recovery," Proc. 3D Data Processing, Visualization and Transmission (3DPVT), 2006.

[2] H. Blum, "A Transformation for Extracting New Descriptors of Shape," Science, vol. 156, no. 3774, pp. 1535 1538, 1967.

[3] S. C. Choi and B. J. Min, "An Algorithm for the Medial Axis Transform of 3D Polyhedral Solids," Computers & Graphics, vol. 21, no. 3, pp. 387–397, 1997.

[4] Y. Liu, Y. Wu, and C. L. Tai, "A Generalized Medial Axis Approach for 3D Shape Representation," arXiv preprint arXiv:1603.03363, 2016.

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," Journal of the Royal Statistical Society: Series B, vol. 39, no. 1, pp. 1–38, 1977.

[6] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: A Skinned Multi-Person Linear Model," ACM Transactions on Graphics, vol. 34, no. 6, pp. 1–16, 2015.

[7] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. Osman, D. Tzionas, and M. J. Black, "Expressive Body Capture: 3D Hands, Face, and Body from a Single Image," Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[8] J. Wang and Y. Wu, "Deep 3D Human Pose Estimation: A Review," Information Fusion, vol. 67, pp. 23–40, 2021.

[9] Z. Shen et al., "Learning Human Mesh Recovery in 3D Scenes," Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[10] S. Salisu et al., “Review of Models for Estimating 3D Human Pose Using Volumetric Representations,” PeerJ Computer Science, vol. 7, e1231, 2025.
[11] S. Hignett and L. McAtamney, “Rapid Entire Body Assessment (REBA),” Applied Ergonomics, vol. 31, no. 2, pp. 201–205, 2000.
[12] A. Parakkat et al., “Automated Posture Recognition for Ergonomic Risk Analysis,” Human Factors, vol. 61, no. 3, pp. 380–389, 2019.