# Joint-Sphere: Intuitive and Detailed Human Joint Motion Representation

Seonghun Kim, Adithya Balasubramanyam, Dubeom Kim, Young Ho Chai, and Ashok Kumar Patil

Chung-Ang University, Seoul, South Korea

**Abstract**

*A motion comparison method using images allows for the motions to be easily recognized and to see differences between each action. However, when using images, orientation differences between similar motions cannot be quantified. Although many studies have been conducted on methods to represent the data and apply detailed motion comparisons, these representations are difficult to understand because the relationship between the motion and the representation is not clear. This paper introduces a novel motion representation method called the Joint-Sphere that enables detailed motion comparisons and an intuitive understanding of each joint movement. In each Joint-Sphere, the movement of a specific joint part is represented. Several Joint-Spheres can be used to represent a full-body motion. The results from a dance motion pattern show that each joint movement can be compared accurately even when several joints are moving quickly.*

**CCS Concepts**

• ***Human-centered computing*** → ***Information visualization;*** *Human computer interaction (HCI); Visual analytics; Visualization toolkits;* • ***Computing methodologies*** → ***Motion capture;***

## 1. Introduction

Recent advances in motion-sensor equipment have made it possible to obtain high-accuracy motion data that can be used to easily analyze human movement. Researchers have been emphasizing motion capture, recognition, and representation of human actions. Understanding a delicate human motion in detail from three-dimensional (3D) models, key frames lined up in a timeline sequence, representing them as a trajectory, and so on are challenging tasks. Peng et al. [PKM*18] trained 3D avatars through YouTube clip images to achieve almost the same behavior even when the terrain conditions and modeling body conditions were different. Huang et al. [HKA*18] showed that with just six sensors, the entire body's movements can be implemented very naturally. However, video-based methods are still used to compare how identical the implemented motions are.

In a motion comparison method using images, the differences in motion are recognized by the observer's eyes. It is easy to understand at a glance the meaning of a motion, but if the movements are similar or complex, it is difficult to identify and assess or measure the differences. On the contrary, in a comparison using data, the differences can be accurately identified, but it is difficult to grasp the meaning of the motion. Therefore, in this paper, we propose a novel motion representation method called the Joint-Sphere. Its use makes it possible to compare the exact movements of each body joint, which are difficult to understand while viewing an image, and it also makes it possible to understand the meaning of the motion intuitively without viewing an image. An overview of the proposed motion representation is as follows:

- Acquisition of motion data of each human body joint using inertial measurement unit (IMU) sensors and reconstruction of the motion by applying it to the 3D avatar.
- Visualization of the 3D avatar's joints motion on the Joint-Sphere and accumulated movements of specific body parts.
- Tracking of bone axis rotations that do not change position but only change direction.
- Defining rules between patterns and motions.

## 2. Related work

Human joint motion is very difficult to record and represent, unlike words, because it moves freely in 3D space and also includes variations in speed. Thus, various studies have been conducted on how to record and express movements. The *Labanotation* is one of the most representative motion recording systems, and it describes dance movements using a combination of symbols and lines [Wik19b]. In one study [RLG*18], a combination of *Labanotation* and Kinect sensors was proposed and the joint data of the entire body were measured and analyzed. Then, *Labanotation* symbols were used by dancers to practice dance movements. That study used the *Labanotation* symbol for training users to ensure that their joint movements exactly match those of the original dancer, in real-time, which helps them learn the movements correctly. However,





DOI: 10.2312/evs.20201065

due to the characteristics of *Labanotation*, an angle of 30° or less cannot be distinguished and expressed, and thus, it is difficult to express an operation accurately, and a user's effort is required to understand the meaning of a symbol.

Also, various researchers have attempted to express human movement as a combination of key poses rather than symbols. *Motion belts* [YKSN08] lists key poses in continuous motion on the time axis that are color-coded according to the direction in which the pelvis rotates. This method makes it easy to grasp the overall meaning of the movements, but it is difficult to distinguish the detailed differences among the movements and the direction differences solely based on the color. *MotionFlow* [JER15] classifies postures in continuous movements into similar postures and different postures and connects each posture according to a tree structure. This method is very useful for composing new movement patterns by separating or merging connected poses, but the exact details of each movement cannot be distinguished.

There have also been attempts to compare movement by visualizing the quaternion rotation value itself, instead of comparing the key poses of the motion. In the study of D. Peszor et al. [PMD*14], they visualized quaternion data using orthogonal/stereographic projections and Hopf fibration to distinguish between healthy and injured people's movements. Also, some researchers have visualized quaternion data over a sphere surface and transformed it into a 2D planar image of a lattice region to identify characteristics according to each motion [PKB*19]. However, in both of these studies, we can recognize the changes in the pattern depending on the motion, but the relationship between the patterns and motion is not clear.

## 3. Motion representation using Joint-Sphere

In this study, motion data were acquired by using 12 IMU sensors attached to human body joints and each limb joint with three kinematic degrees-of-freedom (3 DoF) orientation for human motion. These joints are hierarchically connected with a parent-child relationship on the 3D avatar as follows:

- Upper limb: $FixedPelvis \rightarrow UpperArm \rightarrow LowerArm \rightarrow Hand$.
- Lower limb: $FixedPelvis \rightarrow UpperLeg \rightarrow LowerLeg \rightarrow Foot$.

Before recording the motion data, the IMU sensors need to be calibrated and applied to the 3D avatar for reconstruction. Then, using the proposed Joint-Sphere, the movement of each joint can be intuitively visualized in the form of a trajectory and analyzed in detail by selecting any specified joint in the tool.

### 3.1. Joint-Sphere

We attached a Joint-Sphere to each joint of the avatar to visually identify the differences in the joint movements. Each Joint-Sphere represents the range of rotation that the joint can move through in 3D space. Therefore, the contact point will always occur between the surface of the Joint-Sphere and the bones connecting each joint, and these points will change with the movement of the joint. Connecting each contact point and expressing it as a trajectory allows us to identify the direction in which the joint has moved in 3D space. This method of representation has the advantage of being able to differentiate each joint movement, even if it is a complex movement.

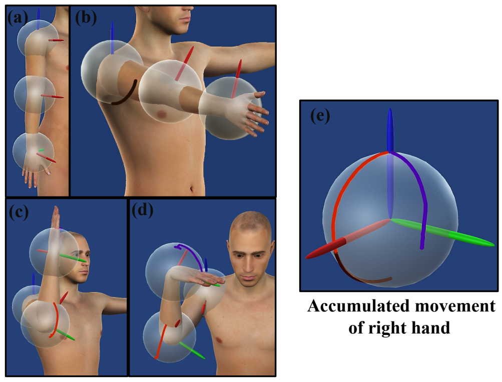


**Figure 1:** *Example of rotating the right arm (in this order): (a) initial pose, (b) rotate the upper arm upward, (c) rotate the lower arm upward, (d) rotate the hand to the left, (e) accumulated pattern of arm movement in a single Joint-Sphere.*

A Joint-Sphere represents the movement of a single joint. For instance, Figure 1 shows a movement of the right arm: (a) Initial pose of the right arm. (b) Only the upper arm is rotated, while the lower arm and hand are not rotated, so only the brown color trajectory is generated on the Joint-Sphere of the upper arm joint. (c) This is followed by only the lower arm rotating without any upper arm or hand movement, and the representation is shown as an red trajectory on the Joint-Sphere of the elbow. Finally, (d) the hand rotates via the wrist to the left, shown as a purple trajectory. Thus, we can observe no movement in the other two Joint-Spheres. The actual position of the hand in 3D space has changed because of the movements of (b) and (c). Therefore, multiple Joint-Spheres are required to understand the overall movement. In order to express all of the joint movements, the movement of the upper parts are connected according to the hierarchical structure and accumulated into a single Joint-Sphere representation (Figure 1(e)).

### 3.2. Bone-Axis rotation

Although the Joint-Sphere can distinguish most of the movements of each joint, there are some cases where the pattern does not change even when the movement occurs. These cases occur when the axis of rotational movement is parallel to the bone axis of the joint. Dobrowolski [Dob15] divided joint motion into two types, swinging and twisting, and rotational motion parallel to the bone axis corresponds to the twist. To track the twist movement of each joint we used a joint vector parallel to the bone axis and a front vector that points forward as shown in Figure 2(a). The rotational state calculation using two vectors is as follows:

- **P1**. Calculate the vector points corresponding to the swinging motion by multiplying the joint vector with the accumulated rotation value $q_{acc}$.
- **P2**. Calculate the rotation value $q_{pos}$ between two consecutive points in each frame.

- **P3**. Calculate $v_{pos.f}$ by multiplying the front vector with $q_{pos}$. For orientation $v_{ori.f}$ the front vector is multiplied with the $q_{acc}$.
- **P4**. Calculate the angle of twisting motion through the dot product between $v_{pos.f}$ and $v_{ori.f}$, and calculate the axis of rotation through the cross product.
- **P5**. Estimate the direction of rotation by comparing the calculated rotation axis with the joint vector.

In P1, the vector value does not change when the axis of rotation is parallel. Thus, each vector point on the Joint-Sphere represents only a change due to a swinging motion. Therefore, the $q_{pos}$ calculated from each vector point on the Joint-Sphere corresponds to the swinging rotation. The information about twisting can be estimated using $q_{acc}$ and $q_{pos}$ as indicated in the steps above (P2–P5).

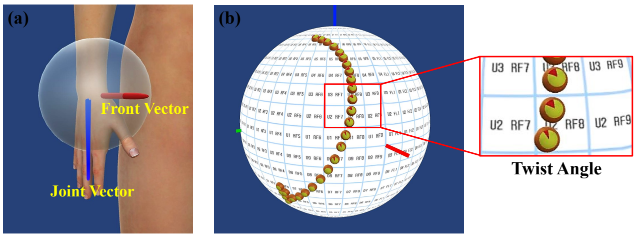


**Figure 2:** *(a) Joint vector and front vector of the right hand. (b) A motion trajectory pattern on a unit sphere.*

### 3.3. Pattern representation

The relationship between the trajectory pattern and the actual motion must be precisely defined to understand how the motion represented by the trajectory pattern is significant. We have calibrated the user's front direction to match the X-axis of the sensor global coordinate system. Afterward, to define the rule between the pattern and the actual motion, three planes used in anatomy were applied to relatively divide the surrounding space [Wik19a]. Based on the attention pose, all body parts are placed on the same vertical plane and can be divided into front and rear areas around the user's torso. However, in the horizontal direction, the left and right regions do not coincide because the rotation value is based on the joint position. For the right arm, the user's torso is in the left area, but for the left arm, it corresponds to the right. Therefore, relative direction rules are applied according to the joint part to be represented, not to all of the joints.

By applying this rule, we represented the meaning of the pattern on the sphere as a combination of letters and numbers. The position areas of each joint part divided into three planes and each area are 90° apart, which corresponds to the point where the $XYZ$-axis sign on the sphere changes. Therefore, the entire sphere is divided into eight areas, and each area can be marked with a symbol with the first letters of Up, Down, Left, Right, Front, and Back. Also, joint movements can be compared in detail depending on where the pattern is located by dividing each area by 10°. A bead yields the positional and orientation information of a single frame, as shown in Figure 2(b). For example, the code $U2\ RF8$ (highlighted bead) means 20° to the Upper side and 80° to the Right to the Front.

## 4. Evaluation

The proposed Joint-Sphere motion representation was evaluated using three variant biceps curl exercises (involving changes in the arm's swing and twist) and a sequence of dance movements. In Figure 3, the first row shows the right arm end poses of biceps curl (standard, rotational, and wide) exercises and its respective trajectory pattern for each movement in the lower row. Parts (d) and (e) show that the trajectory patterns regarding arm swinging movements appear to be similar, since the two vertical movement are similar in action. However, the standard curl has only a vertical swinging movement, whereas in the rotational curl they occur together with a vertical swing and twist of the bone. Thus, we can observe and compare the highlighted mid-pose of both exercises beads in the patterns ((d) and (e)), where the filled red area on the beads of pattern (e) indicate the approximate twisting angle of the lower arm clockwise by 45°, unlike the absence of filled beads in pattern (d).

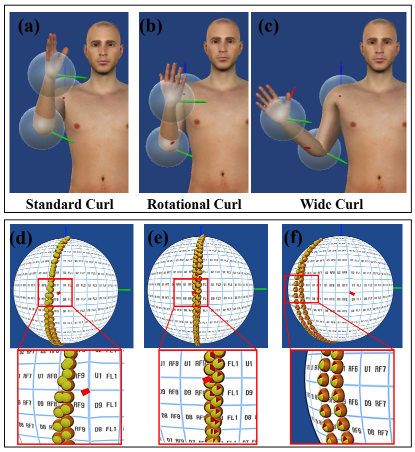


**Figure 3:** *Three types of bicep curl motions and their trajectory pattern representations: (a) standard curl, (b) rotational curl, (c) wide curl, (d) trajectory pattern of (a), (e) trajectory pattern of (b), (f) trajectory pattern of (c).*

The wide curl in Figure 3(c) is a variation of the standard curl and it has a similar vertical swing but it differs in the direction of the motion. Therefore, comparing the results of the two motions, the pattern (f) is located entirely between $RF5$ and $RF6$, and the pattern (d) is located between $RF8$ and $RF9$ so we can see that the wide curl was performed at an angle of ≈30–40° to the right compared to the standard curl.

Figure 4 shows the differences in the right upper limb's direction of swing and bone twist accuracy for three variations of the biceps curl (see Figure 3). In Figure 4(a) we can clearly observe that the twist angles are ≈0–5° and ≈40° for the standard and wide biceps curls, respectively, and it remains constant until the end, whereas the twist angle increases gradually from 0–90° in case of the rotational biceps curl. The filled bead representation (Figure 3) shows the twist angle can be visualized efficiently for each frame. Likewise, the swing results of the three variations of biceps curls shown in Figure 4 (b) clearly reveal that the direction of swinging in the wide biceps curl is different than the other two variations of biceps curls. The representation of human motion using labels provides more detail and understanding of any delicate motion.

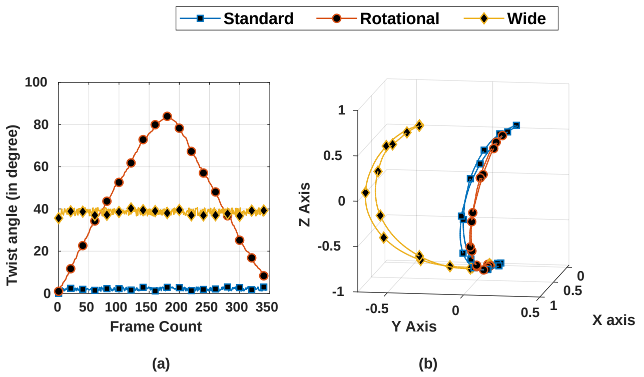


**Figure 4:** *Differences in the rotation and direction for three biceps curls: (a) twisting, (b) swinging.*

Motion patterns of dance sequence as shown in Figure 5 can individually identify joint movements according to the characteristics of each motion. The main feature of the corresponding dance moves is that the left hand moves twice from top to bottom and the left foot moves backward in the last movement. In the overall motion, the right hand does not move much after it is located on the chest area, and in the case of the right foot, there is no special movement other than moving to the right. Thus, the right-hand pattern in Figure 5 shows that the pattern started in $D1$ and rises to $D9$, resulting in a height increase of about 80°, and the pattern of the right foot shows that most patterns are formed in the lower area and then it continues to the right area. The left-hand movements from Poses 2–4 and Poses 4–6 look similar on the avatar, but the trajectory patterns on the Joint-Sphere look different. The trajectory pattern from 4–6 indicates lower movement compare to the trajectory pattern 2–4, distinguishing the variations of both of the movements. In the left foot pattern indicating the left foot has an upward movement and it is located in the area rotated from the left to the rear.

## 5. Discussion

Through our Joint-Sphere motion representation method, we have shown that we can accurately compare complex and similar movements and also how certain joint parts move during fast movements of the entire body. The main goal of our method is to compare the exact movements of the joint parts, using one sphere to represent specific joint movements. In general, several joint parts move together during each motion, and they can be divided into relatively more important parts and not as important parts.

Therefore, our method has the advantage of extending the comparison from the movement of a specific part to the movement of the whole body using several spheres according to the user's choice. In addition, the grid area of the sphere where the pattern is drawn is defined so that the meaning of the pattern can be clearly understood by a combination of letters and numbers, so that anyone can easily understand it without needing specialized knowledge.

However, as the motion continues over time, the overlapping parts of the trajectory patterns increase and it becomes more difficult to understand the motion. Besides, the time sequence between each pattern is unknown, making it difficult to reconstruct continuous motion through pattern analysis, which is a problem that needs to be overcome. If these shortcomings can be addressed, it is expected that the entire pattern area can be transformed into a 2D flat image, making it easier to reconstruct the motion through image recognition.

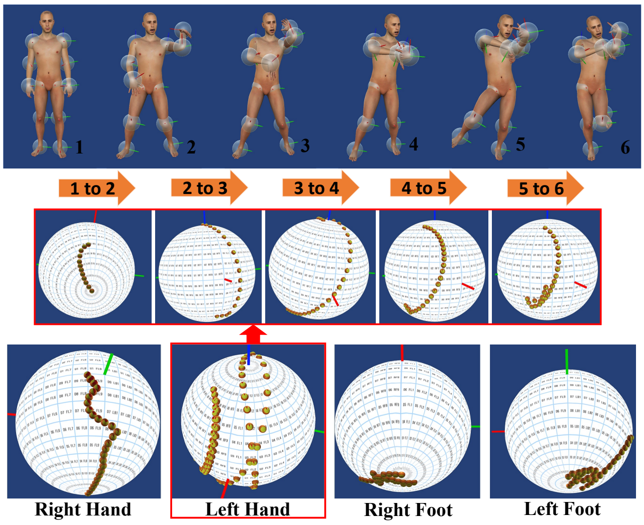


**Figure 5:** *(Top) Six key poses of a dance movement in sequence, (Middle) left hand trajectory pattern, (Below) Result of both upper and lower limbs complete pattern.*

## 6. Conclusion

In this paper, we proposed a novel motion representation method that allows for a detailed comparison of each joint using the Joint-Sphere. It can be used to compare the exact movements of joints that are difficult to recognize in the image.

Human motion data were measured using IMU sensors and the motion was visualized in the form of a trajectory pattern on a sphere corresponding to the rotational range of the joints. Each circular pattern that constitutes a trajectory means the relative area in which the corresponding joint part is located in 3D space and the direction of the joint in that position can be estimated through the area in which the pattern is filled.

In future work, we will improve the representation method so that the temporal order of each pattern can be distinguished, and develop an algorithm to reconstruct the movement on 3D avatars that will receive pattern images as input.

## 7. Acknowledgements

This work was supported by the Ministry of Science & ICT of Korea, under the Software Star Lab. program (IITP-2018-0-00599) supervised by the Institute for Information & Communications Technology Planning & Evaluation.

## References

[Dob15] DOBROWOLSKI P.: Swing-twist decomposition in clifford algebra. *arXiv preprint arXiv:1506.05481* (2015). 2

[HKA*18] HUANG Y., KAUFMANN M., AKSAN E., BLACK M. J., HILLIGES O., PONS-MOLL G.: Deep inertial poser: learning to reconstruct human pose from sparse inertial measurements in real time. *ACM Transactions on Graphics (TOG) 37*, 6 (2018), 1–15. 1

[JER15] JANG S., ELMQVIST N., RAMANI K.: Motionflow: Visual abstraction and aggregation of sequential patterns in human motion tracking data. *IEEE transactions on visualization and computer graphics 22*, 1 (2015), 21–30. 2

[PKB*19] PATIL A. K., KIM S. H., BALASUBRAMANYAM A., RYU J. Y., CHAI Y. H.: Pilot experiment of a 2d trajectory representation of quaternion-based 3d gesture tracking. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (2019), pp. 1–7. 2

[PKM*18] PENG X. B., KANAZAWA A., MALIK J., ABBEEL P., LEVINE S.: Sfv: Reinforcement learning of physical skills from videos. *ACM Transactions on Graphics (TOG) 37*, 6 (2018), 1–14. 1

[PMD*14] PĘSZOR D., MAŁACHOWSKI D., DRABIK A., NOWACKI J. P., POLAŃSKI A., WOJCIECHOWSKI K.: New tools for visualization of human motion trajectory in quaternion representation. In *Asian Conference on Intelligent Information and Database Systems* (2014), Springer, pp. 575–584. 2

[RLG*18] RALLIS I., LANGIS A., GEORGOULAS I., VOULODIMOS A., DOULAMIS N., DOULAMIS A.: An embodied learning game using kinect and labanotation for analysis and visualization of dance kinesiology. In *2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)* (2018), IEEE, pp. 1–8. 1

[Wik19a] WIKIPEDIA CONTRIBUTORS: Anatomical plane — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Anatomical_plane&oldid=930986097, 2019. [Online; accessed 18-February-2020]. 3

[Wik19b] WIKIPEDIA CONTRIBUTORS: Labanotation — Wikipedia, the free encyclopedia, 2019. [Online; accessed 18-February-2020]. URL: https://en.wikipedia.org/w/index.php?title=Labanotation&oldid=920547544. 1

[YKSN08] YASUDA H., KAIHARA R., SAITO S., NAKAJIMA M.: Motion belts: Visualization of human motion data on a timeline. *IEICE TRANSACTIONS on Information and Systems 91*, 4 (2008), 1159–1167. 2