Human Joint Orientation Descriptor Based on Geometric Algebra and Its Application

Wen-ming CAO and Yi-tao LU*

Shenzhen Key Laboratory of Media Security, Shenzhen University, Shenzhen, China

*Corresponding author

Keywords: Motion recognition, Gaussian Naïve Bayes, Geometric algebra, Motion calibration.

Abstract. Motion recognition is becoming more and more widely used in various applications. In this paper we propose a novel descriptor to describe human skeleton based on geometric algebra (GA) that decomposes the skeleton posture into the rotations of skeleton parts. In this model, all body bones are rotated from the same original states. We formulate the rotation operator in 3D GA space, which can be used to describe the rotations of human body bones. Then we select the most informative rotations of body bones and joint angles to represent the skeleton. We train a Gaussian Naïve Bayes classifier which can recognize the motion type of a single input frame captured from video sensors. After the motion type is determined, we find the most similar posture in the motion sequence database using the distance based on posture orientations and joint angles. And finally, we calculate the posture difference to give users the calibration advice. Our experimental results have shown the high accuracy and effectiveness of our method.

Introduction

Human motion recognition and estimation have become more and more universal in computer vision domain due to its wide applications in recent years, such as motion sensing game [1], human-computer interaction [2] etc. In order to recognize the human motions, the critical element is to fetch consistent and discriminative features from the acquired data about human body postures using video sensors or other devices. Several different types of raw data are used for motion recognition, such as RGB data [3], gray color data, depth data [4] and RGB-D data [5] etc.. 3-D skeleton data gets more and more popular due to the easy availability of the low-cost and high-precision RGB-D devices. The depth information can help extract better features such as human body joint positions.

In this work, we propose a new descriptor to describe human skeleton based on geometric algebra that decomposes human body posture into the rotations of body bones. We choose some key bones of the body to represent human body posture. We implement a system which can calibrate human motions in real time based on motion recognition. Figure 1 shows the framework of our system. First, we construct a standard human motion database which contains human body motion sequences represented by joints orientations. We sample some of the human motion that we want to calibrate as calibrating human motions, we need to recognize which motion the human is doing. In order to solve this problem we train a Gaussian Naïve Bayes classifier using the training dataset. After the model is trained, the classifier uses the input candidate posture and finds the most similar motion sequence from the database, which is the type of the motion. Then we calculate the Euclidean distance between the current posture frame and the frames from the sequence recognized by the classifier. The frame from the reference sequence with the minimum distance is the reference frame used to calibrate human motion. With this, we can give the calibration advice based on the angle difference of key human body bones in the two posture frames.

In summary, the major contributions of our work are three aspects: propose a novel body posture descriptor based on geometric algebra to describe human body posture and implement a real-time calibrate system to use the angle difference of body skeleton and give users the calibration advice.
Related Work

With the release of Kinect sensors, the skeleton body data it provided has appealed a lot of researchers. The Kinect sensor can give us lots of information about the people that in the view of the Kinect sensor. There are two methods to represent a human body data. One is using the joint positions which are the absolute positions in terms of the Kinect camera coordinates, the other one is joint orientations that indicate the orientation of body bones and how those joints are rotated for motions. Motion recognition based on 3D skeleton especially has been broadly studied with the release of Microsoft Kinect. With the rapid development of machine learning, deep learning algorithms and training techniques in recent years, the recognition accuracy has been greatly improved. Most of these methods used the whole motion data as the input for the classifier. Methods based on the whole motion have the difficulty to recognize the motion in real time manner.

The Geometric Algebra Joint Orientation Descriptor

The Basics of Geometric Algebra. Geometric algebra (GA) namely Clifford Algebra was proposed by William. K. Clifford. In GA theory, the most important operation is geometric product which combines the inner product and outer product. The definition is

\[ uv = u \cdot v + u \wedge v \]  

where \( u \cdot v \) is the vector inner product or dot product and the \( u \wedge v \) is the outer product. The outer product can be used to construct a bi-vector from two vectors. The following equations list the basic properties of outer product, namely, linearity, associativity and negative symmetry.

\[ u \wedge (v + w) = u \wedge v + u \wedge w \]  

\[ u \wedge v \wedge w = (u \wedge v) \wedge w = u \wedge (v \wedge w) \]  

\[ u \wedge v = -v \wedge u \]  

In N-dimensional GA space \( G_n \), there have orthogonal basis \( \{ e_1, e_2, \cdots, e_n \} \), which has the following properties:

\[ e_1^2 = e_2^2 = \cdots = e_n^2 = 1 \]  

\[ e_i \cdot e_j = 0, \ i \neq j \text{ and } i, j \in [1, n] \]  

\[ e_i e_j = e_i \cdot e_j + e_i \wedge e_j = e_i \wedge e_j , i \neq j \text{ and } i, j \in [1, n] \]
Let $M \in G_n$ be a multi-vector in N-dimensional GA space, $M$ can be represented as

$$M = a_0 + \sum_{i=1}^n a_i e_i + \sum_{k=1}^n a_k e_{ij} + \cdots + a_{12\ldots n}$$

where $a_0, a_i, a_k, \ldots, a_l \in \mathbb{R}$, $C^i_n$ is the combination of size $n$ and $I = e_1 \wedge e_2 \wedge \ldots \wedge e_n = e_{12\ldots n}$ is the unit pseudo-scalar of $G_n$.

**The Joint Orientation descriptor in Geometric Algebra.** The GA has a wide range of applications such as kinematics. The geometrically intuitive operations of Geometric Algebra can describe multi-vector’s rotation in GA space in a consistent and easy way. In GA, there is a rotor $R$ that can describe blade rotation with respect to another blade $w$. Let $q, q' \in G_n$, $q'$ is a rotated blade from $q$ by an angle $2\theta$, there have

$$q' = RqR^\dagger = e^{-\theta w} q e^{\theta w}$$

where $R$ and $R^\dagger$ satisfy the condition $RR^\dagger = R^\dagger R = 1$.

According to above theories about GA, we can utilize GA rotor to represent rotation of the human body bone in 3D space. In 3D space, the orthogonal basis is $(e_1, e_2, e_3)$, the rotor can be written as

$$R = r_0 + r_1 (e_2 e_3) + r_2 (e_3 e_1) + r_3 (e_1 e_2)$$

where $r_0, r_1, r_2, r_3 \in \mathbb{R}$. The basis of rotations in the 3D GA space consists of four components $(1, e_2 e_3, e_3 e_1, e_1 e_2)$. So we have the rotated vector $v' = Rv\hat{R}$.

According above, we propose to use the angles between the key bones of human skeleton and the rotation of human bones to describe the postures. In this paper, we decompose human skeleton posture into the rotations of human body bones, assuming that all human body bones have same original state which is straight up, and all bones have the same length. As for a certain body posture, all bones are rotated from the original state. We choose some key bones to represent the skeleton and use the rotation and joint angles of those bones to describe the body posture.

**Real-time Calibrate System**

**Standard Human Motion Database.** We sample some standard human motion that we want to calibrate. These motions will serve as the baseline or reference for the motion calibration. We use Kinect v2 to build a database which contains six types of motions as the standard human motions. The types are: TW extend, side-leg lift, arm roll, swing arm, rocker arm and walk in-place.

**Calibrate Human Body Motion.** In this stage, we have two tasks. The first is to find the reference frame/posture in the database which is the most similar to the current posture. We select the posture with the smallest distance. To measure the similarity/distance between two body postures, we select body bone rotation data and convert to Euler angle $I_{e_{x,y,z}}$. We use 16 joints, each with 3 components, and 2 additional joint angle $I_a$. So the feature length of each posture is $3 \times 16 + 2 = 50$, namely $v_f \in \mathbb{R}^{50}$. The distance between two posture $i, j$ is calculated as

$$d_{ij} = \sqrt{(I_{e_{1,i}} - I_{e_{1,j}})^2 + (I_{e_{2,i}} - I_{e_{2,j}})^2 + \sum_{n=1}^{16} \sum_{m \in e_{x,y,z}} (I_{e_{m,i}} - I_{e_{m,j}})^2}$$

The second one is to calculate the posture difference so that we can give users the advice for calibration. As for the first task, we use the body joint rotation angles and joint angles as the feature to calculate the Euclidean distance between the current body posture and the postures from the reference motion given by the classifier. We select the posture with the smallest distance. After we get the reference body posture, we calculate the posture difference using joints angles. The Figure 2 illustrates how to calculate the left arm joint angle. In Figure 2, the left arm, the vector $\overline{SD} = \overline{SM} \times \overline{SC}$ is the normal of plane SCM. Vector $\overline{SA}, \overline{SB}$ are parallel to vector $\overline{SC}, \overline{CM}$ respectively. The four joint angles we calculated for left arm are $\theta_{SEF}, \theta_{ASE}, \theta_{ESD}, \theta_{ESB}$. For example, angle $\theta_{SEF}$ is calculated as:
\[ \theta_{SEF} = \arccos \frac{\mathbf{E} \cdot \mathbf{F}}{\| \mathbf{E} \| \| \mathbf{F} \|} \]  

The rest of joint angles are calculated similarly. We can see that for each arm and leg we can calculate 4 angles, and 3 angles for head, giving 19 joint angles in total for a specific body posture.

The following algorithm summarizes the system workflow.

**Algorithm 1: the system workflow**

1. **Input**
   - Body frame: \( J(n) = [J_p(x, y, z), J_o(x, y, z)]^T \), \( n \in [1, 25] \), \( J(n) \in \mathbb{R}^{(3+4) \times 25} \)
   - Exponential moving averages: smooth \( J_p \) and \( J_o \) respectively.

2. **Extract joint angle and joint Euler angle**:
   - For the \( J_o(n) \), convert it into Euler angle: \( J_o(x, y, z) \), \( n \in \mathbb{R}^{4 \times 25} \)
   - For the \( J_p(n) \), calculate body joint angle: \( J_p(x, y, z) \), \( n \in \mathbb{R}^{3 \times 25} \)

3. **Feed the feature data** \( J_f(n) \) into Gaussian Naïve Bayes Classifier and get the reference motion:
   - \( J_f(n) = [J_o(n)^T, J_p(n)^T]^T \), \( J_f(n) \in \mathbb{R}^{6+2} \)
   - reference motion type \( M_x \in [1, N] \)

4. **Experiments**

There are many public datasets, these dataset provide abundant data about human actions. However, these dataset only contains body joint positions in terms of the sensor coordinates. But we also need the joint orientation data as our classifier and system input. So, to demonstrate the effectiveness of our method, we prepare our own database to carry out the experiments. We use Kinect v2 to sample the test dataset which contains the corresponding motions in the standard library with the frame rate 30fps. For each motion we sample 2~3 times, and each motion has 100~200 frames. After we train our Gaussian Naïve Bayes Classifier using the training dataset, we test the classifier on the test dataset. The average recognition accuracy is 84.92%. The confusion matrix of the single posture recognition result is shown in Figure 2. Lastly, we present the real time calibration results from the system. Figure 3. shows three examples of different types of motions, each row corresponds to a motion. The real time calibration system can perform the human body postures calibration with the frame rate 30fps.

![Figure 2](image-url)  
**Figure 2.** The illustration of calculating joint angle of the left arm(left) and the confusion matrix of the recognition accuracy on test dataset(right).
Summary

In this paper we propose a novel skeleton posture descriptor using orientations of human bones based on geometric algebra. Firstly we formulate the rotor in 3D GA space, then select the most informative rotations of body bones and joint angles to represent the human body posture. With this, we train a Gaussian Naïve Bayes classifier to recognize the motion types which uses only single input postures. With the motion type determined, we search for the posture in the reference motion database using the distance based on posture orientations and joint angles. And finally, we calculate the posture difference using the joint angles. We implement the method in a system which can recognize human motions and give users the calibration advice in real time manner. Our experimental results have shown the high accuracy and effectiveness of our method and the system. As one of the future work, we are going to prepare more dataset to improve the recognition performance.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61771322 and Grant 61375015 and Shenzhen Foundation fund No. JCYJ20160307154630057.

References


