Using automatic generation of Labanotation to protect folk dance

Jiaji Wang
Zhenjiang Miao
Hao Guo
Ziming Zhou
Hao Wu

Using automatic generation of Labanotation to protect folk dance

Jiaji Wang,* Zhenjiang Miao,* Hao Guo,* Ziming Zhou,* and Hao Wu* 
*Beijing Jiaotong University, School of Computer and Information Technology, Institute of Information Science, Haidian District, Beijing, China
*bUniversity of South Carolina, Computer Vision Lab, Department of Computer Science and Engineering, 3A19 Swearingen Engineering Center, Columbia, South Carolina 29208, United States
*cBeijing Normal University, College of Information Science and Technology, Haidian District, Beijing, China

Abstract. Labanotation uses symbols to describe human motion and is an effective means of protecting folk dance. We use motion capture data to automatically generate Labanotation. First, we convert the motion capture data of the biovision hierarchy file into three-dimensional coordinate data. Second, we divide human motion into element movements. Finally, we analyze each movement and find the corresponding notation. Our work has been supervised by an expert in Labanotation to ensure the correctness of the results. At present, the work deals with a subset of symbols in Labanotation that correspond to several basic movements. Labanotation contains many symbols and several new symbols may be introduced for improvement in the future. We will refine our work to handle more symbols. The automatic generation of Labanotation can greatly improve the work efficiency of documenting movements. Thus, our work will significantly contribute to the protection of folk dance and other action arts.

1 Introduction

Folk dance is a way for working people to feel and express themselves in their production and living. The world has a rich tradition of folk dances, which are a valuable asset and an important part of intangible cultural heritage. However, influenced by inheritance, geographical environment, and regional development imbalance, many folk dances and other action arts are disappearing. Therefore, it is an urgent task to protect these precious art forms.

Dance notation is an effective method for recording human movements and is a useful tool for the protection of folk dances. Similar to music scores, dancers can understand and perform the content from dance notations. Moreover, the symbolic representation makes dance notation intuitive and easy to understand. However, dance notation is difficult to document because this type of work requires manual drawing by experts, and work efficiency is extremely low.

Generating the dance notation automatically using a computer is an effective method for solving this problem. In our work, we choose Labanotation for the experiments, and the action arts are Chinese folk dance and opera. In 2012, the Chinese ministry of culture and ministry of science and technology worked together and established a project called the National Science and Technology Support Program. We took part in a study called “multiple display modes of dynamic digital cultural resources based on human motion capture,” which is intended to use Labanotation as a means to showcase Chinese folk dance and opera, thereby ensuring the inheritance and protection of these types of action arts.

Our work includes five aspects, as follows: (1) acquiring the motion capture data. We obtain the data via two types of motion capture systems: marked and nonmarked. (2) Analyzing the motion capture data. We analyze the structure of motion capture data stored in biovision hierarchy (BVH) files and calculate the spatial information in the data. (3) Motion segmentation. The segmentation is used to divide human motion into element movements. We use three methods to address the motion of gravity center, upper limbs, and lower limbs. (4) Analyzing element movements. Based on whether the movements are characterized as supporting, we use two methods to determine the corresponding symbols for the movements. (5) Software development. By using the motion capture data acquired through (1), via the processes of (2), (3), and (4), we develop the automatic generation software of Labanotation.

The paper is organized as follows. Section 2 describes related work on dance notation and computer technology. Section 3 describes Labanotation and our human motion capture. Section 4 describes in detail the automatic generation of Labanotation. Section 5 presents experiments and evaluations. Section 6 contains the conclusions and future work.

2 Related Work

Dance notation is a set of symbols used for the recording of dance, which is similar to the music score for music. Using dance notation to describe human motion is rigorous, convenient, and easy to save. There are many kinds of dance notations. In China, the oldest dance notation is Chinese Naxi Dongba dance notation. Dunhuang dance notation is another old one, which was used in the Tang Dynasty of China. In England, there exists Maurice notation, Benish...
dance notation, \(^\dagger\) Seiktein shorthand method, etc. Alphabet dance notation is used in North Korea. Labanotation, created by Rudolf Laban, \(^\dagger\) is recognized as one of the most widely used and most accurate notations for the recording of dance. Therefore, in this paper, we choose Labanotation to do our research for the automatic generation.

Labanotation is intuitive, vivid, and logical. It has been widely used in ballet and other western dances. However, drawing the Labanotation is difficult, which is an obstacle for using this notation to describe human movements. To make the drawing easier, a method to simplify the work through computer technology is necessary.

At present, there are three main research directions regarding combining Labanotation and computer technology.

The first direction is to develop software as a tool for the manual drawing of Labanotation. The software provides a variety of symbols, and people only need to drag the symbols to the designated places. This simplifies the manual drawing. The research on developing software is mature. \(^\dagger\) This kind of software includes Calaban, Labanatory, \(^\dagger\) LED, \(^\dagger\) Laban Writer, etc. Among them, the Laban Writer developed by Ohio State University is one of the most widely used software platforms. Laban Writer is based on the Macintosh platform and provides a graphical tool for the user.

The second direction is to drive models by Labanotation. Japanese researchers have developed a software platform called Laban Editor, \(^\dagger\) which uses Labanotation to drive a human model. The software named Life Forms designed by Maranan et al. \(^\dagger\) and the software called Laban Dancer developed by Wilke et al. \(^\dagger\) can convert Labanotation to a section of human body animation.

The third direction is the automatic generation of Labanotation. As an interdiscipline of dance and computer technology, the study is in the beginning stage. At present, the acquisition of Labanotation is mainly by manual drawing, but the speed of manual drawing is too slow to write down the huge amount of folk dance in the world. Therefore, it is necessary to research automatic generation. The first study on using motion capture data to generate Labanotation is by Hachimura and Nakamura at Ritsumeikan University in 2001. They proposed a method to generate Labanotation of upper limb movements based on spatial analysis. \(^\dagger\) However, their research only focused on the upper limbs and did not discuss the movements of lower limbs or the gravity center. In Thailand, Worawat Choensawat cooperated with Hachimura and Nakamura, and they developed GenLaban software for the automatic generation of Labanotation. \(^\dagger\) However, the GenLaban software does not handle well the complex lower limb movements and gravity center movements. Chen et al. \(^\dagger\) analyzed motion capture data by using the rule-based approach. However, in the process of generating Labanotation, they did not consider the pause and segmentation of movements. Thus, the quality of the generated Labanotation is not high.

3 Labanotation and Human Motion Capture

3.1 Labanotation

Labanotation is a recording system designed for analyzing human movements. The notation uses intuitive symbols to represent the movements of the human body.

The dance notation comes in two parts: structure and notation symbols. The basic structure is composed by three vertical lines based on the symmetrical structure of the human body. \(^\dagger\) The symbols are placed between the vertical lines from the bottom to the top.

The structure of Labanotation is shown in Fig. 1. Among the three bold lines, the middle one represents the human spine while the left and right lines represent the left and right sides of the human body, respectively. Around the three vertical lines, Labanotation generally contains 11 or 9 columns. The quantity of columns is determined by the requirement. From the middle to both sides, the 11 columns are as follows: supporting movement column (left and right), leg movement column (left and right), torso movement column (left and right), arm movement column (left and right), hand movement column (left and right), and the rightmost one is the head movement column (or at the most left). Notation symbols are written in a column corresponding to the body parts. The horizontal lines in the notation are

---

**Fig. 1** Distribution of Labanotation columns.
section lines that represent the rhythm. The first section line is a double line, and the others are single lines.

Labanotation has a lot of symbols to describe all kinds of movements in detail. Nevertheless, the basic symbols are not complex. There are 27 basic symbols corresponding to 27 quantized spaces, including nine kinds of horizontal directions and three kinds of vertical levels, as shown in Fig. 2. For the basic symbols, the shape represents horizontal directions, and the filling of symbol represents vertical levels. For each symbol, the length represents time span of the movement.

The directions of nine kinds of horizontal symbols are as follows: place (original position), left, right, forward, backward, left forward, right forward, left backward, and right backward. The shape of the place symbol is a rectangle. The place symbol represents that the body part is in a natural state. The other eight horizontal symbols come into being by cutting off a part of the rectangle.

The three kinds of vertical symbols include low, middle, and high levels. Filling the symbol with black indicates low level. Filling with a solid dot indicates middle level. And slashes indicate high level.

Figure 3 is an example for human movements and the corresponding Labanotation symbols. The first line shows eight kinds of movements of the right arm, and the second line shows the same movements of the right leg. The movements are as follows: ① place, low; ② right, low; ③ right, middle; ④ right, high; ⑤ place, high; ⑥ forward, high; ⑦ forward, middle; ⑧ forward, low.

3.2 Human Motion Capture
Motion capture can accurately note the movements of each body part in three-dimensional (3-D) space. The method of motion capture was first proposed by Johansson. In the beginning, the method was mainly for producing animation. Later, with the maturity of motion capture technology, it has been widely used in many fields, such as film and animation, human–computer interaction, virtual reality, game production, sport analysis, and so on.

3.2.1 Two kinds of motion capture system
Depending on whether there is a need to install markers on the human body, the motion capture system can be divided into a marked one and an unmarked one. In our experiments, we use two kinds of systems. The marked motion capture system we used is called OptiTrack, and the unmarked one is built by ourselves and utilizes the method in Refs. 18–20. We use the marked one as the main capture method and the unmarked one as a supplement.
OptiTrack motion capture system is a mature industrial product, which is sufficiently accurate for our experiments. The devices of OptiTrack are expensive, including a set of infrared cameras and the corresponding equipment.

Unmarked motion capture system is not a mature product. The devices of the unmarked system are relatively simple. The system only needs a computer and several consumer level cameras. Thus, the unmarked system is really portable. The disadvantage is that the accuracy needs to be improved.

3.2.2 Motion capture data
The commonly used motion capture data formats include BVH, acclaim skeleton file/acclaim motion capture data, hierarchical translation rotation, and so on. In our experiment, we choose the BVH format because it is widely accepted.

BVH is a text file with ASCII encoding format. In the file, there are two parts. The first part starts with the keyword hierarchy, which defines the joint structure of a human skeleton. The second part starts with the keyword motion, which stores the movement information of all the joints.

4 Automatic Generation of Labanotation
Using motion capture data to generate Labanotation can be summarized in three parts: analyzing and processing the motion capture data, motion segmentation, and analyzing the element movements. The flow chart is shown in Fig. 4.

First of all, we analyze the structure of body joints in the BVH file. Through the analysis of each joint, we confirm the relationships between body joints and body parts. To judge the movement of each body part, we convert motion capture data into coordinate data.

Second, we cut the human motion into basic units—element movements. We use three methods for motion segmentation. Spatial clustering based on the Laban direction is used for cutting the motion of the human body's gravity center. We cut the motion of the human upper limbs based on velocity threshold, and probabilistic principal component analysis (PPCA) segmentation is used for the motion of the human lower limbs.

Third, we analyze each element movement by the rules of Labanotation. We find the symbol of Labanotation to describe each element movement and write the symbols at the right places corresponding to each body part. Thus, the generation of Labanotation is complete.

4.1 Motion Capture Data Analysis and Format Conversion
4.1.1 Motion capture data analysis
In a BVH file, each node in the human skeleton describes a body joint. However, different BVH files may use different skeletons.

There are two kinds of differences. One kind is the difference in the number of nodes. For example, a BVH file uses 26 nodes to represent a human body, while another file may use 23 nodes. The other kind is the difference in the meaning of nodes. For example, there are two skeletons with 26 nodes, but only 24 nodes have the same meaning. One skeleton use the other two nodes to describe the details of the hands, while the other skeleton use the two nodes for the details of the feet.

![Flow chart of generating Labanotation based on motion capture data.](image-url)
In this paper, we use the skeleton with 26 nodes in BVH files; the node structure is shown in Fig. 5 (the same structure used in our previous work). To set a unique identifier for each node, we define a quadruple \(h; P; D; C_1 \) to represent a node. The meaning of the four elements is as follows.

- \(P\) (part) represents the location of the node in the skeleton. As shown in Fig. 5, the nodes can be divided into three parts: left, right, and center, so the enumeration values of \(P\) are \{left, right, center\}.
- \(D\) (depth) represents the layer of the node. Node “root” is in the first layer, in other words, \(D(\text{root}) = 1\). Assuming the layer sequence of a nonleaf-node is \(d\), then the layer sequence of its child nodes are \(d + 1\).
- \(C_1\) (children) represents the number of child nodes. For leaf nodes, \(C_1 = 0\). \(C_2\) (count) represents the number of a kind of node. Assuming node \(N\) is on a node chain, on the chain, from root node to node \(N\), \(C_2\) is the number of nodes containing three child nodes. Assuming that \(J_0, J_1, \ldots, J_n\) are nodes on the chain, \(J_0\) is the root node and \(J_n\) is the leaf node; thus, for a node \(J_i\), \(0 \leq i \leq n\), the value of \(C_2(J_i)\) is

\[
C_2(J_i) = \sum_{k=1}^{i} u(J_k).
\]

where

\[
u(J) = \begin{cases} 1, & C_1(J) = 3 \land C_1(J) \neq 3 \\ 0, & \text{otherwise} \end{cases}
\]

In our experiments, unique identifiers of the 26 nodes are shown in Table 1.

### 4.1.2 Format conversion

In the BVH file, motion data are recorded by Euler angles. Euler angles describe the orientation of each joint during the movements, but the relative positions between joints are not intuitive. Consequently, we convert the Euler angles to 3-D coordinates in the Cartesian coordinate system. 3-D coordinates are suitable for the judgment of relative positions.

The conversion process is as follows. Considering any nonroot node \(J\) and its parent node \(J_p\) in the BVH file, Euler angles describe the angle rotation around the \(Z\), \(X\), and \(Y\) axes for the nodes. After a movement, there exists an angular displacement of joint \(J_p\). Assuming \((x_c, y_c, z_c)\) coordinates in the Cartesian coordinate system.
is the relative position between J and Jr after a movement and \((x_0, y_0, z_0)\) is the previous relative position, \((x_r, y_r, z_r)\) can be calculated by the angular displacement matrix \(M\) and \((x_0, y_0, z_0)\):

\[
\begin{bmatrix}
x_r \\
y_r \\
z_r
\end{bmatrix} = M \begin{bmatrix}
x_0 \\
y_0 \\
z_0
\end{bmatrix}.
\] (3)

Assuming that the angular displacement matrices of all the precursor nodes of \(J\) are \(M_1, M_2, \ldots, M_r\), where \(M_1\) is the matrix of the direct precursor node (parent node of \(J\)) and \(M_r\) is the matrix of the root node, after a movement, the relative position between node \(J\) and root node \(J_r\) is

\[
\begin{bmatrix}
x_r \\
y_r \\
z_r
\end{bmatrix} = M_r \cdot \left( M_{r-1} \cdot \left( M_{r-2} \cdot \left( \cdots M_1 \cdot \begin{bmatrix}
x_0 \\
y_0 \\
z_0
\end{bmatrix} \right) \right) \right).
\] (4)

Thus, the relative position between any node and its precursor node can be calculated. For a node \(J_0\) on a node chain, all the precursor nodes are \(J_1, J_2, \ldots, J_r\), where \(J_1\) is the direct precursor node and \(J_r\) is the root node. Assuming the relative position of these nodes to their direct precursor nodes are \(O_0, O_1, \ldots, O_{r-1}\), notice that the root node has no predecessor node. Using \(P_{\text{root}}\) to represent the position of the root node \(J_r\), the position of \(J_0\) in the world coordinate system can be calculated by

\[
P = P_{\text{root}} + O_r + \ldots + O_2 + O_1 + O_0.
\] (5)

Thus, to calculate the coordinates of a node, all the relative positions between every adjacent node on the chain are accumulated and the coordinates of the root node are added.

### 4.2 Motion Segmentation

In Labanotation, a symbol represents an element movement of a body part. To generate the dance notation automatically, motion capture data should be cut into element movements. The human motion is composed of a series of element movements.

In our experiments, we use three methods to cut human motion. The movements of the human body’s gravity center are segmented by the spatial clustering of Laban direction. Upper limb movements are segmented by the method of velocity threshold. Lower limb movements are segmented by the PPCA.

#### 4.2.1 Spatial clustering of Laban direction

The movements of human body’s gravity center represent the tendency of human motion. In Labanotation, symbols of this kind of movement are written in the supporting column.

In Fig. 5, the root node is in the middle of skeleton, so the movements of the root node can be roughly seen as the movements of the gravity center. From motion capture data, we can get the trajectory of the root node, and the movement directions of the node can be calculated through changes of its coordinates. Assuming in \(t_1\) and \(t_2\) moments, the coordinates of root node are \((x_1, y_1, z_1)\) and \((x_2, y_2, z_2)\), respectively, the motion vector of the root node between two moments is

\[
\mathbf{m} = (\Delta x, \Delta y, \Delta z) = (x_2, y_2, z_2) - (x_1, y_1, z_1).
\] (6)

Through motion vector \(\mathbf{m}\), we can confirm the movement direction, including a vertical component and a horizontal component. We first analyze the vertical component by calculating the angle between vector \(\mathbf{m}\) and the \(y\)-axis. When the angle is less than a threshold (16 deg, determined by experience), it indicates that the gravity center only does vertical motion. If the angle between vector \(\mathbf{m}\) and the positive direction of \(y\)-axis is between \((-16\ deg, 16\ deg\)), the movement direction of gravity center is upward. If the angle is between \((164\ deg, 180\ deg) \cup (-180\ deg, -164\ deg))\), the movement direction is downward. When the angle between vector \(\mathbf{m}\) and the \(y\)-axis is bigger than the threshold, it indicates that the movement direction of the gravity center is horizontal. We then analyze the angle between the horizontal component of vector \(\mathbf{m}\) and the positive direction of \(z\)-axis. The relationships between the angle and the horizontal direction are shown in Table 2.

<table>
<thead>
<tr>
<th>Range of angle (\alpha)</th>
<th>Horizontal direction of Labanotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>((-22.5\ deg, 22.5\ deg))</td>
<td>Forward</td>
</tr>
<tr>
<td>((22.5\ deg, 67.5\ deg))</td>
<td>Left</td>
</tr>
<tr>
<td>((67.5\ deg, 112.5\ deg))</td>
<td>Left</td>
</tr>
<tr>
<td>((112.5\ deg, 157.5\ deg))</td>
<td>Left back</td>
</tr>
<tr>
<td>((157.5\ deg, 180\ deg) \cup (-180\ deg, -157.5\ deg))</td>
<td>Back</td>
</tr>
<tr>
<td>((-157.5\ deg, -112.5\ deg))</td>
<td>Right back</td>
</tr>
<tr>
<td>((-112.5\ deg, -67.5\ deg))</td>
<td>Right</td>
</tr>
<tr>
<td>((-67.5\ deg, -22.5\ deg))</td>
<td>Right forward</td>
</tr>
</tbody>
</table>

Using the interframe difference method, we can calculate the movement directions between every two adjacent frames. Considering that there are almost no movements that take less than 0.1 s, in our experiments, a suitable frame frequency used in the interframe difference is 10. Through the twice sampling of high frequency motion capture data, the frame frequency can be reduced to 10. For the consecutive vectors that belong to the same direction, they can be classified as a class, as shown in Fig. 6.

#### 4.2.2 Velocity threshold segmentation

The velocity threshold method is suitable for the segmentation of simple human movements, such as upper limb movements. There are discontinuities between the simple movements. For example, one arm moves to the left, then to the right. When the arm turns from left to right, the moving speed decreases first and then increases, so there exists a discontinuity between the two movements. In other words, there is a minimum velocity.

For a BVH file, setting a speed threshold not only can remove the nodes that do not move but also can separate
two adjacent different movements. When the speed of the arm is larger than the threshold, we know that the arm is in the process of moving. When the speed grows from less than the threshold to larger than it (red points in Fig. 7), we know that a new movement is going to start. When the speed changes from larger than the threshold to less than it (green point in Fig. 7), we know that a movement is finished. Between two adjacent movements, it generally indicates that there exists a transition when the speed value is less than the threshold. Figure 7 is an example of our velocity threshold segmentation.

Velocity threshold segmentation is based on the kinematic feature. This segmentation method is effective for movements with obvious pauses; however, the method does not work well when movements are smooth.

**4.2.3 Motion segmentation based on the probabilistic principal component analysis**

The segmentation based on the PPCA is extended to the PCA segmentation method.23

In the PCA method, each frame of motion capture data can be seen as a point in the high dimensional space. In this paper, the human body skeleton in the BVH file contains 26 nodes. For the 26 nodes (see Fig. 5), the root node has six-dimensions, five leaf nodes have no dimensions, and each of the other 20 nodes has three dimensions, for a total of 66-dimensions. Thus, the motion capture data of the \( i \)th frame \( x_i \) \((i = 1, 2, \ldots, n)\) represents a point in the 66-dimensional space. We can calculate the center of \( x_i \):

\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n},
\]

The essence of the PCA segmentation method is to reduce the dimension. Motion capture data of one frame is a 66-dimensional matrix, and we do eigenvalue decomposition of the data with \( n \) frames

\[
M = U \sum V^T,
\]

where \( M, U, \) and \( V \) are matrices with the size of \( n \times 66, n \times n, \) and \( 66 \times 66, \) respectively. \( \sum \) is a covariance matrix with the size of \( n \times 66 \)

\[
\sum = \begin{bmatrix}
\sigma_1 & & \\
& \sigma_2 & \\
& & \ddots \\
& & & \sigma_{65} \\
& & & & \sigma_{66}
\end{bmatrix}.
\]

The diagonal elements of \( \sum \) are arranged in descending order. Assuming that the first \( d \) columns are the principal
In Labanotation, there are 27 basic symbols, shown in Fig. 2. Analysis of Element Movements

4.3.1 Analysis of nonsupporting movements

Non-supporting movements include the upper limb movements and the lower limb movements that do not cause a change of the gravity center. The analysis of non-supporting movements only needs to consider the posture when the movements are finished. Then, we use Labanotation symbols to represent the posture of all the body parts.

For the 27 subspaces, Figs. 9(a) and 9(b) are the partitions of horizontal directions and vertical levels, respectively. The z-axis in Fig. 9(a) is parallel to vector \( \mathbf{n}_z = (x_f, 0, z_f) \), and its positive direction represents the front of the human body in the horizontal plane.

To determine the direction for an element movement, we need a reference direction based on the human body itself. For the skeleton model, a normal vector of the torso plane is generally used to represent the front of the human body, shown in Fig. 8. Assuming that vector \( \mathbf{n} = (x_f, y_f, z_f) \) and \( \mathbf{n}_z \) is the horizontal component of \( \mathbf{n} \), the positive direction of the human body in the horizontal direction is \( \mathbf{n}_z = (x_f, 0, z_f) \).

4.3 Analysis of Element Movements

In Labanotation, there are 27 basic symbols, shown in Fig. 2 in Sec. 3.1. The 27 symbols are related to 27 quantized spaces, in other words, 27 direction intervals. To determine the direction of an element movement, we need a reference direction based on the human body itself. For the skeleton model, a normal vector of the torso plane is generally used to represent the front of the human body, shown in Fig. 8. Assuming that vector \( \mathbf{n} = (x_f, y_f, z_f) \) and \( \mathbf{n}_z \) is the horizontal component of \( \mathbf{n} \), the positive direction of the human body in the horizontal direction is \( \mathbf{n}_z = (x_f, 0, z_f) \).

The vector that represents the front of the human body.

\[
\sigma^2 = \frac{1}{66-d} \sum_{i=d+1}^{66} \sigma_i^2. \tag{10}
\]

Motion data with \( n \) frames can be represented by a high dimensional Gaussian distribution. The mean value of Gaussian distribution is \( \bar{x} \) in Eq. (7), and the covariance matrix of the distribution is

\[
C = \frac{1}{n-1} (WW^T + \sigma^2 I) = \frac{1}{n-1} \mathbf{V} \sum_{2} \sigma^2 \mathbf{V}^T, \tag{11}
\]

where

\[
W = \mathbf{V} d \left( \sum_{2} - \sigma^2 \mathbf{I} \right)^{\frac{1}{2}}. \tag{12}
\]

Matrix \( \mathbf{V} \) is obtained by eigenvalue decomposition with the size of 66 x 66, \( \mathbf{V}_d \) representing the first \( d \) columns of \( \mathbf{V} \), \( \sum_{2} \) representing the submatrix of the upper left corner of \( \sum \), and \( \sum_{2} \) representing the matrix that uses \( \sigma \) to replace the 66-d eigenvalues in \( \sum \).

Assuming that motion capture data of the front \( K \) frames belong to the same element movement and have been represented by a Gaussian distribution, to determine whether the next \( T \) frames belong to the same Gaussian distribution or not, we need to calculate the mean Mahalanobis distance between the data in the next \( T \) frames and the previous Gaussian distribution:

\[
H = \frac{1}{T} \sum_{i=K+1}^{K+T} (x_i - \bar{x})^T C^{-1} (x_i - \bar{x}). \tag{13}
\]

Let \( K = K + 1 \), and calculate the mean Mahalanobis distance again. If the motion capture data from frame 1 to frame \( K + T \) belong to the same movement, with the increase of \( K \), the new Gaussian distribution will match the previous distribution very well. In this situation, from frame \( K + 1 \) to frame \( K + T \), the mean Mahalanobis distance \( H \) will decrease. When the Gaussian model is convergent, the distance \( H \) will achieve the local minimum value. After that, if the frames continue to increase, the new data will not belong to the previous movement and the value of \( H \) will begin to increase. Consequently, when \( H \) reaches the local minimum values, we need to cut the movements into two parts. Assuming the minimum value of \( H \) corresponding to frame \( K_m \), we can draw a conclusion that the data from frame 1 to frame \( K_m \) belong to an element movement and that the data from frame \( K_m + 1 \) to frame \( K + T \) belong to another element movement. In this way, the PPCA method can do a good job in the motion segmentation of human lower limbs.

4.3 Analysis of Element Movements

In Labanotation, there are 27 basic symbols, shown in Fig. 2 in Sec. 3.1. The 27 symbols are related to 27 quantized spaces, in other words, 27 direction intervals. To determine

\[
\sigma^2 = \frac{1}{66-d} \sum_{i=d+1}^{66} \sigma_i^2. \tag{10}
\]

Motion data with \( n \) frames can be represented by a high dimensional Gaussian distribution. The mean value of Gaussian distribution is \( \bar{x} \) in Eq. (7), and the covariance matrix of the distribution is

\[
C = \frac{1}{n-1} (WW^T + \sigma^2 I) = \frac{1}{n-1} \mathbf{V} \sum_{2} \sigma^2 \mathbf{V}^T, \tag{11}
\]

where

\[
W = \mathbf{V} d \left( \sum_{2} - \sigma^2 \mathbf{I} \right)^{\frac{1}{2}}. \tag{12}
\]
vector. We then take the positive direction of the z-axis as a reference and quantify the direction of the vector to one of the 27 subspaces. Therefore, we determine a symbol for the element movement.

Assuming vector \( \mathbf{V} \) is the direction of an element movement, we calculate the angle between vector \( \mathbf{V} \) and y-axis to quantify the vertical level

\[
\theta = \arccos \left( \frac{\mathbf{V}_y}{|\mathbf{V}|} \right),
\]

where \( \mathbf{V}_y \) is the vertical component of vector \( \mathbf{V} \). The quantified results of the vertical level are as follows:

- high, \( \theta < \frac{\pi}{6} \)
- middle, \( \frac{\pi}{6} \leq \theta < \frac{5}{6} \pi \)
- low, \( \theta \geq \frac{5}{6} \pi \).

Quantifying the horizontal directions of vector \( \mathbf{V} \) is similar to the process of vertical level.

In Fig. 9(a), the angles between every two adjacent solid lines are 45 deg, and the dotted lines (z-axis and x-axis) are angular bisectors. In Fig. 9(b), the angles between two adjacent solid lines are 60 deg, and the dotted line (y-axis) is an angular bisector.

### 4.3.2 Analysis of supporting movements

The analysis of supporting movements needs to consider three aspects: determining the type, the direction, and the body part.

First, we need to determine the type of supporting movements: jumping or moving. For jumping movements, all the body parts leave the ground and the gravity center is higher than normal movements. Moving movements describe the horizontal and vertical changes of the gravity center. Two kinds of supporting movements are easy to distinguish based on the BVH file. If the height of the root node is higher than a threshold, no body parts will be in contact with the ground, and that is jumping. If there is no jumping and only the position of root node changes, that is moving.

Second, we determine the moving direction of supporting movements. The direction is the same with the moving direction of the gravity center, and it is determined by the body part that causes the movement. In Sec. 4.2.1, the movements of the gravity center are segmented by the spatial clustering of the Laban direction, and we have quantified the horizontal directions in Table 2. To quantify the vertical level of supporting movements, we need to calculate the height of the gravity center. As shown in Fig. 10, we use low, middle, and high to quantify vertical levels. When a person is squatting, the height of gravity center falls to the blue line. When a person is standing naturally, the gravity center is at the red line. When a person is standing on tiptoe, the gravity center rises to the green line.

Third, we determine the supporting part; in other words, we need to find out the body part that causes the supporting movement. For a jumping movement, we should find out the part that supports the human body before jumping. For a moving movement, we should find out the part whose moving direction is the same as the gravity center.

Jumping movements are usually caused by the force of the legs. There are two kinds of jumping movements: jumping with two legs or with one leg. For jumping with one foot, the other foot that is not used will be the first to leave the ground. Thus, jumping with one or two legs can be judged by the heights of two feet. In the motion capture data, assuming that body parts begin to leave the ground in the \( i \)'th frame, then calculate the heights of two feet in the \( (i + 1) \)'th frame. If both feet are on the ground, it represents that the jump is caused by two legs. If only one foot is on the ground, it represents that the jump is caused by one leg. Taking into account that the human body is a nonrigid structure, we use 5 cm as a threshold to determine whether the feet have left the ground.

Different from jumping, the moving movements cannot be caused by two legs simultaneously. In the process of a moving movement, if one leg moves, the other leg will play the role of supporting. At this time, the movement of the gravity center is caused by the leg that moves. Therefore, in this situation, the supporting part is the leg...
that moves, not the one that stands on the ground, as shown in Fig. 11.

5 Experiments and Evaluations

The automatic generation of Labanotation is realized by developing a software platform of the proposed system. We use the Microsoft Visual Studio 2008 development environment and call functions in the OpenCV 2.3.1 Library. The graphical interface of the software is based on the Microsoft foundation class library.

5.1 System Evaluation

During our experiments, we processed about 70 sections of Chinese traditional folk dance, drama, and other kinds of action art. To validate the accuracy of the results, we selected two automatic generated Labanotation pieces to compare with the original human motion. In our experiments, the information of all the movements was noted in detail.

The information about the first section of movement is as follows:

1. The time of the motion capture: June 21, 2014.
2. The place of the motion capture: Center for Ethnic and Folk Literature and Art Development, Ministry of Culture, Beijing, China.
3. The equipment of the motion capture: OptiTrack motion capture system.
5. The art form of the motion: a section of “Qi Ba,” Beijing opera, Chinese opera.
6. Motion capture data: BVH format, the skeleton of the human body in the BVH file contains 26 nodes, frame rate of the motion capture data is 150, and the total number of the frames is 34,714.

The generated Labanotation of the first section of movement contains 20 pages, as shown in Fig. 12.

We select several key frames from the motion capture data and compare the Labanotation with the human body postures in the frames, as shown in Fig. 13. ① is the preparatory posture. In accordance with the rules of Labanotation, we need to use a middle place symbol to note the posture in a support column, and the posture of upper limbs is noted by a middle left symbol and a middle right symbol. As shown in ①, the automatic generated Labanotation for the preparatory posture is correct. ② represents that the moving of the left leg causes a right front moving of the human body’s gravity center, the left arm is middle left front, and the right arm is low right. ③ represents that the gravity center is not moved.

Fig. 11 The moving of gravity center is caused by the left leg.

Fig. 12 Generated Labanotation of a section of Beijing opera “Qi Ba.”
the left arm is middle left front, and the right arm is middle front. ④ represents that the moving of the right leg causes a left front moving of the gravity center, the left arm is low left front, and the right arm is low right. ⑤ represents that the moving of the left leg causes a middle front moving of the gravity center, the left arm is low left, and the right arm is low front. ⑥ represents that the moving of the left leg causes a middle back moving of the gravity center, the left arm is middle left front, and the right arm is middle front. ⑦ represents that the gravity center is not moved, the left leg and left arm are both middle left front, and the right arm is high front. ⑧ represents that the moving of the right leg causes a middle left moving of the gravity center, the left arm is middle left, and the right arm is middle left front. ⑨ represents that the moving of the left leg causes a middle left front moving of the gravity center, the left arm is middle left front, and the right arm is low right. Through the analysis, the generated Labanotation can match the human

Fig. 13 Comparison of generated Labanotation and the corresponding postures.
body movements in the motion capture data. Furthermore, the automatic generation of Labanotation can correctly express these movements. Thus, the generated Labanotation is correct.

The information of the second section of movement is as follows:

(1) The time of the motion capture: January 6, 2015.
(2) The place of the motion capture: Center for Ethnic and Folk Literature and Art Development, Ministry of Culture, Beijing, China.
(3) The equipment of the motion capture: OptiTrack motion capture system.
(4) The performer of the motion: Xiaoyi Wang, teacher of the Beijing Normal University.
(5) The art form of the motion: the combination of the shoulder, Mongolia dance, Chinese minority dance.
(6) Motion capture data: BVH format, the skeleton of the human body in the BVH file contains 26 nodes, the data frame rate is 30, and the total number of the frames is 3766.

The second section of movement is a piece of Mongolia dance. The corresponding Labanotation is shown in Fig. 14. In this Mongolia dance, the main movements are the moving of shoulders and the moving of arms. The generated Labanotation reproduces most of the movements correctly, and the symbols match the corresponding postures. Therefore, the generated Labanotation of this Mongolia dance correctly expresses the movements.

We also compare our work with the method of Choensawat et al. The approach in Ref. 12 divides the space into 27 subspaces, including nine horizontal directions and three vertical levels. Then, it analyzes the body parts including two arms and two legs. After that, it segments and quantifies the movements of each body part into 27 subspaces and analyzes the movements with weight supporting, jumping, and bending. In the comparison, we use two series of movements: walking and jumping with two feet. The results in Ref. 12 use the motion capture data from the CMU database, which is a widely used benchmark. We imitate the movements used in Ref. 12 and get the capture data by our system with OptiTrack devices. Because the walking and jumping are easy and common movements, our motion capture data are very close to the data used in Ref. 12. The generated Labanotation pieces are shown in Figs. 15 and 16.

Fig. 14 Generated Labanotation of a piece of Mongolia dance.

Fig. 15 The comparison results of “walking” motion. (a), (b), and (c) are pictures of the motion capture data. (d) and (e) are the generated Labanotations of the work in Ref. 12 and our system.
Fig. 16 The comparison results of "jumping" motion. From (a) to (e) are pictures of the motion capture data. (f) and (g) are the generated Labanotations of the work in Ref. 12 and our system.

Fig. 17 Using human movements (video), human motion capture data (BVH file) and generated Labanotation to show a piece of Tibetan dance, a section of "Dong Wang Guo Zhuang."
In the movements of walking, Figs. 15(a)–15(c) are the pictures of the motion capture data; Figs. 15(d) and 15(e) are the dance notations of Ref. 12 and our method, respectively. There is an obvious mistake in the result of Ref. 12. The first section line, shown by the arrows, should be a double line. However, in Fig. 15(d), it is only a single line, which does not conform to the rules of Labanotation. Moreover, the walking is a normal motion of human beings, so we do not need to describe too many details. Two blue rectangles in Fig. 15(d) represent small natural swings of two arms. The notations of the arms are right; however, they are a little redundant. If we do not express the movements in Labanotation, it means that we do the movements naturally. Thus, in Fig. 15(e), we omit this kind of movement by setting a speed threshold, introduced in Sec. 4.2.2. The movements of jumping are shown in Fig. 16; the results of Ref. 12 have the same problems. If the movements are not natural ones, we will write down the corresponding notations. For example, in Fig. 13, the arm movements of Beijing opera are written down in detail.

To show the human movements, motion capture data (BVH file), and the generated Labanotation simultaneously, we put these three things together and make a video. The pictures in the videos of another two sections of movements are shown in Figs. 17 and 20. Figs. 18 and 19 are the respective Labanotations.

The information of the movement shown in Fig. 17 is as follows:

1. The time of the motion capture: January 18, 2015.
2. The place of the motion capture: Center for Ethnic and Folk Literature and Art Development, Ministry of Culture, Beijing, China.
3. The equipment of the motion capture: OptiTrack motion capture system.
4. The performer of the motion: Haozhi Han, student of the Beijing Dance Academy.

Motion capture data: BVH format, the skeleton of the human body in the BVH file contains 26 nodes, the data frame rate is 150, and the total number of frames is 17,520.

The information of the movement shown in Fig. 20 is as follows:

1. The time of the motion capture: January 17, 2015.
2. The place of the motion capture: Center for Ethnic and Folk Literature and Art Development, Ministry of Culture, Beijing, China.
3. The equipment of the motion capture: OptiTrack motion capture system.
6. Motion capture data: BVH format, the skeleton of the human body in the BVH file contains 26 nodes, the data frame rate is 150, and the total number of frames is 14,342.

5.2 User Evaluation

The purpose of user evaluation is to measure the satisfaction with the proposed system and the generated Labanotation. There are two groups of people that help us do the evaluations. One group is the experts and teachers in Laban Research Center of Beijing Normal University, consisting of five people. The other group is the students of Beijing Normal University and Beijing Jiaotong University who have taken the class of Labanotation with the help of the Laban Research Center, consisting of 5 of 20 students. In their study, there are 16 classes in a semester of about 4 months.

The first evaluation focuses on the five experts and teachers. The work experience of the five people is shown in Table 3. We ask them to complete a questionnaire about the usefulness of the proposed system and the accuracy of the generated notation. There are two parts in our questionnaire. One part contains five questions assessed with quantitative evaluations, where 1 = very bad and 10 = very good, shown in Table 4. The other part encourages the five people to write some comments or give feedback for our system, shown in Table 5.

From the first and second items in Table 4, on average, all users approve that the proposed system is easy to use and has a satisfactory response time. It demonstrates that our system is useful and efficient for generating the Labanotation.

For the third and fourth items, the scores are acceptable, but they are much lower than the first and second ones. The reasons are as follows. In Labanotation, the 27 basic symbols (shown in Fig. 2) occupy a large proportion. Our system can handle these basic symbols, so the system is “helpful for writing Labanotation.” However, the system does not do well in the complex movements; thus, the score of the third item is not high. For a same series of movements, the understandings of different experts are not the same, and, of course, the related dance notations are different. Therefore, the forth item has a relatively low score. To solve the problem, we are committed to allowing the system to deal with more symbols and letting the system give more than one result for people to select corresponding to several possible understandings.

For the fifth item, all users agree that our system does help protect the folk dance. It shows that the users are satisfied with our system considering efficiency and accuracy.

From Table 5, the inadequacies will guide us in future work, and we will try to do better on the good sides.

The second evaluation focuses on the students, in other words, for novices. We design another questionnaire for the students to evaluate our system. The questions and results are shown in Table 6. For the first question, given a series of movements, the students can write about 65% to 85% of the Labanotation by themselves.
However, with the help of our system, the ratio will rise to 80% to 90%. All the students agree that the system can help them complete the notation. For the second question, every student admits that our system can provide ideas and reference answers for them; thus, the system helps them improve the efficiency (30% to 70%) for writing the notation. For the third question, most students feel that the accurate rate of the generated Labanotation is about 70% to 90%, which has a high reference value.

![Image of human movements and Labanotation]

**Fig. 20** Using human movements (video), human motion capture data (BVH file) and generated Labanotation to show a piece of Tujia Dance, a section of "Waving Dance." There are two performers doing the same movements.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Work experience of Labanotation</th>
<th>Profession(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>More than 20 years</td>
<td>Dance researcher, choreographer, and retired dance teacher</td>
</tr>
<tr>
<td>Subject 2</td>
<td>More than 20 years</td>
<td>Dance researcher, and retired dancer</td>
</tr>
<tr>
<td>Subject 3</td>
<td>10 to 20 years</td>
<td>Dancer and dance teacher</td>
</tr>
<tr>
<td>Subject 4</td>
<td>Under 5 years</td>
<td>Dance teacher and dance researcher</td>
</tr>
<tr>
<td>Subject 5</td>
<td>Under 5 years</td>
<td>Dancer, dance researcher, and graduate dance student</td>
</tr>
</tbody>
</table>
Table 4 Five questions and average scores.

<table>
<thead>
<tr>
<th>Questions for the proposed system</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Easy to use</td>
<td>8.8</td>
</tr>
<tr>
<td>2. Response time</td>
<td>9.2</td>
</tr>
<tr>
<td>3. Helpful for writing Labanotation</td>
<td>8.2</td>
</tr>
<tr>
<td>4. Accuracy of the generated Labanotation</td>
<td>7.2</td>
</tr>
<tr>
<td>5. Can do some help for protecting folk dance</td>
<td>8.4</td>
</tr>
</tbody>
</table>

Table 5 Comments and feedback.

<table>
<thead>
<tr>
<th>Comments</th>
<th>1. The software is easy to learn and use, and has a good response time.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>2. The software should contain more Labanotation symbols.</td>
</tr>
<tr>
<td>Negative</td>
<td>1. The software can only deal with one dancer.</td>
</tr>
<tr>
<td></td>
<td>2. The software cannot do well in generating complicated and slow movements, such as slow rotational movements.</td>
</tr>
</tbody>
</table>

Table 6 Questions and results in the second evaluation.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. After watching a series of movements on a video, what is the ratio that you can complete the Labanotation?</td>
<td>65% to 85% (not with the help of our system) 80% to 90% (with the help of our system)</td>
</tr>
<tr>
<td>2. Does the system help you improve the efficiency for writing the dance notation?</td>
<td>Yes (100% students) Improves the efficiency by 30%-70%.</td>
</tr>
<tr>
<td>3. What is the accurate rate of the system?</td>
<td>70% to 90%</td>
</tr>
</tbody>
</table>

6 Conclusions
In this study, we proposed technology that uses human motion capture data to automatically generate Labanotation. We analyze the BVH file, which is a general format of motion capture data, and cut the motion capture data into element movements. We then calculate the symbol of Labanotation for every element movement; this process is the main innovative work in this study. The generated Labanotations match the human body movements and express the most basic movements correctly.

The use of motion capture data to automatically generate Labanotation solves the problem in documenting the dance notation for a series of movements. Therefore, the proposed system is suitable for recording and preserving the folk dance and other types of action arts. The system is an important contribution to the protection of action arts in intangible cultural heritage.

At present, the system handles most of the basic symbols that correspond to several fundamental movements. However, the Labanotation is rich in symbols, especially for the details of the complex movements. In future work, we will extend the system to contain more symbols and deal with movements that are more complicated. Furthermore, we will make the system more practical in the protection of folk dance.

Acknowledgments
This work was supported by the NSFC 61273274, 61672089, 61572064, PXM2016_014219_000025, National Key Technology R&D Program of China 2012BAH01F03. We thank Professor Bingyu Luo for instruction on Laban dance notation.

References
6. L. Venable et al., Laban Writer 2.0, The Ohio State University, Department of Dance (1989).
24. C. MoCap, (2003) The data used in this project was obtained from http://mocap.cs.cmu.edu/ the database was created with funding from nsf eia-0196217.

**Jiaji Wang** received his BE degree from Beijing Jiaotong University in 2012. He is currently pursuing his PhD. He is the author of one conference paper. His current research interests include pattern recognition, image processing, and multiview computer vision. His current research mainly focuses on image processing and multiview human motion capture.

**Zhenjiang Miao** received his BE degree from Tsinghua University in 1987, Beijing, China, and his ME and PhD degrees from Northern Jiaotong University, Beijing, in 1990 and 1994, respectively. He joined Beijing Jiaotong University, Beijing, in 2004. He is currently a professor with Beijing Jiaotong University. His current research interests include image and video processing, multimedia processing, and intelligent human-machine interaction.

**Hao Guo** received his BE and ME degrees from Beijing Jiaotong University in 2012 and 2015. He is currently pursuing his PhD at the University of South Carolina. His current research interests include pattern recognition, image processing, and multiview computer vision.

**Ziming Zhou** received his BE degree from Beijing Jiaotong University in 2014. She is currently pursuing her ME degree. Her current research interests include pattern recognition, image processing, and multiview computer vision.

**Hao Wu** received his BE and PhD degrees from Beijing Jiaotong University in 2010 and 2016. He joined Beijing Normal University, in 2016. He is currently pursuing his PhD at the University of South Carolina. His current research interests include image and video processing, multimedia processing, and intelligent human-machine interaction.