Journal of Information Processing Vol. 18 263–283 (Dec. 2010)

Regular Paper

Adaptation of Grouping Structure Analysis in GTTM to Hierarchical Segmentation of Dance Motion

Takeshi Miura,[†1] Kazutaka Mitobe,[†1]
Takashi Yukawa,[†2] Takaaki Kaiga,[†1,†3]
Toshiyuki Taniguchi[†1] and Hideo Tamamoto[†1]

In this paper, the authors propose the adaptation of the rules used in the grouping structure analysis in Lerdahl and Jackendoff's "A Generative Theory of Tonal Music (GTTM)" to dance motion analysis. The success of the adaptation realizes the segmentation of dance motion in a hierarchical fashion. The analysis method obtained by the trial of the above adaptation consists of the following procedures. A motion-capture data stream of a dance is first divided into a sequence of events by piecewise linear regression. The hierarchical structure of groups each of which consists of a sequence of the events is then extracted by applying the grouping rules adapted to dance motion analysis. The above method is applied to motion-data streams acquired by motion capture systems. The obtained results indicate the following advantages: (1) The structure of hierarchical segmentation is precisely extracted in response to the characteristic of an analyzed dance. (2) The extraction of the hierarchical segmentation provides the possibility of the development of a technique distinguishing the oversegmentation from regular boundaries. (3) The possibility of utilizing the information of hierarchical segmentation for the comparison of dance performances is suggested.

1. Introduction

Much recent interest in dance motion analysis has been generated by the demonstration that the utilization of a motion capture system allows us to analyze human-body motions more precisely than previous trials.

It has been recognized that dividing a raw motion-capture data stream of a dance into small motion segments on the time axis is needed. The obtained motion segments can be utilized as information units such as indexes used for data retrieval [1], motion clips representing highlight movements in a dance [2], primitive motions organized to construct the whole choreography [3),4)], etc. There have been many studies on the development of a technique to segment dance motion [1)–4)] and, as a consequence, the automatic segmentation has been realized in some measure.

Nevertheless, the establishment of a more advanced technique is required in the present stage. In particular, the extraction of the hierarchical structure of motion segments have attracted much attention because of its ability to provide the multi-level information such as that of both the physical level representing the biomechanical smallest-scale segmentation and the semantic level giving the higher-level global structure of choreography [2], etc. However, few studies have been reported on a technique for the hierarchical segmentation of motion of a dance, giving successful results for any dance.

Turning now to a discussion about an art style other than dance, many musical idioms give the situation that appreciators listen to melodies, harmonies and rhythms which can be regarded as the temporal variation of some state. The idiom of dance provides the common style, namely the appreciation of the temporal variation of the state. This fact leads to the possibility that an analysis technique which is effective in the parsing of a piece of music on the time axis may also be applicable to the segmentation of dance motion.

Lerdahl and Jackendoff's *A Generative Theory of Tonal Music* (GTTM) [5] is known as the theory for the analysis aiming at the cognitive organization of a piece of music. GTTM consists of four sub-theories: grouping structure analysis, metrical structure analysis, time-span reduction and prolongational reduction. All of them have a common form, namely that a set of "rules" is used to generate the structural description of a piece, the same as the use of a grammar in generative linguistic theory occupying an important position in cognitive science.

It is noted that the rules used in the grouping structure analysis providing the hierarchical segmentation of a piece are constructed on the basis of the psychological principles which are thought to be more idiom-independent than those used in the other sub-theories of GTTM [5]. This fact indicates the possibility that the grouping rules may easily be adapted to art styles other than music, especially to dance having a style similar to that of music.

†1 Graduate School of Engineering and Resource Science, Akita University
†2 Faculty of Law, North Asia University
†3 Digital Art Factory, Warabi-za Co., Ltd.

If the adaptation succeeds, the segmentation of dance motion is realized in a hierarchical fashion which is the representation form required at the present time as already mentioned. It gives not only the positions of boundaries between adjacent segments on the time axis but also the additional information such as the relative intensity, or the priority, of each of the boundaries. This fact may provide the possibility of estimating the aspect of the organization of motion segments in the whole dance motion.

In order to obtain some insight into the above possibility, the authors attempt, in this paper, to adapt the grouping rules of GTTM to the analysis of a motion-data stream of a dance.

According to Lerdahl and Jackendoff's GTTM book[5], their attention was intensively given to an attempt to make the theory as predictive as possible by clearly stating the rules. In contrast, the implementation in a computable way was not considered because of the difficulty in the numerical processing of the relationship among rules.

More recently, however, trials for the computational implementation of GTTM have been reported[6),7)]. Although several problems such as the robustness against the diversity of objects, of which the detail is described in Section 4, still remain, the usefulness of the numerical approach for the analysis of a piece of music has been demonstrated in some measure.

Considering the above situation, the authors try to adapt the grouping rules to dance motion analysis in a numerical way, introducing several approaches similar to those performed in the above studies as the need arises, since a motion-capture data stream has the style of strings of numerical values.

First, the form of a data stream representing human-body motions is described in Section 2. Next, the adaptation of the grouping rules of GTTM to dance motion analysis is presented in Section 3; the concrete implementation process of the adapted rules are shown in Section 4. Then, the results derived by the method developed in Section 4 and the discussion about their validity are shown in Section 5. Finally, conclusions are summarized in Section 6.

2. Description of Dance Motion

In this paper, the whole motion of a dance is described in the form of a time-series data stream of joint angles. The reason for this is that the analysis of joint motions provides the discrete motion-constituents of a dance as described in Section 3.2. The posture of a dancer in each instant corresponds to each of the frames in the data stream. The individual joint angle is represented by the three-degree-of-freedom (3DOF) rotation of a body segment below the joint in order to describe all of the fundamental joint movements: flexion/extension, adduction/abduction and medial/lateral rotation[8)]. The exponential map[9)] providing the three-dimensional Euclidean parameterization of 3DOF rotations is used for the representation of the joint angles[8)].

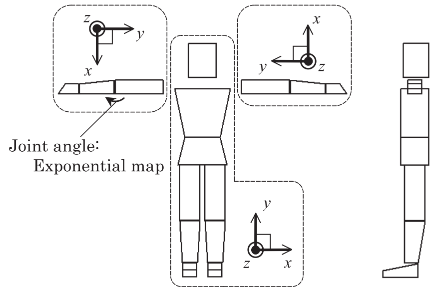


Fig. 1 A model of a human body.

The human-body model used for the description of motion data having the above form is shown in **Fig. 1**. It has 15 body segments connected at 14 joints. Each of the body parts has an individual reference orientation with a local coordinate system in order to eliminate the influence of the singular point appearing in the exponential-map parameterization[8),9)]. The assignment of the local coordinate systems is indicated by the broken lines in Fig. 1. As a result, a motion-data stream is given as the time-series data stream of the 42-dimensional vector $\boldsymbol{u}(n) = [\ u_1(n) \quad u_2(n) \quad \cdots \quad u_{42}(n)\ ]^T$ in which the number of components is obtained as 14 joints × 3 components.

An additional remark which should be made here is that the individual components of the exponential map in this model correspond to the respective fundamental joint movements as follows[8)]: x component−flexion/extension, y component−medial/lateral rotation and z component−adduction/abduction.

3. Application of GTTM to Dance Motion Analysis

3.1 Grouping Rules in GTTM

The grouping structure analysis in GTTM is performed using the grouping rules divided into two types: Grouping Well-Formedness Rules (GWFRs)[5] shown in **Table 1** and Grouping Preference Rules (GPRs)[5] in **Table 2**[⋆1]. GWFRs define the constraints to be satisfied in the description of grouping structures assigned to a piece[5]. GPRs, on the other hand, give the criteria to find the possible grouping structures agreeable to the intuitive judgments of experienced listeners from a number of candidates satisfying GWFRs[5].

3.2 Adaptation of Grouping Rules to Dance Motion Analysis

In GWFRs and GPRs, the rules corresponding to the properties commonly shown in both the data form of music used in GTTM and that of motion-capture data of a dance can directly be applied to dance motion analysis, while some adaptation is necessary for those related to the properties peculiar to music. The above two kinds of properties are extracted in this subsection.

3.2.1 Application of GWFRs

It is easily recognized as one of the common properties that boundary positions between segmented motions are located on the time axis[1)–4)]. In addition, there is a well-known tendency that dance motion has a relatively high correspondence with the musical accompaniment[3] [⋆2]. Hence, it is estimated to be reasonable that the constraints in GWFRs prepared for music are also applicable to dance motion. As a result, the motion of a dance is organized into the hierarchical structure common to that of a piece of music. The detail of the application of GWFRs is shown in Section 4.1.

3.2.2 Derivation of Basic Constituents of Dance Motion

Upon the application of GPRs, a piece of music is organized using the information of the score of it, consisting of basic musical constituents on the time axis,

⋆1 In this paper, GPR 7 is not considered since this rule is prepared to utilize the information sent from the time-span and/or prolongational reductions which are not implemented in the present stage.

⋆2 More detailed examination is necessary to adapt the grouping rules to dances having the different tendency.

Table 1 Grouping Well-Formedness Rules (GWFRs)[5].

GWFR 1	Any contiguous sequence of pitch-events, drum beats, or the like can constitute a group, and only contiguous sequences can constitute a group.
GWFR 2	A piece constitutes a group.
GWFR 3	A group may contain smaller groups.
GWFR 4	If a group G1 contains part of a group G2, it must contain all of G2.
GWFR 5	If a group G1 contains a smaller group G2, then G1 must be exhaustively partitioned into smaller groups.

Table 2 Grouping Preference Rules (GPRs)[5].

GPR 1	Avoid analyses with very smaller groups - the smaller, the less preferable.
GPR 2	(Proximity) Consider a sequence of four notes $n_1n_2n_3n_4$. All else being equal, the transition n_2-n_3 may be heard as a group boundary if **a.** (Slur/Rest) the interval of time from the end of n_2 to the beginning of n_3 is greater than that from the end of n_1 to the beginning of n_2 and that from the end of n_3 to the beginning of n_4, or if **b.** (Attack Point) the interval of time between the attack points of n_2 and n_3 is greater than that between the attack points of n_1 and n_2 and that between the attack points of n_3 and n_4.
GPR 3	(Change) Consider a sequence of four notes $n_1n_2n_3n_4$. All else being equal, the transition n_2-n_3 may be heard as a group boundary if **a.** (Register) the transition n_2-n_3 involves a greater intervallic distance than both n_1-n_2 and n_3-n_4, or if **b.** (Dynamics) the transition n_2-n_3 involves a change in dynamics and n_1-n_2 and n_3-n_4 do not, or if **c.** (Articulation) the transition n_2-n_3 involves a change in articulation and n_1-n_2 and n_3-n_4 do not, or if **d.** (Length) n_2 and n_3 are of different lengths and both pairs n_1, n_2 and n_3, n_4 do not differ in length.
GPR 4	(Intensification) Where the effects picked out by GPRs 2 and 3 are relatively more pronounced, a larger-level group boundary may be placed.
GPR 5	(Symmetry) Prefer grouping analyses that most closely approach the ideal subdivision of groups into two parts of equal length.
GPR 6	(Parallelism) Where two or more segments of the music can be construed as parallel, they preferably form parallel parts of groups.

namely notes and rests [5]; they are the discrete constituents peculiar to music. Hence, the discrete basic motion-constituents of dance should newly be derived.

Each of notes represents a period in which the pitch is kept constant, while each of rests shows a period in which a non-sound state is maintained. According to Ref. 10), on the other hand, the basic constituents of dance motion consist of bending, stretching and twisting. These constituents correspond to the fundamental joint movements as follows: bending−flexion/adduction, stretching−extension/abduction and twisting−medial/lateral rotation. The direction of joint motion is kept constant in each constituent, the same as the cases of notes and rests in each of which a constant state is maintained.

Considering the above situation, we regard the individual fundamental joint movements as the basic motion-constituents of a dance. The detail of the resolution of the whole motion into motion-constituents is shown in Section 4.2.

3.2.3 Application of GPRs 1, 2 and 3

According to the GTTM book [5], GPRs 1, 2 and 3 are prepared to organize the lowest part of the hierarchical structure of a piece (this part is called the “local detail” of a piece [5]), while GPRs 4, 5 and 6 provide its global structure.

As for the rules for the organization of the local detail, GPR 1 is estimated to be applicable to both music and dance since it has little relationship with the idiom of art. On the other hand, some modification is needed for GPRs 2 and 3 since they are related to the properties peculiar to music: proximity and change.

GPR 2 uses the information of proximity. In order to derive the degree of proximity, the distance between the adjacent notes is evaluated on the basis of two criteria: the period in which a non-sound state is maintained and the relative length of a note in which the constant pitch is maintained.

In dance motion, the period of stillness of the body which inserts a rest between motions can be regarded as the part corresponding to the non-sound state in a piece of music. On the other hand, the relative length of a motion-constituent cannot be used for the criterion of proximity since the constituent itself gives the motion without the rest. Hence, we evaluate the proximity of motions only by the stillness of the body. The detail of data processing is shown in Section 4.3.

In GPR 3, four properties are used to evaluated the degree of change between adjacent notes: the interval of pitch, the change of dynamics, the change of articulation and the difference in length. These properties are peculiar to music except for the length of basic constituents. Instead of the remaining three properties, some evaluation item should be prepared.

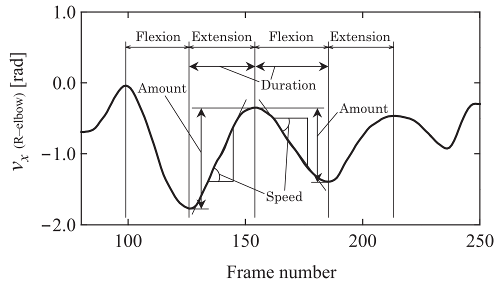


Fig. 2 An example of joint motion (flexion/extension of R-elbow, dance: *Ondo* of *Nishimonai Bon Odori*).

First, it is noted that the direction of motion upon the change of which the whole motion is resolved into motion-constituents can be used as one of the items.

Next, we investigate the aspect of actual joint motion in order to derive other evaluation items. An example is shown in **Fig. 2**; the aspect of flexion/extension of the right elbow in the dance *Ondo* of *Nishimonai Bon Odori* is shown in this figure (the detail of the dance is shown in Section 5). The division points resolving the data stream into motion-constituents are detected by searching the local minimum/maximum points in the temporal variation of a component of the exponential map.

It is seen in Fig. 2 that each motion-constituent has a relatively simple monotonically-increasing/decreasing form. Hence, the three properties indicated in the figure can be regarded as the evaluation items characterizing each motion-constituent: the amount of motion, the speed of motion and the duration of motion which is the counterpart of the length of a note.

Finally, the above four items are adopted as the properties used for the evaluation of the change of motion. The detail of the data processing for the above items is shown in Section 4.3.

3.2.4 Application of GPRs 4, 5 and 6

As already mentioned, GPRs 4, 5 and 6 are used for the derivation of the global structure of a piece of music.

In GPR 4, a criterion in which the intensity of the effect of GPRs 2 and 3 is evaluated is provided. GPR 4 itself, however, is not influenced by the evaluation items used in GPRs 2 and 3. Hence, we directly apply this rule to dance motion analysis. The detail of the application of this rule is shown in Section 4.4.

GPR 5 shows the recommendation of binary-form grouping. This tendency is also valid for many dance categories. However, several dance categories have a ternary form, or a sexenary form which can be given by binary grouping of a ternary form, in their grouping structure [11]. Hence, a procedure which allows not only the derivation of binary grouping but also that of ternary one is added in the implementation of this rule. The detail is shown in Section 4.4.

In order to implement GPR 6 which points out the necessity of processing the parallelism in multiple segments, the definition of parallelism and the means to perform the processing should be prepared by users. In addition, both of them depend on the idiom of the art style. The concrete process of the definition and implementation for dance motion is shown in Sections 4.3 and 4.4.

The adaptation of the grouping rules to dance motion analysis is summarized in **Table 3**.

3.3 Definition of the Terms and Symbols Used for the Description of Hierarchical Segmentation of Dance Motion

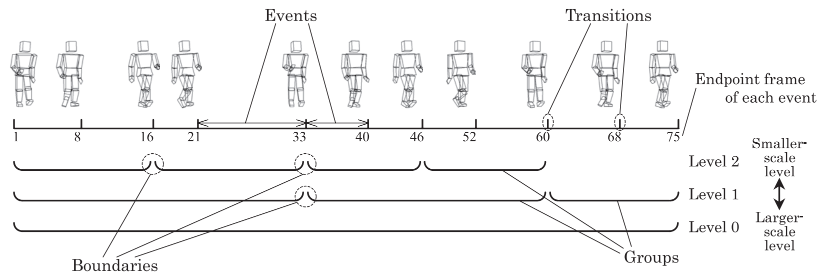
The application of the grouping rules makes the motion of a dance hierarchically organized into segments, called *groups*, each of which consists of either motion-constituents or groups on a lower level. The moment between adjacent groups is called a *boundary*. The other terms and symbols for the description of the hierarchical structure of dance motion are defined as follows.

The term *event* is used as that which refers to each of the motion-constituents. The other terms and symbols are shown in **Fig. 3**; they are based upon the notation used in the GTTM-analysis of a piece of music [5]. The moment between adjacent events is interpreted as a *transition* between events. The hierarchical structure consists of multiple layers each of which is called a *level*. The level number for the largest-scale level including the whole motion is set to zero. It increases with the progress to smaller-scale levels.

In Fig. 3, a motion-data stream of a walk (“02_02” in http://mocap.cs.cmu.edu,

Table 3 Adaptation of Grouping Rules to Dance Motion Analysis.

Basic constituents	The fundamental joint movements are used as motion constituents: **a.** Bending (Flexion/Adduction). **b.** Stretching (Extension/Abduction). **c.** Twisting (Medial/Lateral rotation).
GWFRs	Directly applied.
GPR 1	Directly applied.
GPR 2	(Proximity) The degree of stillness of the body is evaluated.
GPR 3	(Change) The following properties are evaluated: **a.** Direction of motion. **b.** Amount of motion. **c.** Speed of motion. **d.** Duration of motion.
GPR 4	(Intensification) Directly applied.
GPR 5	(Symmetry) The procedure for the derivation of ternary grouping is added.
GPR 6	(Parallelism) Redefined for dance motion analysis.



Event: basic constituent extracted by dividing the whole motion according to the change of fundamental joint movement
Group: a set of contiguous events or groups on the next smaller-scale level
Local boundary: boundary adjoining a group including no boundary
Global boundary: boundary in large-scale levels (selected from local boundaries)

Fig. 3 Definition of grouping structure (sample data: “02_02” in http://mocap.cs.cmu.edu).

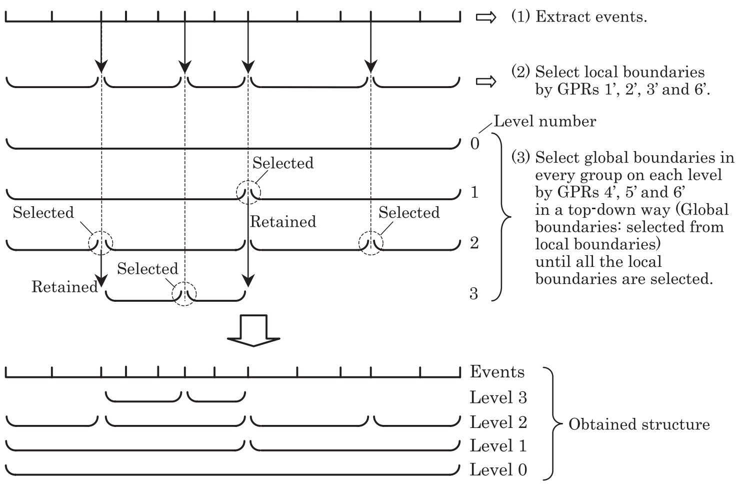


Fig. 4 Algorithm for the extraction of grouping structure.

transformed to the data form shown in Section 2 with the sampling time of 33.3 ms) is used as a sample data stream. This sample is continuously used for the explanation of the data processing in Section 4.

4. Implementation of GTTM in Dance Motion Analysis

4.1 The Whole Implementation Algorithm

As mentioned in Sections 3.2.3 and 3.2.4, GPRs 1, 2 and 3 are prepared to organize the local detail of a piece, namely to detect *local boundaries* in Fig. 3, while GPRs 4, 5 and 6 provide the structure of global levels, namely *global boundaries*. Although GPR 6 is essentially prepared to derive the global organization [5], it is also used to refine the local structure in the practical use of GTTM [5),6)].

The algorithm shown in **Fig. 4** is used for the implementation of GTTM. This top-down algorithm almost conforms to that used in Ref. 6). The notation of the rules adapted to dance motion analysis is distinguished from that of the original ones by adding an apostrophe to the rule number. In this algorithm, a motion-data stream of a dance is first resolved into events. Local boundaries are next selected from the transitions between the events by applying GPRs 1', 2', 3' and 6'. Global boundaries are then selected from the local boundaries at every group on Level n by applying GPRs 4', 5' and 6'; the value of n is initially set to zero and increases by one when all the groups in Level n are finished. This procedure is repeated until all the local boundaries are selected.

No violation of GWFRs occurs in this algorithm. Moreover, it allows the situation that several groups are further subdivided into smaller-scale groups while the others are not, as shown in Fig. 4; this situation is not prohibited by GWFRs [5]. The detail of each procedure in this algorithm is shown below.

4.2 Extraction of Events from Motion Capture Data

First, transitions between events are detected by searching the local minimum/maximum points of the temporal variation of joint motion as already mentioned in Section 3.2.3. The local minimum/maximum points in time-series data are frequently searched by detecting the sign change of the differential coefficient of temporal variation. It is well known, however, that the above method is too sensitive to small fluctuations mainly caused by noises.

In order to guarantee the robustness against the above fluctuations, the technique of piecewise linear regression (PLR) [12] is used in this paper. PLR is a widely used method for dividing a time-series data stream into piecewise approximation functions each of which is a linear regression model. The influence of the fluctuations is absorbed in the regression process.

Some adequate criterion for the decision of the number of division points is needed for the implementation of the PLR algorithm. In the present case, the local minimum/maximum points should be detected as the division points. When this condition is satisfied, the difference between adjacent slopes which are derived at adjacent division points is maximized as shown in **Fig. 5**. The above situation is realized by maximizing the criterion given as the following formula:

$$V = \frac{1}{k_d - 1} \sum_{k=1}^{k_d-1} \sum_{m=1}^{42} |\beta_m(k+1) - \beta_m(k)| \tag{1}$$

where $\beta_m(k)$ is the slope for the mth variable at the kth division point, derived as shown in Fig. 5, and k_d is the number of division points including the first frame of the whole motion, respectively. This formula gives the sum of the differences between adjacent slopes for all the variables, averaged for all the division points.

An example of the application of PLR is shown in **Fig. 6**; the "walk" data

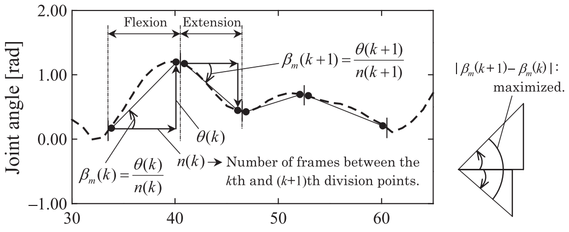


Fig. 5 Detection of transitions.

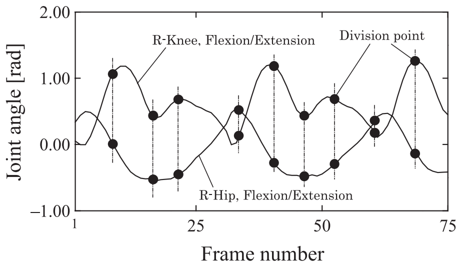


Fig. 6 An Example of detection of transitions (Sample motion: a walk ("02_02" in http://mocap.cs.cmu.edu)).

stream is used. Each of the division points which correspond to the transitions almost agrees with at least one of the local-minimum/maximum points.

4.3 Extraction of Local Boundaries

Next, local boundaries are selected from the transitions. GPRs 1', 2', 3' and 6' are implemented as follows.

(1) GPRs 2' and 3' are applied to events.

(2) If a conflict among GPRs 1', 2' and 3' occurs, the procedure for the solution of conflict is applied.

(3) The existence of parallel dance-motions is investigated.

(4) If parallelism exists, GPR 6' is applied.

The concrete procedure for each of the above processes is shown below.

4.3.1 Implementation of GPRs 2' and 3'

The degree of stillness of the body is evaluated in GPR 2'. It can be assessed by observing the speed of all of the joint motions. The following evaluation index is introduced to assess the sum of the motion speed of all the joints:

$$f_{\mathrm{GPR2'}}(e_i, e_{i+1}) = \omega_{\max} - \omega_{\min}(e_i, e_{i+1})$$

$$\omega_{\max} = \max_{1<i<i_e-1} \omega_{\min}(e_i, e_{i+1}) \tag{2}$$

$$\omega_{\min}(e_i, e_{i+1}) = \min_{k_c(i)<k<k_c(i+1)} \sum_{n=1}^{14} |\boldsymbol{\omega}_n(k)|$$

where e_i is the ith event, $\boldsymbol{\omega}_n(k)$ is the angular velocity of the nth joint at the kth frame, $k_c(i)$ is the center frame of the ith event and i_e is the number of events included in the whole motion, respectively. The value of this index becomes high when the minimum motion speed around the transition becomes small.

As for GPR 3', the following four items are evaluated: the direction of motion, the amount of motion, the speed of motion and the duration of motion. In order to evaluate these items, we use the respective indexes shown below.

$$f_{\mathrm{GPR3'a}}(e_i, e_{i+1}) = \sum_{n=1}^{14} \cos^{-1}\{\boldsymbol{m}_n(i) \cdot \boldsymbol{m}_n(i+1)\} \tag{3}$$

$$f_{\mathrm{GPR3'b}}(e_i, e_{i+1}) = \left| \sum_{n=1}^{14} \theta_n(i+1) - \sum_{n=1}^{14} \theta_n(i) \right| \tag{4}$$

$$f_{\mathrm{GPR3'c}}(e_i, e_{i+1}) = \sum_{n=1}^{14} \big||\boldsymbol{\omega}_n(k_{\max}(i+1))| - |\boldsymbol{\omega}_n(k_{\max}(i))|\big|$$

$$k_{\max}(i) = \arg\max_k \sum_{n=1}^{14} |\boldsymbol{\omega}_n(k)|, \quad k_s(i) \le k \le k_e(i) \tag{5}$$

$$f_{\mathrm{GPR3'd}}(e_i, e_{i+1}) = |k_n(i+1) - k_n(i)| \tag{6}$$

where $\boldsymbol{m}_n(i)$ is the unit vector along the axis of the rotation of the nth joint between the start and end of the ith event, $\theta_n(i)$ is the total rotation angle of the nth joint in the ith event, $k_s(i)$ and $k_e(i)$ are the start and end frames of the ith event and $k_n(i)$ is the number of frames in the ith event, respectively. Since appreciators watch the motions of the whole body, all the degrees of change evaluated in the above indexes are derived by summing the motions of all the joints. All the above indexes have the common tendency that a high degree of

Table 4 GPRs 2' and 3'.

GPR 2'	(Proximity): Consider a sequence of four events $e_1e_2e_3e_4$. All else being equal, the transition e_2-e_3 may be recognized as a group boundary if $f_{\mathrm{GPR2'}}(e_2,e_3)$ is greater than both $f_{\mathrm{GPR2'}}(e_1,e_2)$ and $f_{\mathrm{GPR2'}}(e_3,e_4)$.
GPR 3'	(Change): Consider a sequence of four events $e_1e_2e_3e_4$. All else being equal, the transition e_2-e_3 may be recognized as a group boundary if **a.** (Direction of motion) $f_{\mathrm{GPR3'a}}(e_2,e_3)$ is greater than both $f_{\mathrm{GPR3'a}}(e_1,e_2)$ and $f_{\mathrm{GPR3'a}}(e_3,e_4)$), or if **b.** (Amount of motion) $f_{\mathrm{GPR3'b}}(e_2,e_3)$ is greater than both $f_{\mathrm{GPR3'b}}(e_1,e_2)$ and $f_{\mathrm{GPR3'b}}(e_3,e_4)$, or if **c.** (Speed of motion) $f_{\mathrm{GPR3'c}}(e_2,e_3)$ is greater than both $f_{\mathrm{GPR3'c}}(e_1,e_2)$ and $f_{\mathrm{GPR3'c}}(e_3,e_4)$, or if **d.** (Duration of motion) $f_{\mathrm{GPR3'd}}(e_2,e_3)$ is greater than both $f_{\mathrm{GPR3'd}}(e_1,e_2)$ and $f_{\mathrm{GPR3'd}}(e_3,e_4)$.

change around the transition gives a large value.

As a consequence, GPRs 2' and 3' are summarized as **Table 4**. The transitions satisfying at least one of the criteria in GPRs 2' or 3' are selected as candidates for local boundaries.

4.3.2 Implementation of the Procedure for the Solution of Conflict among GPRs 1', 2' and 3'

The conflict among GPRs 1', 2' and 3' occurs in the case that two candidates selected by GPRs 2' and 3' are located on adjoining transitions, because of the violation of GPR 1' requiring the avoidance of a group containing a single event. Since GPR 1' has the strong priority[5)], the winner-candidate should be selected from the candidates in conflict on the basis of the effect of GPRs 2' and 3'.

The procedure for the solution of the above conflict is shown in **Fig. 7**. In the first half of the procedure ((a)−(c)), quantified scores of the individual rules are given to all candidates. The candidates in conflict are compared in the second half of the procedure ((d)−(e)) using the information of both the obtained scores and the relationship of priority of the rules. The detail is shown below.

First, the values of the evaluation indexes are calculated for every transition by using Eqs. (2) to (6) (Procedure (a) in Fig. 7). Next, for every candidate satisfying at least one of the criteria in GPRs 2' and 3', the degree of saliency for each of the rules contributing to the satisfaction of the criteria is calculated by the following formula (Procedure (b)):

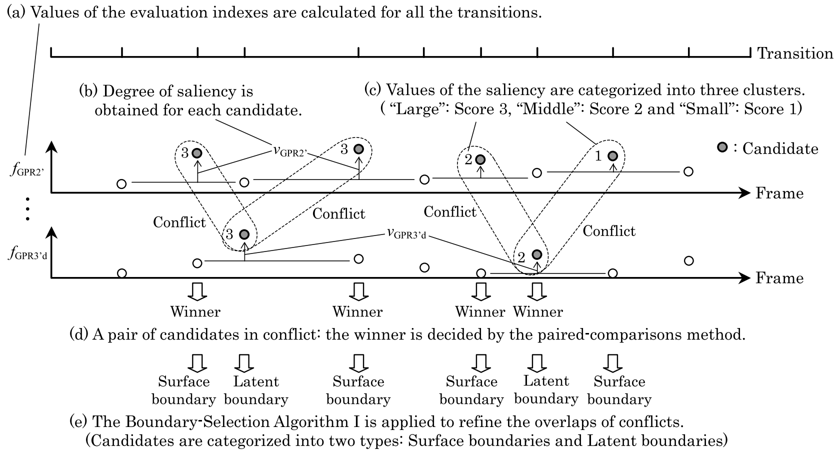


Fig. 7 The procedure for the solution of conflict among GPRs 1', 2' and 3'.

$$v_{\mathrm{GPR}x}(e_i,e_{i+1}) = \sqrt{\Delta_1\Delta_2}$$
$$\Delta_1 = f_{\mathrm{GPR}x}(e_i,e_{i+1}) - f_{\mathrm{GPR}x}(e_{i-1},e_i) \quad (7)$$
$$\Delta_2 = f_{\mathrm{GPR}x}(e_i,e_{i+1}) - f_{\mathrm{GPR}x}(e_{i+1},e_{i+2})$$

where x =2', 3'a, 3'b, 3'c, 3'd. The obtained values are categorized at every rule into "Large", "Middle" and "Small" by k-means cluster analysis[13) ⋆1]. Scores corresponding to the above categories are given to the candidates: 3 for "Large", 2 for "Middle" and 1 for "Small" (Procedure (c)). As a result, all the candidates in conflict obtain the scores of the individual rules on the basis of the relative intensity of the effect of the respective rules.

Then, the candidates in conflict are compared in order to select the winner-candidates. In Ref. 6), the comparison was performed using the values of the weighted mean of the scores; the weight representing the degree of priority was given to each of the rules in this procedure. It was also pointed out that the optimal values of the weights assigned to the respective rules depended on the

⋆1 In the case that at least one of the distances between the cluster-centroids are shorter than the quarter of the mean value of the evaluation index for all the transitions, the cluster analysis in which the number of clusters is reduced by one is performed again.

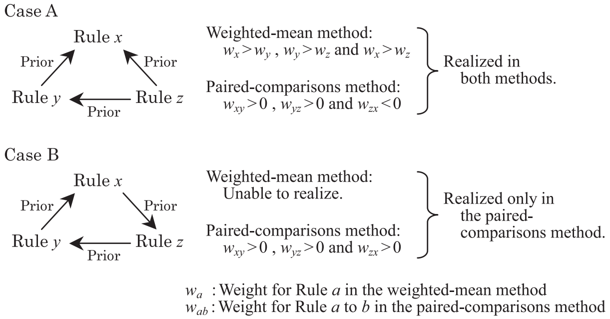


Fig. 8 The advantage of the paired-comparisons method.

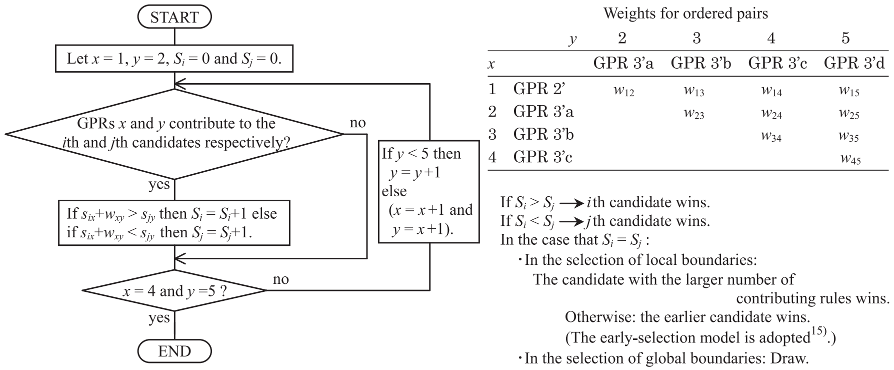


Fig. 9 The procedure for the decision of the winner in a pair of candidates.

selection of a piece. This fact implies that the robustness of this procedure against the diversity of objects may be insufficient.

Considering the above circumstance, we adopt the paired-comparisons method [14] for the comparison of the candidates in conflict (Procedure (d)) ⋆1. This technique has the style that the weight is assigned to each of the ordered pairs of the rules. It gives the following advantage.

The configuration of priority of the rules in Case A of **Fig. 8** can be realized in both the weighted-mean method and the paired-comparisons method. On the other hand, Case B, in which the configuration of priority is given as the rotation style, cannot be realized in the weighted-mean method, while it is possible to be processed by the paired-comparisons method. This leads to the fact that the region of the configuration of priority covered by the paired-comparisons method is wider than that covered by the weighted-mean method.

The paired-comparisons method is performed as follows. Let w_{xy} be the weight assigned to the ordered pair of the rules GPRs x and y. The value of w_{xy} represents the priority of GPR x relative to GPR y. The 7-point scoring system [14] is used in this procedure; w_{xy} is given as an integer satisfying $-3 \leq w_{xy} \leq 3$. Let s_{ix} be the score of GPR x contributing to the ith candidate. For the pair of the ith and jth candidates which are in conflict, the winner-candidate is decided by the procedure shown in **Fig. 9**. In this procedure, each of all the weighted scores for the ith candidate is compared with each of all the scores for the jth one; each candidate obtains the number of times the weighted (or raw) score for it is larger than that for the opponent. The candidate with the larger number wins.

Finally, the overlap of multiple conflicts are refined by the procedure called the Boundary-Selection Algorithm I shown in **Fig. 10** (Procedure (e)). This algorithm categorizes the winner-candidates into two types; one is that of the surface boundaries and the other that of the latent boundaries. If there is no parallelism in the motions of a dance, the surface boundaries are fixed as the local boundaries at the same time the latent ones are deleted.

In this algorithm, the winner-candidates each of which has defeated both the right and left sides are categorized as surface boundaries. The winner-candidates each of which adjoins to the surface boundary are downgraded to latent boundaries at the same time as those adjoining to the latent boundaries become surface boundaries.

4.3.3 Investigation of the Existence of Parallel Dance Motions

In the case that the parallelism of motions in a dance exists, the location of boundaries are refined by applying GPR 6'. The procedure for the investigation of the existence of parallelism is shown below.

A pair of motions parallel to each other can be regarded as two appearances of a common motion. It can be extracted by obtaining the following binary

⋆1 In the case that a more effective comparison method is found, it can be introduced instead of Procedure (d) and/or (e).

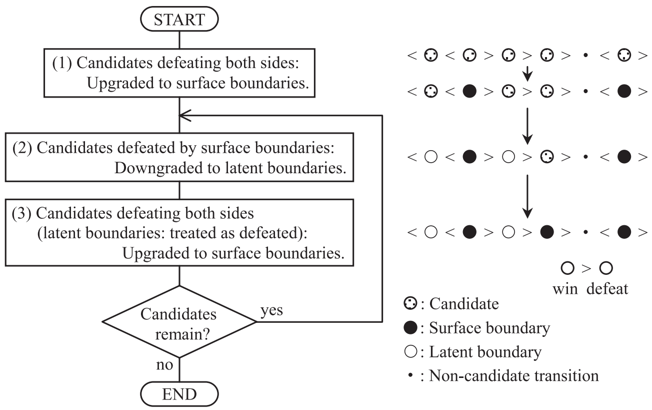


Fig. 10 Boundary-Selection Algorithm I: procedure for the selection of boundaries.

time-series autocorrelation matrix [16]:

$$\boldsymbol{D} = \begin{bmatrix} h(\boldsymbol{u}(1), \boldsymbol{u}(1)) & \cdots & h(\boldsymbol{u}(1), \boldsymbol{u}(N)) \\ \vdots & \ddots & \vdots \\ h(\boldsymbol{u}(N), \boldsymbol{u}(1)) & \cdots & h(\boldsymbol{u}(N), \boldsymbol{u}({}^m N)) \end{bmatrix} \quad (8)$$

$$h(\boldsymbol{u}(i), \boldsymbol{u}(j)) = \begin{cases} 1 & \sqrt{\displaystyle\sum_{k=1}^{42} \{u_k(i) - u_k(j)\}^2} < t_0 \\ 0 & \text{otherwise} \end{cases}$$

where N is the number of frames and t_0 is the adjustable threshold value, respectively. When $h(\boldsymbol{u}(i), \boldsymbol{u}(j)) = 1$, the posture of a dancer at the ith frame is recognized as identical to that at the jth frame. In the case that the postures at the i_{1s}th, $(i_{1s}+1)$th, $\cdots$, $(i_{1e}-1)$th and i_{1e}th frames are identical to those at the i_{2s}th, $(i_{2s}+1)$th, $\cdots$, $(i_{2e}-1)$th and i_{2e}th frames respectively ($i_{1e} < i_{2s}$), the motion in the interval $[i_{1s}, i_{1e}]$ is recognized as parallel to that in $[i_{2s}, i_{2e}]$, in other words a common motion which first appears in $[i_{1s}, i_{1e}]$ and then in $[i_{2s}, i_{2e}]$ is extracted.

The matrix for the "walk" data stream is shown in **Fig. 11**. Since the matrix is a symmetric one, only the analysis for the part enclosed by the triangle PQR is needed. The white parts represent the regions where $h(\boldsymbol{u}(i), \boldsymbol{u}(j)) = 1$, namely

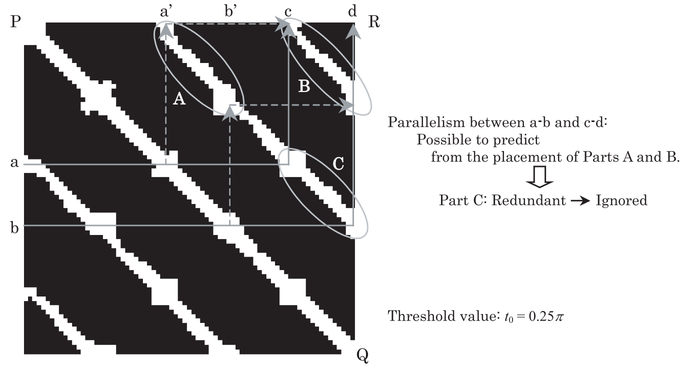


Fig. 11 Extraction of parallel motions by the time-series autocorrelation matrix for a "walk".

common motions. The upper-left apex and lower-right endpoint of a white part represent the startpoint and endpoint of a common motion, respectively. The row numbers in the white part correspond to the frame numbers for the first appearance of the common motion, namely $[i_{1s}, i_{1e}]$, while the column numbers indicate those for the second appearance, namely $[i_{2s}, i_{2e}]$.

In Fig. 11, the parallelism between the motion in $[\mathrm{a}, \mathrm{b}]$ and that in $[\mathrm{c}, \mathrm{d}]$ is recognized by the existence of Part C; this parallelism can also be recognized by the placement of Parts A and B. This fact implies that extracting all these parts causes a redundancy. In order to eliminate the redundancy, the upper parts in the triangle PQR (Parts A and B in this case) are preferentially extracted, at the same time the lower parts (Part C in this case) are ignored.

In order to fit the startpoint and endpoint of each appearance of a common motion with those of events, the procedure shown in **Fig. 12** is performed. In this procedure, the number of events in the first appearance is fitted to that of the second one, keeping the DTW distance [8] between the two appearances short.

4.3.4 Implementation of GPR 6'

In this paper, GPR 6' is implemented by keeping the inside structures of two appearances of a common motion, namely the location of the local boundaries in both the appearances, the same. The concrete procedure is shown below.

First, both the surface and latent boundaries existing in the appearances of a common motion are returned to candidates. On the other hand, the information

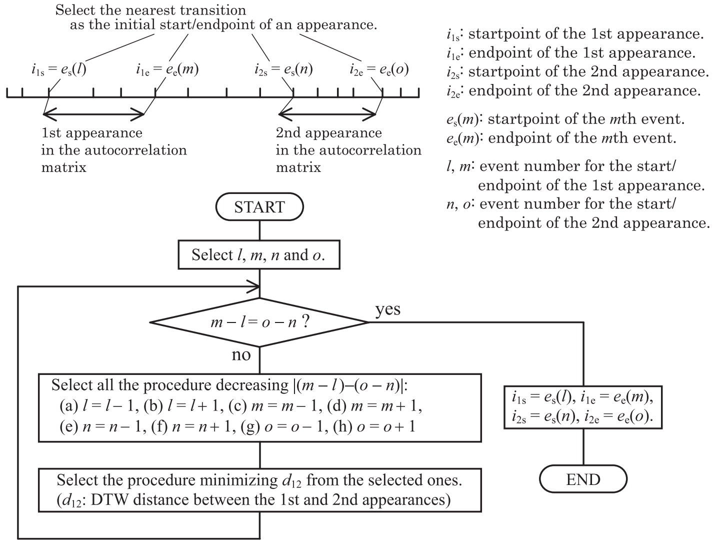


Fig. 12 The procedure to fit the appearances of a common motion with the division of events.

of wins and defeats is retained. In the case that the beginning point or completion of the whole motion of a dance is included in an appearance of a common motion, the procedure shown in **Fig. 13** is applied.

In this procedure, both the beginning point and the completion are essentially regarded as the local boundaries. Therefore, for another appearance of the same common motion, the transition whose location corresponds to that of the beginning point/completion in the former appearance is regarded as a candidate correspoinding to a surface boundary. Hence, the information of wins and defeats is updated such that this candidate defeats both the right and left sides.

Next, the procedure shown in **Fig. 14** is applied to all the pairs of two appearances of common motions. In this procedure, a candidate located in an appearance of a common motion is deleted when a transition on the corresponding position in another appearance has no candidate. This process is introduced to avoid the increase of candidates which may cause the violation of GPR 1'.

Then, the Boundary-Selection Algorithm I is again applied to clarify the new

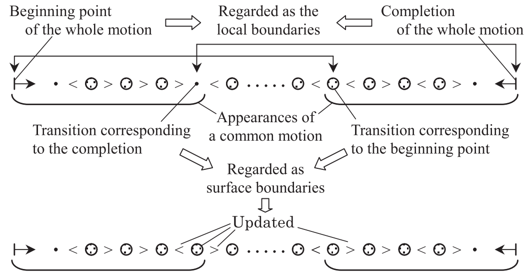


Fig. 13 Procedure for the adjustment of candidates for boundaries in parallel motions (in the case that the beginning or the completion are included in the parallel motions).

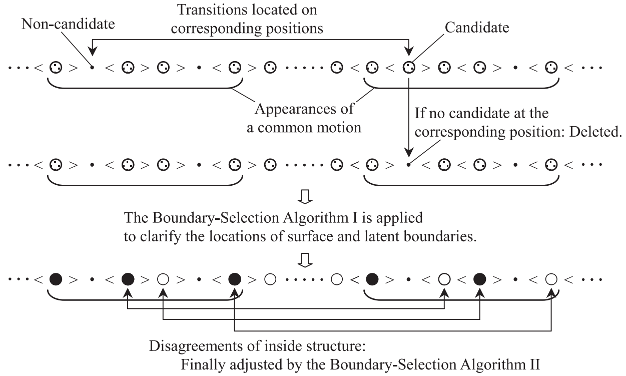


Fig. 14 Procedure for the refinement of boundaries in parallel motions.

location of the surface and latent boundaries. The Boundary-Selection Algorithm II shown in **Fig. 15** is finally implemented to resolve the remaining disagreement of the inside structures in the appearances of the same common motion. This algorithm is constructed to minimize the number of the changes from the surface boundary to the latent one. The selected surface boundaries are fixed as the local boundaries, at the same time latent ones are deleted.

4.4 Extraction of Global Boundaries

Next, the structure of global boundaries is obtained. GPRs 4', 5' and 6' are used in this process as described in Section 4.1. We implement these rules by using the following procedure.

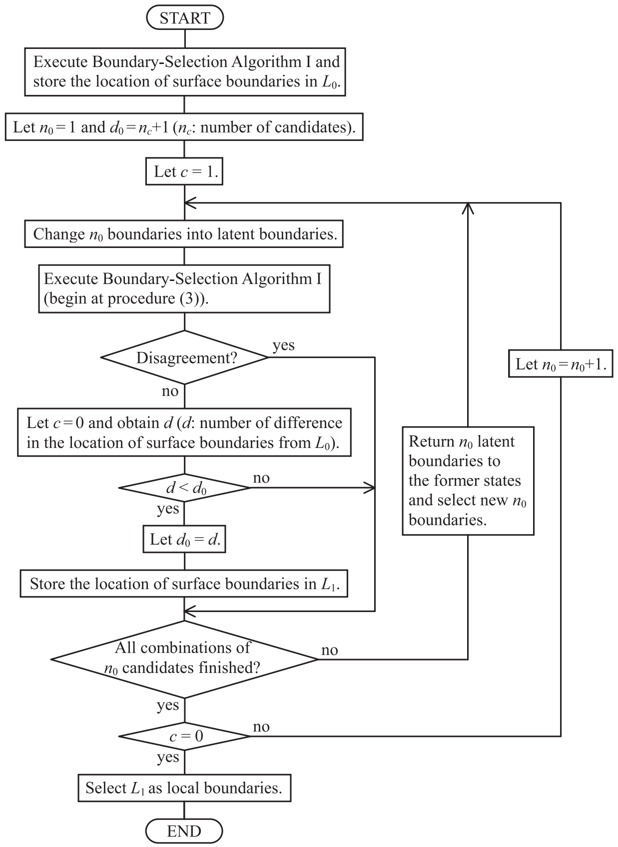


Fig. 15 Boundary-Selection Algorithm II: procedure for the selection of local boundaries.

(1) The intensity of GPRs 2' and 3' for each of the local boundaries is quantified in order to implement GPR 4'.

(2) The values of weight for each local boundary is obtained from the weight functions prepared to implement GPR 5'.

(3) Local boundaries having the largest strength of the weighted intensity at every group are selected as global boundaries.

(4) If the selected boundary is included in an appearance of a common motion, GPR 6' is applied.

These procedures are repeated at every level in a top-down way until all the local boundaries are selected. The use of the weight functions is the means introduced in Ref. 6). The concrete procedures for the above processes are shown below.

4.4.1 Implementation of GPR 4'

In order to quantify the intensity of the effect of GPRs 2' and 3' which is used to implement GPR 4', the procedure for the comparison of two boundaries shown in Fig. 9 is again utilized; the winning percentage of a local boundary given by the comparisons against all the other boundaries included in the same group is adopted as the intensity of the effect of GPRs 2' and 3' on it.

4.4.2 Implementation of GPR 5'

As for GPR 5', the weight function similar to that used in Ref. 6), namely the Gaussian-distribution-type function, is prepared to evaluate the strength of each local boundary for binary grouping. The weight function used is as follows:

$$w_b(k) = \exp\{-(k - k_b)^2/2(\sigma k_g)^2\} \quad (9)$$

where k is the position of a local boundary represented by the frame number, k_b is the position of the center of a group, σ is the adjustable coefficient which governs the broadness of distribution and k_g is the number of frames included in the group, respectively. This function provides the weight distribution which has the maximum value set to unity at the center of a group.

In order to make it possible to allow not only binary grouping but also ternary one as mentioned in Section 3.2.4, we prepare the following weight functions which are used together with Eq. (9) as described later:

$$w_{t_1}(k) = \exp\{-(k - k_{t_1})^2/2(\sigma k_g)^2\} \quad (10)$$

$$w_{t_2}(k) = \exp\{-(k - k_{t_2})^2/2(\sigma k_g)^2\} \quad (11)$$

where k_{t_1} and k_{t_2} are the positions of the first and second partition points for the trisection of a group, respectively.

On the other hand, there is a possibility for the recommendation of binary grouping at smaller-scale levels where dancers will naturally be influenced by the music for which the original GPR 5 is prepared. Hence, we introduce the following weight function for the adjustment of the ternary weight functions:

$$w_L(l) = w_L(l-1)\exp\{-\alpha(l-1)\}, \quad w_L(0) = w_0, \quad 0 \le w_0 \le 1 \quad (12)$$

where l is the level number, α is the adjustable coefficient and w_0 is the adjustable initial value of the weight, respectively. The value of the weight gradually decreases with the progress of top-down parsing.

Finally, we obtain the strengths of a local boundary derived from the effects of both GPRs 4' and 5' as follows:

$$c_{21}(k) = \{1 - w_L(l)\} w_b(k) f_{\mathrm{GPR4'}}(k) \quad (13)$$

$$c_{31}(k) = w_L(l) w_{t_1}(k) f_{\mathrm{GPR4'}}(k) \quad (14)$$

$$c_{32}(k) = w_L(l) w_{t_2}(k) f_{\mathrm{GPR4'}}(k) \quad (15)$$

where $f_{\mathrm{GPR4'}}(k)$ is the intensity of GPR 4' on a local boundary located on the kth frame.

4.4.3 Selection of Global Boundaries

The above strengths are used in the Boundary-Selection Algorithm III shown in **Fig. 16**. This algorithm is prepared to select global boundaries from local ones in a top-down way. On each level, the global boundary or boundaries are selected at every group including at least one unselected local boundary. If the selected boundary belongs to an appearance of a common motion, the local boundary located on the corresponding position in another appearance of the same common motion is also selected. As a result, GPR 6' is satisfied.

An example of the extraction of local and global boundaries from the "walk" data stream has already been shown in Fig. 3. The values of the weights for GPRs 2' and 3' and those of the adjustable coefficients for GPR 4', 5' and 6' are identical to those shown in Section 5.2. It is confirmed that the local boundaries on Level 2 correspond to the moments of the extension of a leg, leading to the fact that each of the local groups represents each individual step in a walk. Each of the groups on Level 1, on the other hand, corresponds to a cycle of walking consisting of a pair of steps of both the legs.

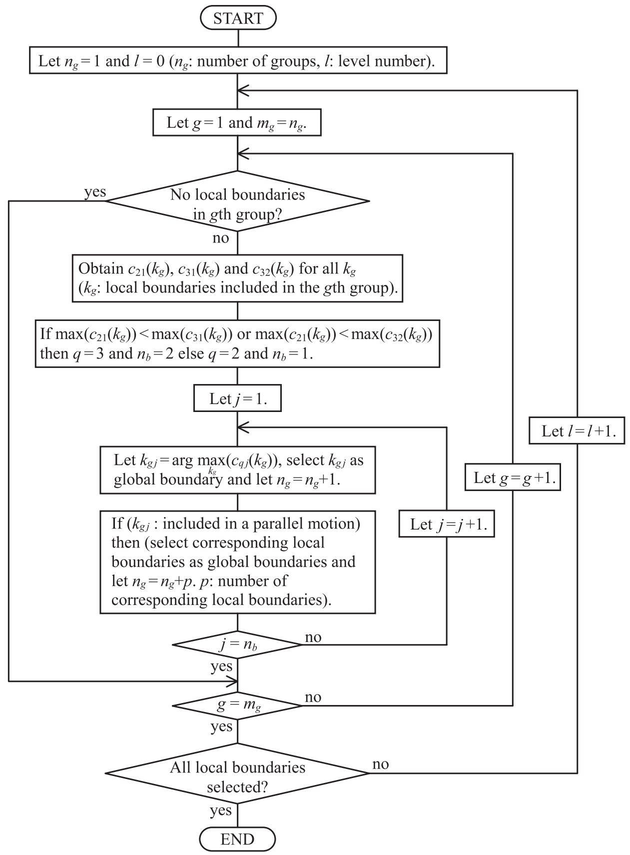


Fig. 16 Boundary-Selection Algorithm III: procedure for the selection of global boundaries.

5. Results of Analysis

5.1 Motion Data

In this section, the results given by implementing GTTM for dance motion analysis using the method shown in Section 4 are presented.

The motion-data streams used in this paper are shown in **Table 5**. These were acquired by motion capture systems with magnetic sensors. The dance numbers shown in Table 5 belong to the category of Japanese folk dances.

5.2 The Adjustment of Parameters

The parameters used in the analysis are set to the values shown in **Table 6**; each of w_{xy}'s is the weight assigned to the ordered pair of the rules used in the procedure shown in Fig. 9, t_0 is the threshold value for the extraction of common motions shown in Fig. 11 and σ, w_0 and α are the adjustable coefficients used in Eqs. (9) to (12), respectively. These values are determined through the preliminary adjustment process in which the data in Table 5 with the "walk" data

Table 5 Motion data (Japanese folk dances, frame interval: 33.3ms).

Dance		Dancer	Data	Frame	System
Nishimonai Bon Odori (*Ondo*, 1st verse)		A (female)	A1	666	A
		B (female)	B1	676	A
Tsugaru Jinku		C (female)	C1	195	A
		C (female)	C2	204	A
Kokiriko	*Sasara Odori*	D (male)	D1	420	B
	Teodori	E (female)	E1	449	B
Kemanai Bon Odori (*Jinku Odori*)		F (female)	F1	293	A
		F (female)	F2	321	A
Hitoichi Bon Odori (*Sankatsu Odori*)		G (female)	G1	356	B
		G (female)	G2	362	B
Akita Obako (1st verse)		H (female)	H1	1,150	C
		I (female)	I1	1,143	A

System A: MotionStar Wireless™ with LIBERTY™ (Polhemus) ×2
System B: MotionStar Wireless™ (Ascension Technology Corporation)
System C: STAR*TRAK™ (Polhemus)

Table 6 Values of the parameters used in the grouping structure analysis.

Rule	GPR 3'a	GPR 3'b	GPR 3'c	GPR 3'd
GPR 2'	2	1	3	2
GPR 3'a		0	3	1
GPR 3'b	(w_{xy})		−2	−2
GPR 3'c				3

$t_0 = 0.25\pi$, $\sigma = 0.05$, $w_0 = 0.4$, $\alpha = 0.8$

are analyzed multiple times in a trial-and-error way. It should be pointed out that the robustness against the variation of data streams is regarded as important in the adjustment process. An additional remark which should be made here is that the obtained priority relationship for several combinations of three rules shows the property which is unable to be realized in the weighted-mean method: e.g., the combination of GPRs 3'a, 3'b and 3'c with the relationship of 3'a=3'b, 3'b<3'c and 3'a>3c', etc.

5.3 Presentation of the Analysis Results

The grouping structures given by applying the present method to the dances *Nishimonai Bon Odori* and *Tsugaru Jinku* are first shown in Sections 5.4 and 5.5, respectively. As for these dances, the following cases are also analyzed in order to confirm the effect of the procedures newly introduced in this paper:

(1) The case in which the weighted-mean method is used instead of the paired-comparisons method.

(2) The case in which the ternary-grouping procedure is not used.

The evaluation of the grouping structures for all the dances shown in Table 5 is then presented in Section 5.6.

5.4 Case I: Analysis of *Nishimonai Bon Odori*

The folk performing art *Nishimonai Bon Odori* belongs to the group of folk performing arts in Akita Prefecture, Japan. Here, the dance *Ondo* which is one of the dances in *Nishimonai Bon Odori* is analyzed.

The choreography of *Ondo* is shown in **Fig. 17** with the structure of rhythm in the musical accompaniment. The division of the whole motion into the unit gestures is quoted from Ref. 17). This division is henceforth regarded as the true unit-segmentation of this dance [⋆1]. The structure of the musical accompaniment is quoted from Ref. 18) in order to examine the validity of the global grouping structure derived by the present method, on the basis of the typical tendency that dance motions have a relatively high correspondence with music [3)]. This opinion is supported by the fact that the sequence of the 12 unit gestures is synchronized with that of measures in the musical form of the accompaniment. Therefore, the ternary and sexenery groupings for Levels 1 and 2 in the hierarchical structure of the whole motion are also inferred from the musical accompaniment having these grouping forms on the highest two levels.

The motion-capture data stream whose data number is given as A1 in Table 5 is used in the analysis. The obtained hierarchical segmentation of this motion-data stream is shown in **Fig. 18**. In Fig. 18 (a), the analysis method presented in Section 4 is directly used. On the other hand, the ternary-grouping procedure, namely Eqs. (14) and (15), is not used in the case shown in Fig. 18 (b). The

⋆1 In this paper, the unit gestures of dances are quoted from Refs. 17), 18), 21) and 22). The reliability and validity of them is regarded as sufficient since they are provided by the researchers who intensively address to the studies of Japanese folk performing arts. On the other hand, it should also be pointed out that most of Japanese folk dances have no strictly-stylized structure in their motions. This fact leads to the possibility of dividing into different unit gestures. Considering these facts, the authors adopt each of the sets of the unit gestures quoted from the references as a typical benchmark used for the evaluation of the present method.

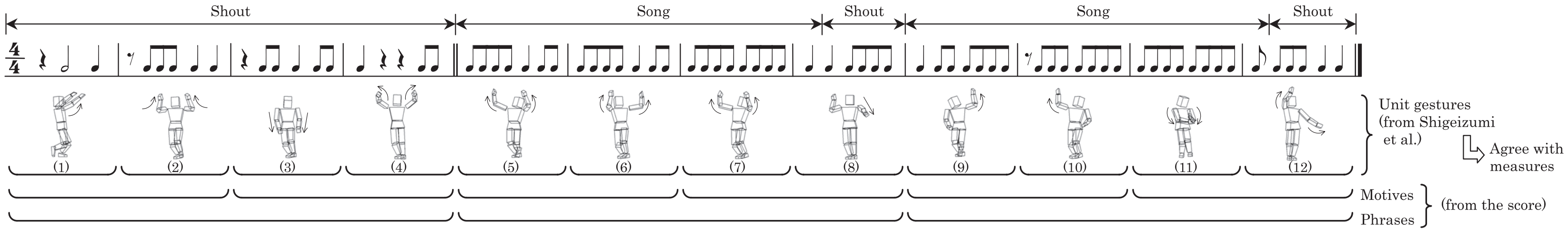
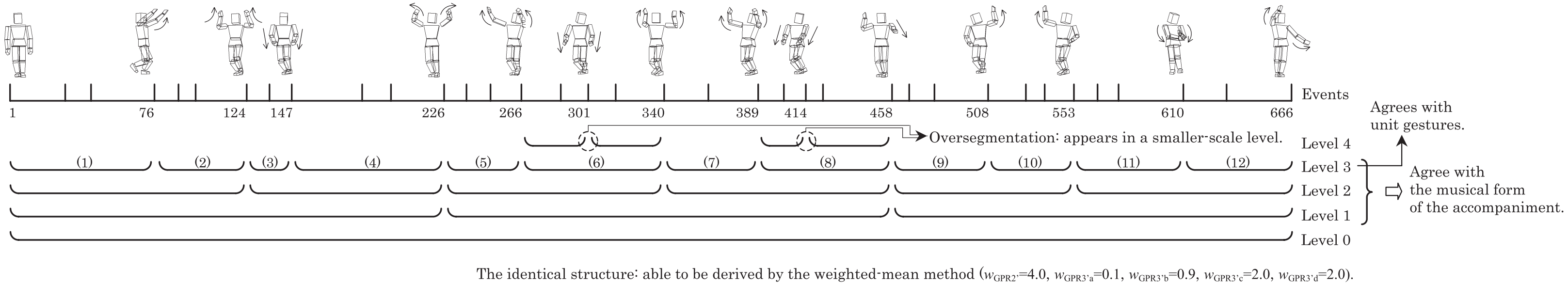
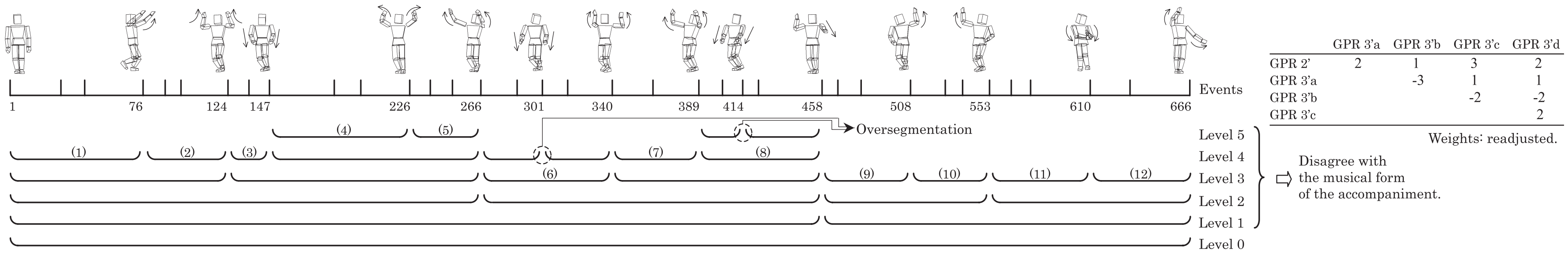


Fig. 17 Choreography and the structure of rhythm of the musical accompaniment for *Ondo* of *Nishimonai Bon Odori*.



(a)



(b)

Fig. 18 Grouping structure of *Ondo* of *Nishimonai Bon Odori* derived by the present method; (a) with the ternary-grouping procedure, (b) without the ternary-grouping procedure.

weights assigned to the ordered pair of the rules are readjusted in (b) such that performing the binary grouping twice on large-scale levels makes the grouping structure as similar as possible to the ternary form shown in Fig. 17.

It is seen in Fig. 18 (a) that the structure of Levels 1 and 2 agrees with that of the highest two levels of the musical form of the accompaniment, while Level 3 has the same structure to the true segmentation. As mentioned in Fig. 18 (a),

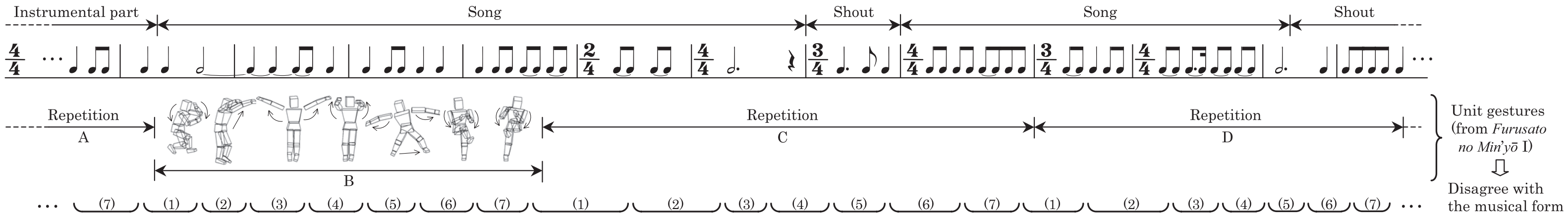


Fig. 19 Choreography and the structure of rhythm of the musical accompaniment for *Tsugaru Jinku*.

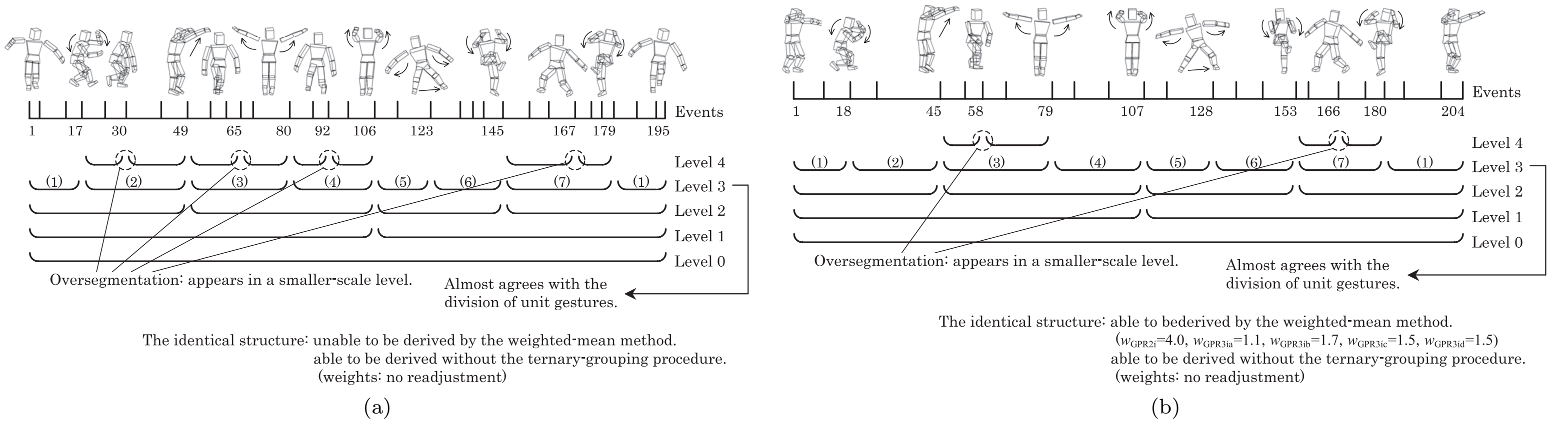


Fig. 20 Grouping structure of *Tsugaru Jinku* derived by the present method; (a) performance in Part B, (b) performance in Part D.

an identical hierarchical structure is also derived by the analysis in which the weighted-mean method is used instead of the paired-comparisons method.

On the other hand, the construction of a ternary form is insufficient in (b), especially in the former part of the whole choreography. This fact indicates that the ternary-grouping procedure newly prepared in this paper is really effective.

An additional remark which should be made here is that the boundaries causing the oversegmentation appear only at the smallest-scale level, namely Level 4 in Fig. 18 (a), where no regular boundary exists. This fact leads to the possibility that the present method may provides the information which can be utilized to estimate the appearance of the oversegmentation. The detail of this possibility is discussed later.

5.5 Case II: Analysis of *Tsugaru Jinku*

In this subsection, the dance *Tsugaru Jinku* which belongs to the group of folk dances in Aomori Prefecture, Japan, is analyzed.

The choreography of *Tsugaru Jinku* is shown in **Fig. 19**. The structure of rhythm in the musical accompaniment is also shown. The division of the whole motion into the unit gestures is quoted from Ref. 18). This division is regarded as the true unit-segmentation of *Tsugaru Jinku* in this paper. The structure of the musical accompaniment shown in Fig. 19 is the one derived by fitting the score in Ref. 18) into the music actually used in the motion-capture experiment.

It is confirmed from Fig. 19 that the sequence of the unit gestures which is repeated any number of times does not synchronize with the musical form of the accompaniment, in contrast to the case of *Nishimonai Bon Odori*.

Considering the above fact, we analyze two motion-data streams each of which corresponds to the different part of the accompaniment. The results are shown in **Fig. 20**. The motion-data stream shown in Fig. 20 (a), whose data number is given as C1 in Table 5, corresponds to Part B in Fig. 19 which appears earlier than Part D corresponding to the data stream shown in Fig. 20 (b) whose data number is given as C2 in Table 5. The beginning of Part B does not synchronize with that of a measure in the accompaniment, while that of Part D coincides.

It is indicated in these results that both the data streams provide the same global grouping structure at Levels 1, 2 and 3. It shows a simple binary fashion, providing the agreement between Level 3 and the true segmentation. This is suggestive of an advantage that the present method is useful for the extraction of the characteristics peculiar to the motion-side when there is little connection between motion and music.

On the other hand, as mentioned in Fig. 20 (a), the structure shown in this figure cannot be derived by the analysis in which the weighted-mean method is used instead of the paired-comparisons method, while that in (b) can be derived. The values of the weights for the weighted-mean method used in (b), which are readjusted to fit the structure with that of Fig. 19, differ from those used in the case of Fig. 18 (a). These facts imply that the paired-comparisons method really has the higher robustness than the weighted-mean method.

As for the effectiveness of the ternary-grouping procedure, the clear difference between the cases in which the procedure is used and those without it is not confirmed. In fact, this dance shows the binary structure even in the cases in which the ternary-grouping procedure is used. These facts suggest that the application of the ternary-grouping procedure does not give a harmful influence in the analysis of dances essentially having the binary structure.

It should also be noted here that the distinction of the oversegmentation from the regular boundaries also occurs in this case.

5.6 Evaluation of Grouping Structure

In this subsection, the results obtained for all the data shown in Table 5 are evaluated. The indexes *precision* and *recall* which are the primary metrics for information retrieval [19),20)] are used for the evaluation of oversegmentation and undetected true boundaries, respectively. The index F_{measure} which is the harmonic mean of the above two indexes [20)] is also used to evaluate the agreement of the extracted boundaries with the true segmentation.

The results given by the present method are compared with those given by two different analysis methods. One is the method proposed in Ref. 3), henceforth Method A, and another the one proposed in Ref. 2), henceforth Method B. These methods divide the whole motion of a dance into unit gestures on a single level.

The procedures and parameters used in the above methods are shown in **Table 7**. In Method A, the series of the boundary-candidates are refined using the information of beat tracking obtained from the sound data of the musical accompaniment. However, this procedure is omitted in this paper in order to compare with the other methods which do not use the sound data.

Table 7 Methods used for the comparison with the present method.

Method A [3)]
Evaluation indexes:
· Local minimums of hands-, feet-, and CM-speeds.
Logical operation for the selection of boundaries:
· (CM Result) AND (Result that more than 2 of Hands and Feet Results match).
Threshold values:
· Minimum- and maximum-speeds for hands (v_{Hmin}, v_{Hmax}).
· Minimum- and maximum-speeds for CM (v_{CMmin}, v_{CMmax}).
· Minimum-speed and motion-distance for feet (v_{Fmin}, d_{F}).
Method B [2)]
Evaluation indexes:
· Local minimums of joints- and end-effectors-speeds.
· Change of the movement direction of CM.
· Change of the volume of a convex hull enclosing the whole body.
Logical operation for the selection of boundaries:
· OR operation among all the candidates.
Threshold values:
· Minimum- and maximum-speeds (v_{min}, v_{max}).
· Cosine of the angle between the directions of successive CM-movements ($t_{\cos}$).
· Change of the volume of a convex hull (V_{CH}).
· Spatial distance between succesive candidates (d_{CM} represented by CM distance).
· Temporal distance between succesive candidates (d_{FN} represented by frame number).

CM: Center of mass of the whole body

The adjustment of the threshold values are needed to implement these methods. They are adjusted at every dance so that the obtained segmentation becomes as parallel as possible to the true segmentation.

The results of evaluation are shown in **Table 8**. It is seen that each result of the present method shown in Table 8 (a) gives the value of precision closer to unity at many levels, compared to the other methods shown in (b). This fact leads to the possibility that the present method has the ability to provide the information of grouping structure from which oversegmentation is almost excluded. It is also confirmed that the values of recall obtained from the present method are close to unity at relatively high levels in the grouping structure, leading to the ability to detect the true boundaries with little exception. In the results of the present method, there are seven cases in each of which a level giving F_{measure} unity exists, while only a few cases show the same situation in both of Methods A and B.

The above results suggest the possibility that the application of the present method may allow us to derive the segmentation of dance motion more precisely than the previous methods. It should also be pointed out that the parameters used for the implementation of the present method, shown in Table 6, are not changed throughout all the cases in Table 8 (a), while the threshold values in both Methods A and B are significantly varied as shown in Table 8 (b). This fact implies that the robustness of the present method against the variation of dance motion is quite stronger than the previous methods.

As for the cases in Table 8 (a), the following property is confirmed; when a level giving F_{measure} unity exists, the whole motion is exhaustively partitioned into groups on this level, at the same time several groups on this level necessarily have no subgroup on the lower levels. In other words, the boundaries causing the oversegmentation are located only on the levels lower than that giving the true segmentation, at the same time the regular boundaries are concentrated on a single level. Typical examples have already been shown in Sections 5.4 and 5.5. Such property is given by the feature peculiar to the present method that the information of the priority of each boundary is provided by deriving the hierarchical structure of the whole motion; this feature is not given by methods providing only a single-level segmentation, such as Methods A and B.

The utilization of the above property may provide a means which enables the distinction of oversegmentation from regular boundaries. In addition, the possibility of positively utilizing the information of oversegmentation for the comparison of dance performances is also suggested.

However, an issue of the confusion of oversegmentation with regular boundaries still exists in the present stage.; examples are shown in the cases of *Teodori* of *Kokiriko*, *Hitoichi Bon Odori* and *Akita Obako* in Table 8 (a). In order to establish a technique to accurately distinguish the oversegmentation from the regular boundaries, both the investigation of the cause of the above issue and the development of the method resolving it should be performed.

6. Conclusion

In this paper, the authors propose the adaptation of the rules used in the grouping structure analysis of the music theory GTTM to dance motion analysis, considering the commonness between musical idioms and that of dance. The success of the adaptation realizes the segmentation of dance motion in a hierarchical fashion. The analysis method obtained by the trial of the above adaptation consists of the following procedures.

A motion-capture data stream of a dance is first divided into events which are the counterparts of notes and rests in a score of a piece of music, by the technique of PLR. The hierarchical structure of groups each of which consists of a sequence of the events is then extracted by applying the grouping rules. Two procedures are used in the above process; one is for the extraction of local boundaries, and another for the selection of global boundaries. The latter procedure is implemented in a top-down way after the application of the former one.

The results obtained from the application of the above method to the analysis of motion-capture data streams indicate the following advantages:

(1) The structure of hierarchical segmentation is precisely extracted in response to the characteristic of an analyzed dance.

(2) The extraction of the hierarchical segmentation provides the possibility of the development of a technique distinguishing the oversegmentation from regular boundaries.

(3) The possibility of utilizing the information of hierarchical segmentation for the comparison of dance performances is suggested.

Table 8 Evaluation of grouping structure; (a) results of the present method, (b) results of Methods A and B.

(a)

Dance	Number of boundaries in true segmentation	Data	Level	Correct boundaries	Over-segmentation	Undetected	Precision	Recall	Fmeasure
Nishimonai Bon Odori (Ondo)	11 [17]	A1	1	2	0	9	1.000	0.182	0.308
			2	5	0	6	1.000	0.455	0.625
			3	11	0	0	1.000	1.000	1.000
			4	11	2	0	0.846	1.000	0.917
		B1	1	2	0	9	1.000	0.182	0.308
			2	5	0	6	1.000	0.455	0.625
			3	11	0	0	1.000	1.000	1.000
			4	11	2	0	0.846	1.000	0.917
Tsugaru Jinku	11 [18]	C1	1	1	0	6	1.000	0.143	0.250
			2	3	0	4	1.000	0.429	0.600
			3	7	0	0	1.000	1.000	1.000
			4	7	4	0	0.636	1.000	0.778
		C2	1	1	0	6	1.000	0.143	0.250
			2	3	0	4	1.000	0.429	0.600
			3	7	0	0	1.000	1.000	1.000
			4	7	2	0	0.778	1.000	0.875
Kokiriko Sasara Odori	7*	D1	1	1	0	6	1.000	0.143	0.250
			2	3	0	4	1.000	0.429	0.600
			3	7	0	0	1.000	1.000	1.000
			4	7	5	0	0.583	1.000	0.737
			5	7	6	0	0.538	1.000	0.700
Kokiriko Teodori	7 [21]	E1	1	1	0	6	1.000	0.143	0.250
			2	2	1	5	0.667	0.286	0.400
			3	5	2	2	0.714	0.714	0.714
			4	7	4	0	0.636	1.000	0.778
Kemanai Bon Odori (Jinku)	5 [18]	F1	1	2	0	3	1.000	0.400	0.571
			2	5	0	0	1.000	1.000	1.000
			3	5	2	0	0.714	1.000	0.833
		F2	1	2	0	3	1.000	0.400	0.571
			2	5	0	0	1.000	1.000	1.000
Hitoichi Bon Odori (Sankatsu Odori)	11 [18]	G1	1	2	0	9	1.000	0.182	0.308
			2	5	1	6	0.833	0.455	0.588
			3	10	1	1	0.909	0.909	0.909
		G2	1	0	1	11	0.000	0.000	0.000
			2	3	1	8	0.750	0.273	0.400
			3	8	1	3	0.889	0.727	0.800
			4	9	2	2	0.818	0.818	0.818
			5	11	2	0	0.846	1.000	0.917
Akita Obako	32 [22]	H1	1	1	0	31	1.000	0.031	0.061
			2	2	1	30	0.667	0.063	0.114
			3	5	2	27	0.714	0.156	0.256
			4	12	3	20	0.800	0.375	0.511
			5	25	4	7	0.862	0.781	0.820
			6	29	7	3	0.806	0.906	0.853
		I1	1	1	0	31	1.000	0.031	0.061
			2	4	0	28	1.000	0.125	0.222
			3	8	1	24	0.889	0.250	0.390
			4	18	1	14	0.947	0.563	0.706
			5	27	2	5	0.931	0.844	0.885
			6	28	3	4	0.903	0.875	0.889

(b)

Dance	Number of boundaries in true segmentation	Data	Method	Threshold values	Correct boundaries	Over-segmentation	Undetected	Precision	Recall	Fmeasure
Nishimonai Bon Odori (Ondo)	11 [17]	A1	A	$v_{\mathrm{Hmin}}=60,\ v_{\mathrm{CMmin}}=50,\ v_{\mathrm{Fmin}}=10$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	10	1	1	0.909	0.909	0.909
			B	$v_{\mathrm{min}}=70,\ v_{\mathrm{max}}=120,\ t_{\mathrm{cos}}=0.0$ $V_{\mathrm{CH}}=30,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=3.0$	10	3	1	0.769	0.909	0.833
		B1	A	$v_{\mathrm{Hmin}}=60,\ v_{\mathrm{CMmin}}=50,\ v_{\mathrm{Fmin}}=10$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	9	2	2	0.818	0.818	0.818
			B	$v_{\mathrm{min}}=70,\ v_{\mathrm{max}}=120,\ t_{\mathrm{cos}}=0.0$ $V_{\mathrm{CH}}=30,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=3.0$	9	3	2	0.750	0.818	0.783
Tsugaru Jinku	11 [18]	C1	A	$v_{\mathrm{Hmin}}=200,\ v_{\mathrm{CMmin}}=80,\ v_{\mathrm{Fmin}}=40$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	7	0	0	1.000	1.000	1.000
			B	$v_{\mathrm{min}}=220,\ v_{\mathrm{max}}==270,\ t_{\mathrm{cos}}=0.2$ $V_{\mathrm{CH}}=200,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=3.0$	7	0	0	1.000	1.000	1.000
		C2	A	$v_{\mathrm{Hmin}}=200,\ v_{\mathrm{CMmin}}=80,\ v_{\mathrm{Fmin}}=40$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	6	1	1	0.857	0.857	0.857
			B	$v_{\mathrm{min}}=220,\ v_{\mathrm{max}}==270,\ t_{\mathrm{cos}}=0.2$ $V_{\mathrm{CH}}=200,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=3.0$	6	2	1	0.750	0.857	0.800
Kokiriko Sasara Odori	7*	D1	A	$v_{\mathrm{Hmin}}=150,\ v_{\mathrm{CMmin}}=40,\ v_{\mathrm{Fmin}}=40$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=1.0,\ d_{\mathrm{F}}=40$	7	7	0	0.500	1.000	0.667
			B	$v_{\mathrm{min}}=100,\ v_{\mathrm{max}}=150,\ t_{\mathrm{cos}}=0.2$ $V_{\mathrm{CH}}=120,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=3.0$	7	4	0	0.636	1.000	0.778
Kokiriko Teodori	7 [21]	E1	A	$v_{\mathrm{Hmin}}=150,\ v_{\mathrm{CMmin}}=40,\ v_{\mathrm{Fmin}}=15$ $v_{\mathrm{Hmax}}=5,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	5	7	2	0.417	0.714	0.526
			B	$v_{\mathrm{min}}=155,\ v_{\mathrm{max}}=160,\ t_{\mathrm{cos}}=0.2$ $V_{\mathrm{CH}}=120,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=3.0$	6	3	1	0.667	0.857	0.750
Kemanai Bon Odori (Jinku)	5 [18]	F1	A	$v_{\mathrm{Hmin}}=60,\ v_{\mathrm{CMmin}}=25,\ v_{\mathrm{Fmin}}=15$ $v_{\mathrm{Hmax}}=10,\ v_{\mathrm{CMmax}}=1.0,\ d_{\mathrm{F}}=10$	3	3	2	0.500	0.600	0.545
			B	$v_{\mathrm{min}}=30,\ v_{\mathrm{max}}=100,\ t_{\mathrm{cos}}=0.0$ $V_{\mathrm{CH}}=150,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=1.0$	4	2	1	0.667	0.800	0.727
		F2	A	$v_{\mathrm{Hmin}}=60,\ v_{\mathrm{CMmin}}=25,\ v_{\mathrm{Fmin}}=15$ $v_{\mathrm{Hmax}}=10,\ v_{\mathrm{CMmax}}=1.0,\ d_{\mathrm{F}}=10$	3	2	2	0.600	0.600	0.600
			B	$v_{\mathrm{min}}=30,\ v_{\mathrm{max}}=100,\ t_{\mathrm{cos}}=0.0$ $V_{\mathrm{CH}}=150,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=1.0$	5	2	0	0.714	1.000	0.833
Hitoichi Bon Odori (Sankatsu Odori)	11 [18]	G1	A	$v_{\mathrm{Hmin}}=270,\ v_{\mathrm{CMmin}}=30,\ v_{\mathrm{Fmin}}=20$ $v_{\mathrm{Hmax}}=20,\ v_{\mathrm{CMmax}}=1.0,\ d_{\mathrm{F}}=5$	5	2	6	0.714	0.455	0.556
			B	$v_{\mathrm{min}}=145,\ v_{\mathrm{max}}=160,\ t_{\mathrm{cos}}=0.2$ $V_{\mathrm{CH}}=180,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=0.2$	9	2	2	0.818	0.818	0.818
		G2	A	$v_{\mathrm{Hmin}}=270,\ v_{\mathrm{CMmin}}=30,\ v_{\mathrm{Fmin}}=20$ $v_{\mathrm{Hmax}}=20,\ v_{\mathrm{CMmax}}=1.0,\ d_{\mathrm{F}}=5$	6	1	5	0.857	0.545	0.667
			B	$v_{\mathrm{min}}=145,\ v_{\mathrm{max}}=160,\ t_{\mathrm{cos}}=0.2$ $V_{\mathrm{CH}}=180,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=0.2$	9	3	2	0.750	0.818	0.783
Akita Obako	32 [22]	H1	A	$v_{\mathrm{Hmin}}=200,\ v_{\mathrm{CMmin}}=80,\ v_{\mathrm{Fmin}}=30$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	20	7	12	0.741	0.625	0.678
			B	$v_{\mathrm{min}}=200,\ v_{\mathrm{max}}=220,\ t_{\mathrm{cos}}=0.0$ $V_{\mathrm{CH}}=350,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=1.5$	26	11	6	0.703	0.813	0.754
		I1	A	$v_{\mathrm{Hmin}}=200,\ v_{\mathrm{CMmin}}=80,\ v_{\mathrm{Fmin}}=30$ $v_{\mathrm{Hmax}}=40,\ v_{\mathrm{CMmax}}=2.5,\ d_{\mathrm{F}}=40$	24	7	8	0.774	0.750	0.762
			B	$v_{\mathrm{min}}=200,\ v_{\mathrm{max}}=220,\ t_{\mathrm{cos}}=0.0$ $V_{\mathrm{CH}}=350,\ d_{\mathrm{CM}}=12,\ d_{\mathrm{FN}}=1.5$	24	14	8	0.632	0.750	0.686

*: Extracted by the authors.

In addition to the above advantages, the robustness against the diversity of dance motion is also suggested.

The issue for the practical use of the present method such as the confusion of oversegmentation with regular boundaries is also pointed out. The resolution of this problem will be the subject of further study. The investigation of the possibility of adapting the other sub-theories of GTTM, namely the metrical structure analysis, the time-span reduction and the prolongational reduction, will also be expected to become the subject of future work.

Acknowledgments This study was supported by the "Strategic Information and Communications R&D Promotion Programme (SCOPE)" of the Ministry of Internal Affairs and Communications of Japan.

We deeply regret the death of Prof. Toshiyuki Taniguchi, one of the authors of this paper, who passed away after a struggle with his disease on the 7th May, 2010.

References

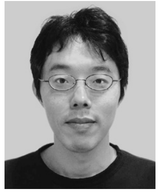
1) Hachimura, K.: Digital Archiving of Dancing, *Review of the National Center for Digitization*, Vol.8, pp.51–66 (2006).
2) Sonoda, M., Tsuruta, S., Yoshimura, M. and Hachimura, K.: Segmentation of Dancing Movement by Extracting Features from Motion Capture Data, *Journal of the IIEEJ*, Vol.37, No.3, pp.303–311 (2008).
3) Shiratori, T., Nakazawa, A. and Ikeuchi, K.: Detecting Dance Motion Structure through Music Analysis, *The 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.857–862 (2004).
4) Yukawa, T., Obara, N. and Tamamoto, H.: Automatic Detection of Motion Primitive Boundaries from Human Motion Capture Data, *IPSJ Journal*, Vol.45, No.4, pp.1198–1201 (2004) (in Japanese).
5) Lerdahl, F. and Jackendoff, R: *A Generative Theory of Tonal Music*, The MIT Press (1983).
6) Hamanaka, M., Hirata, K. and Tojo, S.: Implementing "A Generative Theory of Tonal Music," *Journal of New Music Research*, Vol.35, No.4, pp.249–277 (2006).
7) Stammen, D.R. and Pennycook, B.: Real-time Segmentation of Music Using an Adaptation of Lerdahl and Jackendoff's Grouping Principles, *The 3rd International Conference for Music Perception and Cognition*, pp.269–270 (1994).
8) Miura, T., Mitobe, T., Yukawa, T., Kaiga, T., Taniguchi, T. and Tamamoto, H.: Extraction of Motion Characteristics in Dances by Statistical Analysis of Joint Motions, *Journal of Information Processing*, Vol.18, pp.49–62 (2010).
9) Grassia, F.S.: Practical Parameterization of Rotations Using the Exponential Map, *J. Graphics Tools*, Vol.3, pp.29–48 (1998).
10) Preston, V.: *A Handbook for Modern Educational Dance*, Macdonald & Evans Ltd. (1963).
11) Yamada, A.: Hand Gestures and Steps of the Bon-Dance, *Bulletin of the Faculty of Education, Kochi University*, Vol.69, pp.203–218 (2009) (in Japanese).
12) McGee, V.E. and Carleton, W.T.: Piecewise Regression, *J. Am. Stat. Assoc.*, Vol.65, No.331, pp.1109–1125 (1970).
13) Johnson, R.A. and Wichern, D.W.: *Applied Multivariate Statistical Analysis*, Pearson Education Inc. (2007).
14) Scheffé, H.: An Analysis of Variance for Paired Comparisons, *J. Am. Stat. Assoc.*, Vol.47, pp.381–400 (1952).
15) Reber, A.S., Allen, R. and Reber, E.S.: *Penguin Dictionary of Psychology*, 4th ed. Penguin Books (2009).
16) Yamane, R., Todaka, C., Kawashima, K. and Shakunaga, T.: Analysis of Dance Motion by Correlations between Motion Data, *IEICE Trans. Inf. & Syst.*, Vol.J88-D-II, No.8, pp.1652–1661 (2005) (in Japanese).
17) Shigeizumi, Y. and Kudo, E.: The History of Folkdance "AKITA ONDO," *Memoirs of the Faculty of Education, Akita University, Educational Science*, Vol.31, pp.114–126 (1981) (in Japanese).
18) National Folk Dance Federation of Japan (Ed.): *Furusato no Min'yō I*, National Folk Dance Federation of Japan (2007) (in Japanese).
19) van Rijsbergen, C.J.: *Information Retrieval*, 2nd ed. London: Butterworth (1979).
20) Hripcsak, G. and Rothschild, A.S.: Agreement, the F-measure, and Reliability in Information Retrieval, *Journal of the American Medical Informatics Association*, Vol.12, No.3, pp.296–298 (2005).
21) National Folk Dance Federation of Japan (Ed.): *Furusato no Min'yō III*, National Folk Dance Federation of Japan (2007) (in Japanese).
22) Suzuki, Y.: Creativity in Dance Expression – A Case Study of "Akita Obako," *Bulletin of the Faculty of Education, Fukushima University, Education-Psychology*, Vol.37, pp.17–28 (1985) (in Japanese).

(Received April 8, 2010)
(Accepted September 17, 2010)
(Released December 8, 2010)

Takeshi Miura received his D.Eng. degree in electrical engineering from Hokkaido University in 1998. He is currently an associate professor in the Department of Electrical and Electronic Engineering, Graduate School of Engineering and Resource Science, Akita University.

Kazutaka Mitobe received his D.Eng. degree in biomedical engineering from Hokkaido University in 1996. He is currently an associate professor in the Department of Electrical and Electronic Engineering, Graduate School of Engineering and Resource Science, Akita University. His research work is on biomedical engineering such as human motion analysis and psychophysical experiments for older adults.

Takashi Yukawa received his D.Eng. degree in computer systems engineering from Akita University in 2009. He is currently a professor at the Faculty of Law, North Asia University. His research interests include human motion analysis and computer graphics. He is a member of the IPSJ and VRSJ (Virtual Reality Society of Japan).



Takaaki Kaiga received his M.E. degree in mechanical engineering from Ibaraki University in 1995. Since 1996, he has been with the computer division of Digital Art Factory, Warabi-za Co., Ltd.

Toshiyuki Taniguchi received his D.Eng. degree in electrical engineering from Hokkaido University in 1979. He was a professor in the Department of Electrical and Electronic Engineering, Graduate School of Engineering and Resource Science, Akita University. He died after a struggle with his disease in May, 2010.

Hideo Tamamoto received his D.Eng. degree in electrical engineering from the University of Tokyo in 1976. He is currently a professor in the Department of Computer Science and Engineering, Graduate School of Engineering and Resource Science, Akita University. His research interests include design-for-testability of logic circuits, archiving and handing-down technique for traditional folk dances, and e-learning system.