Regular Paper

Development of a Cartogram Construction Method for Visualizing Japanese Folk Dance Distribution

Takeshi Miura[1,a]   Takeshi Shibata[2]   Madoka Uemura[3]   Katsubumi Tajima[1]   Hideo Tamamoto[4]

Abstract: This paper proposes a cartogram construction method to visualize the relevance of the motion-characteristic distribution of Japanese folk dances to the geographic elements of regional communities in which the dances have been passed down. We use motion capture data of the dances to quantitatively extract their motion characteristics. To systematically organize the cartogram construction process, we adopt a hierarchical model representing the relationship among motion capture data, folk dances, and settlements in which the dances have been passed down. Different cartogram types are selected for different levels in the hierarchical model, and, thereby, a hybrid of circle and distance cartograms is provided. We show that an algorithm to locate the constituents of the hierarchical model in the above hybrid cartogram can be obtained by slightly modifying the existing circle cartogram construction algorithm. On the other hand, we develop another new algorithm to locate geographic elements other than the constituents of the hierarchical model in the hybrid cartogram. The results obtained by analyzing the Furyu type folk dances passed down in Akita Prefecture demonstrated the effectiveness of the proposed method.

Keywords: Japanese folk dance, motion characteristic, cartogram, motion capture

1. Introduction

Folk dances are one of the important constituents of Japanese folk performing arts, along with dramatic, narrative, and musical presentations [1]. In most cases, Japanese folk dances are performed in local events held in respective regional communities [1], [2]. Hence, each of the dances has been strongly affected by the geographic and cultural conditions of each region [3]. For example, the *Bon Odori* [*1] dances of Akita Prefecture are categorized into different dancing-style groups, and the dancing style of each group has its own cultural background such as industrial and transportation conditions, which strongly depend on the geographic features of each region, e.g., the distribution of mines and river basins [4]. By investigating regional dancing-style variation and analyzing its relevance to the geographic elements of corresponding regions, therefore, we can clarify what kind of influence the natural or social environment, such as the distribution of river basins or the condition of a transportation system, has had on the culture of the respective community or at least yield a slight clue.

Visualizing the above information will help understand the overview. Using a technique of map deformation, i.e., constructing a cartogram, is known as one of the approaches to effectively visualize geographically-referenced quantitative data [5], [6]. In the case of investigating folk dances, their motion characteristics can be quantitatively extracted by using motion capture (Mocap) techniques [7]. This allows us to collect information on folk dances in the form of geographically-referenced quantitative data, i.e., Mocap data acquired at each regional community. Consequently, it becomes possible to adopt the cartogram construction approach to visualize the relationship between the dancing-style variation of folk dances and geographic elements.

Taking the above situation into consideration, in this paper we propose a cartogram construction method to visualize the relevance of the motion-characteristic distribution of Japanese folk dances to the geographic elements of regional communities. As already pointed out, we use Mocap data of the dances to quantitatively extract their motion characteristics. Each of the Mocap data streams acquired at each dance can be quantitatively characterized by using a feature vector having a specified dimensionality [8]. This makes it possible to quantitatively evaluate the motion-characteristic similarity of multiple dances. As a result, the configuration of the dances in a cartogram can be determined on the basis of information on quantitative motion-characteristic similarity.

To systematically organize cartogram construction process, we use a structured model representing the relationship among Mocap data, folk dances, and settlements in which respective regional communities have been formed and the dances have been passed down. We adopt the folk dance distribution model proposed in Ref. [9], which represents the above relationship in a systematic hierarchical structure.

[1]   Graduate School of Engineering Science, Akita University, Akita 010–8502, Japan
[2]   Collage of Information and Systems, Muroran Institute of Technology, Muroran, Hokkaido 050–8585, Japan
[3]   Faculty of Education and Human Studies, Akita University, Akita 010–8502, Japan
[4]   Tohoku University of Community Service and Science, Sakata, Yamagata 998–8580, Japan
a)   miura@mail.ee.akita-u.ac.jp

[*1]   *Bon Odori* is a type of Japanese folk dance performed during the annual Buddhist festival called *O-Bon* (or simply *Bon*) [1].

As for cartogram construction, on the other hand, it is well known that there are two types of cartograms: area cartograms and distance cartograms [5], [6]. Area cartograms are constructed by changing the surface area of each spatial unit in step with the corresponding specified value, whereas distance cartograms are constructed by changing the distance of each of the selected point pairs in step with the corresponding specified value. Area cartograms are further classified into multiple types, e.g., contiguous cartograms, non-contiguous cartograms, proportional symbol cartograms such as circle cartograms, etc [6]. We select different cartogram types for different levels in the hierarchical folk dance distribution model, considering the property of each level. As a result, a cartogram constructed becomes a hybrid of multiple cartogram types. Specifically, a hybrid of circle and distance cartograms is provided (details will be described in Section 3).

It is desirable to ensure objectivity and reproducibility in the cartogram construction process. This can be satisfied by providing an algorithm in which a specific procedure to process a given data set is described. We show that an algorithm to construct a hybrid of circle and distance cartograms is easily obtained by slightly modifying the circle cartogram construction algorithm proposed in Ref. [10].

The above algorithm determines the configuration of the constituents of the hierarchical folk dance distribution model, using dancing-style characteristics obtained from Mocap data. As for geographic elements such as roads, rivers, lakes, boundary lines between prefectures none of which is included in the model, however, their configuration cannot be determined. Therefore, we introduce another new algorithm to locate the above elements in a cartogram. The algorithm is developed so as to be capable of displaying the elements in a hybrid of circle and distance cartograms. In the algorithm, the process of realizing a desired geometrical relationship between a given geographic element and the constituents of the model is formulated as a set of linear least squares optimization problems, and numerically solved. By using the developed algorithm, the elements are appropriately deformed according to the change of the locations and sizes of the constituents of the hierarchical model.

To explain each of the cartogram construction algorithms in detail and evaluate the effectiveness of the proposed method, we use the data of actual Japanese folk dances. Specifically, the case of the *Furyū* [*2] type folk dances (mainly, the *Bon Odori* dances) passed down in Akita Prefecture is used as a sample case. As already mentioned at the beginning of this section, the *Bon Odori* dances of Akita Prefecture can be categorized into multiple groups based on regional dancing-style variations, with each region having its own cultural background associated with the distribution of geographic elements. This means that analyzing the above dances is suitable for evaluating the proposed method that visualizes the relationship between the regional dancing-style variation and the distribution of geographic elements.

2.   Related Work

As for the quantitative analysis of Japanese folk dances, attempts have been made to introduce Mocap techniques to extract their motion characteristics. For example, Shiratori et al. [11] presented a trial to extract the motion structure of Japanese folk dances. In their study, it is assumed that the structure of dance motion is represented as a sequence of "primitive motions." They proposed a method to automatically detect keyposes, i.e., boundaries between primitive motions, based on motion-speed characteristics. The obtained results showed that the proposed method gave a different value for each dance category. Usui et al. [12] attempted to use the Mocap data of the folk dances *Minbu* [*3] and *Kagura* [*4] for dance practice. They found that students could have gained an objective perspective by deliberately reducing the information in computer animations of Mocap data. Terada and Fukuhara [13] developed a method to evaluate the dancing skill of the *Awa Odori* [*5] dance. They used Mocap data obtained by using a stereo-camera system, and succeeded at quantifying the degree of skillfulness.

In the above examples, each dance was individually and separately analyzed without considering the relevance to other dances. In other words, the relevance of multiple dances such as similarities and/or differences among dances was not examined. As a result, only the difference among dance performances of the same dance was evaluated, and the distribution of multiple dances, i.e., information on mutual relationship among a plurality of motion characteristics of multiple dances, was not revealed. As an example in which the above motion-characteristic distribution was considered, a study of Ref. [9] can be given. By using the method proposed in Ref. [9], a set of multiple folk dances passed down in a certain area is analyzed altogether, and the distribution of their motion characteristics is schematically visualized. However, their relevance to the geographic elements of regional communities in which the dances have been passed down, the visualization of which is the main subject of this paper, is not considered.

With regard to cartogram construction approaches, on the other hand, there has been an extensive effort to visualize geographically-referenced data using cartograms. Most area cartograms have been used to visualize the mapping of population distribution [14]. In several previous studies, however, area cartograms have been used to schematically represent the distribution of statistics other than ordinary population distribution. For example, Mislove et al. [15] used area cartograms to visualize the distribution of Twitter users in U.S. Han et al. [5] examined the usability of area cartograms for visualizing the GlobeLand30 dataset containing ten major types of land cover data such as vegetation, water body and tundra, and clarified the effectiveness and limitations of using cartograms to a certain extent. As for distance cartograms, most of them has been used as time-space maps that alter the distance of each point pair with respect to its

---

[*2]   *Furyū* is a type of Japanese folk performing art performed to invoke divine assistance to counter the harm from insects or lack of rainfall that may befall crops and the diseases that strike people already enervated by heat and humidity in summer [1].

[*3]   *Minbu* is a type of Japanese folk dance created in accordance with the folk music passed down in each region [12].

[*4]   *Kagura* is a type of Japanese folk dance performed as a ritual to pray for good harvest, good fish and good health [12].

[*5]   The *Awa Odori* dance is a folk dance passed down in Tokushima Prefecture [13], and categorized as a *Bon Odori* dance [2].

travel time [16], with a few exceptions such as the case of Bouts et al. [17] in which the geographic map of Australia was deformed based on the distribution of house prices.

In both the area and distance cartogram studies, to our knowledge, there has been no attempt to use a cartogram for visualizing the regional distribution of the characteristics of folk performing arts. As mentioned in Section 1, on the other hand, a hybrid of circle and distance cartograms is used in this study. However, there are few reports presenting a trial to combine multiple cartogram types. For example, Inoue and Shimizu [10] introduced a technique for constructing the distance cartogram to improve the algorithm of constructing the circle cartogram.
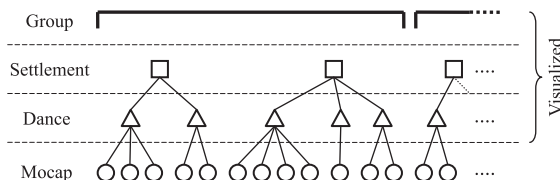
3. Cartogram Construction Method

3.1 Modeling of Folk Dance Distribution

As already mentioned in Section 1, we adopt the folk dance distribution model proposed in Ref. [9]. In the model, the relationship among Mocap data, folk dances, and settlements in which the dances have been passed down is represented in a hierarchical structure. **Figure 1** shows the structure of the model. The hierarchical structure consists of four levels. In the highest "Group" level, each of the constituents of this level, i.e., groups, is a set of settlements each of which is a constituent of the second "Settlement" level. Settlements belonging to the same group are the ones regarded as similar to each other in terms of the dancing style of folk dances passed down in them. Each of the settlements provides a set of dances, constituents of the third "Dance" level, passed down in it. In the lowest "Mocap" level, there are sets of Mocap data streams, each corresponding to any of the dances.

The grouping of settlements in the "Group" level is determined by using information provided in previous studies (many previous studies have reported on the grouping of settlements with respect to the dancing-style similarity of folk dances [18], [19]). We visualize the higher three levels, i.e., the "Group," "Settlement" and "Dance" levels. The lowest "Mocap" level is not visualized because we do not consider visualizing the direct influence of differences generated on multiple performances of the same dance. However, the above difference is indirectly reflected in the analysis of motion characteristics as will be mentioned in Section 3.4.

Table 1 and **Fig. 2** show an example of the model. As already mentioned in Section 1, this is the case of the *Furyū* type folk dances (mainly, the *Bon Odori* dances) passed down in Akita Pre-

fecture. The geographic map shown in Fig. 2 is that projected on a plane with the Japan Plane Rectangular Coordinate System X [20]. The classification into three groups in the "Group" level is that proposed in Ref. [18]. The above grouping is schematically displayed as closed dotted-line loops in Fig. 2 (settlements enclosed in a single loop are those belonging to the same group). The total number of settlements included in the model is ten, and that of the dances is nineteen. Hereinafter, this model is used for

Table 1 *Furyū* type folk dances (mainly, *Bon Odori* dances) of Akita Prefecture.

Group	Settlement	Dance (Symbol used in Figs. 5 and 12)	Mocap (No., Time, System)
Kazuno-Odori System	Kemanai	*Dainosaka* (D)	3, 10.8 s, (a)
		Jinku (J)	3, 9.9 s, (a)
Nanshū-Odori System	Hitoichi	*Dendenzuku* (De)	3, 6.6 s, (b)
		Kitasaka (K)	6, 5.5 s, (b)
		Sankatsu (S)	4, 11.6 s, (b)
	Yamada	*Kitasaka* (K)	5*, 5.5 s, (c)
		Dagasuko (Da)	4*, 6.9 s, (c)
		Sankatsu (S)	5*, 11.8 s, (c)
Akita-Ondo System	Kubota	*Akita Ondo*	1, 67.8 s, (d)
	Nakasen	*Donpan Odori* (Do)	1, 33.8 s, (b)
		Emazō Jinku (E)	1, 32.7 s, (b)
	Kakumagawa	(no name)	1, 59.3 s, (b)
	Nishimonai	*Ondo* (O)	3, 44.5 s, (a)
		Ganke (G)	3, 41.1 s, (a)
	Iwasaki	*Otoko Odori* (Ot)	1, 72.6 s, (b)
		Onna Odori (On)	1, 73.8 s, (b)
	Masuda	(no name)	2, 69.2 s, (b)
	Innai	*Innai Ginzan Odori* (Od)	4, 29.3 s, (b)
		Innai Ginzan Ondo (Ond)	2, 39.3 s, (b)

Time
  Mean time of all the Mocap data streams in a single dance.
System
  (a): MotionStar Wireless (Ascension Technology Corporation)
      with LIBERTY (Polhemus, 2 sets) (30 fps)
  (b): MotionStar Wireless (Ascension Technology Corporation) (30 fps)
  (c): MVN (Xsens) (120 fps)
  (d): STAR*TRACK (Polhemus) (30 fps)
Mocap data: provided by Warabi-za Co., Ltd. (no symbol) and the Center of Community (COC) Project of Akita University (with '*').
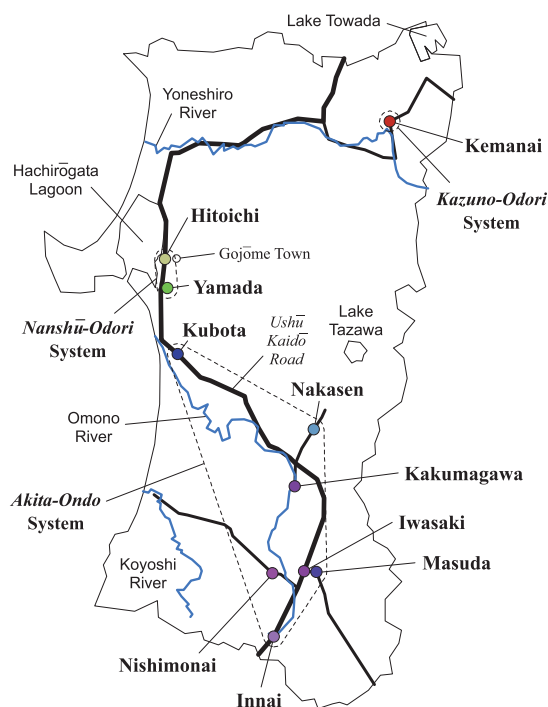

Fig. 1 Hierarchical model of folk-dance distribution.

Constituents of each level
  Group: a set of settlements in which folk dances of the same type have been passed down.
  Settlement: a village or town in which a folk dance (or a set of folk dances) has been passed down.
  Dance: a folk dance passed down in each settlement.
  Mocap: a time-series Mocap data stream acquired in a folk-dance performance.

Fig. 2 Distribution of the *Furyū* type folk dances (mainly, *Bon Odori* dances) in Akita Prefecture.

(1) Setup of circles
(Visualization of the "Dance" level)
A circle and a set of points represent the motion-characteristic distribution of dances in each settlement. A circle represents the motion-characteristic variety of a settlement and a set of points included in the circle represents the distribution of dances.

(2) Arrangement and grouping of settlements
(Visualization of the "Settlement" and "Group" levels)
Circles are arranged as elements of a hybrid of circle and distance cartograms. Circles belonging to the same group are enclosed by a closed loop.

(3) Arrangement of geographic elements
(Addition of geographic information)
Roads, rivers, lakes, boundary lines between prefectures, etc. are added by using a point location algorithm.
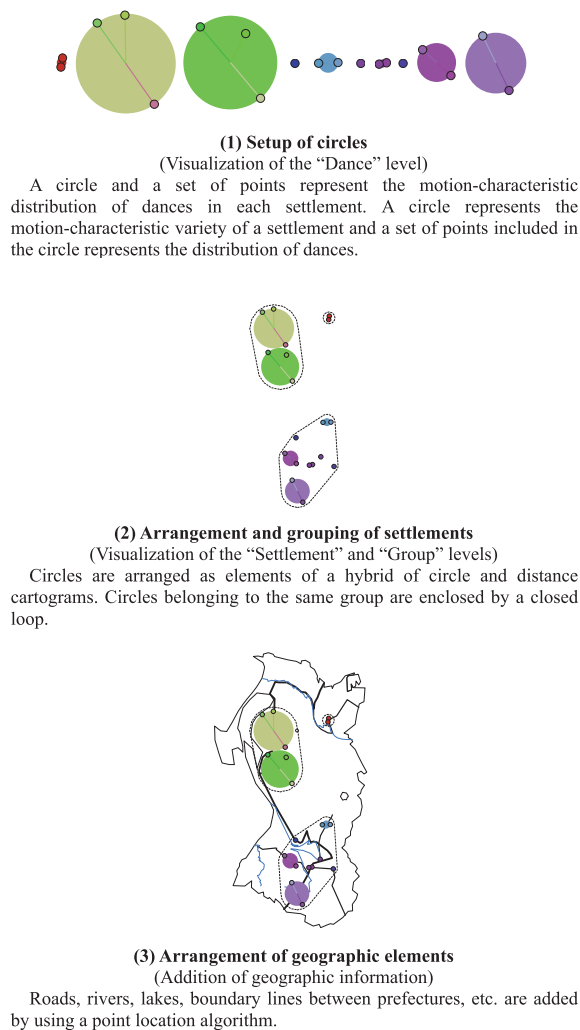
Fig. 3   Overview of cartogram creation process.

explaining details of the cartogram construction process proposed in this paper.

3.2   Overview of Cartogram Construction Process

In this section, we present the overview of the cartogram construction method proposed in this paper, in which the structure of the above folk dance distribution model is taken into account. As shown in Fig. 2, the distribution of the constituents of the "Group" and "Settlement" levels can be geographically illustrated. On the other hand, the distribution of the constituents of the "Dance" level, i.e., dances passed down in respective settlements, cannot be well illustrated in the geographic map, because dances passed down in the same settlement concentrate at a single position, regardless of their motion-characteristic difference.

Considering the above situation, we visualize the distribution of the constituents in different ways for each of the levels. **Figure 3** shows the overview of the visualization procedures executed in the cartogram construction process. For the "Dance" level (Fig. 3 (1)), the tendency of motion-characteristic distribution is represented by a circle at every settlement. The radius of the circle represents the motion-characteristic variety of the multiple dances passed down in a single settlement, and points added inside the circle represent the motion-characteristic distribution

of the dances passed down in the settlement. Colors of the circles and those of the points give information on the motion characteristics of the dances (details will be described in Sections 3.3 and 3.4). The above structure in which each spatial unit is represented as a circle and its radius corresponds to a specified quantity is the fundamental style of a circle cartogram.

Then, the circles, i.e., settlements each of which is a constituent of the "Settlement" level, are arranged on a plane (Fig. 3 (2)). Both the geographical locations and motion-characteristic similarity or distance of dances passed down in respective settlements are reflected in the configuration of them. In this process, the distances of settlement pairs are changed using the infomation on between-settlement motion-characteristic distances obtained by Mocap data analysis. Changing the distances of spatial-unit pairs in accordance with specified quantity values is the fundamental procedure in distance cartogram construction. In other words, the constituents of the "Settlement" level, i.e., circles, are arranged so as to construct a distance cartogram. As a result, the obtained schematic representation becomes a hybrid of circle and distance cartograms. After arranging all the constituents of the "Settlement" level, settlements belonging to the same group in the "Group" level are enclosed by a closed dotted-line loop (Fig. 3 (2)), thereby completing the schematic representation of all the "Dance," "Settlement" and "Group" levels (details will be described in Section 3.5).

In order to make it possible to grasp the relevance of the above constituents to geographic elements such as roads, rivers, lakes and boundary lines between prefectures, information on these elements are additionally displayed in the cartogram (Fig. 3 (3)). These elements are deformed according to the change of the locations and sizes of settlements. Concrete deformed shapes are determined by a point location algorithm (details will be described in Section 3.6).

3.3   Motion Characteristic Extraction from Motion Capture Data

In this section and the next section, we describe a procedure to setup circles each of which represents the distribution of folk dances in a settlement. The above distribution is obtained based on the similarity of the motion characteristics of the dances. In this section, we first describe details of the motion characteristic extraction method.

We adopt the technique proposed in Ref. [8] for extracting motion characteristics from the Mocap data of folk dances. This technique provides motion-characteristic quantities in the form of a simple low-dimensional feature vector having only six components. This allows us to intuitively and easily grasp the motion characteristics of each Mocap data stream from the component values of the obtained feature vector. In addition, the similarity of Mocap data streams can be easily evaluated by calculating the distance between their feature vectors. In fact, the above feature vector showed a high performance comparable to other high-dimensional feature vectors (having over two hundred dimensions) in the Mocap-data classification analysis for multiple dance categories originated in various countries [8].

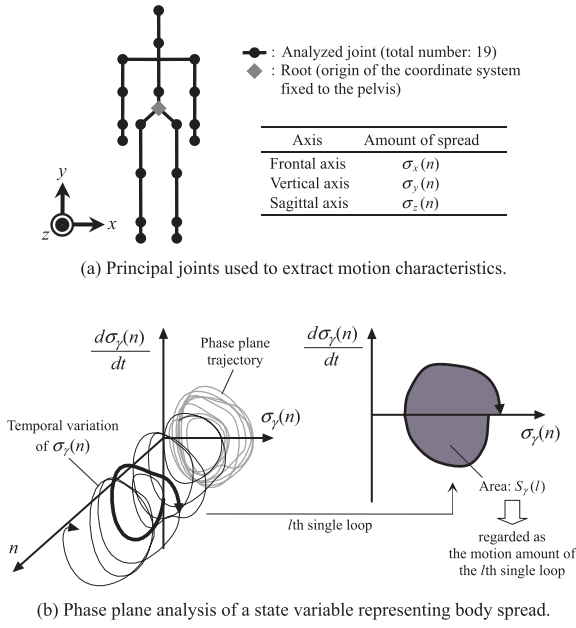The above quantities are obtained by analyzing time-series data

(a) Principal joints used to extract motion characteristics.

(b) Phase plane analysis of a state variable representing body spread.

Fig. 4   Concept of phase plane analysis.

streams of the state variables representing the spatial arrangement of the body segments. Specifically, the body-segment spread along each of the axes of movement, i.e., frontal, vertical and sagittal axes [21], is formulated as a state variable, and evaluated by means of a phase plane analysis method [22] as follows.

First, a set of phase plane trajectories with respect to the temporal variation of the state variables $\sigma_x(n)$, $\sigma_y(n)$ and $\sigma_z(n)$ is drawn:

$$\sigma_\gamma(n) = \sqrt{\frac{1}{J} \sum_{j=1}^{J} \{p_{j,\gamma}(n) - \bar{p}_\gamma(n)\}^2} \tag{1}$$

$$\bar{p}_\gamma(n) = \frac{1}{J} \sum_{j=1}^{J} p_{j,\gamma}(n) \qquad (\gamma: x, y, \text{or } z)$$

where $p_{j,\gamma}(n)$ (γ: x, y or z) is the γ-coordinate of the jth joint at the nth frame (coordinate system: fixed to the pelvis, x: left-ward, y: upward and z: forward as shown in **Fig. 4** (a), the same hereinafter), J is the number of the principal joints used in the analysis (shoulders, elbows, wrists, fingers, hips, knees, ankles, toes, waist, neck and head, $J = 19$). Each of the variables $\sigma_x(n)$, $\sigma_y(n)$ and $\sigma_z(n)$ represents the amount of body-segment spread along each of the frontal, vertical and sagittal axes, respectively, at each instant. A phase plane consists of two axes (a state variable and its time derivative) [22], and the trajectories of $\sigma_x(n)$, $\sigma_y(n)$ and $\sigma_z(n)$ are drawn on their respective phase planes as shown in Fig. 4 (b). The time derivative of each state variable, $d\sigma_\gamma(n)/dt$, is numerically obtained by a finite-difference calculation [23], and the time series of $\sigma_\gamma(n)$ is filtered by a Gaussian filter (cut-off frequency: 10 Hz) before the finite-difference calculation to eliminate noise.

Next, two types of feature quantities are extracted from the phase plane trajectories. The first type is a set of three quantities each of which represents the average motion amount in the γ-direction throughout the whole trajectory, and obtained as follows:

$$q_{\mathrm{MA}\gamma} = \log\left\{\frac{1}{L} \sum_{l=1}^{L} S_\gamma(l) + C\right\} \qquad (\gamma: x, y, \text{or } z) \tag{2}$$

where L is the number of single loops included in the whole trajectory (Fig. 4 (b), the same hereinafter), $S_\gamma(l)$ is the area of the lth single loop and C is a small constant to avoid log(0) (we set $C = e^{-10}$). Here, we define a locus from a negative-direction zero-cross point to the next point as a single loop as shown in Fig. 4 (b). As a result, the greater the body-segment spread or the faster the motion speed, the larger the value of $q_{\mathrm{MA}\gamma}$.

On the other hand, the second type is a set of three quantities each of which represents motion complexity in the γ-direction. This is quantified by using the value of approximate entropy [24], [25] as follows:

$$q_{\mathrm{MC}\gamma} = \Phi^m - \Phi^{m+1} \qquad (\gamma: x, y, \text{or } z) \tag{3}$$

$$\Phi^m = \frac{\sum_{n=1}^{N-(m-1)\tau_\gamma} \log D_n^m}{N - (m-1)\tau_\gamma}$$

$$D_n^m = \frac{\sum_{j=1}^{N-(m-1)\tau_\gamma} H(r_d - d(\boldsymbol{\Sigma}_\gamma(n), \boldsymbol{\Sigma}_\gamma(j)))}{N - (m-1)\tau_\gamma}$$

$$d(\boldsymbol{\Sigma}_\gamma(n), \boldsymbol{\Sigma}_\gamma(j))$$
$$= \max_{i=1,2, \, k=1,2,\cdots,m} (|\mu_{i\gamma}(n + (k-1)\tau_\gamma) - \mu_{i\gamma}(j + (k-1)\tau_\gamma)|)$$

$$\boldsymbol{\Sigma}_\gamma(n) = \begin{bmatrix} \mu_{1\gamma}(n) & \mu_{1\gamma}(n + \tau_\gamma) & \cdots & \mu_{1\gamma}(n + (m-1)\tau_\gamma) \\ \mu_{2\gamma}(n) & \mu_{2\gamma}(n + \tau_\gamma) & \cdots & \mu_{2\gamma}(n + (m-1)\tau_\gamma) \end{bmatrix}$$

where $\mu_{1\gamma}(n)$ and $\mu_{2\gamma}(n)$ are the standardized $\sigma_\gamma(n)$ and $d\sigma_\gamma(n)/dt$ (with zero mean and unity variance, standardized throughout the overall frames), N is the number of frames, $H(x)$ is the Heaviside function, and τ_γ is one fifth of the weighted mean of single-loop periods (weight: $S_\gamma(l)$ for each loop) obtained as follows:

$$\tau_\gamma = \mathrm{round}\left[\frac{0.2}{\sum_{l=1}^{L} S_\gamma(l)} \sum_{l=1}^{L} \{S_\gamma(l)(n_{\mathrm{E}\gamma}(l) - n_{\mathrm{S}\gamma}(l) + 1)\}\right] \tag{4}$$

where $n_{\mathrm{S}\gamma}(l)$ and $n_{\mathrm{E}\gamma}(l)$ are the start and end frames of the lth γ-direction single loop, respectively. We set the parameters $m = 4$ and $r_d = 0.5$ according to Ref. [8]. The value of $q_{\mathrm{MC}\gamma}$ becomes large when a trajectory on the phase plane shows a complex shape. In actual calculations, we use a fast algorithm [26] to reduce the calculation time.

As a result, the following six-dimensional feature vector is obtained at every Mocap data stream:

$$\boldsymbol{F} = \begin{bmatrix} q_{\mathrm{MA}x} & q_{\mathrm{MA}y} & q_{\mathrm{MA}z} & q_{\mathrm{MC}x} & q_{\mathrm{MC}y} & q_{\mathrm{MC}z} \end{bmatrix}^{\mathrm{T}}$$
$$= \begin{bmatrix} f_1 & f_2 & f_3 & f_4 & f_5 & f_6 \end{bmatrix}^{\mathrm{T}}. \tag{5}$$

The distance between a pair of Mocap data streams can be obtained by calculating the Euclidean distance between their feature vectors as follows:

$$d_{\mathrm{M}}(i_l, j_m) = \sqrt{\sum_{k=1}^{6} \{f_k(i_l) - f_k(j_m)\}^2} \tag{6}$$

where $d_{\mathrm{M}}(i_l, j_m)$ is the distance between the i_lth Mocap data stream in the lth dance and the j_mth data stream in the mth dance, $f_k(l_m)$ is the kth component of the feature vector for the l_mth Mocap data stream in the mth dance. In actual calculations, the component values are standardized throughout the overall Mocap data

streams used in the analysis (with zero mean and unity variance) to avoid underestimating (or overestimating) the variation of a particular component.

3.4   Circle Setup to Represent Folk Dance Distribution in a Settlement

As mentioned in Section 3.1, we visualize the distribution of folk dances in a settlement, using information on motion-characteristic similarity. The similarity between a pair of dances is evaluated by using the motion-characteristic distance between them. In the folk dance distribution model adopted in this paper, a single dance can include multiple Mocap data streams as shown in Table 1. To evaluate the between-dance similarity we must obtain the distance between a pair of *sets* of Mocap data streams. Here, we use the Earth Mover's Distance (EMD) [27] as a between-set distance as follows:

$$d_\mathrm{D}(i, j) = \frac{\sum_{k_i=1}^{N_{\mathrm{D}_i}} \sum_{l_j=1}^{N_{\mathrm{D}_j}} d_\mathrm{M}(k_i, l_j) u(k_i, l_j)}{\sum_{k_i=1}^{N_{\mathrm{D}_i}} \sum_{l_j=1}^{N_{\mathrm{D}_j}} u(k_i, l_j)} \tag{7}$$

where $d_\mathrm{D}(i, j)$ is the distance between the ith and jth dances, $u(k_i, l_j)$ is the "flow" from the k_ith Mocap data stream in the ith dance to the l_jth data stream in the jth dance (obtained by solving a transportation problem [*6] [27]), and N_{D_i} is the number of the Mocap data streams belonging to the ith dance. We solve the transportation problem under the following conditions:

$$\sum_{l_j=1}^{N_{\mathrm{D}_j}} u(k_i, l_j) = \frac{1}{N_{\mathrm{D}_i}}, \quad \sum_{k_i=1}^{N_{\mathrm{D}_i}} u(k_i, l_j) = \frac{1}{N_{\mathrm{D}_j}}. \tag{8}$$

This means that every dance is evaluated with the same weighting, regardless of the difference in the Mocap-data-stream numbers of the respective dances. These conditions make $d_\mathrm{D}(i, j)$ a true metric [27]. In the calculation of Eq. (7), on the other hand, the influence of motion difference generated on multiple performances of the same dance is indirectly taken into account, because the information on individual actual performances is included in the component values of the feature vectors used in the calculation of $d_\mathrm{M}(k_i, l_j)$.

After calculating the distances of all the dance pairs, the distribution of all the dances are obtained by mapping them on a plane. We use a technique of metric multidimensional scaling (metric MDS) [29], because $d_\mathrm{D}(i, j)$ is metric as already mentioned. An example for the model of Table 1 is shown in **Fig. 5**. In the obtained scatter plot, the assignment of its axes is selected from the

---

*6   The EMD is obtained by solving a transportation problem in which each of the elements in each set is weighted. In the present case, each Mocap data stream in the ith dance is weighted with $1/N_{\mathrm{D}_i}$ to equally weight all the Mocap data streams in a single dance and unify the total weight of each dance into unity. As a result, the following transportation problem is solved to obtain the flows of the weights between two dances:

$$\text{Minimize} \sum_{k_i=1}^{N_{\mathrm{D}_i}} \sum_{l_j=1}^{N_{\mathrm{D}_j}} d_\mathrm{M}(k_i, l_j) u(k_i, l_j)$$

$$\text{Subject to} \sum_{l_j=1}^{N_{\mathrm{D}_j}} u(k_i, l_j) = \frac{1}{N_{\mathrm{D}_i}}, \quad \sum_{k_i=1}^{N_{\mathrm{D}_i}} u(k_i, l_j) = \frac{1}{N_{\mathrm{D}_j}}, \quad u(k_i, l_j) \geq 0.$$

The above constraints are described in Eq. (8). There are several methods to solve this type of transportation problems [28].

---

following eight combinations:

(1)   Horizontal: Axis 1 (L:−, R:+), Vertical: Axis 2 (B:−, T:+)
(2)   Horizontal: Axis 1 (L:−, R:+), Vertical: Axis 2 (B:+, T:−)
(3)   Horizontal: Axis 1 (L:+, R:−), Vertical: Axis 2 (B:−, T:+)
(4)   Horizontal: Axis 1 (L:+, R:−), Vertical: Axis 2 (B:+, T:−)
(5)   Horizontal: Axis 2 (L:−, R:+), Vertical: Axis 1 (B:−, T:+)
(6)   Horizontal: Axis 2 (L:−, R:+), Vertical: Axis 1 (B:+, T:−)
(7)   Horizontal: Axis 2 (L:+, R:−), Vertical: Axis 1 (B:−, T:+)
(8)   Horizontal: Axis 2 (L:+, R:−), Vertical: Axis 1 (B:+, T:−)
       (Axis 1: axis giving the largest variance of coordinate values, Axis 2: axis giving the second largest variance of coordinate values, L: Left, R: right, B: bottom, T: top)

The combination minimizing the cost function below is selected:

$$g = \sum_{i=1}^{N_\mathrm{S}-1} \sum_{j=i+1}^{N_\mathrm{S}} |\theta(i, j) - \phi(i, j)| \tag{9}$$

where $\theta(i, j)$ is the bearing angle between the ith and jth settlements (grid north [20], the same hereinafter) in the original geographic map, $\phi(i, j)$ is that between the centers of the smallest enclosing circles of the ith and jth settlements in the scatter plot (details of the circles are explained in the next paragraph), and N_S is the number of settlements. Consequently, the difference between the bearing angles of settlement pairs in the original geographic map and those in the scatter plot is minimized. In the case of Fig. 5, the combination (5) is selected.

In the scatter plot shown in Fig. 5 (a), each set of dances belonging to the same settlement is enclosed by their smallest enclosing circle. Its radius is used in a cartogram as that of a circle representing the motion-characteristic variety of the dances belonging to the settlement. In the case that only a single dance has been passed down in a settlement, its radius becomes zero, i.e., the settlement is represented as a point. Otherwise, dance points are plotted in the circle to represent their distribution in each settlement.

To "code" the motion characteristics of the dances, a color is given to each of them. As shown in Fig. 5 (b), we regard the inside of the smallest enclosing circle including all the dance points as the HSV color space [30] ($V = 0.8$), and give each dance a color corresponding to its position as an index representing its motion characteristics. As for each settlement that includes multiple dances, a color corresponding to the centroid of the dances is given as an index representing their average motion-characteristic tendency, and the circle of the settlement is filled with its color.

In order to grasp the relationship between colors and motion characteristics, the axes of the motion-characteristic feature quantities are drawn in the color space. The above axes are obtained by using correlation coefficients between the components of the scatter plot and the feature quantities of the Mocap data streams as follows:

$$\boldsymbol{A}_{f_i} = R_{q_1, q_2}^{f_i} \frac{[\; r_{f_i, q_1}^{q_2} \quad r_{f_i, q_2}^{q_1} \;]^\mathrm{T}}{\left| [\; r_{f_i, q_1}^{q_2} \quad r_{f_i, q_2}^{q_1} \;]^\mathrm{T} \right|} \tag{10}$$

where $\boldsymbol{A}_{f_i}$ is the vector representing the direction and magnitude of the axis of the feature quantity f_i, q_1 and q_2 are the components of the horizontal and vertical axes of the scatter plot, $R_{q_1, q_2}^{f_i}$
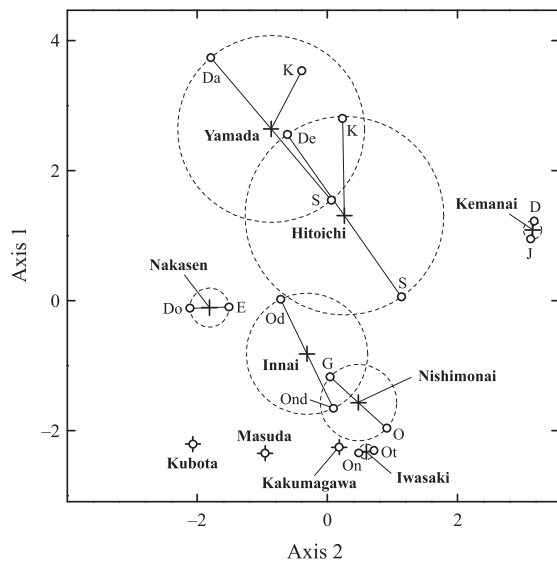
(a) +: center of the smallest circle enclosing all the dances belonging to the same settlement. Its radius is used as that of the circle of the settlement (coefficient of determination: $R^2 = 0.853$).

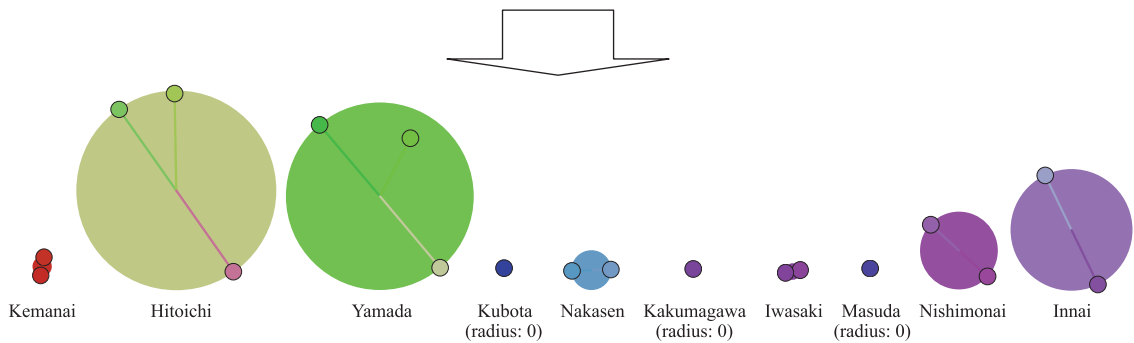Radius → Variety of multiple dances.
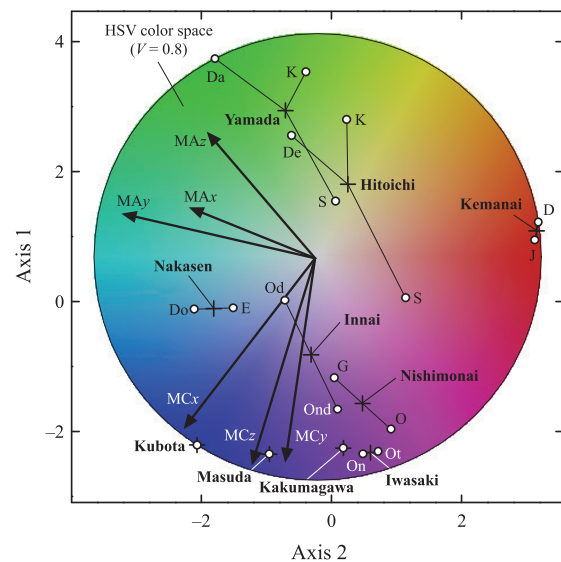
(b) +: centroid of all the dances belonging to the same settlement. The color corresponding to its position is used as that of the settlement. Each dance is also colored in accordance with its position.
(Arrow: axis of the motion-characteristic quantity, MA → Motion amount, MC → Motion complexity.)

Color → Motion characteristics.

(c) Obtained circles. Each of them represents the motion-characteristic distribution of dances in a settlement.

Fig. 5 Determining of the radii and colors of circles (decided by using the MDS scatter plot of the settlements).

is the multiple correlation coefficient between f_i and a set of q_1 and q_2, and $r^b_{f_i,a}$ is the partial correlation coefficient between f_i and the variable a holding the variable b fixed. Arrows in the scatter plot shown in Fig. 5 (b) are the obtained axes (the length of each axis vector is normalized by the radius of the color-space circle). In this case, the axes of q_{MAx}, q_{MAy} and q_{MAz}, i.e., a set of the feature quantities representing motion amount, are turned toward almost the same direction (upper left). On the other hand, the axes of q_{MCx}, q_{MCy} and q_{MCz}, i.e., a set of the feature quantities representing motion complexity, show a similar tendency in that they are turned in almost the same direction (downward). As a result, one can grasp the following relationship between colors and motion characteristics:

　Motion amount　　　Large: green ⟺ Small: red
　Motion complexity　　Complex: blue ⟺ Simple: yellow
The circles of the respective settlements obtained by the above coding procedure are shown in Fig. 5 (c).

　As for computational complexity, that of the calculation of Eq. (6) for all Mocap-data-stream pairs is $O(N_M^2)$ where N_M is

the number of Mocap data streams. On the other hand, the computational complexity of metric MDS in which the number of points, i.e., settlements, is N_S becomes $O(N_S^2)$ [31]. Furthermore, the assignment of colors to dances requires $O(N_D)$ where N_D is the number of dances. Since the relationship $N_S \leq N_D \leq N_M$ is satisfied in the folk dance distribution model of Fig. 1, the computational complexity of the overall circle-setup procedure becomes $O(N_M^2)$. The actual calculation time for the model of Table 1 ($N_S = 10$, $N_D = 19$ and $N_M = 53$) was 0.281 s (CPU: Intel Core i3-350M, the same hereinafter). As will be mentioned in Sections 3.5 and 3.6, the shape of the finally obtained cartogram can be changed as the need arises by adjusting several parameters, even when the same Mocap and geographic data are used. To allow users to adjust the parameters by trial and error, it is desirable to enable interactive processing. The calculation time for the above Mocap data analysis is small enough to perform interactive processing.

3.5 Arrangement and Grouping of Settlements

As mentioned in Section 3.2, the obtained circles each of which represents a settlement are arranged on a plane so as to construct a hybrid of circle and distance cartograms. In this section, we describe an algorithm to obtain the above schematic representation.

To arrange the circle cartogram construction, Inoue and Shimizu [10] proposed an effective algorithm in which the arrangement of circles was adjusted using a technique of distance cartogram construction. In the algorithm, a distance between centers of circles in a cartogram is adjusted in accordance with the distance between the corresponding points in the original geographic map. This approach can be used in the present hybrid-cartogram construction by replacing the geographic distance data with between-settlement motion-characteristic data.

Specifically, a distance cartogram construction algorithm [32] is applied to a set of the following target distances [10]:

$$d'_S(j,k) = \beta\{r(j) + r(k)\} + (1 - \beta)A d_S(j,k) \qquad (11)$$

where $d'_S(j,k)$ is the target distance between the centers of the jth and kth circles (i.e., the jth and kth settlements) in the cartogram, $r(j)$ is the radius of the jth circle, $d_S(j,k)$ is the motion-characteristic distance between the jth and kth settlements (details will be described in the next paragraph), β is the weight of the radius sum term ($0 \le \beta \le 1$, set by users), and A is the coefficient given as $A = \max[\{r(j) + r(k)\}/d_S(j,k)]$ in all the links between circles. The above target distance is the weighted mean of the radius sum of neighboring circles and the between-settlement motion-characteristic distance.

The value of $d_S(j,k)$, i.e., the between-settlement motion-characteristic distance, is given as follows:

$$d_S(j,k) = \frac{\sum_{l_j=1}^{N_{S_j}} \sum_{m_k=1}^{N_{S_k}} d_D(l_j, m_k) v(l_j, m_k)}{\sum_{l_j=1}^{N_{S_j}} \sum_{m_k=1}^{N_{S_k}} v(l_j, m_k)} \qquad (12)$$

where $d_S(j,k)$ is the motion-characteristic EMD between the jth and kth settlements, $v(l_j, m_k)$ is the "flow" from the l_jth dance in the jth settlement to the m_kth dance in the kth settlement (obtained by solving the transportation problem), and N_{S_j} is the number of the dances belonging to the jth settlement. The transportation problem is solved under the following conditions:

$$\sum_{m_k=1}^{N_{S_k}} v(l_j, m_k) = \frac{1}{N_{S_j}}, \quad \sum_{l_j=1}^{N_{S_j}} v(l_j, m_k) = \frac{1}{N_{S_k}}. \qquad (13)$$

As already mentioned, the distance cartogram construction algorithm [32] is applied to a set of the target distances. In the obtained cartogram, circles can be placed apart from each other in response to a given $d_S(j,k)$ data set by appropriately adjusting the value of β. Consequently, we can obtain a hybrid of circle and distance cartograms representing the distribution of settlements in terms of dancing style.

Figure 6 shows the overall algorithm. This algorithm is the same as that proposed in Ref. [10], except for the setup of $d_S(j,k)$ (the motion-characteristic EMD is used instead of the geographic distance as already mentioned). First, the values of the following parameters are input: $x_G(i)$, $y_G(i)$ (coordinates of the ith settlement in the geographic map, $x_G(i)$ for the horizontal axis and
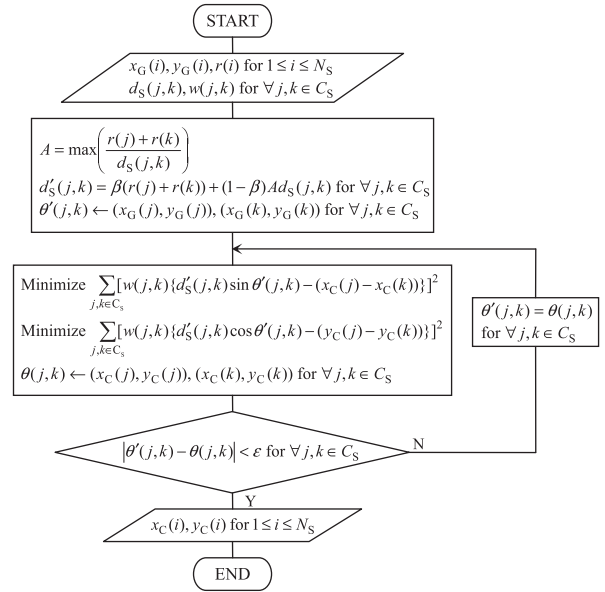

Fig. 6 Algorithm for hybrid cartogram construction.

$y_G(i)$ for the vertical axis [*7], $1 \le i \le N_S$), $r(i)$ ($1 \le i \le N_S$), $d_S(j,k)$ ($j,k \in C_S$, C_S: set of settlement links taken into account, set by users), and $w(j,k)$ (weight of the link between the jth and kth settlements, $j,k \in C_S$, set by users). Next, the following parameters are calculated: $A = \max[\{r(j) + r(k)\}/d_S(j,k)]$, $d'_S(j,k)$ ($j,k \in C_S$), and $\theta'(j,k)$ (bearing angle between the jth and kth settlements in the cartogram, $j,k \in C_S$, the value in the original geographic map obtained below is used as the initial value):

$$\sin \theta'(j,k) = \frac{x_G(j) - x_G(k)}{d_G(j,k)},$$
$$\cos \theta'(j,k) = \frac{y_G(j) - y_G(k)}{d_G(j,k)} \qquad (14)$$
$$d_G(j,k) = \sqrt{(x_G(j) - x_G(k))^2 + (y_G(j) - y_G(k))^2}.$$

Then, the coordinate values of the circle centers ($x_C(i)$ and $y_C(i)$, $1 \le i \le N_S$) are obtained by iterations of the following linear least squares optimizations and the renewal of the $\theta'(j,k)$ values:

$$\text{Minimize} \sum_{j,k \in C} [w(j,k)\{d'_S(j,k) \sin \theta'(j,k) - (x_C(j) - x_C(k))\}]^2 \qquad (15)$$

$$\text{Minimize} \sum_{j,k \in C} [w(j,k)\{d'_S(j,k) \cos \theta'(j,k) - (y_C(j) - y_C(k))\}]^2 \qquad (16)$$

$$\sin \theta(j,k) = \frac{x_C(j) - x_C(k)}{d_C(j,k)}, \cos \theta(j,k) = \frac{y_C(j) - y_C(k)}{d_C(j,k)} \qquad (17)$$

$$d_C(j,k) = \sqrt{(x_C(j) - x_C(k))^2 + (y_C(j) - y_C(k))^2}$$

where $\theta(j,k)$ is the renewed $\theta'(j,k)$. In Eqs. (15) and (16), the squared errors between the actual between-circle-center distances ($x_C(j) - x_C(k)$ and $y_C(j) - y_C(k)$) and the target values ($d'_S(j,k) \sin \theta'(j,k)$ and $d'_S(j,k) \cos \theta'(j,k)$) are minimized on all axes (numerically solved by singular value decomposition [33]). After the convergence of the iteration loop (determined by the

---

[*7] Note that the roles of the variables x and y in this paper are contrary to those in the Japanese surveying and mapping community (x: northing, y: easting) [20]. We assign x and y the horizontal and vertical axes, respectively, in accordance with mathematical conventions.

———— : $w(j,k) = 1.0$, − − − − : $w(j,k) = 0.02$

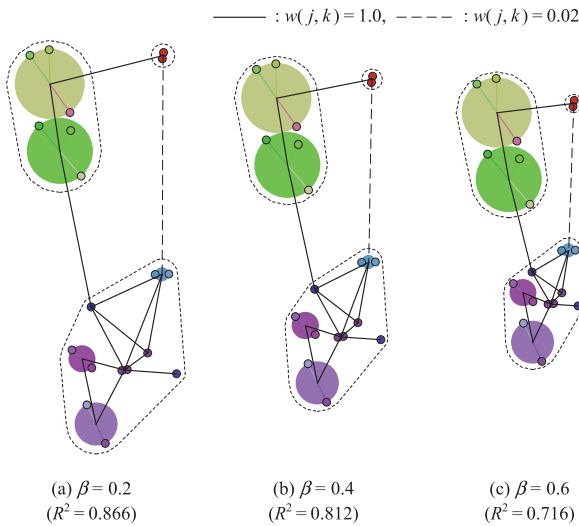| (a) $\beta = 0.2$	(b) $\beta = 0.4$	(c) $\beta = 0.6$
($R^2 = 0.866$)	($R^2 = 0.812$)	($R^2 = 0.716$)

Fig. 7 Settlement configurations obtained by the proposed algorithm (R^2: coefficient of determination for the $w(j,k) > 0$ links).

condition $|\theta'(j,k) - \theta(j,k)| < \epsilon$ for all $j,k \in C_S$), the obtained $x_C(i)$ and $y_C(i)$ values are output.

After giving the configuration of all the circles, each group is enclosed by a closed dotted-line loop. The loop is obtained by drawing a convex hull [34] enclosing all the circles belonging to the same group. Specifically, a convex hull enclosing all the auxiliary points added to each of the circles is used. The auxiliary points are added to intelligibly display the region of each group. Sixteen auxiliary points are regularly plotted around each circle (these points are not displayed in an actual cartogram). The distance between the point and the center of the ith circle is given as $r(i) + 0.015\max[x_W, y_W]$ (x_W and y_W: widths of the existing regions of coordinate values in the horizontal and vertical axes of the obtained cartogram, respectively).

Figure 7 shows examples for the model of Table 1. The weights of the settlement links corresponding to the roads each of which gives a direct link between a settlement pair (13 links) are set as $w(j,k) = 1.0$. To consider the relationship between the group of the *Kazuno-Odori* System and that of the *Akita-Ondo* System between which no direct connection exists, the link between the Kemanai and Nakasen settlements is weighted with a small value ($w(j,k) = 0.02$). The parameter for determining the convergence of iteration is set as $\epsilon = 10^{-5}\pi$. Figure 7 shows three cases differing in the value of the parameter β in Eq. (11). One can recognize that distances between circles vary in response to the variation of the β value. Hereinafter, we adopt the $\beta = 0.4$ case (Fig. 7 (b)) giving the best balance between circle sizes and between-circle distances.

The computational complexity of the algorithm of Fig. 6 is $O(N_S^2)$ [32], and the actual calculation time for the case of Fig. 7 (b) ($N_S = 10$) was 1.406 s. This is short enough to allow users to perform a trial-and-error adjustment of the parameters $w(j,k)$ and β to obtain a cartogram with a more preferable shape.

3.6   Arrangement of Geographic Elements

The final stage of cartogram construction is to display geographic elements such as roads, rivers, lakes and boundary lines between prefectures in the obtained cartogram. In this section,
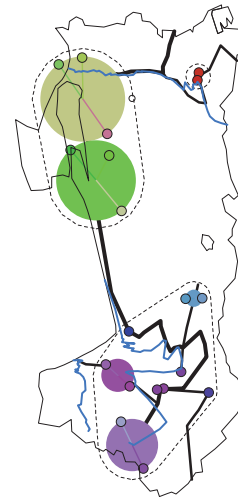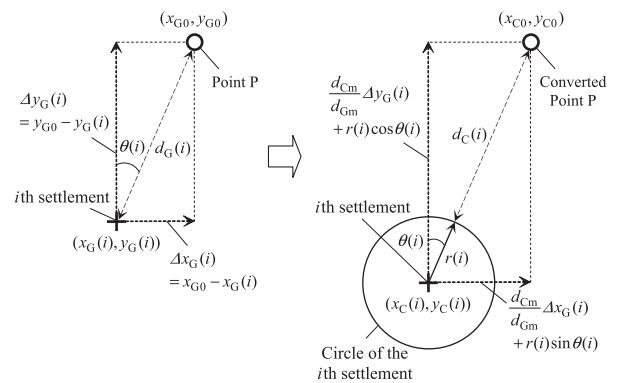

Fig. 8 Cartogram obtained by the Delaunay triangulation and barycentric interpolation ($\beta = 0.4$).

$\begin{pmatrix} d_{Gm} : \text{Mean between-settlement distance in the geographic map} \\ d_{Cm} : \text{Mean between-settlement distance in the cartogram} \end{pmatrix}$

$$\text{Minimize} \sum_{i=1}^{N_S} \left[\frac{1}{d_C(i)^n} \{(x_{C0} - x_C(i)) - (\frac{d_{Cm}}{d_{Gm}} \Delta x_G(i) + r(i)\sin\theta(i))\} \right]^2 \rightarrow x_{C0}$$

$$\text{Minimize} \sum_{i=1}^{N_S} \left[\frac{1}{d_C(i)^n} \{(y_{C0} - y_C(i)) - (\frac{d_{Cm}}{d_{Gm}} \Delta y_G(i) + r(i)\cos\theta(i))\} \right]^2 \rightarrow y_{C0}$$

Fig. 9 Concept of point-location conversion.

we describe a point location algorithm to determine the location of the elements in a cartogram.

In distance cartogram construction, an interpolation technique consisting of triangulation and barycentric interpolation is often used to locate the above elements [16]. **Figure 8** shows an example in which the Delaunay triangulation [35] and barycentric interpolation are applied to the case of Fig. 7 (b). In the obtained cartogram, several circles overlap with rivers or prefecture boundary lines none of which overlaps with any of the settlement in the original geographic map shown in Fig. 2. This means that applying triangulation and barycentric interpolation to a hybrid of circle and distance cartograms in which a settlement point can be converted into a circle rather than a point is inappropriate.

To resolve the above issue, we develop a new point location algorithm. **Figure 9** shows the basic concept of the developed algorithm. Consider the case that the point $P(x_{G0}, y_{G0})$ in the geographic map is given as an input to be mapped in the cartogram. As shown in the upper left of Fig. 9, the relative po-

sition of P to the ith settlement whose position is $(x_G(i), y_G(i))$ is given as $(\Delta x_G(i), \Delta y_G(i))$ where $\Delta x_G(i) = x_{G0} - x_G(i)$ and $\Delta y_G(i) = y_{G0} - y_G(i)$, and the bearing angle between them, $\theta(i)$, is given as follows:

$$\sin\theta(i) = \frac{\Delta x_G(i)}{d_G(i)}, \cos\theta(i) = \frac{\Delta y_G(i)}{d_G(i)} \tag{18}$$

$$d_G(i) = \sqrt{\Delta x_G(i)^2 + \Delta y_G(i)^2}.$$

We assume that the ith settlement is converted into the ith circle having the center $(x_C(i), y_C(i))$ and the radius $r(i)$ in the cartogram as shown in the upper right of Fig. 9. We also assume that the distance between P and the ith settlement in the geographic map, $d_G(i)$, is converted into the following $d_C(i)$ in the cartogram:

$$d_C(i) = \frac{d_{Cm}}{d_{Gm}} d_G(i) \tag{19}$$

where d_{Gm} and d_{Cm} are respectively the mean between-settlement distance in the geographic map and that in the cartogram and given as follows:

$$d_{Xm} = \frac{2}{N_S(N_S - 1)}$$
$$\times \sum_{i=1}^{N_S-1} \sum_{j=i+1}^{N_S} \sqrt{(x_X(j) - x_X(i))^2 + (y_X(j) - y_X(i))^2}$$
$$(\text{X: G or C}). \tag{20}$$

This conversion is adopted considering that the scale conversion from the geographic map into the cartogram can be evaluated by the changing ratio between the mean between-settlement distance in the geographic map and that in the cartogram. We here define that the distance between the converted $P(x_{C0}, y_{C0})$ and the ith circle in the cartogram, i.e., $d_C(i)$, is given as that between P and the nearest point on the circumference of the circle as shown in the upper right of Fig. 9. We assume that the bearing angle between the converted P and the center of the ith circle in the cartogram is kept the same as $\theta(i)$ in the original geographic map. As a result, the relative position of the converted P to the center of the ith circle in the cartogram is given as $(\Delta x_C(i), \Delta y_C(i))$ where $\Delta x_C(i) = x_{C0} - x_C(i)$, $\Delta y_C(i) = y_{C0} - y_C(i)$ and at the same time

$$\Delta x_C(i) = \frac{d_{Cm}}{d_{Gm}} \Delta x_G(i) + r(i) \sin\theta(i) \tag{21}$$

$$\Delta y_C(i) = \frac{d_{Cm}}{d_{Gm}} \Delta y_G(i) + r(i) \cos\theta(i) \tag{22}$$

as shown in the upper right of Fig. 9. In an actual cartogram, the configuration of the centers of the circles generally becomes different from that of the settlements in the geographic map. Therefore, the conditions of Eqs. (21) and (22) are not necessarily satisfied at every settlement.

Taking the above situation into account, we perform the following linear least squares optimizations in which $\Delta x_C(i)$ and $\Delta y_C(i)$ of Eqs. (21) and (22) are used as target values as shown in the bottom of Fig. 9:

$$\text{Minimize} \sum_{i=1}^{N_S} \left[\frac{1}{d_C(i)^n} \left\{ (x_{C0} - x_C(i)) \right.\right.$$
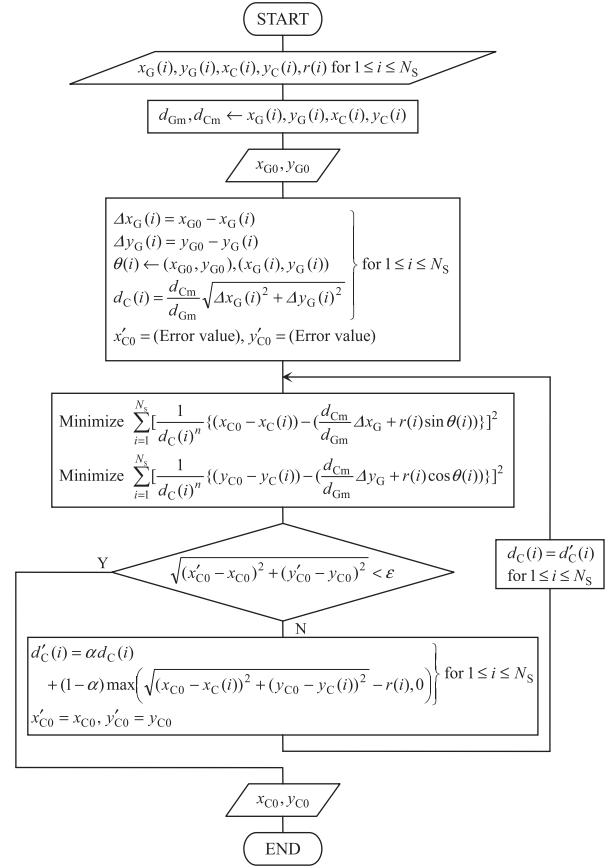$$\left.\left. - \left(\frac{d_{Cm}}{d_{Gm}} \Delta x_G(i) + r(i) \sin\theta(i) \right) \right\} \right]^2 \tag{23}$$

Fig. 10 Algorithm for point location.

$$\text{Minimize} \sum_{i=1}^{N_S} \left[\frac{1}{d_C(i)^n} \left\{ (y_{C0} - y_C(i)) \right.\right.$$
$$\left.\left. - \left(\frac{d_{Cm}}{d_{Gm}} \Delta y_G(i) + r(i) \cos\theta(i) \right) \right\} \right]^2 . \tag{24}$$

These are numerically solved by singular value decomposition [33]. In the above optimizations, each settlement is weighted with the value of $1/d_C(i)^n$ (n: parameter to adjust the strength of the weight values). This means that the closer a settlement is to the point P, the more the settlement is weighted. In actual calculations, however, the true value of $d_C(i)$ in the cartogram cannot be obtained unless the values of x_{C0} and y_{C0} are determined by solving Eqs. (23) and (24). Therefore, we adopt a point location algorithm shown in **Fig. 10** in which the above optimizations are iteratively performed with the renewal of $d_C(i)$.

In the algorithm, the values of the following parameters are first input: $x_G(i), y_G(i), x_C(i), y_C(i)$ and $r(i)$ ($1 \le i \le N_S$), and d_{Gm} and d_{Cm} are calculated. Next, the coordinate values of the point P in the geographic map (x_{G0}, y_{G0}) are input and the following values are calculated: $\Delta x_G(i) = x_{G0} - x_G(i)$, $\Delta y_G(i) = y_{G0} - y_G(i)$, $\theta(i)$ (Eq. (18)), and $d_C(i) = (d_{Cm}/d_{Gm}) \sqrt{\Delta x_G(i)^2 + \Delta y_G(i)^2}$ ($1 \le i \le N_S$). Then, the coordinate values of the converted point P in the cartogram (x_{C0}, y_{C0}) are obtained by iterations of the linear least squares optimizations (Eqs. (23) and (24)) and the following renewal of the $d_C(i)$ value:

$$d_C'(i) = \alpha d_C(i) + (1 - \alpha)$$
$$\times \max\left(\sqrt{(x_{C0} - x_C(i))^2 + (y_{C0} - y_C(i))^2} - r(i), 0 \right) \tag{25}$$

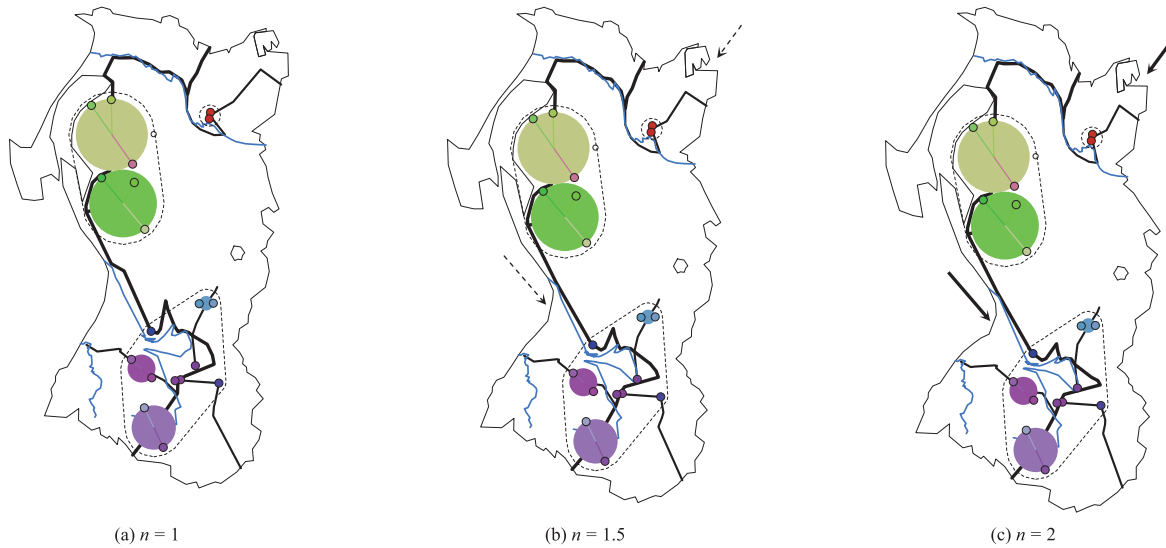(a) $n = 1$ (b) $n = 1.5$ (c) $n = 2$

Fig. 11   Cartograms obtained by the proposed point location algorithm ($\beta = 0.4$).

where $d'_C(i)$ is the renewed $d_C(i)$. This is the weighted mean of the value just before the renewal and that newly obtained from the configuration in the cartogram ($0 \leq \alpha \leq 1$, the "max" operation is used to avoid a negative value). This renewal procedure is introduced to prevent the occurrence of numerical instability in iterative calculations (we set $\alpha = 0.9$ through trial-and-error calculations). After the convergence of the iteration loop (determined by the condition $\sqrt{(x'_{C0} - x_{C0})^2 + (y'_{C0} - y_{C0})^2} < \epsilon$ where x'_{C0} and y'_{C0} are the values of x_{C0} and y_{C0} just before the renewal), the obtained x_{C0} and y_{C0} values are output.

Figure 11 shows examples in which the proposed point location algorithm is applied to the case of Fig. 7 (b). The parameter for determining the convergence of iteration is set as $\epsilon = 10^{-5} d_{Cm}$. Three examples differing in the value of the parameter n in Eqs. (23) and (24) are shown. Although the conversion performed in the proposed algorithm is not strictly homeomorphic [16], the homeomorphism of mapping is not violated in any of the obtained cartograms. In addition, overlap between the circles and the elements that do not geographically overlap with any of the settlements is not seen in the obtained cartograms. These facts suggest the utility of the proposed algorithm in practical use. On the other hand, one can recognize that the degree of deformation (indicated by arrows in Fig. 11) becomes larger as the n value increases. This means that users can adjust the degree of deformation by varying the n value. Hereinafter, we adopt the $n = 1.5$ case (Fig. 11 (b)) that shows a moderate degree of deformation.

In the case that the algorithm of Fig. 10 is applied to N_P input points, its computational complexity becomes $O(N_S^2 + N_S N_P)$. This is because that the calculation of d_{Gm} and d_{Cm} (requiring $O(N_S^2)$) can be executed independently from searching N_P sets of x_{C0} and y_{C0} (requiring $O(N_S N_P)$). The actual calculation time for the case of Fig. 11 (b) ($N_S = 10$, $N_P = 388$) was 0.250 s. This is short enough to allow users to perform a trial-and-error adjustment of the parameter n in interactive processing.

4. Results and Discussion

This section describes details of the comparative examina-

tion of the original geographic map and the cartogram that resulted from the application of the proposed method. As already mentioned in Section 3.6, we select the cartogram shown in Fig. 11 (b), i.e., the version with $\beta = 0.4$ and $n = 1.5$, as the definitive version obtained by the proposed method. **Figure 12** shows the original geographic map (left, identical to Fig. 2) and the cartogram of the definitive version (right). A colored circle displayed in the bottom of Fig. 12 is that of the color space shown in Fig. 5 (b). White arrows in the circle are the motion-characteristic axes. Their directions were obtained from the axis vectors shown in Fig. 5 (b) as follows:

MA: Large ⇐ Mean of MAx, MAy and MAz.
MA: Small ⇐ Opposite to MA: Large.
MC: Complex ⇐ Mean of MCx, MCy and MCz.
MC: Simple ⇐ Opposite to MC: Complex.

The above setup was made considering the tendency of the motion-characteristic distribution mentioned in Section 3.4. In the obtained cartogram, the three *Bon Odori* groups, i.e., the *Kazuno-Odori*, *Nanshū-Odori* and *Akita-Ondo* Systems, are clearly classified by color. In other words, it is clearly indicated in the cartogram that the above three groups have different motion characteristics, respectively, as follows:

Kazuno-Odori: red (MA: Small)
Nanshū-Odori: yellowish green (MA: Large, MC: Simple)
Akita-Ondo: blue (MC: Complex)

These characteristics are almost consistent with those pointed out in previous studies (*Kazuno-Odori*: showing elegant and refined choreography [36], *Nanshū-Odori*: characterized by a simple but intense dancing style [37], and *Akita-Ondo*: having a sophisticated dancing style including a variety of elaborate performance forms [38]).

On the other hand, two distinctive features with respect to deformation are seen in the cartogram. One is a large expansion of the Hitoichi and Yamada settlements, and the other is a remarkable southward movement of the Kubota settlement. First, we focus on the first feature. Both the Hitoichi and Yamada settlements belong to the group of the *Nanshū-Odori* System. As
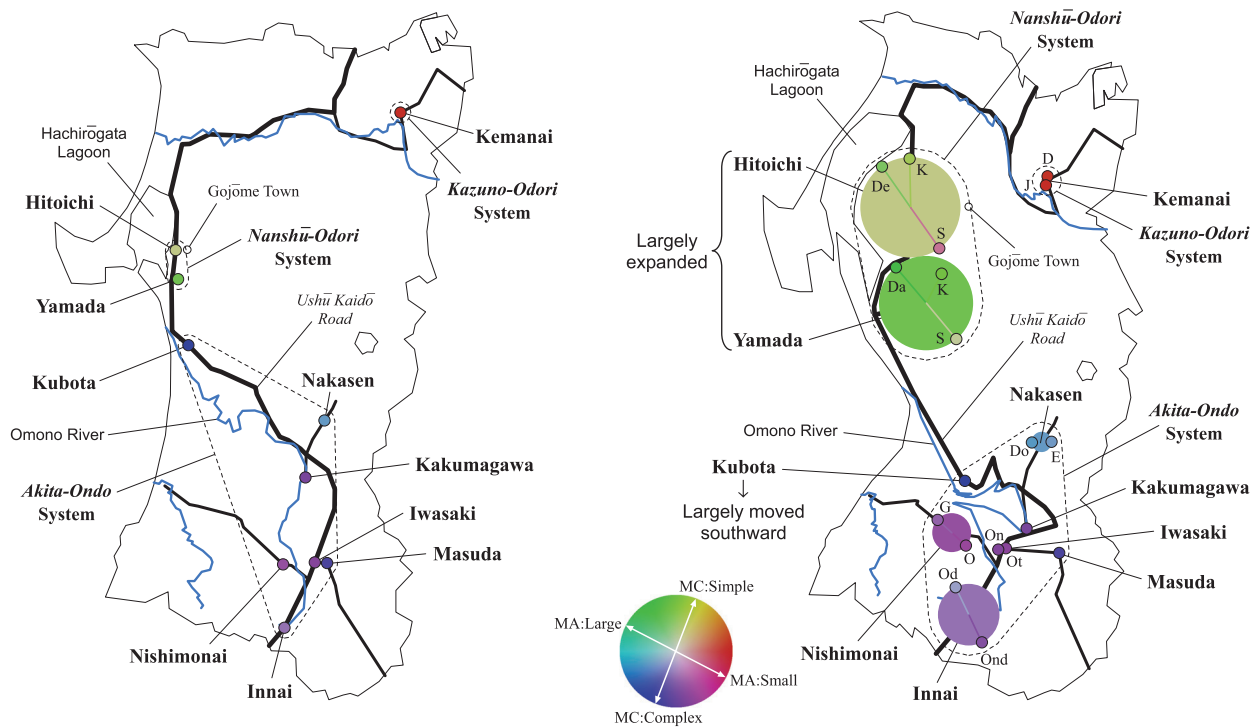
Fig. 12 Comparison between the original geographic map (left) and the cartogram (right, $\beta = 0.4$, $n = 1.5$).

shown in Table 1, the number of the dances passed down in a single settlement belonging to the *Nanshū-Odori* System is generally larger than that of the other groups. In fact, the region of the *Nanshū-Odori* System is known as a place where *Bon Odori* is most popular in Akita Prefecture [18], and each of the multiple dances passed down in this region is characterized by different types of dancing styles [37]. In the obtained cartogram, the circle size of these settlements (large radius, i.e., large variety of motion characteristics in multiple dances) were a good match with the above tendency.

In addition, it is seen that the above two settlements exist in the area of the eastern coast of Hachirōgata Lagoon. This area is known as that on which Gojōme Town (indicated at the right side of the Hitoichi settlement in the cartogram of Fig. 12) had exerted a great influence as a commercial center in the ages from the medieval times to the early-modern times [39] (i.e., the period *Bon Odori* dances spread around the country [40]). It has been pointed out that the development of Gojōme Town brought about cultural variety in the area of the eastern coast of Hachirōgata Lagoon, and this situation might have affected the diversity of the motion characteristics of folk dances in this area [4]. The obtained cartogram can provide at least a slight clue to the above opinion by schematically associating the motion characteristics of the *Bon Odori* dances with their geographic configuration.

Next, we focus on the second feature, i.e., the southward movement of the Kubota settlement. The Kubota settlement is one of the settlements belonging to the group of the *Akita-Ondo* System, and was the capital of Akita Domain in the early-modern times (now, part of Akita City, the capital of Akita Prefecture) [41]. As shown in the geographic map (left of Fig. 12), the Kubota settlement is connected with the Innai settlement by the Ushū Kaidō Road and Omono River that was a water route in the

early-modern times [41]. The Innai settlement was a mining town neighboring the Innai Silver Mine, and its amusement quarter attracted many professional dancers while the mine prospered in the early-modern times [42]. It has been pointed out that their sophisticated dancing style was distributed throughout the region of the *Akita-Ondo* System [43], and the traffic network consisting of the Ushū Kaidō Road and Omono River played an important role in the distribution, especially in the connection between the Kubota settlement and the Innai Silver Mine [4]. In the obtained cartogram, the Kubota settlement moved toward the Innai settlement, as if it was pulled by a set of two springs consisting of the Ushū Kaidō Road and Omono River. One can thereby easily imagine the influence of the above traffic network from the configuration of the obtained cartogram. Consequently, the motion-characteristic distribution of the *Akita-Ondo* System is recognized as concentrating in a narrow region, in contrast to the wide-range distribution of the *Nanshū-Odori* System.

The above regional variation of the relationship between the motion characteristics of the *Bon Odori* dances of Akita Prefecture and the geographic elements of corresponding regions has been poined out in several previous studies [4], [9], [40], [43]. However, it has never been schematically displayed until now. The proposed method makes it possible to visualize the above variation on the basis of quantitative grounds. Although it is not necessarily easy to recognize the influence of the cultural background of each region without preliminary knowledge, the obtained cartogram would help attract renewed attention to the study of folk dances with a visual impact, as in the case that an area cartogram is used to visualize the population distribution of the world [6].

As mentioned above, the deformation in the cartogram obtained by the proposed method supported previous studies on the

Bon Odori dances of Akita Prefecture. In particular, schematically providing information on the relationship between the motion characteristics of folk dances and their geographic configuration was considerably convenient to intuitively understand their features. If the Mocap and geographic data of the old days of the above dances are obtained, for example, it becomes possible to visualize the transition of these dances from old days to current days. In addition, it is possible to apply the proposed method to any other dance category as long as its regional distribution can be represented as the hierarchical model shown in Fig. 1.

It should also be pointed out that the proposed method has a limitation. The feature quantities extracted from Mocap data are limited to those representing only two types of motion characteristics, i.e., motion amount and motion complexity. As already mentioned in Section 3.3, the feature vector in which only the above quantities are used has a high dance-category classification ability. However, the possibility that the regional distribution of some other motion characteristic might provide further valuable information cannot be ruled out. Therefore, additional work is needed to use more diverse types of information on motion characteristics without excessively increasing the number of feature quantities. Moreover, the usability of the obtained hybrid-type cartogram is not sufficiently examined at this moment. More detailed work is necessary to clarify the usability and disadvantage of this type of cartogram.

5. Conclusion

The main contribution of this paper is to provide a cartogram construction method to visualize the relevance of the motion-characteristic distribution of Japanese folk dances to the geographic elements of regional communities in which the dances have been passed down. Mocap data of the dances are used to quantitatively extract their motion characteristics. To systematically organize cartogram construction process, we adopt a hierarchical model representing the relationship among Mocap data, folk dances and settlements. Different cartogram types are selected for different levels in the hierarchical model, and, thereby, a hybrid of circle and distance cartograms are provided. The results obtained by analyzing the *Furyū* type folk dances passed down in Akita Prefecture demonstrated the effectiveness of the proposed method to a certain extent. The clarification of the application range of the proposed method will be discussed in a future work.

References

[1] Thornbury, B.E.: *The Folk Performing Arts*, State University of New York Press (1997).
[2] Lancashire, T.A.: *An Introduction to Japanese Folk Performing Arts*, Ashgate (2011).
[3] Mimuro, K.: An Introduction to a Study on the Traditional Folk Dance of Japan: With Special Reference to Traditional Folk Dances in the Kibi District, *Bulletin of School of Education, Okayama University*, Vol.18, pp.154–173 (1964) (in Japanese).
[4] Miura, T., Kaiga, T., Shibata, T., Katsura, H., Uemura, M., Tajima, K. and Tamamoto, H.: Motion Characteristics of Bon Odori Dances in Areas along Ushu Kaido Road in Akita Domain, *IPSJ Symposium Series*, Vol.2015, No.2, pp.269–276 (2015).
[5] Han, R., Li, Z., Ti, P. and Xu, Z.: Experimental Evaluation of the Usability of Cartogram for Representation of GlobeLand30 Data, *International Journal of Geo-Information*, Vol.6, No.6-180, pp.1–13 (2017).
[6] Markowska, A. and Korycka-Skorupa, J.: An Evaluation of GIS tools for Generating Area Cartograms, *Polish Cartographical Review*, Vol.47, No.1, pp.19–29 (2015).
[7] Kitagawa, M. and Windsor, B.: *MoCap for Artists*, Focal Press (2008).
[8] Miura, T., Kaiga, T., Shibata, T., Tajima, K. and Tamamoto, H.: Low-dimensional Feature Vector Extraction from Motion Capture Data by Phase Plane Analysis, *Journal of Information Processing*, Vol.25, pp.884–887 (2017).
[9] Miura, T., Kaiga, T., Shibata, T., Uemura, M., Tajima, K. and Tamamoto, H.: Development of a Visualization Method for Motion-characteristic Distribution of Japanese Folk Dances - A Case Study of the Bon Odori Dance, *Journal of Information Processing*, Vol.26, pp.74–84 (2018).
[10] Inoue, R. and Shimizu, E.: A New Method for Constructing Circle Area Cartogram, *Theory and Applications of GIS*, Vol.13, No.1, pp.43–50 (2005) (in Japanese).
[11] Shiratori, T., Nakazawa, A. and Ikeuchi, K.: Detecting Dance Motion Structure through Music Analysis, *Proc. FGR 2004*, pp.857–862 (2004).
[12] Usui, Y., Sato, K. and Watabe, S.: The Effect of Motion Capture on Learning Japanese Traditional Folk Dance, Herrington, J., Couros, A. and Irvine, V. (Eds.): *Proc. EdMedia: World Conference on Educational Media and Technology 2013*, pp.2320–2325 (2013).
[13] Terada, K. and Fukuhara, T.: Quantitative Evaluation of Motion of Awa Odori Dance by Three Dimensional Sensor, *IEEJ Trans. EIS*, Vol.129, No.5, pp.876–884 (2009) (in Japanese).
[14] Nusrat, S. and Kobourov, S.: The State of the Art in Cartograms, *EuroVis 2016 Proc. Eurographics/IEEE VGTC Conf. Visualization*, pp.619–642 (2016).
[15] Mislove, A., Lehmann, S., Ahn, Y.-Y., Onnela, J.-P. and Rosenquist, J.N.: Understanding the Demographics of Twtter Users, *Proc. 5th Intr. AAAI Conf. Weblogs and Social Media*, pp.554–557 (2011).
[16] Bies, S. and van Kreveld, M.: Time-Space Maps from Triangulations, Didimo, W. and Patrignani, M. (Eds.): *GD 2012, LNCS 7704*, pp.511–516 (2013).
[17] Bouts, Q.W., Dwyer, T., Dykes, J., Speckmann, B., Goodwin, S., Riche, N.H., Carpendale, S. and Liebman, A.: Visual Encoding of Dissimilarity Data via Topology-preserving Map Deformation, *IEEE Trans. Vis. Comput. Graph.*, Vol.22, No.9, pp.2200–2213 (2015).
[18] Japan Broadcasting Corporation (Ed.): *Tōhoku Min'yōshū Akita Ken (A Collection of Folk Songs of Tohoku Region: Akita Prefecture)*, Japan Broadcast Publishing Association (1957) (in Japanese).
[19] Mimura, Y.: *Hiroshima no Kagura Tanbō (An Inquire into Kagura of Hiroshima)*, Nannansha (2004) (in Japanese).
[20] Okazawa, H., Kubodera, T., Sasada, K., Tasumi, M., Hosokawa, Y., Matsuo, E. and Mihara, M.: *Newer Surveying -Fundamentals and Applications with Newest Technology*, Corona Publishing Co., Ltd. (2014) (in Japanese).
[21] Bartlett, R.: *Introduction to Sports Biomechanics*, 2nd ed., Routledge (2007).
[22] Slotine, J.-J.E. and Li, W.: *Applied Nonlinear Control*, Prince-Hall, Inc. (1991).
[23] LeVeque, R.J.: *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-state and Time-dependent Problems*, SIAM (2007).
[24] Pincus, S.M.: Approximate Entropy as a Measure of System Complexity, *Proc. Natl. Acad. Sci. USA*, Vol.88, pp.2297–2301 (1991).
[25] Kaffashi, F., Foglyano, R., Wilson, C.G. and Loparo, K.A.: The Effect of Time Delay on Approximate & Sample Entropy Calculations, *Physica D*, Vol.237, No.23, pp.3069–3074 (2008).
[26] Manis, G.: Fast Computation of Approximate Entropy, *Computer Methods and Programs in Biomedicine*, Vol.91, No.1, pp.48–54 (2008).
[27] Rubner, Y., Tomasi, C. and Guibas, L.J.: The Earth Mover's Distance as a Metric for Image Retrieval, *International Journal of Computer Vision*, Vol.40, No.2, pp.99–121 (2000).
[28] Stojanović, V., Spalević, L. and Božinović, M.: Software Application for Solving the Transportation Problem, *Proc. ERK 2014*, pp.B:23–26 (2014).
[29] Hair, Jr., J.F., Black, W.C., Babin, B.J. and Anderson, R.E.: *Multivariate Data Analysis*, 7th ed. Pearson Education Inc. (2010).
[30] Ibraheem, N.A., Hasan, M.M., Khan, R.Z. and Mishra, P.K.: Understanding Color Models: A Review, *ARPN Journal of Science and Technology*, Vol.2, No.3, pp.265–275 (2012).
[31] Brandes, U. and Pich, C.: Eigensolver Methods for Progressive Multidimensional Scaling of Large Data, Kaufmann, M. and Wagner, D.

(Eds.): *GD 2006, LNCS 4372*, pp.42–53 (2007).

[32] Shimizu, E. and Inoue, R.: A Generalized Solution of Time-distance Mapping, *Journal of Japan Society of Civil Engineers*, Vol.2004, No.765, pp.105–114 (2004) (in Japanese).

[33] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P.: *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press (1992).

[34] Skiena, S.S.: *The Algorithm Design Manual,* 2nd ed., Springer (2008).

[35] Dinas, S. and Bañón, J.M.: A Review on Delaunay Triangulation with Application on Computer Vision, *IJCSE*, Vol.3, No.2, pp.9–18 (2014).

[36] Yanagisawa, T.: *Kemanai no Bon Odori* (*Bon Odori of Kemanai*), privately published (1999) (in Japanese).

[37] Akita-Ken Hachirōgata-Machi Kyōiku Iinkai (Education Board of Hachirogata Town, Akita Prefecture) (Ed.): *Hitoichi Bon Odori Chōsa Hōkokusho* (*The Search Report on Hitoichi Bon Odori*), Akita-Ken Hachirōgata-Machi Kyōiku Iinkai (2005) (in Japanese).

[38] Kakizaki, T.: *Nishimonai Bon Odori ni Kansuru Shiken* (*Saihan*) (*A Personal Opinion about Nishimonai Bon Odori* (2nd ed.)), privately published (1974) (in Japanese).

[39] Gojōme Chōshi Hensan Iinkai (Gojome Town History Compilation Committee) (Ed.): *Gojōme Chōshi* (*The History of Gojome Town*), Gojome Town (1975) (in Japanese).

[40] Odashima, S.: Akita Kennan no Bon Odori, Sono Rekishi to Genzai (Bon Odori Dances of the Sourthern Part of Akita Prefecture, Their History and Present State), *Akita Minzoku*, Vol.36, pp.24–41 (2010) (in Japanese).

[41] Kuniyasu, H.: *Omono Gawa to Ushū Kaidō* (*Omono River and the Ushu Kaido Road*), Yoshikawa Kōbunkan (2001) (in Japanese).

[42] Ogachi-Machi Kyōdoshi Hensan Iinkai (Ogachi Town Local History Compilation Committee) (Ed.): *Ogachi Chōshi* (*The History of Ogachi Town*), Akita-Ken Ogachi-Gun Ogachi-Machi Kyōiku Iinkai (Education Board of Ogachi Town, Ogachi County, Akita Prefecture) (1961) (in Japanese).

[43] Kosaka, T.: *Nishimonai Bon Odori: Waga Kokoro no Genfūkei* (*Nishimonai Bon Odori: The Landscape of My Heart*), Kageshobō (2002) (in Japanese).

Katsubumi Tajima received his D.Eng. degree in electrical engineering from Tohoku University in 1998. He is a professor in the Cooperative Major in Life Cycle Design Engineeirng, Graduate School of Engineering Science, Akita University.

Hideo Tamamoto received his D.Eng. degree in electrical engineering from the University of Tokyo in 1976. He is currently a researcher at Tohoku University of Community Service and Science. His research interests include design-for-testability of logic circuits, archiving and handing-down technique for traditional folk dances.

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-Electronic-Computer Engineering, Graduate School of Engineering Science, Akita University.

Takeshi Shibata received his D.Eng. degree in electronic and computer system engineering from Akita University in 2012. He is currently an assistant professor in the Collage of Information and Systems, Muroran Institute of Technology. His research interests include virtual reality technique and archiving handing-down technique for traditional folk dances.

Madoka Uemura received her Ph.D. from the University of Tokyo in 2013. She is currently an associate professor in the Faculty of Education and Human Studies, Akita University.