Electronic Preprint for Journal of Information Processing Vol.30

Technical Note

Extraction of Feature Quantities Suitable for Distribution Visualization of Motion Capture Data

Takeshi Miura[1,a)]

Received: May 9, 2022, Accepted: July 6, 2022

Abstract: To handle a motion-capture (Mocap) data set such as a Mocap database, it is beneficial to grasp the overview of the motion-characteristic distribution of the set in advance. To easily grasp the overview, concisely visualizing the distribution using a scatter plot is effective. In this paper, we propose a new method to extract the feature quantities suitable for the above visualization from each of the Mocap data. The one-dimensional motion-speed time series is analyzed in the frequency domain. Consequently, two feature quantities representing the motion intensity and motion complexity are derived. It is shown in the scatter-plot-construction experiments that explicitly weighting each frequency value in the frequency domain is effective for extracting the characteristics specific to each motion category.

Keywords: motion capture, motion characteristic, visualization, scatter plot, frequency domain

1. Introduction

Nowadays, motion-capture (Mocap) data are used for many applications such as motion analysis, creating CG animations, etc [1]. There are some research examples focusing on handling Mocap data sets each of which includes the data of multiple motion categories, such as Mocap databases or Mocap archives [2], [3]. In such cases, it is beneficial to grasp the overview of a given Mocap data set in advance.

To easily grasp the overview of the motion-characteristic tendency of a given Mocap data set, concisely visualizing their distribution in a simple manner is effective. A scatter plot is known as one of the methods suitable for the above purpose. It is designed to encode the values of the variables of each data unit in the set as the vertical and horizontal coordinates of a data point, and visually emphasizes and characterizes the distribution of all the data points in a two-dimensional space [4], [5].

There are several examples in which scatter plots are used for visualizing the motion-characteristic distribution of given Mocap data sets [6], [7]. In most of them, the multidimensional feature vector representing the motion characteristics of each of the Mocap data is used. To encode the information on the multiple feature-vector coordinates as two scatter-plot coordinates, techniques of dimensionality reduction, such as principal component analysis (PCA) or multidimensional scaling (MDS), are used [7]. However, two problems occur in the dimensionality-reduction process. The first one is that the positions of all the data points change even when only a small part of the data set is changed (e.g., a small number of data are added or removed, or partially replaced). The second one is that the meanings of the vertical and horizontal axes cannot be known until the analysis is completed.

The above problems can be solved by essentially using only two motion-characteristic variables. In this paper, we propose a new method to extract two feature quantities suitable for visualizing the distribution of Mocap data in a scatter plot. We use the one-dimensional motion-speed time series proposed in Ref. [8] to extract the above quantities. This time series emphasizes only the information on the temporal motion-speed variation of the whole body, without considering the spatial arrangement of the body parts. By using this time series, therefore, motion characteristics can be concisely summarized with a small amount of information, i.e., into only two feature quantities.

We analyze the above time series in the frequency domain. As will be shown in Section 3, explicitly weighting each frequency value in the frequency domain is effective for extracting the motion characteristics specific to each motion category. Finally, the two feature quantities representing the motion intensity and motion complexity are derived.

We conduct experiments in which a Mocap data set including multiple motion categories is used to construct scatter plots. We compare the proposed method with the other conventional methods. The experimental results show that the proposed method can provide better characteristics for the grouping of the motion categories than those provided by the other methods.

2. Derivation of Feature Quantities

First, we obtain a whole-body motion-speed time series from the temporal variation of the positions of the principal joints (16 joints including end effectors: shoulders, elbows, wrists, fingers, knees, ankles, toes, neck and head) [8]. We apply a technique of the fast Fourier transform (FFT) [9] to the time series to extract the frequency-domain characteristics. In this case, the length of the time series is required to be a power of 2 [9]. To satisfy this condition, we pad the time series with zeros as follows [9]:

[1] Graduate School of Engineering Science, Akita University, Akita 010–8502, Japan
[a)] miura@mail.ee.akita-u.ac.jp

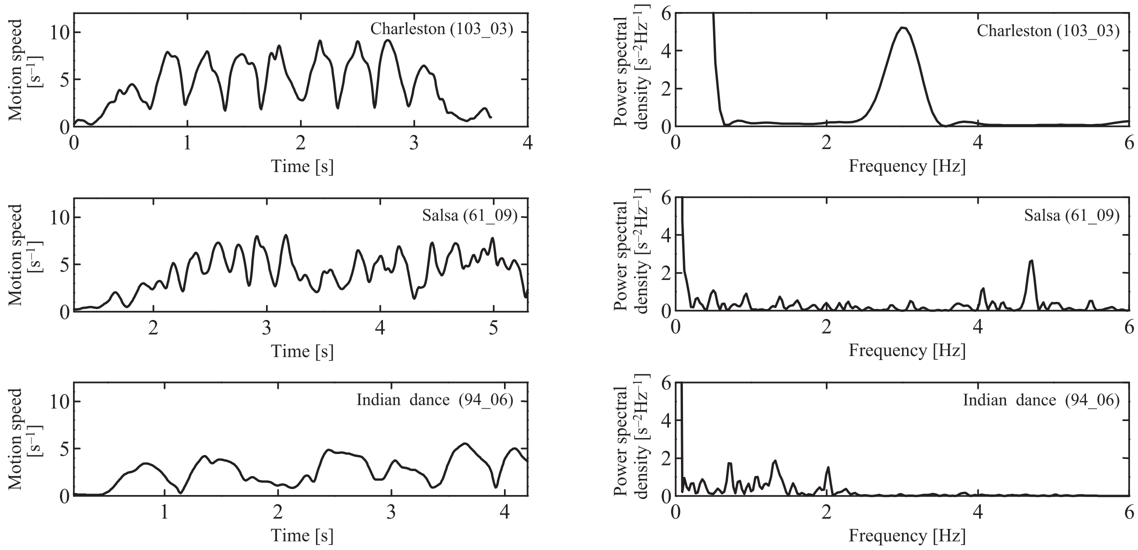


Fig. 1 Examples of motion-speed time series and their power-spectrum characteristics (Mocap data: downloaded from "CMU Mocap Database" [11]).

$$v(n) = \begin{cases} \dfrac{\sqrt{\sum_{i=1}^{J} \sum_{\gamma=x,y,z} \{p_{\gamma,i}(n+1) - p_{\gamma,i}(n)\}^2}}{\Delta t} & \\ & (1 \le n \le N) \\ 0 & (N < n \le N_{\mathrm{ZP}}) \end{cases} \quad (1)$$

where $p_{\gamma,i}(n)$ is the γ-cooridnate of the ith joint in the coordinate system fixed to the pelvis at the nth frame (γ: x, y or z, value of $p_{\gamma,i}(n)$: normalized by the body height [*1], time series of $p_{\gamma,i}(n)$: filtered by a Gaussian filter to eliminate jitter), J is the number of the principal joints ($J = 16$), Δt is the sampling time, N is the number of the frames of a given time series [*2] and N_{ZP} is the minimum power-of-2 integer satisfying $N_{\mathrm{ZP}} \ge N$ and $N_{\mathrm{ZP}} \ge 2{,}048$ [*3].

The power spectrum of a given motion-speed time series in the frequency domain is obtained as follows [10]:

$$P(m) = \begin{cases} \dfrac{\Delta t}{S_2}|V(m)|^2 & \left(m = 1 \text{ or } m = \dfrac{N_{\mathrm{ZP}}}{2} + 1\right) \\ \dfrac{2\Delta t}{S_2}|V(m)|^2 & \left(2 \le m \le \dfrac{N_{\mathrm{ZP}}}{2}\right) \end{cases} \quad (2)$$

$$V(m) = \sum_{n=1}^{N_{\mathrm{ZP}}} v(n) w(n) \exp\left\{-2\pi j \frac{(m-1)(n-1)}{N_{\mathrm{ZP}}}\right\}$$

$$w(n) = \begin{cases} \dfrac{1}{2}\left[1 - \cos\left\{\dfrac{2\pi(n-1)}{N-1}\right\}\right] & (1 \le n \le N) \\ 0 & (N < n \le N_{\mathrm{ZP}}) \end{cases}$$

$$S_2 = \sum_{n=1}^{N} w(n)^2$$

where $P(m)$ is the power spectral density, $w(n)$ is the Hanning window [*4] and $j = \sqrt{-1}$. Practically, $V(m)$ is obtained by FFT.

The physical dimension of $P(m)$ is $(\text{speed})^2/(\text{frequency})$ [10]. This can be interpreted as that $P(m)$ represents the degree of motion intensity at each frequency, and the total intensity can be obtained by integrating $P(m)$ over the whole frequency domain. In the integration of $P(m)$, we weight each $P(m)$ with the corresponding frequency value as follows:

$$q_{\mathrm{I}} = \log\left\{\sum_{m=1}^{N_{\mathrm{ZP}}/2+1} P(m) f_m^{k_1} \Delta f\right\} \quad (3)$$

$$f_m = (m-1)\Delta f, \quad \Delta f = \frac{1}{N_{\mathrm{ZP}} \Delta t}$$

where $f_m^{k_1}$ is the weight function for $P(m)$ and k_1 is the user parameter to adjust the strength of weight at each frequency [*5]. The above weighting is introduced to evaluate the occurrence of repetitive rapid motion-speed change, which enhances the impression of being intense. We adopt q_{I} as the first feature quantity.

Next, we derive the second feature quantity based on actual frequency-domain characteristics. **Figure 1** shows examples of motion-speed time series and their power-spectrum characteristics in the frequency domain. In the case of Charleston, the time series shows a relatively simple and regular speed variation. Its power spectrum gives a smooth shape having a single distinctive peak. On the other hand, the time series of Salsa shows a complex and irregular speed variation, and its power spectrum shows an irregular uneven pattern. Although the time series of the Indian dance also shows a complex and irregular waveform, the motion-speed value varies more slowly. As a result, its unevenness in the power spectrum is shown only in the low-frequency region.

The above characteristics suggest that motion complexity is reflected in the smoothness/unevenness of the power-spectrum curve. The smoothness/unevenness of a given curve can be evaluated by its curvature values. Therefore, we introduce the feature quantity below, in which the curvature of the power-spectrum curve is integrated over the whole frequency domain, to evaluate

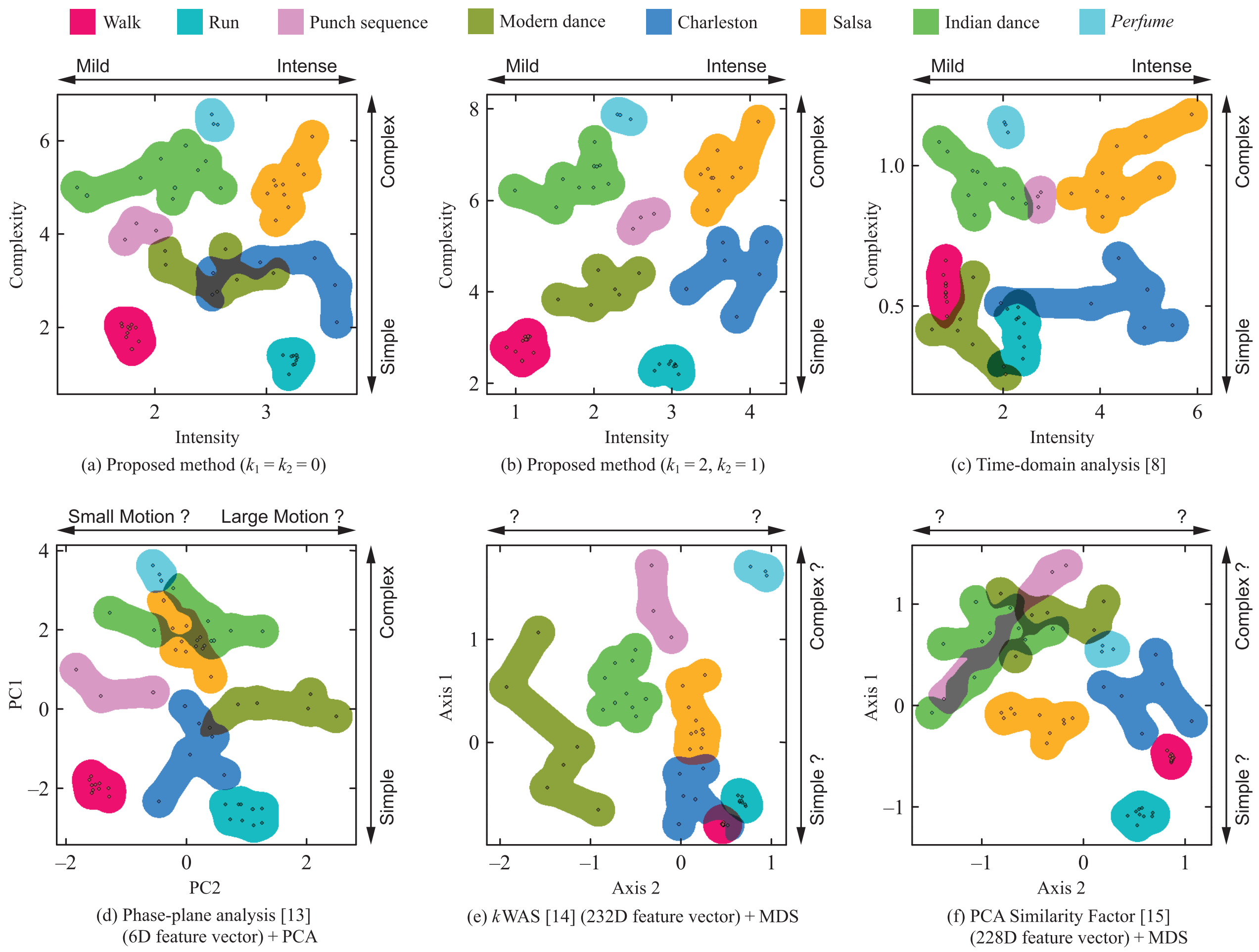
[*1] The original coordinate values are divided by the body height.

[*2] In Eq. (1), $v(n)$ is obtained in the region $1 \le n \le N$. In actual calculations, $v(n)$ is obtained only in the part of actual performance as will be mentioned in Section 3. Therefore, $v(N)$ can be obtained by using $p_{\gamma,i}(N+1)$ existing at the instant just after the performance. Otherwise, $p_{\gamma,i}(N+1)$ can be obtained by, e.g., linear extrapolation.

[*3] The condition $N_{\mathrm{ZP}} \ge 2{,}048$ is introduced to ensure enough frequency resolution for short time series. The frequency resolution is given as $\Delta f = 1/(N_{\mathrm{ZP}}\Delta t)$.

[*4] Another window, e.g., Hamming [10], can be used as the need arises.

[*5] The log transformation is used in the calculation of q_{I} to cover a wide range.



Motion capture data downloaded from "CMU Mocap Database" [11]
Walk (35_01 – 10), Run (09_01 – 09), Punch sequence (144_13, 14, 21), Modern dance (05_04, 07, 10, 11, 13, 14), Charleston (103_03 – 08), Salsa (61_01 – 10), Indian dance (94_01 – 10).
Motion capture data downloaded from "Perfume Global Site" [12]: *Perfume* (aachan, kashiyuka, nocchi).

Fig. 2 Feature-quantity distribution of Mocap data. In (d), (e) and (f), number of joints: $J = 19$ (waist and hips are added), number of PCs in (e) (or SVs in (f)): $k = 4$, dimensionality of the feature vector: 6 for (d), $(3J+1)k = 232$ for (e) and $3Jk = 228$ for (f), and the symbol "?" means "difficult to interpret."

the degree of motion complexity:

$$q_C = \log\left[\sum_{m=2}^{N_{ZP}/2} \frac{|c_2(m)|}{\{1 + c_1(m)^2\}^{3/2}} f_m^{k_2} \Delta f\right] \tag{4}$$

$$c_1(m) = \frac{P(m+1) - P(m-1)}{2\Delta f}$$

$$c_2(m) = \frac{P(m+1) - 2P(m) + P(m-1)}{(\Delta f)^2}$$

As in the case of q_I, we weight each of the curvature values in q_C [*6] at each frequency with the weight function $f_m^{k_2}$ (k_2: user parameter to adjust the strength of weight at each frequency) to enhance the difference in the distribution of uneven regions in the frequency domain. It is reasonable to think that the complexity and difficulty of motion sequence are higher as the motion frequency is higher, i.e., as the motion sequence is more quickly performed, and the above weighting conforms to this.

3. Results

This section presents the experimental results of the proposed method. We use Mocap data open to the public [11], [12]. In some data, periods in which the whole body is kept in a still state are included before and after the actual performance. We remove these periods and use only the part sandwiched between the n_1th and n_2th frames (n_1 and n_2: frames first and finally satisfying $v(n) > 0.75v_m$, v_m: mean speed). We compare the obtained results with those obtained by the other four conventional methods: time-domain analysis [8], phase-plane analysis [6], kWAS [13] and PCA Similarity Factor [14].

Figure 2 shows the obtained scatter plots of the Mocap data. The feature-quantity distribution of 57 Mocap data selected from 8 motion categories was visualized. Each of the colored areas represents a motion-category region (estimated by the Bubble Sets method [15]). (a) and (b) were obtained by the proposed method. (a) was obtained by setting $k_1 = k_2 = 0$ in Eqs. (3) and (4), i.e., without the frequency weighting, whereas (b) by setting $k_1 = 2$ and $k_2 = 1$, i.e., using the adjusted frequency weighting (adjusted by trial and error). In (a), the areas of several motion categories overlap. On the other hand, there is no overlap in (b). This suggests that the frequency weighting is actually effective in extracting motion characteristics specific to each motion category.

*6 The log transformation is used in the calculation of q_C to cover a wide range, as in the case of q_I.

Table 1 Evaluation of grouping charadteristics in the scatter plots.

	Proposed method ($k_1 = 2, k_2 = 1$)	Time domain analysis	Phase plane analysis	kWAS	PCA Similarity Factor
DB Index	0.494	0.758	1.293	0.651	1.801
Emprical Accuracy	1.000	0.877	0.754	0.965	0.825

(c) in Fig. 2 was obtained by the time-domain analysis, which provides two feature quantities almost identical to those of the proposed method (i.e., intensity and complexity) [8]. Multiple motion-category overlaps are seen in (c). This incomplete grouping was probably caused by the fact that each of the frequency values was not explicitly weighted in the time-domain analysis, and thereby motion characteristics specific to each category was not sufficiently evaluated.

(d), (e) and (f) in Fig. 2 were obtained by the phase-plane analysis, kWAS and PCA Similarity Factor, respectively. These methods extract a multidimensional feature vector from each of the Mocap data (dimensionality: (d): 6, (e): 232 and (f): 228). The distribution in a multidimensional space was projected on a two-dimensional scatter plot by a dimensionality-reduction technique (PCA or MDS). In all of (d), (e) and (f), multiple motion-category overlaps, i.e., incomplete groupings, are seen. As for the coordinate axes of the obtained scatter plots, it was difficult to explicitly interpret their meanings, especially in the cases of (e) and (f) in which extremely high-dimensional feature vectors were used. Such a problem does not occur in the application of the proposed method that provides only two feature quantities.

We quantitatively evaluate the grouping characteristics in each of the obtained scatter plots. Specifically, we use the Davies-Bouldin (DB) index defined as a function of the ratio of the within-category scatter to the between-category separation [16]. A lower DB-index value means that the grouping is better. **Table 1** shows the results. The proposed method (with $k_1 = 2$ and $k_2 = 1$) gave the lowest DB-index value, i.e., the best grouping.

To further verify the effectiveness of the proposed method, we perform the leave-one-out cross validation [17] by applying the 1-nearest-neighbor classifier to each scatter plot in Fig. 2. In the cases of (d), (e) and (f), the validation is done not in the multidimensional space but in the two-dimensional scatter plot. The results are shown in Table 1. The proposed method gave the highest, i.e., best, empirical-accuracy value.

As shown in the above results, the proposed method gave better characteristics than those of the other methods. This suggests that the proposed method is more suitable for concisely visualizing the motion-characteristic distribution of multiple motion categories in a two-dimensional scatter plot, which helps easily grasp the overview of the distribution.

4. Conclusion

The main contribution of this paper is to provide the feature quantities suitable for concisely visualizing the overview of the motion-characteristic distribution of a given Mocap data set, using a scatter plot. It is shown in the experiments that the derivation of the feature quantities by the frequency-domain analysis, in which each frequency value is explicitly weighted, is effective. The derived feature quantities give better characteristics in the grouping of multiple motion categories than the other conventional methods. To clarify the application range of the proposed method will be the subject of future work.

Acknowledgments This study was supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) Grant Number JP18K11981.

References

[1] Kitagawa, M. and Windsor, B.: *MoCap for Artists*, Focal Press (2008).
[2] Tanco, L.M. and Hilton, A.: Realistic Synthesis of Novel Human Movements from a Database of Motion Capture Examples, *Proc. IEEE Workshop on Human Motion*, pp.137–142 (2000).
[3] Mahmood, N., Ghorbani, N., Troje, N.F., Pons-Moll, G. and Black, M.J.: AMASS: Archive of Motion Capture as Surface Shapes, *Proc. 2019 IEEE/CVF Intl. Conf. Computer Vision*, pp.5441–5450 (2019).
[4] Munzner, T.: *Visualization Analysis & Design*, CRC Press (2015).
[5] Sarikaya, A. and Gleicher, M.: Scatterplots: Tasks, Data, and Designs, *IEEE Trans. Visualization and Computer Graphics*, Vol.24, No.1, pp.402–412 (2017).
[6] 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).
[7] Miura, T., Kaiga, T., Shibata, T., Uemura, M., Taijma, 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).
[8] Miura, T., Kaiga, T., Matsumoto, N., Katsura, H., Shibata, T., Tajima, K. and Tamamoto, H.: Characterization of Motion Capture Data by Motion Speed Variation, *IEEJ Trans. Electronics*, *Information and Systems*, Vol.133, No.4, pp.906–907 (2013).
[9] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P.: *Numerical Recipes in Fortran 77*, *The Art of Scientific Computing*, 2nd ed., Cambridge University Press (1992).
[10] Heinzel, G., Rüdiger, A. and Schilling, R.: Spectrum and Spectral Density Estimation by the Discrete Fourier Transform (DFT), Including a Comprehensive List of Window Functions and Some New Flat-top Windows, *Technical Report, Max Planck Institute for Gravitational Physics*, pp.1–84 (2002).
[11] CMU Mocap Database, available from ⟨http://mocap.cs.cmu.edu⟩.
[12] Perfume Global Site, available from ⟨http://perfume-global.com⟩.
[13] Li, C. and Prabhakaran, B.: A Similarity Measure for Motion Stream Segmentation and Recognition, *Proc. 6th Intl. Workshop on MDM*, pp.89–94 (2005).
[14] Krzanowski, W.J.: Between-groups Comparison of Principal Components, *Journal of the American Statistical Association*, Vol.74, No.367, pp.703–707 (1979).
[15] Collins, C., Penn, G. and Carpendale, S.: Bubble Sets: Revealing Set Relations with Isocontours over Existing Visualizations, *IEEE Trans. Visualization and Computer Graphics*, Vol.15, No.6, pp.1009–1016 (2009).
[16] Davies, D.L. and Bouldin, D.W.: A Cluster Separation Measure, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.PAMI-1, No.2, pp.224–227 (1979).
[17] Konishi, S.: *Introduction to Multivariate Analysis*, CRC Press (2014).

Takeshi Miura received his D.Eng. degree in electrical engineering from Hokkaido University in 1998. He is currently an associate professor in the Cooperative Major in Sustainable Engineering, Graduate School of Engineering Science, Akita University.