Perceptual Characteristics by Motion Style Category

Hye Ji Kim and Sung-Hee Lee
{kimhyejee923, sunghee.lee}@kaist.ac.kr

Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea

Abstract

Motion style is important as it characterizes a motion by expressing the context of the motion such as emotion and personality. Yet, the perception and interpretation of motion styles is subjective and may vary greatly from person to person. This paper investigates the perceptual characteristics of motion styles for a wide range of styles. After categorizing the motion styles, we perform user studies to examine the diversity of interpretations of motion styles and the association level between style motions and their corresponding text descriptions. Our study shows that motion styles have different interpretation diversity and association level according to their categories. We discuss the implications of these findings and recommend a method of labeling or describing motion styles.

CCS Concepts

• ***Computing methodologies*** → ***Computer graphics; Animation; Graphics systems and interfaces; Perception;***

Keywords: Motion recognition, Motion description, Motion style, Motion style perception

1. Introduction

Human motion is generally regarded as a combination of two aspects: Content and style [THB06]. Motion content is associated with the nature of a motion, while style represents a particular way that a motion is performed. Motion style is important in defining the characteristics of a motion, as it shows the context of a motion, such as the mood and behavior of a character. Yet, the perceiving and interpreting of motion styles is rather subjective, and thus may vary greatly from person to person.

In this paper, the perceptual characteristics of a wide range of stylized motions are investigated through user studies. First, we grouped overall motion styles into six categories according to the principles of Stanislavski's method acting. Some categories concern extrinsic aspects of motion style, such as verbal description of how a motion is made, while the others are intrinsic, such as the inner state of an actor. Then, through user studies, we examined the diversity of interpretations of motion styles by category and the association level between style motions and their corresponding text descriptions.

The results show that 1) motions of intrinsic style categories (e.g., character, personality, emotion) have higher interpretation diversities than extrinsic style categories (e.g., action, objective,motivation; 2) extrinsic style categories have a stronger association level with respect to their original text labels than intrinsic style categories, but one particular intrinsic category also shows a strong association level; 3) the association level of extrinsic categories drops sharply with respect to user descriptions, while it is more or less the same for intrinsic categories.

2. Related Work

Motion styles have been explored in computer graphics mostly from the perspectives of generation, retrieval, and perception. Among them, we discuss previous work on perceiving stylistic motions.

A number of studies have viewed motion styles as individual characteristics [NWAW10] while others have investigated emotion as a driver of motion style [RCB98]. For example, [NLK*13] defined style as subtle details used to detect and interpret human motion, analyzing the correlation between the perception intensity of emotion and angles and velocities of motion.

In avatar-mediated communication, motion style has been studied as a means of non-verbal expression. [HRZ*13] proposed the concept of distinctiveness and attractiveness that make motions more recognizable. [YOHO17] explored the visual realism and plausibility of body shapes with respect to understanding motion style variations. They regarded the motion style as a critical factor to differentiate motions and tried to find proper criteria.

[KHC*13] suggested that human motion conveys social traits and intentions in communication, and that the visual features of the human body and motion are crucial for the perception of social information. [ONB17] investigated the expression ability of non-verbal communication across agent embodiment.

Researchers have also investigated the connection between motion styles and personality. [SN17] defined motion style as gesture performance, and suggested that gesture performance is relevant to expressing or recognizing personality. A major problem in the style

© 2019 The Author(s)
Eurographics Proceedings © 2019 The Eurographics Association.

delivered by EUROGRAPHICS DIGITAL LIBRARY www.eg.org diglib.eg.org

DOI: 10.2312/egs.20191000

Figure 1: *Example motion styles in Parry's database.*

Table 1: *100 labels of the dataset used in our study*

100 Ways of Walking	
CHAR	chimpanzee, dancer, boxer, Ballerina, swat, scuba diver, royal guard, pirate,detective, penguin, runway model, cowboy, Charlie Chaplin, clown, tourist,zombie, robot, mummy, Frankenstein's monster, waiter, marionette, ultimate fighter, mime, Broadway musical
PER	feminine, the old man, hippy, teenage hooligan, old timey fighter, generic,sassy,timid, confident,strongman
EMO	angry, fainting, paralyzed by fear, all is lost, collapsing
ACT	skip, pushing, pulling, march, crawl, leg asleep,backwards, kicking, double bounce, lunging, creeper, stumble, strut
MOT	on vacation, tight rope, on hot coals, cold, hot, on the phone, stepping over stuff, hurried, emergency, looking at art, in a wind storm, limbo, trying to hide, wedgie, sick, in the rain, look at me,drunk, in the dark, lost in thought, being chased, chasing, carry weights
OBJ	sneak, sword fight, moonwalk, getting married, fancy feet, speed Walk,toddler,sweeping,ducking, football training, sleepwalk
excluded	bad posture, good posture, sore back,hot tea, something in shoes, through a crowd, stepped in something, secret run, dance fight, line dance, flower child, helicopter, stomp clap, generic

perception study is how to quantify style factors of a human so that it is generated in an expected way. [DKD*16] attempted to formally define the mapping between the characteristic parameters of human movement and different personality traits in an effort to synthesize motions with personality. The OCEAN personality model was used as a reference to categorize personality.

Compared to studies that have focused on particular aspects of motion styles, such as emotion or personality, we deal with a broader range of motion styles, categorize them, and investigate the difference in perceptual characteristics across the category.

3. Style Categorization and Dataset

Artists and researchers in the fields of performing arts, movie, and computer animation have explored ways of acting to create interesting and expressive characters, and have proposed a number of rules and concepts. Among them, *method acting* created by Stanislavski, is a technique of acting by which an actor can effectively express the thoughts and emotions that a character has, and thus create a lifelike performance [HW13]. Based upon the idea that in order to develop an emotional and cognitive understanding of their roles, method acting stresses that actors should use their own experience to identify personally with their characters. In computer animation, animators, like actors, also apply the method acting principles to create animation. Therefore, the criteria of method acting provide good references for categorizing motion styles.

According to Stanislavski's system, we categorized motion styles into six groups: ***character (CHAR), personality (PER), emotion (EMO), action (ACT), objective (OBJ), motivation (MOT)***. These six groups are required to build a character's identity and action, thus are directly linked to motion style. Specifically, **character** refers to the subject that inhabits, shapes, and determines the story. It can be a vocation, a public figure, or even an animal. **Personality** is what characterizes a character, and includes gender, age, and the nature of the character. **Emotion** is the internal feeling of a character, while **action** is the external, physical movement. **Objective** is the goal that a character wants to achieve. Emotion and action are both the result of the pursuit of an objective. **Motivation** is what causes a character to act in certain way. Character, personality and emotion categories are internal attributes that define the nature of the subject, while action, objective and motivation are external attributes for a particular style of a motion. Therefore, we grouped CHAR, PER, EMO as the extrinsic categories, and ACT, OBJ, MOT as the intrinsic categories.

To investigate the characteristics of human perception of motion styles, we used the video dataset from "100 Ways of Walk" by Kevin Parry (youtu.be/HEoUhlesN9E). Figure 1 shows sample motion styles from the dataset. The duration of each video clip is 2 to 3 seconds. Unlike other popular motion databases such as CMU dataset, HDM05 and UCF 101, this dataset focuses on very large possibilities of style variations for a single category of motion content: Walking. Table 1 shows the entire list of 100 walking styles. The dataset contains simple styles that can be described by only a single verb or a short phrase as well as more subtle styles that are related with character, personality, emotion, and situation (e.g., Charlie Chaplin, timid, all is lost). The wide range of motion styles in Parry's dataset provides motion styles in the above six groups.

4. User Study 1: Interpretation Diversity of Motion Styles by Category

Our first question was whether the diversity of interpreting motion styles differed by category. To quantitatively measure the interpretation diversity, we used a method that estimates the semantic distance between two text descriptions on the same video clip. Specifically, we measured the semantic distance between the original label (dubbed *intended motion style* hereafter) and the participants' description (dubbed *perceived motion style*).

In the user survey, participants were asked to observe a video motion clip and "write two to four descriptions that can be used as querying text for the motion." All 100 video clips in the dataset were evaluated. In the video, the facial regions as well as the text labels were hidden to exclude the effect of facial expression so that the participants could only judge from the body motions. Twenty four graduate students participated in the survey and wrote descriptions in their local language, which were later translated into English. The descriptions could include duplicate expressions proposed by different participants.

For calculating the semantic distance, we used the word2vec model, which embeds words to a high dimensional space and identifies semantic relationships between words [MCCD13]. The

© 2019 The Author(s)
Eurographics Proceedings © 2019 The Eurographics Association.

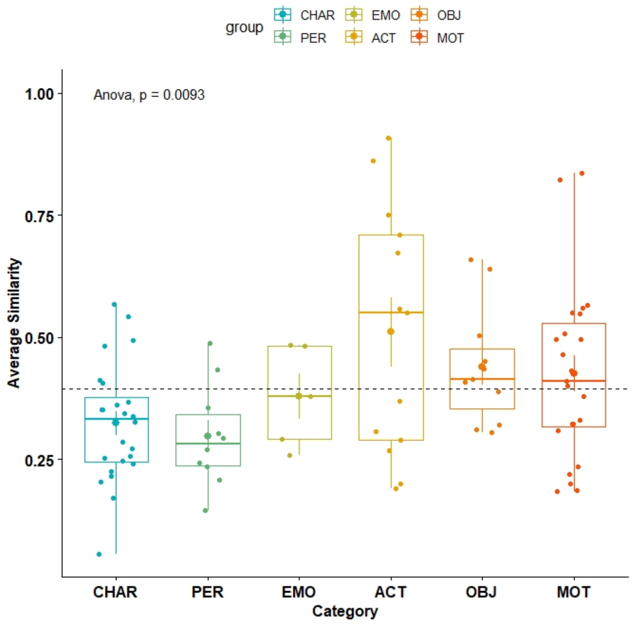

Figure 2: *Average similarity of each category.*

word2vec algorithm learns the semantic distance between two words based on their distance and the co-occurrence frequency in a corpus used for the training. Specifically, we used a tool provided by Google Code Archive. We used the cosine similarity for the semantic distance between words, thus 1 indicates the highest similarity, and 0 is the lowest. In the cosine similarity, the dot product of two words was calculated, and the similarity value ranges [0, 1], with 1 indicating the highest similarity.

To apply the word2vec model, we replaced the labels and descriptions with unitary words. For example, the *robot* style shows walking motion at a short and quick steps. Some participants describe the style as "playful walk", which was replaced with the word *playful*. In case of *march*, most of the participants described the style as "walk bravely", and it was replaced with *bravely*. However, some labels and descriptions could not be represented with a single word included in the training corpus. A short phrase such as *through a crowd* and *stepped in something* and a compound word such as *bad posture* and *flower child* could not be represented with unitary words. We removed such motions and descriptions altogether from the analysis. After this filtering, a total of 86 motions were used for the first description survey.

Figure 2 shows the average similarities between the intended and the perceived styles in each category. The result shows that ACT has the highest similarity of 0.51, followed by OBJ (0.44) and MOT (0.43), while EMO (0.38), CHAR (0.32), and PER (0.3) show lower average similarities. In order to verify statistical significance, we performed one-way Analysis of Variance (ANOVA) test on the average similarity results with *the predictor: six categories* and *the dependent variable: different perception.* Resulting group p-value (p = 0.01) is less than the significance level 0.05, suggesting that there are significant differences between the groups, i.e., the perceptual characteristics are different according to category.

These results showed that participants perceived extrinsic categories as intended by the actor better than the intrinsic categories. As the ACT category styles directly concern the way a motion is taken, the perceived styles turn out to be similar to the intended styles. OBJ and MOT are also extrinsic factors that drive movement. For example, the video clip for *in the dark* (MOT) shows a person walking and groping blindly. By contrast, CHAR, PER, and EMO are rather intrinsic features that concern the inner states of the performer and are thus more indirectly related to styles.

5. User Study 2: Association Level between Motion Style and Text Description by Category

The second user study investigates another area of style perception. Given a style motion and its text description, what is the degree of approval of the description by people? Is it different per style category? This question has practical importance for the satisfaction of text-based style motion retrieval, and also for the recognition or prediction of the motion in distinctive styles. We will call the extent that people approve a given description as *the association level* between the two.

To examine this tendency, we executed two sets of experiments. First, we examined the association level between the original label and its motion, followed by the second part to investigate the association level between the user descriptions obtained in user study 1 and its motion. In this survey, participants (24 graduate students) viewed both a motion and its label without being informed that it is the intended label. They were asked to rate the question "Do you approve of the label on this motion?" on a 5-point Likert scale {1: strongly disagree, 5: strongly agree}. This range gives the weight of the responses. We randomly selected five motions from each category, then chose ten descriptions from each motion. The selected motions were: *on hot coals, hurried, limbo, trying to hide, sick* (MOT); *angry, fainting, paralyzed by fear, all is lost, collapsing* (EMO); *skip, pushing, march, crawl, double bounce* (ACT); *getting married, fancy feet, speed walk, toddler, ducking* (OBJ); *ballerina, broadway musical, pirate, detective, penguin* (CHAR); *feminine, the old man, hippy, teenage hooligan, old timey fighters* (PER). For the second experiment, we made five question sets, each set containing two descriptions per motion (i.e., a total of 60 questions per set), and each of the 24 participants evaluated two sets.

Table 2: *Approval scores on original labels and user descriptions. Top two highest scores are marked in bold.*

Category	Original Label Ave. (std)	User Description Ave. (std)
CHAR	**3.96 (0.65)**	**4.00 (0.33)** Δ: +0.04
PER	3.64 (0.43)	3.61 (0.43) Δ: -0.03
EMO	3.84 (0.35)	**3.78 (0.49)** Δ: -0.06
ACT	**4.30 (0.62)**	3.66 (0.76) Δ: -0.34
OBJ	3.92 (0.38)	3.64 (0.17) Δ: -0.28
MOT	3.92 (0.46)	3.74 (0.48) Δ: -0.18

Table 2 shows the average approval scores in each category. Approval scores on the original label are similar with or higher than those on the user descriptions, showing that original labels, or intended styles, are better approved than other people's perceived styles. For the original labels, extrinsic styles (ACT, OBJ,

© 2019 The Author(s)
Eurographics Proceedings © 2019 The Eurographics Association.

MOT) have higher approval scores than PER and EMO. This tendency is well aligned with the interpretation diversity, showing that categories of smaller interpretation diversity have higher approval scores on the labels. A remarkable difference is observed at CHAR. It has high interpretation diversity, but it also has the second highest average approval score. In other words, it is not easy to think of CHAR-type descriptions when humans view a motion, but they approve its descriptions if provided. For example, with respect to Charlie Chaplin-style walking, people tend to interpret it as a waddling walk, but they approve the description of "Charlie Chaplin" if it is provided.

The average approval scores on the perceived styles show a noticeable feature as well. They are similar to the average approval scores of the intended styles for the intrinsic style categories, but they drop sharply for the extrinsic style categories. This can be explained in connection with the small diversity in interpretation and high approval score for the original label for extrinsic style categories. Because the original labels are recognized as explaining the motion very well, other descriptions provided by participants tend to receive lower approval scores. In contrast, in the PER and EMO categories, the original labels are only weakly associated with the style motions, making them indifferent from user descriptions.

Table 3: *Percentage of responses of either "Agree" or "Strongly agree."*

	CHAR	PER	EMO	ACT	OBJ	MOT
Original	72	62	74	**84**	**78**	72
User's	72	63	71	64	63	68

Overall, motion style reveals quite a subjective nature. Table 3 shows the percentage of participants who marked either "strongly agree" or "agree" on the original label and participant descriptions for user study 2. It is quite surprising that except for ACT (84%) and OBJ (78%) categories for the original label, more than 25% of people do not agree with the text descriptions. This shows the high subjectivity of style recognition and inherent difficulty of expressing intended styles with motions.

6. Discussion and Future Work

In this paper, we grouped motion styles into the extrinsic and intrinsic categories, and investigated different characteristics of the interpretation diversity and association level between motion and text descriptions according to the motion style category.

This study not only provides us with deeper understanding of human perception of motion styles but it also offers lessons for effectively labeling style motions. A style of a motion can often be described from multiple perspectives. For example, one motion can be labeled as "sneak" (ACT), "thief" (CHAR), or "walk silently" (MOT). Our finding suggests that ACT-type descriptions will be approved more than other categories, and thus, if possible, ACT should be preferred. Another recommended strategy of describing a motion style is to use multiple labels in different categories. Instead of applying only a couple of labels, describing a motion style from different categorical perspectives will increase the ways a style motion is retrieved and utilized.

Our work has some limitations and overcoming them remains interesting future work. The Parry's database was created by one person, and so personal preference might have been included in motions and labels, which is somewhat inevitable due to the subjective nature of motion styles. In addition, the sample sizes of some categories (e.g., EMO) were small. An expanded database containing multiple performers' motions and labels would allow for more objective results. In future work, we aim to build a database focusing on style parameters such as inner drives and investigate further on the perceptual aspects of human motion styles.

Acknowledgement

We thank Kevin Parry for allowing us to use and edit "100 Ways of Walk" dataset. This work was supported by Giga Korea Project (GK17P0200) and Basic Science Research Program (NRF-2017R1A2B2006160) funded by MSIT, Korea.

References

[DKD*16] DURUPINAR F., KAPADIA M., DEUTSCH S., NEFF M., BADLER N. I.: Perform: Perceptual Approach for Adding OCEAN Personality to Human Motion using Laban Movement Analysis. *ACM Trans. Graph. 36*, 4 (Oct. 2016). 2

[HRZ*13] HOYET L., RYALL K., ZIBREK K., PARK H., LEE J., HODGINS J., O'SULLIVAN C.: Evaluating the Distinctiveness and Attractiveness of Human Motions on Realistic Virtual Bodies. *ACM Transactions on Graphics 32*, 6 (2013), 1–11. 1

[HW13] HAYES D., WEBSTER C.: *Acting and Performance for Animation*, 1st ed. Focal Press, 2013. 2

[KHC*13] KIISKI H., HOYET L., CULLEN B., O'SULLIVAN C., NEWELL F. N.: *Perception and Prediction of Social Intentions from Human Body Motion*. SAP '13. 2013. 1

[MCCD13] MIKOLOV T., CHEN K., CORRADO G., DEAN J.: Efficient Estimation of Word Representations in Vector Space. *CoRR abs/1301.3781* (2013). 2

[NLK*13] NORMOYLE A., LIU F., KAPADIA M., N. I. B., JÖRG S.: The Effect of Posture and Dynamics on the Perception of Emotion. *Proceedings of the ACM Symposium on Applied Perception - SAP '13* (2013), 91. 1

[NWAW10] NEFF M., WANG Y., ABBOTT R., WALKER M.: Evaluating the Effect of Gesture and Language on Personality Perception in Conversational Agents BT - Intelligent Virtual Agents: 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings. 222–235. 1

[ONB17] O'SULLIVAN S. A. C., NEFF M., BESKOW J.: Mimebot—Investigating the Expressibility of Non-Verbal Communication Across Agent Embodiments. *ACM Transactions on Applied Perception 14*, 4 (2017), 1–13. 1

[RCB98] ROSE C., COHEN M. F., BODENHEIMER B.: Verbs and adverbs: Multidimensional Motion Interpolation. *IEEE Computer Graphics and Applications 18*, 5 (1998), 32–40. 1

[SN17] SMITH H. J., NEFF M.: Understanding the Impact of Animated Gesture Performance on Personality Perceptions. *ACM Transactions on Graphics 36*, 4 (2017), 1–12. 1

[THB06] TORRESANI L., HACKNEY P., BREGLER C.: Learning Motion Style Synthesis from Perceptual Observations. *Neural Information Processing Systems* (2006), 1393–1400. 1

[YOHO17] YINXUAN SHI, ONDREJ J., HE WANG, O'SULLIVAN C.: Shape up! Perception based Body Shape Variation for Data-driven Crowds. *2017 IEEE Virtual Humans and Crowds for Immersive Environments (VHCIE)* (2017), 1–7. 1

© 2019 The Author(s)
Eurographics Proceedings © 2019 The Eurographics Association.