2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)
Beijing, China, November 6-9, 2018

A robot with style: can robotic attitudes influence human actions?

Vannucci F., Di Cesare G., Rea F., Sandini G., Sciutti A.

Abstract— **The style of an action, i.e. the way it is performed, has a strong influence on interaction between humans. The same gesture has different consequences when it is performed aggressively or kindly, and humans are very sensitive to these subtle differences in others' behaviors. In this work we investigated how to endow a humanoid robot with behaviors expressing different vitality forms, by modulating robot action kinematics and voice intonation. Drawing inspiration from human voice and motion, we modified a passing action and a passing voice command performed by the robot to convey an aggressive or kind attitude. In a series of experiments we demonstrated that the humanoid was consistently perceived as aggressive or kind. Human behavior changed slightly in response to the different robot attitudes and was characterized by faster responses to robot utterances than to robot actions. The opportunity of humanoid behavior to express vitality enriches the array of nonverbal communication that can be exploited by the robots to foster seamless interaction. Such behavior might be crucial in emergency and in authoritative situations in which the robot should instinctively be perceived as assertive and in charge, as in case of police robots or teachers.**

I. INTRODUCTION

In human-human interaction, fast and efficient collaboration is promoted by non-verbal communication[1], [2]. Implicit signals such as details in the motion of the body, gaze direction or voice features are typically exchanged during a joint task between two humans. The correct exchange of these signals, greatly enhances the quality of the interaction by revealing the partner's goals, intentions, desires[3], [4] or effort [5] and even discloses interactant`s emotional status [6], [7]. Actions can be performed gently, vigorously, or rudely, in general, as a function of the mood driving them [1], [8]. These different ways of moving are called *vitality forms* and play a crucial role in social relations.

Vitality forms influence the perception and behaviour of the interacting partner in human-human exchanges, communicating immediately and intuitively the attitude of the other, enabling for a fast and immediate adaptation. The same action, i.e. passing an object, acquires a different meaning, also in terms of urgency or importance, when associated with different vitalities. Recent neurophysiological findings show correlation in the activation of the dorso-central insula in presence of the perception of different vitality forms. More specifically, this brain area discriminates rude from gentle behavior – both from action observation or from the perception of different voice intonations [9]. In addition, this brain area is activated, not only when people perceive different vitality forms, but also when themselves express these forms of communication towards others. During social interactions, vitality forms expressed by an agent modulate the motor behavior of the receiver. In particular, when someone asks us something through voice or gestures, his/her positive or negative attitude modulates our subsequent motor response [10]. The expression and the recognition of vitality forms allow people to be socially connected with others, communicating their own mood or attitude and, consequently, to understand those of others.

Research in robotics has often focused on generation and execution of human-like movements in the attempt of creating communicative actions (e.g., [11][12], [13]. Although the emotional aspect has been reproduced mainly through the use of facial expressions [14], [15] there are also several attempts to communicate affective states with motion[1], [16]. For example different authors proposed to generate humanoid motions on the basis of the Laban Movement Analysis [13], that describes the emotion conveyed by movement using features such as velocity, curvature and acceleration (e.g. [14]). However, there are no studies addressing the issue of expressing vitality in a humanoid behaviors. The challenge we address in this work is to create robot behaviors that can achieve a goal while communicating various vitality forms, by exploiting a modulation of the kinematics of the motor act or a of the robot's voice.

The aim of the present study is twofold: first, to generate goal-oriented movements and voices in a humanoid robot conveying gentle and rude vitality forms; second, to assess the subjective evaluation and the behavioral responses of human participants to these robot behaviors during two experiments. In the first step, by using a motion tracking system, we recorded the human kinematic relative to a gesture i.e. passing an object, performed in gentle and rude ways and we remapped it into the joint space of the robot. Additionally, using a text to speech synthesizer, we produced a robotic voice pronouncing an action verb (take it) in gentle and rude way. Then we carried out a subjective evaluation experiment showing videos of both a human actor and the iCub performing the same movements with gentle or rude vitality forms or pronouncing the same action verbs with the same vitality forms. Last, we carried out a kinematic study aiming to analyze the effect of vitality forms expressed on participants during a real human-robot interaction.

II. METHODS

To study how humans perceive action vitality forms expressed by a robot, we developed a *motor control system* that moves the robot end-effector with a velocity profile respecting the regularities of the human motion [19]. Exploiting this system, we generated a human-like movement for the humanoid iCub robot, refining it to

978-1-5386-7282-2/18/$31.00 ©2018 IEEE

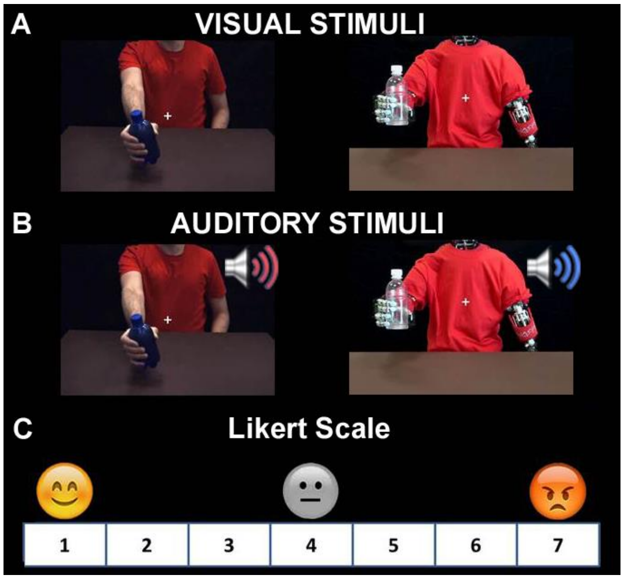

Figure 1. Here are displayed sample frames of the videos shown to participants during action (A) and speech (B) vitality forms perception. The robot is wearing a red shirt to make its appearance as close as possible to the actor's. At the bottom, the rating scale for the experiment (C).

include vitality forms. To create this movement, first, we recorded the kinematic data of a trained actor passing an object with gentle and rude vitality forms. Second, we remapped the captured motion into a movement of a kinematic chain of two links, 4 degrees of freedom (1 at the elbow and 3 at the shoulder). This kinematic representation resembles the kinematic chain of the humanoid iCub with a reasonable level of approximation [20]. The movement is remapped into robot kinematics by keeping into account the difference between actor's and robot's arm length. We obtain a list of joint positions that are reached in constant interval time (5ms). We compared the movement remapped from actor's action with the same movement generated by the control module, realizing that 1) the resulting gentle movement approximated well the human action 2) the aggressive movement performed by the robot and generated by the control module involved smaller elbow displacement than the original human motion to achieve similar end-effector movement. This led to a seemingly unnatural humanoid behavior. Therefore, we opted to replicate the actor's movement by directly mapping the motions at the joints from the human to the robot. This allowed to generate iCub's motion, which maintained accurate timing while exhibiting the natural communicative aspects associated with the movement of the elbow also in the aggressive condition.

To generate a robotic voice we recorded the voice of a human actor that pronounces the following motor command: "take it" in rude and gentle forms. We then manipulated some physical properties e.g. the pitch and duration (see Results for more details, Cool Edit Pro Software). Finally, the intensity of action verbs was equated for loudness in order to match the corresponding gentle and rude vitality forms of the human voice.

To verify that action and speech vitality forms would be perceived in the same manner as those produced by a human actor, we performed a subjective evaluation experiment. 20 participants (5 males, mean age: 24, SD: 2 years) were presented with video-clips showing human or robot expressing gestures or voice with vitality forms. Both visual and auditory stimuli were presented in two forms: gentle or rude. After visual or auditory stimulus perception, participants were immediately requested to indicate on a Likert scale (7 points: 1 very gentle - 7 very rude; Fig 1C) the vitality forms perceived.

After this subjective evaluation, we decided to carry out a behavioral study to test the same stimuli during a real human-robot interaction. The experimental setting is depicted in Fig 2. After the iCub action execution (passing the object) or word pronunciation ("take it") participants had to take a ball held by the robot at approximately 30 cm from their right hand. The participant comfortably sat in front of the robot with small headphones to hear the robotic voice indications, covered by hearing protectors to avoid experimental biases due to the noise of the moving motors. Between the participant and the robot, we placed a small table with marks indicating the starting position of the right hand and two different targets (yellow and orange) on which the ball had to be placed by the participant (Fig. 4B). The robot performed one action (pass the object) or pronounced one action verb in Italian language ("prendi" that is "take it"), with two different vitality forms. The two actions and voices were designed to show rude and gentle robot behavior towards the participant, inspired by the protocol by Di Cesare et al. [10]. The face of the robot was covered since the salience of the information had to be conveyed only either by the action or by the speech. Participants were instructed to replace the ball in the hand of the robot with their left hand.

Before the beginning of the experiment, each participant

Figure 2. Experimental Setup included the working table with two target areas (yellow and orange) and the humanoid robot iCub standing in front of the subject. The face of the robot was covered to avoid bias regarding eyes or facial information.

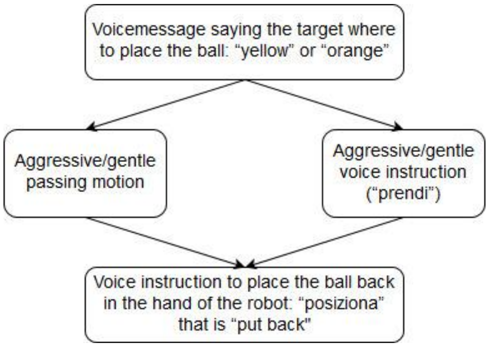

Figure 3. Flowchart showing the sequence of one trial with the possible variations

performed a first training phase of ten repetitions, in which they had to take the ball and place it on the target indicated by the robot. During this phase, the robot posture was fixed in the final passing position. After that, another short training allowed participants to familiarize with the behavior of the robot. In this part, the robot showed gentle action, gentle voice, rude action and rude voice one after the other. After the training, we presented the stimuli in three blocks of 16 repetitions. Each one of the 4 conditions (gentle action, gentle voice, rude action, aggressive voice) was performed 4 times per block in a randomized order. The sequence of the conditions in each block was identical for every participant. Each repetition had the structure depicted in Fig. 3.

The voice messages that indicated the target or the repositioning of the ball, were designed to be neutral and to avoid any influence on the participant. Also the return movement was programmed to respect biological motion but peak velocity speed was reduced to show neutral vitality.

We gathered kinematic data of each participant through an Optotrak tracking system (NDI) with 5 active infrared markers (sampling rate 100 Hz) placed on the right hand, as shown in Fig. 4A.

We then examined different features of the motion:

- From marker 4, at the center of the back of the participant’s hand, we analyzed the speed and acceleration of hand motion.
- With markers 1 (tip of the thumb) & 2 (tip of the index) we analyzed the dynamics of finger aperture

Marker 3 and 5 were used as a backup for reconstruction of the possible gaps in the motion tracking recordings.

To assess the potential effect of different vitality on participants’ behavior during the interaction, we analyzed a few kinematic features of their movement drawing inspiration from a previous study on vitality forms in human-human interaction [10]: maximum hand speed during reaching and return – i.e. positioning of object in the target position; maximum hand acceleration during reaching and return; maximum finger aperture, as the maximum 3D distance between marker 1 & 2 (see Fig.4A); maximum finger opening speed and maximum finger closing speed.

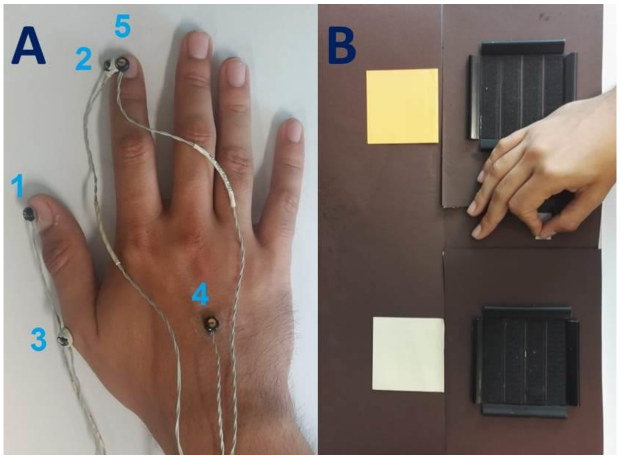

Figure 4. Panel A shows the position of the markers used to record kinematic data. Panel B shows the table and initial position of the hand.

At the end of the experiment, we asked the participants how they would describe the audio and movements of the robot and we left the possibility to give open comments about the experiment.

We tested a total of 10 participants (all right-handed, 4 females, mean age: 27, SD: 3 years). All of them already had some form of experience in interaction with the iCub. This selection was performed to minimize the influence of a potential novelty effect on participants’ behavior.

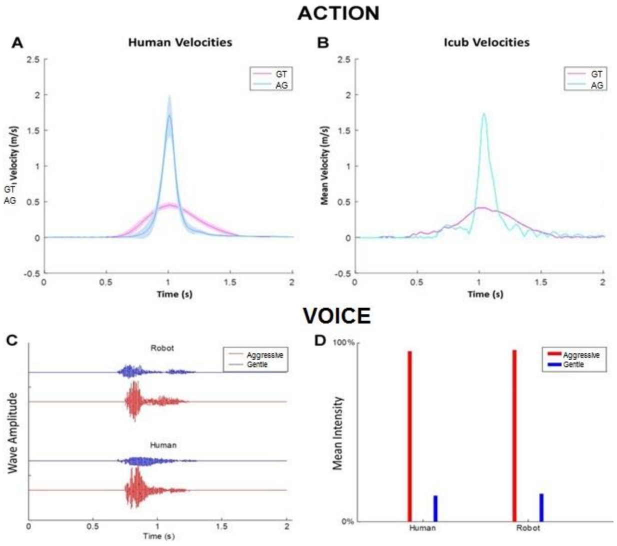

Figure 5. Action: velocity profiles of the movements performed by the human actor (A) by robot (B). Voice: graphs show the wave amplitude (C) and the mean intensity (D) relative to the action verb “take it” pronounced by human and iCub robot in gentle and rude way.

III. RESULTS

The first main result is the successful porting of the human actions to the joint space of the robot with accurate timing. The iCub could execute very precise movements, closely resembling the ones of the human actor (see Fig.5A and B). Analogously the robot voices consistently replicated the differences in the features of the gentle and rude human voice (see Fig. 5C and D).

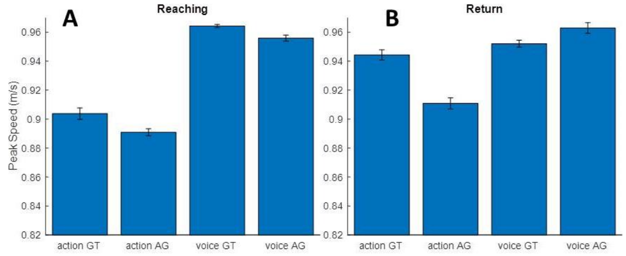

Figure 7. Mean among all participants of the maximum hand speed for the reaching (A) and return (B) phase of the movement. The error bars represent standard error of the mean. GT stands for gentle, AG for aggressive

We conducted two experiments: the first to test the subjective evaluation of vitality forms that we created for the iCub, and the second, to investigate the effect of robot behaviors on participants during a real human-robot interaction. In the first experiment participants were asked to watch video of actions performed either with a rude or a gentle style or to listen to instructions given with a rude or gentle voice modulation and to rate from 1 (very gentle) to 7 (very rude) the actor's behavior. The agent could be either a human male actor or the iCub robot, exhibiting the voices and kinematics we have generated. The results, plotted in Fig. 6, indicate that the vitality of all the stimuli was correctly recognized for both the human and the robot stimuli. This is confirmed by two Two Way Repeated Measures ANOVA on the ratings for the different voices and movies, with Agent (human/robot) and Style (gentle/aggressive) as factors. For both type of stimuli (auditory and actions) the difference between Style is highly significant (voice: F(1,19) = 682.5, p<0.001; action: F(1,19)=298.7, p<0.001), whereas the difference between Agents is not significant (voice: F(1,19)=0.76, p=0.39; action: F(1,19)=0.84, p=0.37). From the analysis it emerges also a significant interaction between Agent and Style (voice: F(1,19)=63.1, p<0.001; action: F(1,19)=14.1, p=0.001). A Bonferroni post-hoc highlights a significant difference between the ratings for aggressive styles between human and robot (p = 0.002) and a significant difference between the ratings for gentle styles in the voice modality only (p<0.001). Robot voices then tend to appear on average more gentle than the corresponding human voices. Conversely, the actions exhibit the same level of aggressiveness and kindness for the two agents. In summary the distinction between aggressive and gentle behavior is highly evident for both agents. In the second experiment, we assessed whether participants action kinematics was affected by the vitality expressed by the robot during a passing interaction. Fig. 7A show the maximum speed averaged across all participants for the reach-to-grasp phase of the movement in response to the robot gentle action (action GT) or aggressive action (action AG) and to the robot gentle or aggressive vocal instruction (voice GT, voice AG, respectively). The motion seems slightly faster when the stimulus was gentle, but the main variation can be found between voice and action. In fact, through an two-Way Repeated Measures ANOVA, we did not find a significant difference between the two vitality forms, but only between action and voice (F(1,9) = 16.13, p<0.01) and no significant interaction. Fig. 7B shows a chart with data about the second part of the action: from the grasping to the placement on the table. It is clear that findings here are approximately equivalent to the ones about the "reaching" part of the movement, just described: a Two-Way Repeated Measures ANOVA confirms again significant difference only between action and voice (F(1,9) = 16.83, p<0.01).

Results regarding peak acceleration are displayed in the charts of Fig. 8. In this case, contrary to what happens for speed, acceleration is higher when the robot is aggressive, a difference which reaches significance in the return phase of the movement (Fig 8B, F(1,19) = 5.95, p<0.05). In the reaching phase insteadthe main difference is between voice and action and is confirmed by a Two-Way Repeated Measures ANOVA (Fig. 8A - F(1,19) = 13.33, p<0.01). . Peak finger aperture was not significantly influenced by the different robot's styles as can be seen in Fig. 9. Additionally,

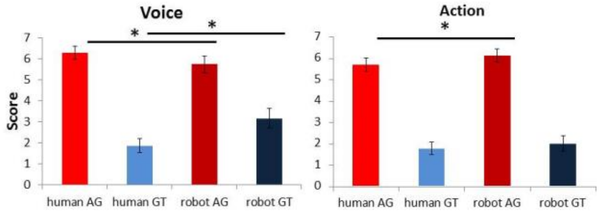

Figure 6. Subjective evaluation of robot behaviors with different vitality, conveyed with different vocal (left) or action features (left) . Gentle and rude between the robot and actor seem to be perceived almost in the same manner by participants.

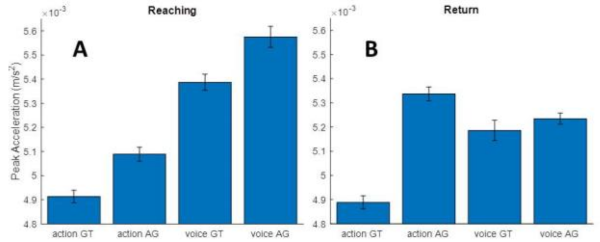

Figure 8. Mean among all participants of the maximum acceleration for the reaching (A) and return (B) phase of the movement. The error bars represent standard error of the mean.

the peak speed of finger opening and closing tends to be faster in response to the aggressive than to the gentle robot behavior, in particular after robot actions (Fig 10A and B). This is confirmed by a significant interaction between style and type of behavior in a Two-Way Repeated Measures ANOVA on peak aperture speed (F(1,19) = 5.32 , p=0.046). No differences are significant in the analysis of speed of closure.

The last results regard subjective answers to the question "how would you describe the voices and movement of the robot?" asked after the experiment. All participants described the rude stimuli using at least one of these words: "aggressive", "commanding", "angry", "rude", while the gentle stimuli were defined "kind", "calm", "relaxed". Moreover, 8 out of 10 declared that, in their opinion, the difference between the rude and gentle audio was approximately the same if compared to the rude and gentle motion. These comments extend the findings of the first subjective evaluation experiment, the robot was able to convey vitality forms. In particular, the vitality exhibited by the robot communicated the correct attitudes also when participants did not have only to provide a binary evaluation (rude/gentle, as in the first experiment) but were free to choose an arbitrary description. As additional note, in the open comments, three of them affirmed that, while the end of the voice command was easy to recognize, sometimes they were not sure about the end of the motion, therefore they waited for a little while before stating their action.

IV. DISCUSSION

In this study we investigated how to endow a humanoid robot with vitality forms, that is to convey gentleness or rudeness through robotic action and speech (robotic vitality forms). Moreover we assessed if, during a human-robot interaction in shared real operating workspace, these vitality forms may influence the subsequent action of participants. The results showed that kinematics parameters of the robot motion and properties of its voice are adequate to express different attitudes, that are consistently perceived rude or gentle by human partners. Interestingly, the subjective recognition of the "style" of the robot is similar for both modalities of communication: action and voice. The two vitality forms are also clearly perceived during an actual interaction with the robot, conveying the perception of the robot being "calm" or "aggressive/commanding". The behavioral reactions, however, show only minor differences between the two types of robot actions, with just a tendency to show an increase in hand acceleration and speed of grasp aperture in response to an aggressive rather than a gentle robot behavior. Moreover, participants' motor response was significantly faster and more accelerated in response to robot's voice rather than robot's action.

These findings do not entirely replicate what has been shown in human-human interaction, where the style and emotion conveyed by voice and movement of the agent, influenced similarly the motor response of the receiver [10]. In particular, participants were more rude – significantly faster and exhibiting higher acceleration – when they watched a video showing rude behavior, on the contrary, they were slower and more relaxed when the video displayed gentle behavior. There can be different reasons for this discrepancy. First, these human-human experiments were not performed in presence of another person, but by showing videos. In our experiment, the physical presence of the robotic partner might have influenced the kinematics responses of participants. Indeed, when a person is aggressive towards another one in a shared workspace, the response is not always equally aggressive, but other feelings like caution or fear can arise, resulting in a more careful answer.

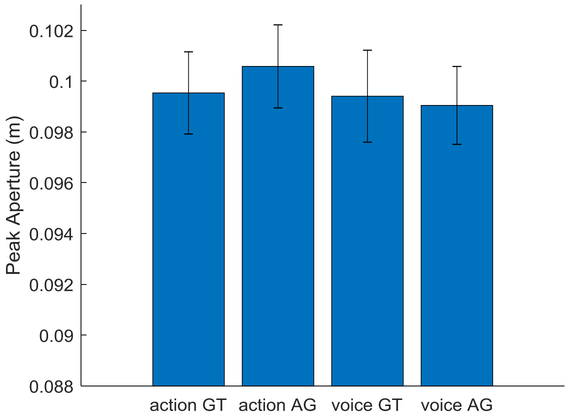

Figure 9. Mean among all participants of the max distance between index and thumb during the reaching phase. The error bars represent standard error of the mean

Second, a possible cause of the difference between reactions to robot's voice and action is that, according to the annotations of some of the participants, the moment indicating the end of robot motion and the start of the human response was not clearly communicated. Conversely, the timing of the vocal instruction was precisely detectable. This might have slowed down the response in the "action" condition, also making the human response less automatic and natural.

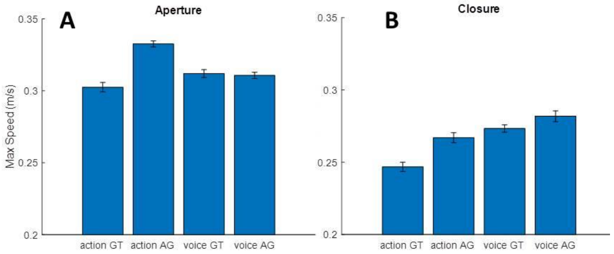

Figure 10. Mean among all participants of the maximum aperture speed (A) and closure speed (B) for the reaching phase of the movement.

Third, we cannot exclude the fact that the interaction partner was a humanoid robot might have influenced the participants' action planning, making their behavior less susceptible of changes due to the partner's actions. However, we believe that this last possibility is less likely, as humanoid robot actions have often been proven to be able to significantly modulate human partner's behavior and kinematics[21], [22]. Future studies with larger sample sizes will enable us to establish which of these factors has influenced our current results.

V. CONCLUSION

In summary, it is possible to endow a humanoid robotic action and voice with different vitality forms, conveying the impression of a rude or a gentle action. This could prove of particular relevance in all contexts in which an action might have particular urgency (e.g., in emergency scenarios) or when the robot should assume an authoritative role (e.g, in teaching or for police robots). The mere kinematics or even just the voice of the agent can become a valuable tool to make the robot appear more "commanding" or "assertive" to the eyes of the partners.

In future work, we would like to deepen the analysis, finding the specific kinematics and vocal features that convey aggression or gentleness in human motion and voice, in order to be able to selectively manipulate them, enriching our current possibility to reproduce human-like motion.
This generalization would allow recreating vitality forms in different actions or speech, starting from any kind of movement, not to mention the possibility to transfer this skill to other robots.

ACKNOWLEDGMENT

This research has been conducted in the framework of the European Project ROBOCOM++ (FLAG-ERA Joint Transnational Call (JTC) 2016). We are grateful to all study participants for their contributions. We also would like to thank Valentina Baruzzi and Linda Lastrico for the help with the generation of the motion for the iCub, and Settimio Ziccarelli for the help with the subjective evaluation experiment.

REFERENCES

[1] G. Sandini, A. Sciutti, and F. Rea, "Movement-Based Communication for Humanoid-Human Interaction," in *Humanoid Robotics: A Reference*, 2017, pp. 1–29.

[2] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2005.

[3] O. Palinko, F. Rea, G. Sandini, and A. Sciutti, "A Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration," in *IEEE International Conference on Intelligent Robots and Systems*, 2016.

[4] A. Sciutti and G. Sandini, "Interacting with Robots to Investigate the Bases of Social Interaction," *IEEE Trans. Neural Syst. Rehabil. Eng.*, pp. 1–1, 2017.

[5] A. Sciutti, L. Patanè, F. Nori, and G. Sandini, "Understanding object weight from human and humanoid lifting actions," *IEEE Trans. Auton. Ment. Dev.*, 2014.

[6] S. Planalp, "Varieties of Cues to Emotion in Naturally Occurring Situations," *Cogn. Emot.*, 1996.

[7] F. E. Pollick, H. M. Paterson, A. Bruderlin, and A. J. Sanford, "Perceiving affect from arm movement," *Cognition*, 2001.

[8] J. Montepare, S. B. Goldstein, and A. Clausen, "The identification of emotions from gait information," *J. Nonverbal Behav.*, vol. 11, no. 1, pp. 33–42, 1987.

[9] G. Di Cesare, M. Marchi, A. Errante, F. Fasano, and G. Rizzolatti, "Mirroring the Social Aspects of Speech and Actions: The Role of the Insula," *Cereb. Cortex*, pp. 1–10, 2017.

[10] G. Di Cesare, E. De Stefani, M. Gentilucci, and D. De Marco, "Vitality Forms Expressed by Others Modulate Our Own Motor Response: A Kinematic Study," *Front. Hum. Neurosci.*, 2017.

[11] A. Sciutti, M. Mara, V. Tagliasco, and G. Sandini, "Humanizing human-robot interaction: On the importance of mutual understanding," *IEEE Technol. Soc. Mag.*, vol. 37, no. 1, pp. 22–29, 2018.

[12] M. J. Gielniak, C. K. Liu, and A. L. Thomaz, "Generating human-like motion for robots," *Int. J. Rob. Res.*, 2013.

[13] L. Takayama, D. Dooley, and W. Ju, "Expressing thought: improving robot readability with animation principles," *Proc. 6th ACM/IEEE Int. Conf. Human-Robot Interact.*, 2011.

[14] T. Hashimoto, S. Hitramatsu, T. Tsuji, and H. Kobayashi, "Development of the face robot SAYA for rich facial expressions," in *2006 SICE-ICASE International Joint Conference*, 2006.

[15] T. Fukuda, J. Taguri, F. Arai, M. Nakashima, D. Tachibana, and Y. Hasegawa, "Facial expression of robot face for human-robot mutual communication," *Proc. 2002 IEEE Int. Conf. Robot. Autom. (Cat. No.02CH37292)*, 2002.

[16] M. Karg, A. A. Samadani, R. Gorbet, K. Kuhnlenz, J. Hoey, and D. Kulic, "Body movements for affective expression: A survey of automatic recognition and generation," *IEEE Trans. Affect. Comput.*, 2013.

[17] R. Laban and F. Lawrence, *Effort*. London, UK, 1947.

[18] E. I. Barakova and T. Lourens, "Expressing and interpreting emotional movements in social games with robots," *Pers. Ubiquitous Comput.*, 2010.

[19] N. Noceti, F. Rea, A. Sciutti, F. Odone, and G. Sandini, "View-invariant robot adaptation to human action timing," in *IEEE Technically Sponsored Intelligent Systems Conference (IntelliSys)*, 2018.

[20] S. Ivaldi, M. Fumagalli, M. Randazzo, F. Nori, G. Metta, and G. Sandini, "Computing robot internal/external wrenches by means of inertial, tactile and F/T sensors: Theory and implementation on the iCub," in *IEEE-RAS International Conference on Humanoid Robots*, 2011.

[21] F. Vannucci, A. Sciutti, M. Jacono, G. Sandini, and F. Rea, "Adaptation to a humanoid robot in a collaborative joint task," in *RO-MAN*, 2017.

[22] A. Bisio *et al.*, "Motor contagion during human-human and human-robot interaction," *PLoS One*, vol. 9, no. 8, 2014.