24

Assessment of the Impressions of Robot Bodily Expressions using Electroencephalogram Measurement of Brain Activity

A. Khiat[1], M. Toyota[2], Y. Matsumoto & T. Ogasawara
Nara Institute of Science and Technology – NAIST
JAPAN

1. Introduction

Recently, robotics research has focused on issues surrounding the interaction modalities with robots, how these robots should look like and how their behavior should adapt while interacting with humans. It is believed that in the near future robots will be more prevalent around us. Thus it is important to understand accurately our reactions and dispositions toward robots in different circumstances (Nomura et al., 2006). Moreover, the robot's correct production and perception of social cues is also important. Humans have developed advanced skills in interpreting the intentions and the bodily expressions of other human beings. If similar skills can be acquired by robots, it would allow them to generate behaviors that are familiar to us and thus increase their chances of being accepted as partners in our daily lives.

The expressiveness of a gesture is of great importance during an interaction process. We are often required to give special attention to these signs in order to keep track of the interaction. Humans have learned to adapt their behavior and to react to positive and negative bodily expressions (Bartenieff & Lewis, 1980). Although there has been remarkable work on the design issues of sociable robots (Breazeal, 2002) and affective autonomous machines (Norman et al., 2003), there has not been much work on investigating the real impact of robot bodily expressions on the human user in the context of human-robot interaction. Knowing the effect of a generated gesture, a robot can select more accurately the most appropriate action to take in a given situation. Besides, computer-animated characters have been used to evaluate human perception of the significance of gestures. However, animated characters and embodied ones should be treated differently since the latter are tangible entities (Shinozawa et al., 2005).

In this article we report a study on the relation between bodily expressions and their impacts on the observer. We also attempt to understand the effect that expressions have on the observer's brain activity. Its sensitivity to bodily expressions can be used during an interaction task since the brain is the source of every cognitive and emotional effort.

---

[1] Corresponding author: **ABDELAZIZ KHIAT**, *Robotics laboratory, Graduate School of Information Science, Nara Institute of Science and Technology, Keihanna Science City, 630-0192 JAPAN*. Email: khiat@ieee.org
[2] MASATAKA TOYOTA is currently with Canon Corporation.

Fig. 1. Considered scenario for robot bodily expressions and its perceived impression.

In this work, we have conducted an experimental study where several users were asked to observe different robot bodily expressions while their brain activity was recorded. The results suggest the existence of a relation between the type of bodily expressions and the change in the level of low-alpha channel of brain activity. This result helped in the selection of features that were used to recognize the type of bodily expression an observer is watching at a certain time. The recognition rate was of about 80% for both cases of robot bodily expressions and of human bodily expressions. Potential applications include customized interface adaptation to the user, interface evaluation, or simple user monitoring.

2. Bodily expressions and their impressions

The considered scenario for this study is depicted in Fig. 1. First, we have a robot that is executing a series of movements. It transmits to the observer a meaningful expression which is called bodily expression ①. Second, we have a human observer that perceives the expression and interprets it using his/her a priori knowledge ②. Then, the observer gets an impression, which means that bodily expression affects him/her to a certain level, depending on its strength, his/her awareness or attention and his/her state of mind or mentality ③. It is important to emphasize the difference between how the observer perceives and interprets a bodily expression, and what impact this expression evokes in the observer. It is expected that the two are related, but there is no information about the nature of this relation or how it evolves and changes over time. One of the goals of this work is to clarify and explain certain aspects of this relation to open the possibility of generating an adaptive robot behavior based on this information.
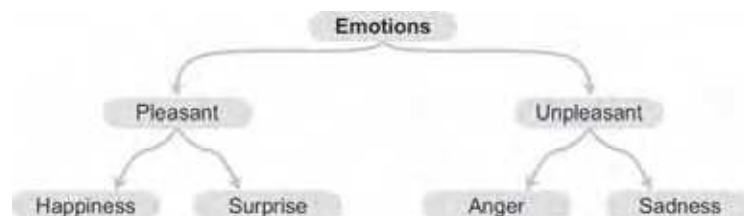

Fig. 2. The subset of Shaver's classification of emotions used in the categorization of Bodily Expressions.

2.1 Classification of bodily expressions

There is a need to classify bodily expressions generated by a robot in order to investigate their effects on the user. For this reason, salient differences among motions should be implemented. During an interaction process, humans go through different affective states, depending on several conditions such as degree of engagement, degree of awareness, and degree of interest among others. It is thus possible to classify every action taking place during an interaction process into the emotional effects that it would have on the observer. We adopted a simplified version of Whissel's wheel of activation-evaluation space described in (Whissel, 1989). We used the fact that we have two primary states for emotions: positive and negative ones, also known as pleasant and unpleasant emotions. The considered emotions are the following: happiness, surprise, sadness, and anger. In order to categorize these emotions we used a subset of Shaver's classification (see Figure 2), where happiness and surprise represent pleasant emotions while sadness and anger represent unpleasant emotions (Shaver et al., 1987). Bodily expressions were classified using one of the specified four emotions as pleasant or unpleasant.

2.2 Generation of robot bodily expressions

The humanoid robot ASKA (Ido et al., 2002) used in this study is shown in Figure 3. The body has a mobile platform and two arms and is based on the commercial robot TMSUK-4[3]. The head is a replica of the Infanoid robot (Kozima, 2002). This humanoid robot with its mobile platform has the advantage of being able to generate relatively fast motions compared to the currently available biped humanoid robots.

Since the pioneering work of (Johansson, 1973) on visual perception of biological motion, it has been known that humans can perceive a lot of information from body movements including the emotional state of the performer (Allison et al., 2000; Pollick et al., 2001). Recently, there is a growing interest in mathematically modeling emotion-based motion generation for real-world agents such as robots (Lim et al., 2004) and for virtual agents such as animated characters (Amaya et al., 1996). To be able to generate bodily expressions that reflect the selected emotions we rely on Laban features of movements (Bartenieff & Lewis, 1980). It has been shown by (Tanaka et al., 2001) that the qualitative Laban features of Effort and Shape correlate with the four basic emotions we have selected in section 2.1.

Based on the mathematical description of Laban features, shown in the Appendix, it is relatively easy to classify bodily expressions that reflect a certain emotion. Although there is no unique solution to this problem, the goal is to be able to generate a representative bodily expression for each one of the selected emotions.

The generated bodily expressions (BE) which reflect one of the basic emotions of happiness, surprise, sadness, anger or none are the following:

- **BE1:** The robot raises both arms and turns its body to the left, then to the right, twice. The goal is to show an expression of happiness.
- **BE2:** The robot raises its right hand and moves it in an arc toward the right side, then goes back to its initial position and lowers its right arm, the goal is to show an expression of no particular emotion.
- **BE3:** The robot raises both arms and its head, then moves backward for some distance, the goal is to show an expression of amazement or surprise.

---

[3] *TMSUK-4* is a trademark of tmsuk Co. Ltd, Kitakyushu.

- **BE4:** The robot lowers both arms and its head, then moves backward at low speed for some distance, the goal is to show an expression of sadness.
- **BE5:** The robot raises both arms gradually while advancing before stopping, then it lowers and raises its arms progressively for two cycles; the goal is to show an expression of happiness.
- **BE6:** The robot advances quickly, then goes back and raises its right arm while turning its head a bit to the right. It then lowers its arm and returns its head to the original position; the goal is to show an expression of anger.
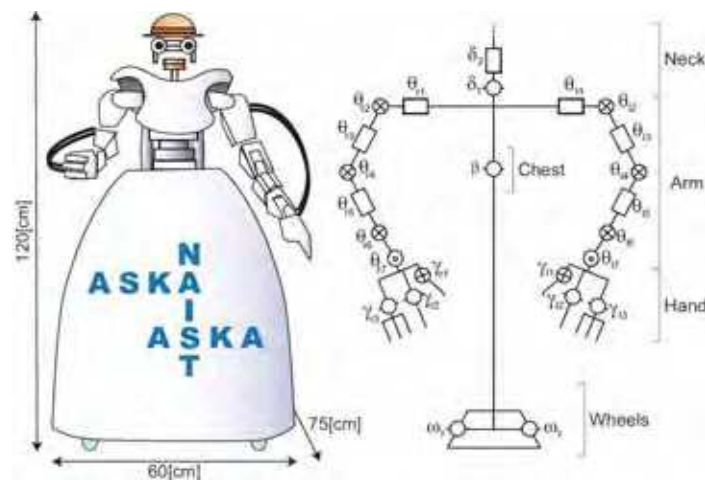
Fig. 3. Overview of the receptionist robot ASKA and its joints-link model.

The duration of each of these BEs was about 14[sec]. Their appropriateness and their expressiveness was tested experimentally using questionnaires (see section 3.1).

2.3 Assessment of impression and expressiveness of bodily expressions

There are mainly two types of methods to assess the effects of a particular action on a human. The classic self-reporting approach is widely used, while the assessment from measured physiological information is still an open research problem. The first type of methods gives subjective evaluation results; whereas the second type of methods is deemed to be more objective but suffers from inaccuracies. For our case, in order to assess expressiveness we adopted a self-reporting approach and asked the subjects to answer questionnaires. However, in order to assess impression the subjects answered questionnaires and their brain activity was also recorded.

Summarizing the subject's answers to questionnaires was used in order to assess expressiveness. Every subject had to select from: expression of happiness, expression of surprise, expression of sadness, expression of anger, or no meaningful expression. The subject also had to specify the degree of the expression in a scale of five: **1** for impertinent, **2** for slight, **3** for medium, **4** for strong and **5** for very strong. This selection of the degree of expression is a redundancy that was meant to confirm the subject's choice and assess the degree of confidence in his/her answer. These answers were then categorized into pleasant or unpleasant expressions using the subset of Shaver's classification shown in Figure 2.

As for impression assessment, spectral analysis method of electroencephalogram (EEG) data was used. A short EEG segment can be considered as a stationary process, which can be characterized by an autoregressive (AR) model. Let us denote $s(n)$ as a sample of EEG data of N points. We calculate $r_f(n)$ and $r_b(n)$, respectively the forward and backward prediction errors, as follows:

$$r_f(n) = \sum_{k=0}^{p} a(k)s(n+p-k) \tag{1}$$

$$r_b(n) = \sum_{k=0}^{p} a(k)s(n+k) \tag{2}$$

where $a(k)$ is the AR parameters and p is the order of the model. The order p is based on the "goodness of fit" criterion. We use the relative error variance (REV) criterion (Schlögl et al., 2000), defined as:

$$REV(p) = \frac{MSE(p)}{MSY} \tag{3}$$

$MSE(p)$ is the mean square error or variance of the error process of order p, and MSY is the variance of the total power of the signal sample. The optimal p is the one that minimizes $REV(p)$. In our case we take $p = 14$.
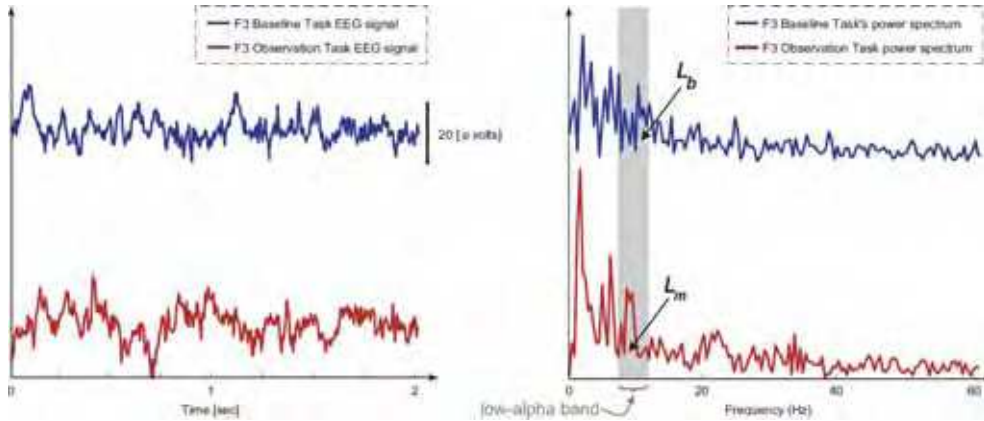

Fig. 4. An example illustrating the calculation of the power of low-alpha band for a 2[sec] data segment taken from electrode placement F3. The graph to the left shows the raw EEG signal for the baseline period and the observation period. The graph to the right shows the power spectra of the EEG signals, where low-alpha frequency band is highlighted.

We apply (1) and (2) to calculate an approximate estimation of the power spectrum $PS(f)$ of the signal s as follows:

$$PS(f) = \frac{V_p T}{\left|1 + \sum_{k=0}^{p} a(k) e^{-j2\pi fkT}\right|^2} \tag{4}$$

$$V_p = \frac{1}{2} \sum_{n=0}^{N+p-1} \left([r_f(n)]^2 + [r_b(n)]^2\right), \tag{5}$$

where V_p is the averaged sum of the forward and backward prediction error energies and T is the sampling period.

Research in cognitive neuroscience has shown that the power of low-alpha frequency band is the most reactive band to social cues such as movements (Allison et al., 2000; Cochin et al., 1998). We suppose that this frequency band reacts in a similar way to robot bodily expressions (Khiat et al., 2006). The next step in assessing the impression is to observe the amount of change in the power of low-alpha frequency band compared to the whole spectrum. The power L of a band between frequencies a and b is defined by:

$$L(a,b) = \frac{\int_a^b PS(f) df}{\int_0^\infty PS(f) df} \tag{6}$$

Using **(6)**, we calculate the power of low-alpha band frequency L_b for the data taken during the baseline period and L_m for the data taken during the period of the execution of a bodily expression. An example illustrating this calculation is shown in Fig. , where raw 2 seconds EEG signals collected during the baseline period and the observation period is shown to the left. The power spectrum of these signals is shown to the right, and the low-alpha frequency band is highlighted. A comparison between L_b and L_m would indicate the effect of a particular bodily expression on the user. This is used as the main evaluation criterion for impression.

3 Experimental study

3.1 Expressiveness of robot bodily expressions

The goal of this experiment is to evaluate the expressiveness of every generated robot bodily expression. Since this quality is highly subjective, the self-reporting approach is used.

Subjects. Seventeen (17) participants (two females and fifteen males aged between 20 and 50 years old) volunteered to take part in this experiment. They were either students or faculty members at the Graduate School of Information Science. They were all familiar with robots and had previous experiences of dealing with robots similar to the one used in the experiment.

Procedure. Every subject was shown a total of six bodily expressions, which were described in section 2.2. The execution of each of the bodily expressions by the humanoid robot ASKA lasted 14 seconds. After observing each robot bodily expression, enough time was given to the subject to answer two questions about the expressiveness of that robot bodily expression, and one more question about their impression after the observation. These answers were then summarized as explained in section 2.3 to assess their expressiveness.

BEs	Pleasant	Unpleasant	Neither
BE1	**100**%	0%	0%
BE2	6%	35%	**59**%
BE3	**94**%	6%	0%
BE4	0%	**94**%	0%
BE5	**65**%	12%	23%
BE6	0%	**82**%	18%

Table 1. Users' evaluations of the expression of each generated robot bodily expression (BE).

Results. Table 1 shows the results about bodily expressions after categorization into pleasant expressions, unpleasant expressions, or neither, clearly indicating the expressiveness as evaluated by the observers. The result about impressions is presented in Table 2 after categorizing the answers into pleased or unpleased.

These results demonstrate the existence of a strong correlation between the expressiveness of the robot bodily expressions as seen by the subjects and the target expression when these bodily expressions were generated (see section 2.2). BE1, which was created to express happiness, was classified as having a 100% pleasant expression. BE2, which was created to express a neutral emotion, was classified by 59% as neither pleasant nor unpleasant, and by 35% as unpleasant, suggesting that neutral bodily expressions can have a negative connotation. BE3, which was created to express surprise, was classified by 94% as a pleasant expression. BE4, which was generated to express sadness, was classified by 94% as being an unpleasant expression. Similarly, BE6 which was created to express anger was also classified by 82% as an unpleasant expression. The special case of BE5 was classified to a great extent as a pleasant expression by up to 65%. However, 23% said it did not express anything in particular and 12% claimed it was unpleasant.

BEs	Pleased	Unpleased	Neither
BE1	**65**%	35%	0%
BE2	30%	**70**%	0%
BE3	**68**%	32%	0%
BE4	19%	**81**%	0%
BE5	**100**%	0%	0%
BE6	47%	**53**%	0%

Table 2. Users' evaluations of their impressions after observing each robot bodily expression (BE).

The expressiveness of the generated BEs is confirmed to be in accordance with the target expressions for which they were created. BEs generated to express happiness and surprise expressions were classified as pleasant, and the BEs generated to express sadness and anger expressions were classified as unpleasant. Among the generated BEs we could choose one that is representative of each category in order to use it in the evaluation of its impressions on the observer.

3.2 Impressions of robot bodily expressions

The goal of this experiment is to evaluate the impression on the observer of the generated bodily expressions using a hybrid approach that combines the results of self-reporting and the analysis of brain activity.

Subjects. Seven (7) participants (one female and six males, 23~43 years old) volunteered to take part in this experiment. They were all students or faculty members at the Graduate School of Information Science, and only two of them had the experience of using electroencephalography to measure brain activity. Before starting the experiment each participant was fitted with electrodes and allowed to spend more than 20 minutes reading books of interest to familiarize and condition them to the electrodes' presence.

Procedure. During each session, 12 EEG channels (using sintered Ag/AgCl electrodes) were recorded by the 5200 Series DigitalAmp System[4]. The recording was performed from 10 placements, namely: Fp1, Fp2, F3, F4, T3, T4, P3, P4, O1, and O2 according to the international 10-20 standard (see Fig. 5). The placement Fz was used as the ground, and the signal from the left ear placement A1 was used as the reference signal. The contact impedance between all electrodes and the skull was kept below 5[kΩ]. The subjects were shown a total of six motions lasting 14 seconds each by the humanoid robot ASKA while their brain activity was recorded with 16-bit quantization at a sampling frequency of 200[Hz].
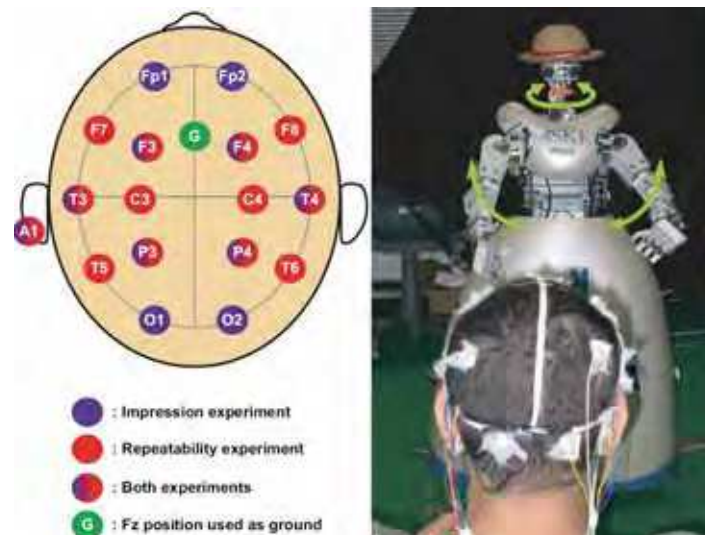

Fig. 5. The experimental setup where brain activity was measured according to the 10-20 international standard (Jasper, 1958).

The subjects were asked to relax as much as possible and think of nothing in particular when recording the baseline period, which lasted for 14[sec]. They were also told that they would be asked about the robot's movements and that they had to watch carefully when the robot was moving. This was important because we needed to make sure that the subjects attended to the task. After the observation of each bodily expression, the subjects described

---

[4] The 5200 Series DigitalAmp System is a trademark of NF Corporation, Yokohama.

their impression in their own words. Having no constraints to express themselves, the subjects gave more details about their impressions. These answers were used in categorizing the impressions into pleased or unpleased based on Shaver's original classification of emotions (Shaver et al., 1987).

Results. Table 2 shows the self-reporting result about the subjects' impressions after observing every robot bodily expression. There is a strong correlation between these results and the expression results, reported previously in section 3.1, with a coincidence level of 71%. For example, BE4 impression was considered to be unpleasant by up to 81% and its expressiveness was considered unpleasant by 94%. This is also the case for BE1 where its impression of being pleasant is 65%, and it expression of being pleasant is 100%. The same could be said for BE3, with a pleasant impression of 68% and a pleasant expression of 94%.

The case of BE6 is different from the previous ones. While its expression was considered unpleasant by 82%, its impression shows the small rate of 53% for being unpleasant and 47% for being pleasant. It is still inclined to the unpleasant side. However, its pleasant effect cannot be explained knowing that this bodily expression was created to express anger. The last case of BE2 shows a big difference between its 59% neutral expression and its 70% unpleasant impression.

Subject	Category	Electrodes									
		Fp1	Fp2	F3	F4	T3	T4	P3	P4	O1	O2
1	Pleasant	–	–	–	+	–	–	–	–	+	–
	Unpleasant	–	+	–	+	+	–	–	–	+	–
2	Pleasant	–	–	+	–	–	–	–	–	–	–
	Unpleasant	–	–	–	–	–	–	+	+	–	–
3	Pleasant	–	–	–	–	+	+	–	+	–	–
	Unpleasant	–	–	–	–	–	–	–	–	–	–
4	Pleasant	–	–	–	–	–	+	+	+	–	–
	Unpleasant	–	–	–	–	–	–	+	+	+	–
5	Pleasant	–	–	–	–	–	+	–	–	–	–
	Unpleasant	–	–	–	–	+	+	+	+	–	–
6	Pleasant	+	+	–	–	+	+	–	–	–	–
	Unpleasant	–	–	–	–	+	+	–	–	–	–
7	Pleasant	–	–	+	–	+	+	–	–	–	+
	Unpleasant	–	–	+	–	–	–	–	+	–	–

Table 3. Significant change in low-alpha power according to observed motion categories at every electrode and for each subject. (+: significant change p<.05; –: no significant change).

This suggests that bodily expressions with a neutral expression can be perceived negatively and can generate an unpleasant impression. The analysis of EEG data using the method described in section 2.3 allowed the calculation of the power L_m of low-alpha frequency band in each electrode channel and for each bodily expression. It also allowed the calculation of the power L_b of the same frequency band for the baseline period. Comparing L_m and L_b revealed the effect of observing a bodily expression in the change in the power of low-alpha frequency band for each electrode channel. Table 3 summarizes the results of this change in power, where only statistically significant change is indicated with the symbol +.

It can be seen that significant effect is mostly present at locations T3 and T4, then at P3 and P4, and finally at F3 and F4. Knowing that these positions are located above the superior temporal sulcus (STS) and above some specific parts of the prefrontal cortex (PFC) confirms previous research findings about the activation of STS in the perception of social cues (Cochin et al., 1998; Allison et al., 2000), and the activation of the mirror neurons located in the PFC during learning and imitation tasks (Rizolatti & Craighero, 2004). Some reaction can also be seen at other locations, for instance O1 and Fp2 for subject 1, O2 for subject 7, Fp1 and Fp2 for subject 6. The reaction at locations Fp1 and Fp2 are thought to be the result of blinking activity during the recording process, since these electrode positions are the closest to the eyes. It is important to assert that no preprocessing was done to avoid data with eye blinking artifacts. This approach was adopted because the goal is to conduct this investigation in natural conditions, where blinking activity is possible and should be considered. The reaction at locations O1 and O2 could be explained by the fact that during the vision process the visual cortex gets activated and this activation is usually captured at locations O1 and O2.

Nevertheless, the reactive locations were not always the same among different observers, suggesting high individual differences. A generalization cannot be derived at this point about the reaction of brain locations according to the category of the bodily expression that is being observed. However, the presence of a reaction is confirmed and another approach is necessary to achieve a more comprehensive result. On the other hand, there is a need to assess the repeatability of similar reactions from the same observer when he/she is shown the same bodily expression.

3.3 Repeatability of reaction in brain activity

The goal of this experiment is to confirm that the results obtained in the impression experiment (see section 3.2) are consistent over time for the same person. In other words, to make sure that brain reaction does happen all the time and at the same set of electrodes if a subject observes the same bodily expression several times.

Subject. One (1) student (male, 32 years old) volunteered to take part in this experiment. Similar to the previous experiment, the subject was fitted with an electro-cap and was given about 30 minutes to familiarize and get used to the presence of the cap.

Procedure. The subject participated in ten recording sessions. In each session, he was shown two bodily expressions, one for each category of bodily expressions, executed by the humanoid robot ASKA. Showing only representative bodily expressions is sufficient since the goal is to confirm the repeatability of brain reaction. Each bodily expression lasted for 14[sec], and the baseline period was recorded during the 14[sec] before the execution of each bodily expression. BE1 was chosen as a representative of pleasant bodily expressions, and BE4 was chosen as a representative of unpleasant bodily expressions. On one hand, BE1 was chosen because its expressiveness evaluation as pleasant (100%) was the highest among all the bodily expressions. Its impression evaluation (65%) was high enough to make sure it will have the desired effect on the observer, even though its impression was evaluated as the lowest among all the pleasant bodily expressions. In this case, the advantage was given to the expressiveness evaluation over the impression evaluation. On the other hand, BE4 was chosen because, similarly to BE1, its expressiveness evaluation as unpleasant (94%) was the highest among all the bodily expressions. Its impression evaluation (81%) was also the highest among all the bodily expressions, making it the

perfect candidate to represent unpleasant bodily expressions. The recording was performed from 12 placements of an electro-cap[5], namely: F3, F4, C3, C4, P3, P4, F7, F8, T3, T4, T5, and T6 (see Fig. 5), using the Polymate AP1132[6] EEG recording device. The sampling frequency was set to 200[Hz] and the impedance was kept below 5[kΩ]. As a result of the experiment in section 3.2, the placements Fp1, Fp2, O1, and O2 were omitted, since they were shown to be not of a big importance. On the other hand, new placements were introduced, namely: F7, F8, C3, C4, T5, and T6, in order to get a more detailed coverage of the prefrontal and temporal cortices.

During the recording of the baseline period the subject was asked to relax as much as possible and to think of nothing in particular. To confirm the attendance to the task, the subject was told that he would be asked about the robot's movements and that he had to observe carefully. After the observation of each bodily expression, he was asked to explain the difference between the recently observed bodily expression and the one just before.
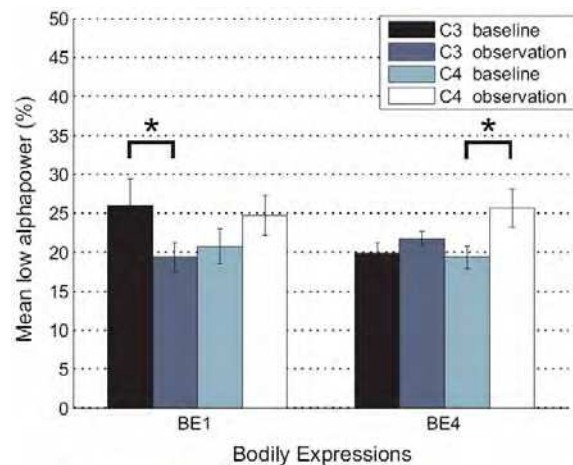
Fig.6. Mean alpha power calculated for the ten trials and for electrodes C3 and C4 (✱: p< .05).

Results. Analysis of the collected data, using the method described in section 2.3, resulted in identifying the electrode channels of placements C3 and C4 as the most reacting for this subject. Figure 6 shows the mean power of low-alpha frequency band calculated from the 10 trials for the electrode placements C3 and C4: where C3 reacted significantly to the pleasant bodily expression and C4 reacted significantly to the unpleasant bodily expression.

Fig. shows the overall result for the two bodily expressions by averaging the power change for all the electrodes over the 10 trials. The difference in means is significant between BE1 and BE4. Since BE1 is representative of pleasant bodily expressions and BE4 is representative of unpleasant ones, this result suggests an overall significant decrease in the power of low-alpha frequency band for pleasant motions, and a significant increase in power of low-alpha frequency band for unpleasant ones. This

---

[5] Electro-Cap is a trademark of Electro-Cap International Inc., USA.
[6] Polymate AP1132 was designed by Digitex Lab. Co. Ltd, and is commercialized by TEAC Corporation, Tokyo.

confirms that the change in low-alpha power happens every time the observer watches a bodily expression, and that this change is inversely proportional to the category of the observed bodily expression.
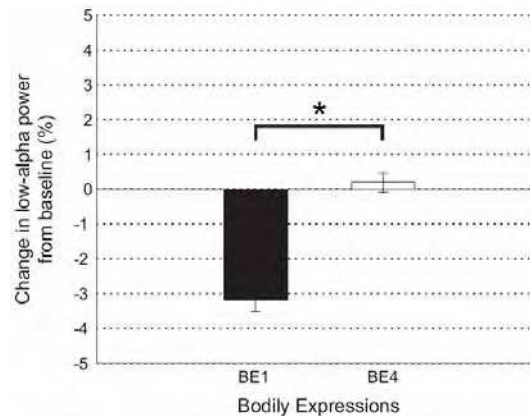

Fig. 7. Change in alpha power from the baseline using all electrodes, for each of the considered unpleasant and pleasant bodily expressions (✱: p < :05).

3.4 Discussion

The results presented in Table 1 confirmed the appropriateness of the expressiveness of the generated bodily expressions used in the experimental study. They show that the unpleasant bodily expressions were classified as unpleasant, and the pleasant bodily expressions were also classified as pleasant. During every experiment, the order of which the bodily expressions were shown to the observer was random so as not to allow the prediction of the nature of the next bodily expression. Although the subjects were not told anything about the bodily expressions, their answers agreed with the hypothesis. This implies that people tend to see bodily expressions in similar ways, which facilitates their interpretations and use in interactions. There exist a shared basic knowledge that allows humans the proper interpretation of similar expressions, although this knowledge is highly affected by the environment factors of culture and local customs. The bodily expressions were treated by the observers as if they were performed by a human even though it was the robot ASKA which actually performed them. It would be interesting to compare the interpretation of the same bodily expressions executed by both humans and robots to evaluate the existence of interpretation differences.

On the other hand, Table 2 correlates to a high extent with the results of Table 1. Here we can infer that observing a pleasant bodily expression will result in a pleasant impression on the observer and vice versa. This means that the observer is affected by what he sees even though the performer is just a robot. This effect on the observer is shown to be present in his/her brain activity with the results of section 3.2. Although, a generalization could not be concluded from the obtained results, the presence of a reaction in brain activity was proven.

It is important to acknowledge that the most reactive electrode positions were F3, F4, T3, T4, P3 and P4, which are located above the STS and PFC. This supports the claims that STS and mirror neurons get activated during the perception of social cues and the

observation of movements (Allison et al., 2000; Cochin et al., 1998), and that this can be used effectively in the implementation of Brain-Machine Interfaces (Nicolelis, 2001; Wessberg et al., 2000).

Finally, it was necessary to confirm the repeatability or the reproducibility of the same reaction in similar conditions. The results showed that the power level of low-alpha frequency band over all brain activity was inversely proportional to the category of the observed bodily expression. Particularly the most significant reaction was present at electrode positions C3 and C4 for the considered subject. These positions are close to the premotor and motor cortices. Due to the low spatial resolution of EEG, it is difficult to assert precisely which part of the brain is responsible for these reactions. However, current research findings confirm that the STS has an important role in the interpretation of social cues (Allison et al., 2000), and that mirror neurons are important during learning and imitation tasks (Rizolatti & Craighero, 2004).

4. Recognition of the impressions using Self-Organizing Maps

Self-Organizing Maps or Kohonen Networks are well suited to represent and generalize input data with an underlying structure that is not easily grasped (Kohonen, 1982). The unsupervised learning process tries to give a representation to the high-dimensional data rather than only a classification. Both the metric relation and the probability density of the data is approximated during this process, allowing the classification of newly measured data. Once the learning process has terminated, a labeling of the learned map can be done by using a relatively small set of labeled data items. The resulting map could be used to monitor the topographic patterns related to specific events as in (Joutsiniemi et al., 1995), or it could be used in the recognition of newly observed data. The similarity between a d-dimensional feature vector $X = (x_1, \ldots, x_d)$ and a prototype vector $M = (m_1, \ldots, m_d)$ in the learned map is calculated using the weighted Euclidean distance defined as:

$$D^2(X, M) = \sum_{j=1}^{d} w_j (x_j - m_j)^2 \tag{7}$$

where w_j is the weighting factor which can be used to give preference to certain features over others. This proved helpful and important in the semi-assisted learning of the data structure that was necessary for our EEG data. The usual approach in using SOM starts by preprocessing the selected data. Then, a feature extraction method is specified and used. After that, the map is calculated using competitive learning. Finally, the resulting map is labeled and used for recognition. It is important to note that in practical applications the selection and preprocessing of data is of extreme importance, because unsupervised methods only illustrate the structure in the data set, and the structure is highly influenced by the features chosen to represent data items (Kaski, 1997). In the following we will show how we used SOM in the recognition of bodily expressions executed by a humanoid robot and of similar bodily expressions executed by a human.

4.1 Recognition of the impressions of robot bodily expressions

4.1.1 Data acquisition

In the recognition of the impressions of robot bodily expressions we use the same data as in section 3.2. The data consists of all the signals collected at a sampling rate of 200[Hz] from the 10 EEG placements for twice 14 seconds and for each one of the seven subjects.
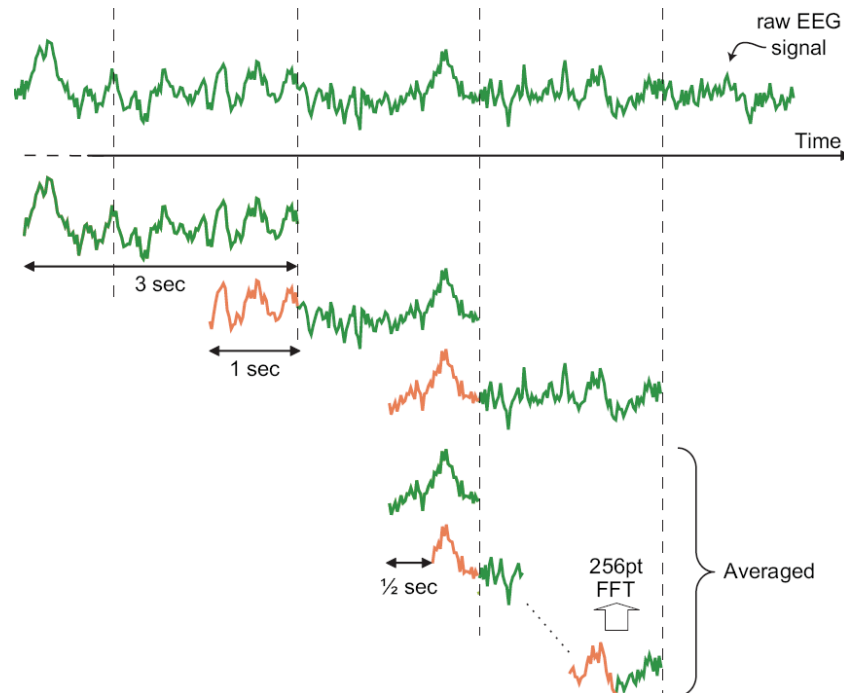

Fig. 8. Preprocessing EEG data for features extractions is done by calculating a moving average of overlapping windows of predefined length (3[sec]).

4.1.2 Data preprocessing

To prepare the collected data for the training task, we calculate the moving averaged power spectra of 10×2×7=140 signal sources. As shown in Fig. , we apply a 3[sec] (600-point) Hanning window on the signal with 1[sec] (200-point) overlap. The windowed 3[sec] epochs are further subdivided into several 1[sec] (200-point) sub-windows using the Hanning window again with 1/2[sec] (100-point) overlap, each extended to 256 points by zero padding for a 256-point fast Fourier transform (FFT). A moving median filter is then applied to average and minimize the presence of artifacts in all the sub-windows. The resulting moving averaged power spectrum is then reduced to six (6) features by integrating the spectral values weighted by six raised-sine shaped windows with an area normalized to unity. Consequently, the feature components do not need to be normalized. In a similarly way to (Joutsiniemi et al., 1995), the weighting windows are overlapping to ensure a smooth change of the features in accordance with the change in power spectrum. These windows cover the following frequency bands:

- ♦ Delta :   00~04[Hz]
- ♦ Theta :   04~08[Hz]
- ♦ Low-Alpha :   08~11[Hz]
- ♦ High-Alpha: 11~13[Hz]
- ♦ Beta :   13~30[Hz]
- ♦ Gamma :   30~50[Hz]

Notice that the alpha frequency band is divided into low and high. This will allow us to give different importance coefficients to each frequency band, according to each one's contribution in recognizing the effect of bodily expressions. The resulting time series of EEG power spectrum features consists of a vector of 10×6=60 features every 2[sec] (400-point) time intervals.

4.1.3 Map training and recognition results

The 2D map to learn using the collected features is arranged as a 2D lattice, with each location containing a 60-dimensional prototype vector. During the learning process or the self-organization, the importance coefficients w_j, used in the similarity metric **(7)**, were taken as 0.5, 0.5, 1.0, 0.9, 0.5, and 0.3 for the features delta, theta, low-alpha, high-alpha, beta, and gamma, respectively. Higher importance was given to the alpha bands with an emphasis on the low-alpha band, as a result of its sensitivity to the category of bodily expression being observed by a subject (see section 3.2). On the other hand, the higher frequency gamma band was given the lowest importance coefficient, since it was not proven to react significantly to social cues (Cochin et al., 1998).

After the training of the map, an approximation of the probability density of the input data is reached generating clusters which can be identified as associated to one of the following experimental conditions: observing unpleasant bodily expression, observing pleasant bodily expression, or baseline condition. 80% of the data was used for the training and the remaining 20% was used for the evaluation. The resulting recognition rate was of 65.2%, divided as 62.8% for data associated to unpleasant bodily expressions, 59.3% for data associated with pleasant bodily expressions, and 73.5% for data associated to baseline. Clearly, the rate of 65.2% is not satisfactory since this is a low rate to rely on when making a decision. However, when individual data was used separately in the learning process the recognition rate jumped high to the 79.5%. The previously realized low rate is explained by the existence of conflicting data items collected from different subjects. In order to understand the effect on the recognition rate of using data from different sources (subjects), all possible combinations of data sources were used to learn several maps and the recognition rates calculated. The change in recognition rates is shown in Fig. . It can be noticed that the addition of a new source of data decreased the recognition rate by about 5% for the first two additions. However, there is certain stability in the rate when the number of data sources was 3, 4 and 5. But again the rate decreased by a lower factor when adding more sources. This decline is explained by the existence of individual differences in the reaction to bodily expressions and probably even the interpretation. This outcome was also observed in the previous experiment about impressions (see section 3.2). To cope with this problem it is recommended to train several SOMs with a small number of data sources and use the totality in the recognition process.

4.2 Recognition of the impressions of human bodily expressions

There is little knowledge of the difference in the effect left on a person when observing humans and when observing robots. Most of the literature reports either cases (Ito & Tani, 2004; Allison et al., 2000; Schaal, 1999) and there is, to our knowledge, no previous work that tried to draw on the parallels between the two cases. Therefore, it is worthwhile to investigate similarities and differences in the recognition of the category of the impressions left on the observer for both cases.

In order to collect brain activity data when subjects are observing human bodily expressions, we conducted an experiment similar the one for the expressiveness of robot bodily expressions (see section 3.2). The goal is to evaluate the impression on the observer of the bodily expressions described in section 2.2.
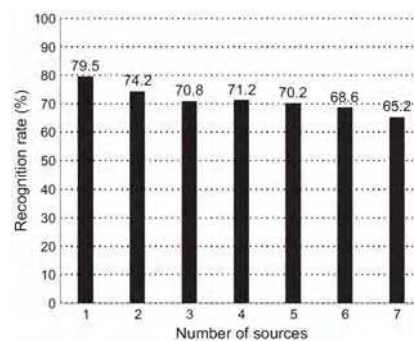
Fig. 9. Recognition rates when using data from different sources (subjects), when observing robot's bodily expressions.

4.2.1 Data acquisition

Subjects. Three (3) participants (males aged between 21 and 23 years old) volunteered to take part in this experiment. They were students at the graduate school of information science. They were all familiar with the experiment since they did participate in the previous one about the impressions of robot bodily expressions. Their brain activity was collected with an EEG measurement device and they were familiarized with the presence of the electrodes by letting them spend about 20 minutes reading books or surfing the Internet.
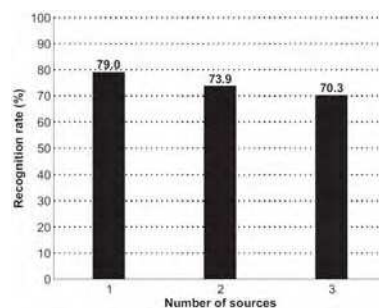

Fig. 10. Recognition rates when using data from different sources (subjects), when observing a human performer.

Procedure. The human bodily expressions to be shown to the subjects were prepared beforehand. A volunteer in a black tight suit performed these bodily expressions. He had a black cover on his head to make sure that no facial expressions get to the observer. These bodily expressions were a reproduction of the bodily expressions described in section 2.2 and were recorded on HDTV[7] video tapes. During the experiment, the bodily expressions were projected on a display big enough to ensure that the projected images of human performer is as close as possible to real size. Similar to the experiment in section 3.2, the recording was obtained from 12 electrode locations, namely: F3, F4, F7, F8, C3, C4, P3, P4, T3, T4, T5, T6, using the Polymate AP1132 EEG recording device. The sampling frequency was 200[Hz] and the impedance was kept below 5[kΩ]. The subjects were shown a total of six bodily expressions lasting 10[sec] each. During the recording of the baseline period, the subjects were asked to relax as much as possible and to think of nothing in particular. To help them achieve this state of mind, they were shown for 10[sec] the empty space of the room where the recording of the bodily expressions was performed. In addition, to confirm the subjects attended to the task they were told that they would be asked about the bodily expressions and that they had to observe carefully. After each observation, they were asked to explain the difference between the bodily expressions they just observed and the previous ones.

Performer of BEs	Human		Robot	
Data sources	Individual	Combined (3)	Individual	Combined (7)
Recognition rate (%)	79.0	70.3	79.5	65.2

Table 4. Comparison of the recognition rates for the cases of human bodily expressions and robot bodily expressions.

4.2.2 Recognition results

The preprocessing of the data was the same as in the previous case (section 4.1.2). Moreover, 80% of the data was used for training and the remaining 20% was used for evaluation. The resulting recognition rate was 70.3%, divided into 68.0% for data associated to the observation of unpleasant bodily expressions, 67.5% for data associated to the observation of pleasant bodily expressions, and 80.4% for data associated to the baseline. This rate is much better than the 65.2% rate achieved with the data of robot bodily expressions. However, the number of sources in this case is only 3, while it was 7 for the robot case. Thus, it is more accurate to compare this result to the 70.8% rate achieved when using only 3 data sources for the robot case, as shown in Fig. .

Using data from different sources showed degradation of the recognition rate, similar to the result of the robot, see Fig.. The addition of one source resulted in a decrease of 5.1%, then a decrease of 3.6% after adding a third source. This result supports the fact that individual differences remain present even in the case of observing bodily expressions performed by a human.

4.3 Discussion

Regardless of the performer, whether a robot or a human, the recognition rates of the category of the observed bodily expression was about 80% when using subject's data

---

[7] HDTV stands for High Definition television a.k.a High-Vision which allows the recording of a high resolution video stream.

individually. However, this rate decreased significantly when additional data from other subjects was used in the training process. To cope with this problem it would be interesting to train one SOM for each data source, and then combine the resulting SOMs into a bigger structure for the recognition task. Adopting this approach could result in keeping a high recognition rate while taking into consideration all the data that was collected so as not to loose the generality of the solution.

It is interesting to note that the average difference between the recognition rates for robot and human cases is relatively small as summarized in Table 4. This proves that SOMs are suitable to the generalization of the effect in the input EEG signals regardless of whether this effect is generated by a human or a robot. Even if differences appear clearly when analyzing raw EEG data, the SOMs succeed in eliminating these differences and keep only the important information. SOMs also succeed in separating the noise and artifacts from signals reflecting brain activity. This is another powerful characteristic that helps in the online processing of EEG signals for applications related to Brain-Machine Interfaces (BMI).

5 Conclusions and future work

In this paper, we investigated the relation between bodily expressions and their impressions on the observer. We started by generating six bodily expressions, then we classified them to belong to two categories according to their expressiveness, namely: pleasant and unpleasant. Their expressiveness was confirmed statistically by a self-reporting experiment where a number of volunteers answered questionnaires about the bodily expressions. Afterwards, we conducted an experiment to assess the impressions on the observers while watching the considered bodily expressions by collecting the observer's brain activity using electroencephalogram (EEG). The method adopted for spectral analysis revealed a correlation between the power level of low-alpha frequency band (8~11[Hz]) and the category of the observer bodily expression. The reproducibility or repeatability of this band's reaction was confirmed with a third experiment where a subject observed repeatedly candidate bodily expressions for each category. These results have opened the opportunity to utilize the change in the power level of low-alpha frequency band to examine the capacity of a humanoid robot in activating the social perception system in a human observer. A challenging problem that rises from this result is about the degree to which such reaction appears when observing robots with different human-like physical and behavioral characteristics. The understanding of which robot properties are necessary or sufficient to activate the social perception system in an observer is of particular interest.

Another important direction was to define a method which can assess and recognize the impression category from the observer's brain activity. For this reason, we presented a computational method to use for the recognition of the impression of bodily expressions.

The self-organizing maps (SOM), which allows at the simultaneous reduction of the amount of data and its projection to a lower dimensional space, was used for the recognition. It was shown that SOMs achieve relatively high recognition rates considering that the data we used is not filtered for noise elimination. The existence of high individual differences in the considered data was handled successfully with the SOM, because of its ability to separate signals resulting from different processes.

Future research directions should focus on improving the recognition rate to over 90% and to try to recognize a refined classification of bodily expressions. A link with human motion

styles (Hsu et al., 2005) would be interesting to provide more details about the bodily expressions. There is an extensive work on generating emotional motions (Amaya et al. 1996; Lim et al. 2004) that could be incorporated in this research. Reaching this goal would enable us to build an adaptive interface that makes the robot judge the effect of its own gestures on the users and allows it to change its behavior accordingly.

6. Acknowledgements

7. References

T. Allison, A. Puce & G. McCarthy (2000). Social perception from visual cues: Role of the STS region. Trends in Cognitive Sciences, 4(7):251–291.

K. Amaya, A. Bruderlin & T. Calvert (1996). Emotion from motion. In Graphics Interface '96, pages 222–229.

I. Bartenieff & D. Lewis (1980). Body Movement: Coping with the Environment. Gordon and Breach Science Publishers, USA.

C. Breazeal (2002). Designing Sociable Robots. The MIT Press, USA.

S. Cochin, C. Barthlemy, B. Lejeune, S. Roux & J. Martineau (1998). Perception of motion and qEEG activity in human adults. Electroencephalography and Clinical Neurophysiology, 107:287–295.

E. Hsu, K. Pulli, & J. Popovic (2005);. Style translation for human motion. ACM Transactions on Graphics, 24(3):1082–1089.

J. Ido, K. Takemura, Y. Matsumoto & T. Ogasawara (2002). Robotic receptionist ASKA: a research platform for human-robot interaction. In IEEE Workshop of Robot and Human Interactive Communication, pages 306–311.

M. Ito & J. Tani (2004). On-line imitative interaction with a humanoid robot using a dynamic neutral network model of a mirror system. Adaptive Behavior, 12(2):93–115.

H. H. Jasper (1958). The ten-twenty electrode system of the international federation. Electroencephalography and Clinical Neurophysiology, 10:371–375.

G. Johansson (1973). Visual perception of biological motion and a model for its analysis. Perception and Psychophysics, 14(2):201–211.

S.-L. Joutsiniemi, S. Kaski & T. A. Larsen (1995). Self-organizing map in recognition of topographic patterns of EEG spectra. IEEE Transactions on Biomedical Engineering, 42(11):1062–1068.

S. Kaski (1997). Data Exploration using Self-Organizing Maps. PhD thesis, Helsinki University of Technology.

A. Khiat, M. Toyota, Y. Matsumoto & T. Ogasawara (2006). Brain activity in the evaluation of the impression of robot bodily expressions. In IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 5504–5508.

T. Kohonen (1982). Self-organized formation of topologically correct feature maps. Biological Cybernitics, 43:59–69.

H. Kozima (2002). Infanoid: A babybot that explores the social environment, pages 157–164. Socially Intelligent Agents: Creating Relationships with Computers and Robots. Kluwer Academic Publishers, The Netherlands.

H. Lim, A. Ishii & A. Takanishi (2004). Emotion-base biped walking. Robotica, 22:577–586.

M. L. Nicolelis (2001). Actions from thoughts. Nature, 409(18):403–407.

T. Nomura, T. Kanda & T. Suzuki (2006). Experimental investigation into influence of negative attitudes toward robots on human-robot interaction. AI & Society, 20(2):138–150.

D. A. Norman, A. Ortony, and D. M. Russell (2003). Affect and machine design: Lessons for the development of autonomous machines. IBM Systems Journal, 42(1):38–44.

F. E. Pollick, H. M. Paterson, A. Bruderlin & A. J. Sanford (2001). Perceiving affect from arm movement. Cognition, 82:B51–B61.

G. Rizolatti & L. Craighero (2004). The mirror-neuron system. Annual Reviews of Neuroscience, 27:169–192.

A. Schlögl, S. J. Roberts & G. Pfurtscheller (2000). A criterion for adaptive autoregressive models. In IEEE International Conference of Engineering in Medicine & Biology Society, pages 1581–1582.

S. Schaal (1999). Is imitation learning the route to humanoid robot? Trends in Cognitive Science, 3(6):233–242.

P. Shaver, J. Schwartz, D. Kirson & C. O'Connor (1987). Emotion knowledge: Further exploration of a prototype approach. Journal of Personality and Social Psychology, 52(6):1061–1086.

K. Shinozawa, F. Naya & K. Kogure (2005). Differences in effect of robot and screen agent recommendations on human decision-making. International Journal of Human Computer Study, 62(2):267–279.

T. Tanaka, T. Mori & T. Sato (2001). Quantitative analysis of impression of robot bodily expression based on laban movement theory. Journal of Robotics Society of Japan, 19(2):252–259. *(in Japanese)*

J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan & M. L. Nicolelis (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature, 408:361–365.

C. M. Whissel (1989). The Dictionary of Affect in Language, volume 4 of Emotion: Theory, research and experience. The measurement of emotions. Academic Press, USA.

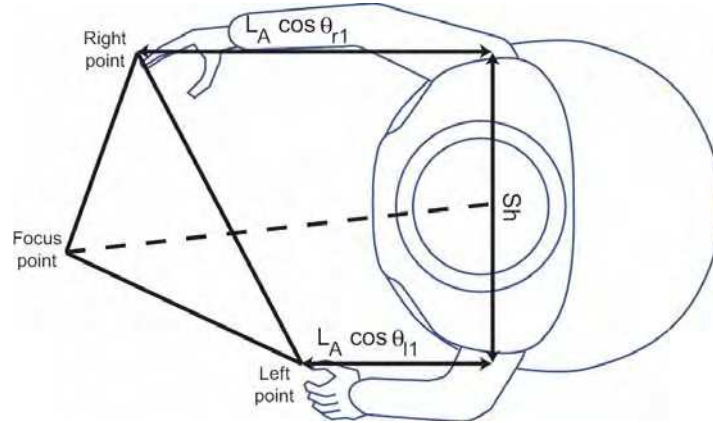## Appendix: Mathematical expression of Laban features of the robot ASKA

Fig. 11. Diagram of table plane superposed on a top view of the robot ASKA.

The mathematical definition of Laban features (Shape and Effort) using the robot's kinematic and dynamic information is given such that larger values describe fighting movement forms and smaller values describe indulging ones (Tanaka et al., 2001). Bartenieff and Lewis stated in (Bartenieff & Lewis, 1980) that the Shape feature describes the geometrical aspect of the movement using three parameters: table plane, door plane, and wheel plane. They also reported that the Effort feature describes the dynamic aspect of the movement using three parameters: weight, space, and time. The robot's link information which will be used in the features definitions is given in Fig. . In order to simplify the mathematical description, a limited number of joint parameters were considered in this definition, namely: the left arm θ_{l1}, the right arm θ_{r1}, the neck δ_1, the face δ_2, the left wheel ω_l, and the right wheel ω_r. The remaining parameters were fixed to default values during movement execution.

Using the diagram shown in Fig. , the table parameter of feature Shape represents the spread of silhouette as seen from above. It is defined as the scaled reciprocal of the summation of mutual distances between the tips of the left and the right hands along with a focus point, as shown in **(8)**.

$$Shape_{table} = s\big/\big(T_{LF} + T_{RF} + T_{LR}\big) \tag{8}$$

where,

$$T_{LF} = \sqrt{\left(L_F \cos\delta_1 \cos\delta_2 - L_A \cos\theta_{l1}\right)^2 + \left(Sh\big/2 - L_F \cos\delta_1 \sin\delta_2\right)^2} \ ,$$

$$T_{RF} = \sqrt{\left(L_F \cos\delta_1 \cos\delta_2 - L_A \cos\theta_{r1}\right)^2 + \left(Sh\big/2 - L_F \cos\delta_1 \sin\delta_2\right)^2} \ ,$$

$$T_{LR} = \sqrt{Sh^2 + \left(L_A \cos\theta_{l1} - L_A \cos\theta_{r1}\right)^2} \ ,$$

The point of focus is set at the fixed distance $L_F = 44$ [cm] in the gaze line of the robot's head. $Sh = 33$ [cm] is the distance between the shoulders; $L_A = 44$ [cm] is the arm's length during movement execution and s is a scaling factor. The door parameter of feature Shape represents the spread of the silhouette as seen from the font. It is defined as the weighted sum of the elevation angles of both arms and the head as shown in **(9)**. The sine is used to reflect how upward or downward is each joint angle. The weights d_l, d_r, d_n were fixed empirically to 1.

$$Shape_{door} = d_l \sin \theta_{l1} + d_r \sin \theta_{r1} + d_n \sin \delta_{n1} \qquad (9)$$

The wheel parameter of feature Shape represents the lengthwise shift of the silhouette in the sagittal plane. It is defined as the weighted sum of the velocities of the robot and the velocities of the arm extremities as shown in **(10)**. Weights were set empirically to -8 for w_t , to -1 for w_l and w_r .

$$Shape_{wheel} = w_t v_{tr} + w_l L_A \frac{d}{dt} \cos \theta_{l1} + w_r L_A \frac{d}{dt} \cos \theta_{r1} \qquad (10)$$

The weight parameter of feature Effort represents the strength of the movement. It is defined in **(11)** as the weighed sum of the energies exhibited during movement per unit time at each part of the body. Weights were adjusted with respect of to the saliency of body parts. Relatively large weights $e_{rt} = e_{tr} = 5$ were given to the movement of the trunk and smaller weights were given to the movements of the arms $e_l = e_r = 2$ and the neck $e_n = 1$.

$$Effort_{weight} = e_l \dot{\theta}_{l1}^2 + e_r \dot{\theta}_{r1}^2 + e_n \dot{\delta}_{n1}^2 + e_{tr} v_{tr}^2 + e_{rt} v_{rt}^2 \qquad (11)$$

where $v_{tr} = \dot{\omega}_l + \dot{\omega}_r$ is the translation velocity and $v_{rt} = \dot{\omega}_l - \dot{\omega}_r$ is the rotation velocity.

The space parameter of feature Effort represents the degree of conformity in the movement. It is defined in **(12)** as the weighed sum of the directional differences between elevation angles of the arms and the neck as well as the body orientation. Weights are also defined empirically by giving advantage to the arms' bilateral symmetry $s_{lr} = -5$ and body orientation $s_{rt} = -5$ over the other combinations $s_{ln} = s_{rn} = -1$.

$$Effort_{space} = s_{rt}|\omega_{rt}| + s_{lr}|\theta_{l1} - \theta_{r1}| + s_{ln}|\theta_{l1} - \delta_{n1}| + s_{rn}|\theta_{r1} - \delta_{n1}| \qquad (12)$$

The time parameter of feature Effort represents the briskness in the movement execution and covers the entire span from sudden to sustained movements. It is defined in **(13)** as the ratio indicating the number of generated commands per time unit.

$$Effort_{time} = \frac{number\ of\ generated\ commands}{time\ span} \qquad (13)$$

Humanoid Robots: New Developments

Edited by Armando Carlos de Pina Filho

For many years, the human being has been trying, in all ways, to recreate the complex mechanisms that form the human body. Such task is extremely complicated and the results are not totally satisfactory. However, with increasing technological advances based on theoretical and experimental researches, man gets, in a way, to copy or to imitate some systems of the human body. These researches not only intended to create humanoid robots, great part of them constituting autonomous systems, but also, in some way, to offer a higher knowledge of the systems that form the human body, objectifying possible applications in the technology of rehabilitation of human beings, gathering in a whole studies related not only to Robotics, but also to Biomechanics, Biomimmetics, Cybernetics, among other areas. This book presents a series of researches inspired by this ideal, carried through by various researchers worldwide, looking for to analyze and to discuss diverse subjects related to humanoid robots. The presented contributions explore aspects about robotic hands, learning, language, vision and locomotion.

INTECH
open science | open minds