Relationship between Speakers' Nonverbal Expressions and Features of HRV and EMG during Poster Session

MOTOKI SAKAI[†1], HIROKO TOKUNAGA[†1], NAOKI MUKAWA[†1], MASAKI SHUZOU[†1], and EISAKU MAEDA[†1]

[†1]Tokyo Denki University, School of System Design and Technology

Abstract: A poster session requires the high-level communication skills of a speaker. Thus, it is desirable that these skills are improved with training. For this reason, objective evaluation indices are required. In a previous research, verbal/nonverbal expressions were reported as promising evaluation indices. Following the previous work, this study investigates a relation between biological signals and nonverbal expressions during a poster presentation. If a relationship exists between them, we can automatically assess the performance of the poster presentation by analyzing biological signals. In an experiment of a poster session, the electrocardiogram (ECG) and electromyogram (EMG) signals were measured from one practician of poster presentation and two beginners. Additionally, nonverbal expressions of three students were recorded using a video camera. As features of the ECG and EMG signals, the integrated values of the heart rate variability (HRV) signal's continuous wavelet transformation (CWT) components, RMS of EMG, and integrated values of EMG spectral powers were calculated, which were compared between a practician and beginner. Consequently, significant differences were found in several features. Similarly, differences occurred between the practician and beginners in nonverbal expressions. Therefore, we conclude that a relationship might exist between the difference in nonverbal expression and that in a biological feature.

1. Introduction

A poster session is a type of presentation style, often conducted in many occasions, including academic conferences. A poster session requires a speaker to present information while focusing on the audience's reactions or queries; the required level of communication skill can be high for speakers. Similarly, the audience in a poster session must also acquire effective communication skills to be part of an audience, to ask effective questions, and subsequently to obtain valuable information.

These communication skills required in a poster session are versatile, and are also useful in many circumstances such as a job interview. Thus, it is desirable that these communication skills are improved by effective training. Hence, objective and quantitative evaluation indices are imperative.

To realize a productive poster session, an easily viewable poster and logical explanations are first required. Meanwhile, [1] asserts that the verbal/nonverbal expressions of speakers affect the audience's impression or assessment of a presentation. In summary, verbal/nonverbal expressions not involved in the explanation content itself can be conductive to the audience's understanding and observation of the presentation. In the previous research [1], the authors manually counted the numbers of meaningful verbal/nonverbal expressions of speakers from recorded video images, and then found that the appearance frequencies of the specific verbal/nonverbal expressions were correlated to the evaluative consequence for a presentation. Following this previous research, we investigate the relationship between nonverbal expressions and biological information. This approach has the following two significances. One is that the amount of time to manually analyze a video image for counting meaningful behaviors can be reduced if the relationship between nonverbal expressions and biological information is revealed. That is, a presentation might be able to automatically be evaluated by analyzing biological information. The other is that a speaker can objectively obtain information about his/her condition during a presentation through biological information. If the effective features of the biological information can be obtained, the speakers can utilize these information to assess themselves and to improve their presentation performance. In summary, this approach is a kind of biofeedback from visualizing his/her physical and mental status during a poster presentation.

Initially, this research investigates whether a relation exists between the features of biological signals and nonverbal expressions of speakers. Among the biological signals, this study adopts the electrocardiogram (ECG) and electromyogram (EMG). From the ECG signal, the heart rate variability (HRV) signal is computed, and meaningful features such as the VLF, LF, and HF can be obtained, which reflects the mental conditions. In fact, the ECG signal is often used to assess acute mental stresses [2], [3]. Likewise, it is known that the EMG signal measured from the trapezius or frontalis muscle can reflect mental stress [4], [5].

In our study, two types speakers, i.e., an experienced speaker (called practician) and an undeveloped one (called beginner) are adopted to compare the polarized evaluatee's presentation performance, and the differences in the features of the ECG, EMG, and the nonverbal expressions between the two speaker types are investigated.

2. Experiment

A poster session experiment was conducted on 23 students: one

master student and 22 junior students. A master student and junior student are regarded as a practician and beginner, respectively. For the junior students, this poster session was held as a rehearsal for a workshop, which is part of the university's academic program. The master student was adopted to compare the junior students' behaviors, execution of presentation, and biological signals with those of the master student's.

In this experiment, two poster sessions were conducted. In session one, five students (four junior and one master students) were speakers, and the remaining 18 junior students were the audience. In session two, three junior students who were not the same students in session one were the speakers, and the remaining 20 students were the audience.

The duration of one session is 40 minutes. During that time, the speakers wait for the audience to come by their posters, and present information when the audience stops in front of them. Meanwhile, the audience moves about freely in the session room; if they are interested in a speaker's poster, they can audit a presentation. Additionally, the audience can ask a speaker.

During a poster sessxion, TSND-151 and AMP-151 produced by ATR promotions were attached to the speakers and audience, and the ECG and EMG signals were measured. The ECG and EMG were measured at 1000-Hz sampling rate, 16-bit resolution. In this experiment, the left and right trapezius EMGs were measured. The electrode placement of the trapezius EMG is shown in Fig. 1. (Hereafter, only the left trapezius EGM signal will be used). Meanwhile, the ECG signal is obtained by the modified CM5 lead. As shown in Fig.2, a positive electrode was located on the 5-th ribs (this position corresponds to V5 in the 12-lead ECG) and the negative one is attached below the right clavicle. The grand electrode is attached to the lower abdomen. In addition, the utterances and behaviors of the speakers and audience during the poster session were recorded with digital video cameras. As described above, 23 students participated in the poster sessions, but we herein present the evaluation results of only three students (one master and two juniors). Henceforth, this master student is called the "practician" and the two junior students selected in session two are called "beginner A" and "beginner B." Moreover, in the following feature extraction step, the measured data within approximately 25 min of the launch is evaluated.

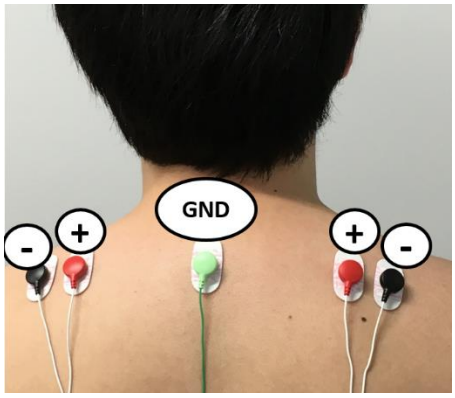

Fig.1: Electrode positions of trapezius EMG

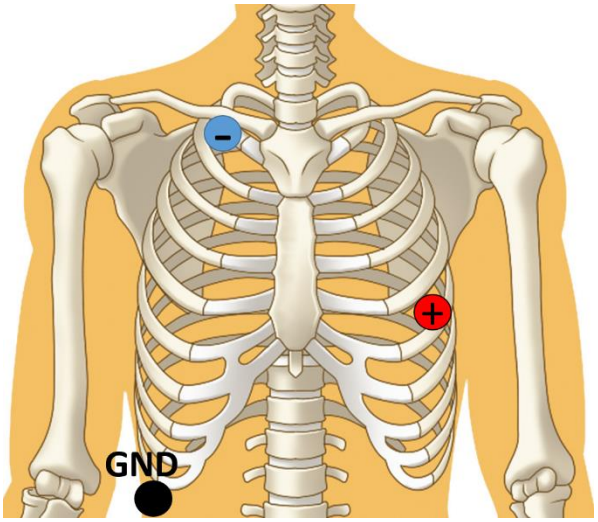

Fig.2: Electrode position of ECG

3. Feature extraction method

From the ECG and EMG signal, many features can be extracted to assess the participant's (or patient's) physical or mental conditions. In both signals, three major classes of features exist: time domain, frequency domain, and time–frequency domain ones. In this research, these three feature types calculated from the ECG and EMG signals are used to evaluate the speakers' condition.

For the ECG signal, a continuous wavelet transformation (CWT)-based time–frequency domain features are used. For the EMG signal, the root mean square (RMS) and integrated values of the spectral power are used as the features. Concrete feature extraction methods are described in 3.1 and 3.2.

3.1 Feature extraction of ECG signal

In this study, the ECG signal features are calculated as follows:

a) The measured ECG signal is filtered by a 0.05–100 Hz bandpass filter.
b) The HRV signal is calculated from the filtered ECG signal.
c) The HRV signal is transformed into the time–frequency domain components using the CWT (scale levels of CWT correspond to 0.0011 Hz through 1.7517 Hz).
d) In each scale, the time–frequency components are integrated every 10 s.

After the processing above, $[105 \times N]$ features are obtained (the number of scales is 105, and N means the number of 10 s-epochs. For a practician, $N = 161$, and for two beginners, $N = 146 \times 2 = 292$).

3.2 Feature extraction of trapezius EMG signal

3.2.1 RMS feature

As with the ECG signal, the RMS is computed every 10s after filtering with the 50–500 Hz bandpass filter; subsequently, $[1 \times N]$ RMS features are obtained.

3.2.2 Spectral features

To extract the frequency domain feature, the following processes are performed.

a) The measured EMG signal is filtered by a 50–500 Hz bandpass filter.
b) The fast Fourier transformation (FFT) is performed for the filtered EMG signal every 10 s.
c) For every 10-s signal processed by the FFT, the integrations of spectral power of $(x \pm 5)$ Hz are computed $(x = 101, 111, 121, 131, \cdots 481)$.

For the spectral feature of the EMG, $[39 \times N]$ features are obtained (the number of spectral features is 39).

3.3 Counting nonverbal expressions in video data

To compare the ECG and EMG features with the speaker's behaviors, meaningful nonverbal expressions such as emblems, deictic gestures, descriptive gestures, and beat gestures were manually counted from the recorded video images of the poster session. These expressions are typical hand gestures, and we often use these gestures to complement a perplexing conversational content. The definitions of these hand gestures are shown in Table. 1.

Table. 1: Definitions of four hand gestures

Emblem	A gesture substituting for words (For example, v-sign, ok-sign, etc.)
Deictic gesture	Pointing to an objective for an explanation.
Descriptive gesture	Gesture to express a location, behavior, or figuration.
Beat gesture	Action related to a thought process.

4. Evaluations and results

The purpose of this evaluation is to compare the practician's features of the ECG and EMG signals with those of the beginner's, and to obtain the effective ones to distinguish between a practician and beginner. As described above, the computed features were the EMG's RMS and integrated value of the spectral power, and the integrated value of the HRV's CWT components. To test the difference between those of the practician and beginners, Welch's t-test was performed for every feature type, RMS, 39 integrated values of the spectral power, and 105 integrated values of the HRV's CWT components (null hypothesis: both the practician's and beginner's features are equal without assuming equal variances).

In regard to the integrated value of the HRV's CWT components, calculations of Welch's t-test showed that there were statistically significant between the practician and beginners in 87 of the 105 frequency bands ($p < 0.001$) (there are no statistically significant values in the following frequency bands: 0.001965, 0.002106, 0.002593, 0.002779, 0.002978, 0.003192, 0.003667, 0.006384, 0.009677, 0.012769, 0.01385, 0.014667, 0.01572, 0.016848, 0.18058, 0.019354, 0.20743, 0.02737 Hz). Regarding the EMG's integrated value of the spectral power, the t-test results showed that they were statistically significant in most frequency bands except 441 Hz at the $p = 0.001$ level. Regarding the EMG's RMS, they were statistically significant at the $p = 0.001$ level.

Figs. 3 and 4 show the p-values computed in each frequency feature (p-values were expressed as $-log_{10}(p - value)$). Fig. 3 illustrates the p-values for the integrated values of the HRV's CWT component. Fig. 4 shows the ones for the EMG's integrated value of spectral power. Additionally, Fig. 4 includes the p-value for the RMS.

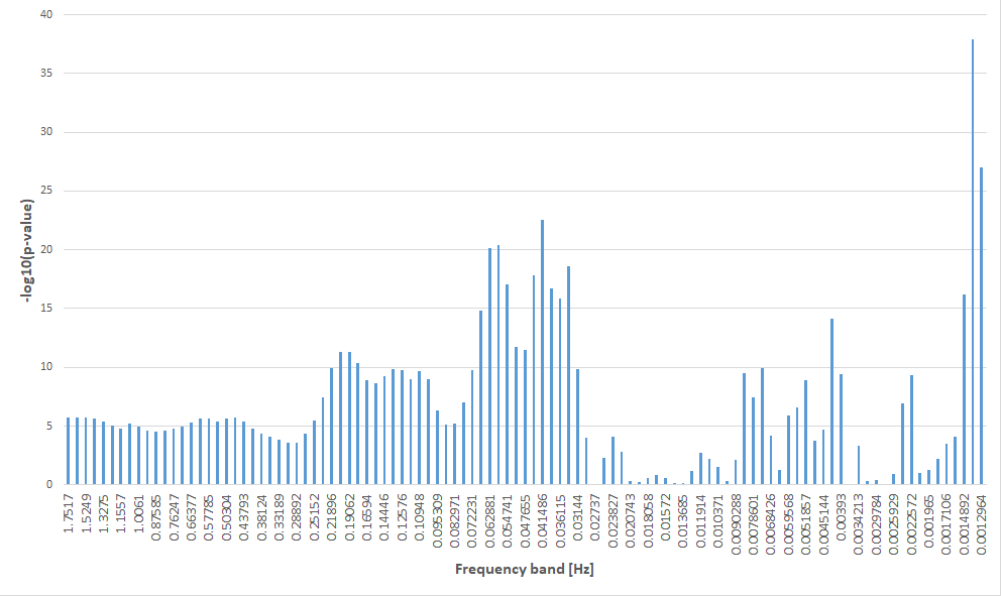

Fig. 3: P-values calculated in each frequency band's feature for the HRV's CWT component

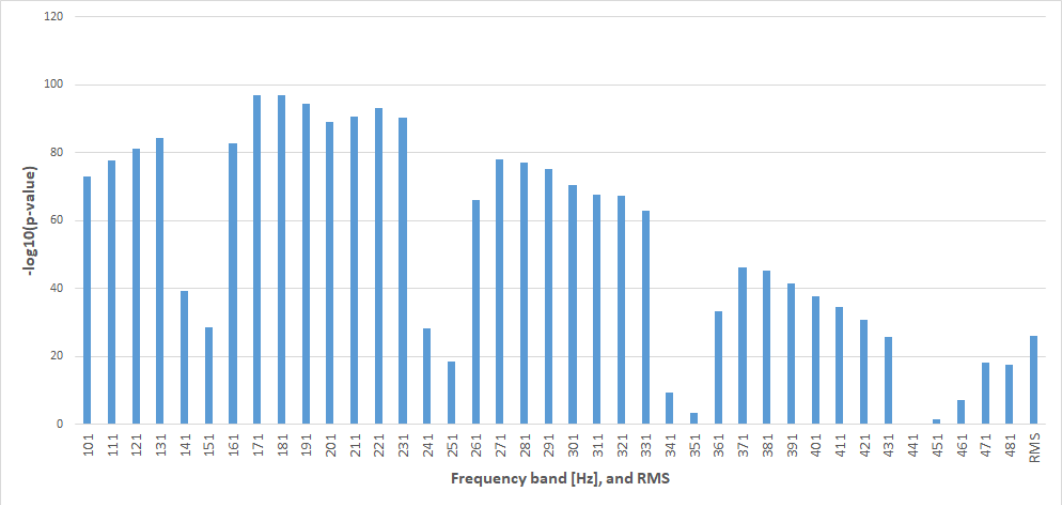

Fig. 4: P-values calculated in each frequency band's feature for the EMG's spectral power and that of the RMS

Fig. 5 illustrates the frequencies of the nonverbal expressions' appearances counted manually in the practician, beginner A, and beginner B.

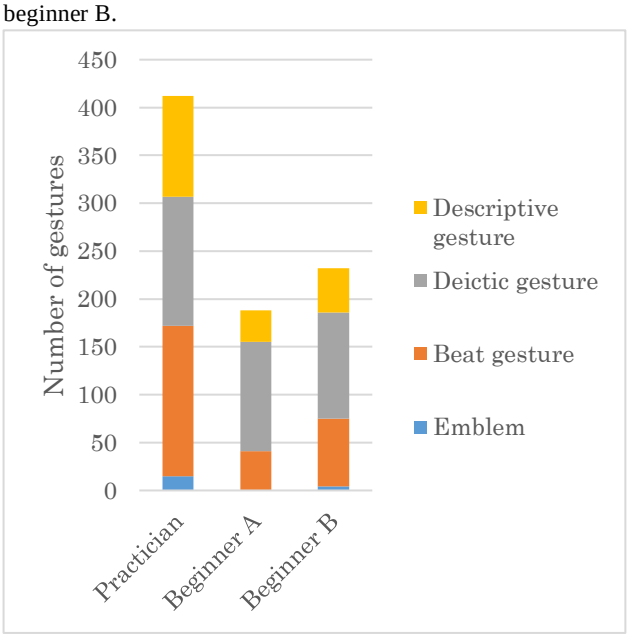

Fig. 5: Appearance frequencies of nonverbal expressions

As an example, the scatter plots between the HRV and EMG features are shown in Figs 6 and 7. Fig. 6 shows the scatter plot when the integrated value of the EMG's spectral power is approximately 171 Hz and the integrated value of the HRV's CWT component corresponding to approximately 0.04 Hz (LF) were selected as the features. In these two features, the p-values were adequately low values. Meanwhile, Fig. 7 shows the scatter plot when the integrated value of the EMG spectral power is approximately 341 Hz and the integrated value of the HRV's CWT component corresponding to approximately 0.15 Hz (HF) were selected as the features. In the two features shown in Fig. 7, their p-values were higher than those of the features shown in Fig. 6, as much as p-values < 0.001.

5. Discussion

As shown in Fig. 5, the appearance frequencies of nonverbal expressions between the practician and beginners were notably different apart from the deictic gesture. In particular, the appearance frequency of the practician's beat gesture was approximately twice or thrice higher than those of the beginners'. This fact implies that the practician could make a logical argument in his head compared to the beginners. Meanwhile, the beginners might not extemporarily explain their investigations and only recited what they wrote in their poster, because the number of deictic gestures were larger than those of the descriptive gestures. Hence, the presentation performance can be quantitatively evaluated with the analysis of nonverbal expressions. In this case, we could verify that the master student (practician) was certainly accustomed to giving a poster presentation than the junior students (beginners).

Meanwhile, several practician's and beginners' biological features are shown to be statistically significant in Figs. 3 and 4. In particular, the significant differences are notable at approximately 0.04–0.07 Hz (LF band) and 0.0013–0.0014 Hz (VLF band) in the HRV features. In the EMG, frequency band features between 171 and 231 were notably significant. These imply that a causal relationship exists between the difference in nonverbal expression and that in the biological feature. For example, Fig. 6 shows that the LF values of the practician tend to be somewhat larger than those of the beginners'. This means that the practician gave a presentation in a more relaxed manner compared to the beginners, because the LF component reflects a parasympathetic activity. Likewise, Fig. 6 shows that the beginners' trapezius is more tensed than that of the practician, because the beginners' features around the 171-Hz frequency band were higher than that of the practician's (high EMG value means high muscle tension). In summary, the beginners' bodies were stiffened during their poster presentations. However, it is easy to assume that a relaxed state leads to the desired result. Therefore, a synthetic interpretation for a number of ECG and EMG features, which were statistically significant between the practician and beginners, must be considered with a sufficient number of measurement dataset.

Hence, the effective feature set must be investigated to explain the favorable conditions for a presentation. In this research, many features between the practician and beginners were statistically significant. However, as shown in Figs. 6 and 7, the significant features could not always distinguish a practician from a beginner (the discrimination between a practician and beginner shown in Fig. 7 was less successful than that shown in Fig. 6). Therefore, more effective features will be determined by feature selection methods such as the genetic algorithm in the future work.

In addition, an advantageous method to present effective biological features for an estimator or students will be studied.

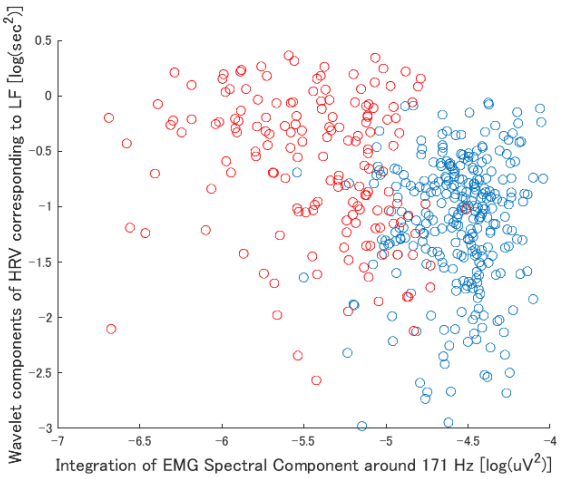

Fig. 6: Scatter plot between HRV and EMG features resulting in adequately low p-values; red: practician, blue: beginner

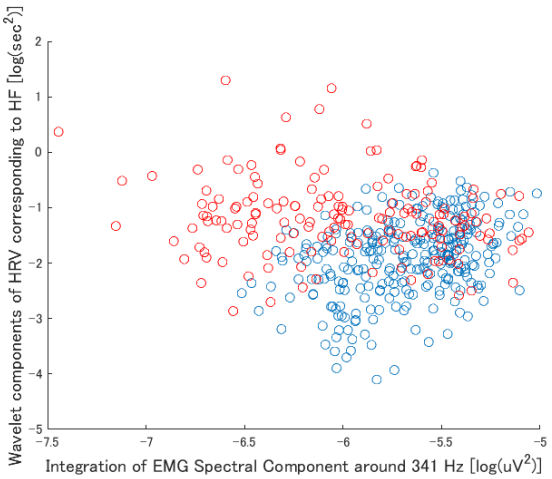

Fig. 7: Scatter plot between HRV and EMG features resulting in higher p-values, as much as $p < 0.001$

6. Conclusion

The goal of this study was to investigate whether a relationship exists between biological signals and nonverbal expressions during a poster presentation. In this study, the ECG and EMG signals were adopted as the biological signal, which were measured from one master student (practician) and two junior students (beginners) during the poster sessions. At a certain time, the nonverbal expressions of the abovementioned students were recorded using a video camera. From the measured ECG and EMG, some integrated values of the HRV's CWT components, RMS of EMG, and some integrated values of the EMG's spectral power were computed as features, and were compared between a practician and beginners. The results show that they were statistically significant in some features. Similarly, a difference occurred between a practician and beginners in nonverbal expressions such as beat and descriptive gestures. Hence, we concluded a causal relationship might exist between the difference in nonverbal expression and that in the biological feature. In the future work, a feature selection method to distinguish a practician from a beginner will be investigated and the evaluation result will be presented.

References

[1] N. MUKAWA, T. NAKAYAMA, H. TOKUNAGA, J. YAMATO, and N. YAMASHITA, "Analysis of Verbal / Nonverbal Expressions of Speakers and Evaluators'. Evaluations in Group Discussions, " Institute of Electronics, Information and Communication Engineers Journal D, Vol.J101-D, No.2, pp.284–293, 2018.

[2] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, L. Pecchi, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis, " *Biomedical Signal Processing and Control*, vol. 18, pp. 370–377, 2015.

[3] K. Patil1, M. Singh, G. Singh, and A. N. Sharma, "Mental Stress Evaluation using Heart Rate Variability Analysis: A Review," *International Journal of Public Mental Health and Neurosciences*, vol. 2, Issue 1, pp. 10–16, 2015.

[4] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "*Trapezius* Muscle *EMG* as Predictor of *Mental Stress*, " *Proceedings of WH '10 Wireless Health 2010*, pp. 155–163, 2010.

[5] K. Yanagisawa, T. Sugahara, M. Kamijo, T. Sadoyama, Y. Shimizu, "The proposal of the measurement method of the mental stress using the surface electromyogram," *Proceedings of The Japan Joint Automatic Control Conference*, vol. 48(0), pp. 257–257, 2005.

Acknowledgments This work was supported by the Grant-in-Aid for Cutting-Edge Research, aiming for the social implementation of the next-generation artificial intelligence technology of the New Energy and Industrial Technology Development Organization (NEDO) of Japan.