Multimodal BigFive Personality Trait Analysis Using Communication Skill Indices and Multiple Discussion Types Dataset

Candy Olivia Mawalim[1]([✉]), Shogo Okada[1,3], Yukiko I. Nakano[2,3], and Masashi Unoki[1]

[1] Graduate School of Advanced Science and Technology,
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
{candyolivia,okada-s,unoki}@jaist.ac.jp
[2] Seikei University, Musashino, Tokyo, Japan
y.nakano@st.seikei.ac.jp
[3] RIKEN AIP, Tokyo, Japan

Abstract. This paper focuses on multimodal analysis in multiple discussion types dataset for estimating BigFive personality traits. The analysis was conducted to achieve two goals: First, clarifying the effectiveness of multimodal features and communication skill indices to predict the BigFive personality traits. Second, identifying the relationship among multimodal features, discussion type, and the BigFive personality traits. The MATRICS corpus, which contains of three discussion task types dataset, was utilized in this experiment. From this corpus, three sets of multimodal features (acoustic, head motion, and linguistic) and communication skill indices were extracted as the input for our binary classification system. The evaluation was conducted by using F1-score in 10-fold cross validation. The experimental results showed that the communication skill indices are important in estimating agreeableness trait. In addition, the scope and freedom of conversation affected the performance of personality traits estimator. The freer a discussion is, the better personality traits estimator can be obtained.

Keywords: Multimodal analysis · Multiple discussion ·
BigFive personality traits · Communication skill · Task type

1 Introduction

Personality traits are important in reflecting the way humans think, feel and act. In many cases, knowing the personality traits of an individual can give several advantages. For instance, in hiring new staff, someone with a good personality is more preferable. Consequently, having a general measurement of personality

traits is crucial. The BigFive factors [1] are well-known as the most general personality traits measurement. This measurement consists of five traits, including openness, conscientiousness, extraversion, agreeableness, and neuroticism. Psychologists usually evaluate these traits by using standardized factor analysis of personality description questionnaires. However, this manual personality traits evaluation is time-consuming and expensive. Therefore, measuring these traits automatically has become a great interest in the computing field.

In recent years, prior studies focus on using automatic nonverbal analysis for numerous sorts of applications, including the estimation of personality traits. The nonverbal features were obtained from audio and visual data based on the knowledge of social science. For instance, in [2], personality trait was modeled by using audiovisual data of intrapersonal communication. In addition, modeling personality trait by dyadic interactions from body language and speech information was also conducted in [3].

As the time slipped by, investigation on interpersonal communication has been considered to predict either the personality traits or the other functional roles of the participant. It has been reported that using interpersonal communication (such as a group discussion), in which there exists group interaction, can achieve promising performance for detecting some speaker-related variables. For instance, [4] investigated the speaker role in group discussion. [5] attempted to detect the functional roles of each participant in group conversation. Furthermore, personality traits have also been investigated by using co-occurrent multimodal event discovery approach [6]. In this research, we conducted a multimodal analysis from multiple discussion datasets to estimate the BigFive personality traits. The group discussion approach was used since the way a person expresses their opinion and their response in group discussion have a close relationship with their personality traits.

This paper has two novel points. First, we investigated the effectiveness of the communication skill for predicting BigFive personality traits. As we know, social communication skill helps humans exchange their thought in a more convincing way. Furthermore, people with good communication skill tend to have an impressive personality. Second, we investigated whether the discussion task type affected the personality traits of the participants. MATRICS corpus introduced in [7] which consists of three discussion tasks were employed in this research. The discussion tasks were varied with regard to the scope and freedom on dialog structure of the conversation.

2   Related Work

The aim of automatic personality computing is to model the relationship between stimuli (everything observable people do) and the outcomes of the social perception processes (how we form impressions about others). There have been many studies of on a multimodal analysis of the personality trait inference. For instance, Pianesi et al. [8] conducted a personality prediction for each participant using self-reported questionnaires. Aran et al. [9] presented an analysis

of personality prediction in small groups on the basis of trait attributes from external observers. Jayagopi et al. [10] proposed a mining approach for finding context features to link to group performance and personality traits. Okada et al. [6] proposed another mining approach to extract co-occurrent events between multimodal time-series data for personality classification. Batrinca et al. [11] conducted a comparative analysis to investigate the difference in the recognition accuracy of personality traits between a human-machine interaction (HMI) setting and a human-human interaction (HHI) setting. Valente et al. [12] conducted personality modeling using dialog acts with speaking activity, prosody, and n-gram distributions.

Besides, several works also focused on improving the accuracy of BigFive prediction. Fang et al. [13] conducted BigFive prediction by using three different nonverbal feature categories, i.e intrapersonal features, dyadic features, and one-vs-all features. On the other hand, Lin et al. [14] attempted to predict the BigFive by modeling vocal behaviors of participants using the interaction-based mechanism in BLSTM.

According to literature [15], several experiments have successfully confirmed the influence of personality traits towards numerous human behavior aspects, such as leadership and job performance. communication skill is also one of the most important human behavior aspects which can lead to creating a successful global relationship. Hence, the association between communication skill and human personality trait has not been investigated yet. Utilizing the communication skill indices for the personality trait inference is one of the main differences between this research and the previous works.

In addition, a comparative analysis of task types varied in the scope and freedom of conversation was conducted for classifying the personality traits. This is also the distinctive point of this research. The prior work of Okada et al. [16] suggested that depending upon the assessed task, people show different manner (different effective multimodal features) in group communication. In contrast, this research aims to investigate the relationship between the assessed task type and the predictive level of BigFive personality traits.

3   Multimodal Data Corpus

The MATRICS multimodal data corpus presented in [7] was employed in this research. This corpus consists of head motion data, audio data, and video data. The head motion data was obtained by an accelerometer and the recorded audio were used to form the acoustics and linguistic features. Previously, in [16], the communication skill indices which assessment by human resource management experts by using video data was the target of the inference. As for now, we aim to confirm how is the relationship between these indices and the BigFive personality traits. The BigFive personality traits scores were annotated by using the self-questionnaire survey (as the standard method in physiology domain).

The MATRICS corpus is a Japanese group discussion corpus which contains 10 discussion groups with 4 participants each. For every discussion groups, three

tasks were set for the discussion. The tasks were varied in regards to the scope and freedom on dialog structure of the conversation. The first task is defined as an in-basket task. In this task, the participants acted as the executive committee members who required to select an invited guest for a school festival. Most prior information was provided in this task. The second task is defined as a case study with prior information. In this task, the participants required to create a food and beverage booth for a school festival. Some information was provided with regard to the booth. Lastly, the third task is defined as a case study without prior information. In this task, the participants had to create a two-day itinerary plan in Japan for their foreign friends. Every participant can express their thought freely without time limit per each individual.

4   Feature Representation

We extracted self-context features, including three sets of multimodal features (acoustic features, linguistic features, and head motion features) and communication skill indices. The acoustic features and linguistic features were extracted from audio data and the manual transcription of the discussion dialog. The head motion features were extracted from head accelerator data. The communication skill features were assessed manually by human resource management experts. All the features were normalized by using z-score normalization. The feature sets were summarized in Table 1.

Table 1. Summary of feature sets for the BigFive personality traits estimation

	Variables
Acoustic feature (AFs)	4 energy related LLDs
	54 spectral LLDs
	6 voicing related LLDs
Linguistic features (LFs)	5 PoS tags (number of noun, verb, new noun, interjection, filler)
	12 dialog tags
	3 speech act tags
	2 semantic tags
Head motion (HMs)	Mean of movement
	Deviation of movement
	Mean of movement while speaking
	Deviation of movement while speaking
	Difference of movement while speaking
Communication skill (CSs)	Listening attitude
	Smooth interaction
	Aggregation of opinions
	Communication own claim
	Logical and clear presentation
	Total communication skill

4.1   Acoustic Features

The acoustic features were extracted from each participant speech by using the speech features extractors openSMILE [17]. The unified test-bed for perceived speaker traits configuration file [18] were used to obtain 6,125 features. These features were derived from 64 low-level descriptors (LLDs) (the detail is shown in Table 2). We used these features since these features are considered as the baseline in speaker trait research [19].

Table 2. 64 LLDs of the INTERSPEECH 2012 speaker trait challenge [18]

4 energy related LLDs
Sum of auditory spectrum (loudness)
Sum of RASTA-style filtered auditory spectrum
RMS energy
Zero-crossing rate
54 spectral LLDs
RASTA-style auditory spectrum, bands 1–26 (0–8 kHz)
MFCC 1–14
Spectral energy 250–650 Hz, 1–4 kHz
Spectral roll off point 0.25, 0.50, 0.75, 0.90
Spectral flux, entropy, variance, skewness, kurtosis
Slope, psycho-acoustic sharpness, harmonicity
6 voicing related LLDs
F0 by SHS + viterbi smoothing, probability of voicing
logarithmic HNR, jitter (local, delta), shimmer (local)

4.2   Linguistic Features

The linguistic features consist of part of speech (PoS), dialog act, and semantic tag. These features were extracted using the same approach in [16]. The PoS features were extracted from the manual transcription by using a Japanese morphological analysis tool, MeCab [22]. The number of nouns, verbs, new nouns (the nouns which are spoken for the first time in the discussion), interjection (the word or phrase to convey emotion or feeling of the speaker), and filler (the word or phrase for filling an interlude in an utterance of conversation) belonged to this feature set. The dialog act and semantic tag set consist of 17 tags. Twelve tags came from DAMSL (Dialog Act Markup in Several Layers) [20] and MRDA (Meeting Recorder Dialog Act) [21] tag set, including "conversational opening", "open question", "suggestion", "backchannel", "open opinion", "partial accept", "accept", "reject", "understanding check", "other question", "WH-question", and "y/n question". the other five tags were defined in [16], which consist three

speech act tag ("plan", "agreement", and "disagreement") and two semantic tags ("describe fact" and "reason").

4.3   Head Motion Features

We utilized five features to represent head motion, i.e. mean and deviation of head movement, mean and deviation of head movement while speaking, and the range of movement while speaking. The head movement was calculated as the norm of the head acceleration at one duration of time. On the other hand, the head accelerator data were joined with the speaking time data to obtain the head movement while speaking.

4.4   Communication Skill Indices

We employed six features for representing the communication skill, i.e. listening attitude, smooth interaction, aggregation of opinions, communicating one's own claim, logical and clear presentation, and total communication skill. Listening attitudes reflect the participant listening manner towards other participants. Smooth interaction captures the efficiency of information exchange of the participant in the group discussion. Aggregation of opinions represents how well a participant could organize and summarize other opinions. Communicating one's own claim reflects how the participant could express appropriate information in every kind of situations. The logical and clear presentation reflects the logic and coherence of a participant in expressing their opinions. Finally, the total communication skill is the total of all five other features.

5   Experimental Setting

The objectives of this experiment are: (1) to clarify the effective features multimodal (verbal and nonverbal) features and communication skill indices to predict the BigFive personality traits and (2) to identify the relationship among multimodal features, discussion type, and the BigFive personality traits. Since the acoustic features were designed in binary classification environment, we also performed binary classification tasks for achieving the objectives. We used 99 out of 120 data samples since there were some missing values in head motion features or the problem with audio files. The target for inference is the BigFive personality traits, including neuroticism, openness, conscientiousness, agreeableness, and extraversion. The assessment scores from experts were classified into high or low (with threshold $= 50$).

Comparative Tasks
In this experiment, we also compared four tasks [16]. Task 1 is defined as the in-basket task (32 samples). Task 2 is defined as a case study with prior information (36 samples). Task 3 is defined as a case study without prior information (31 samples). All tasks (99 samples) is defined as a combination of task 1, task 2,

and task 3. The utilized dataset for classifying the BigFive personality traits in each task refers to the task type explained in Sect. 3.

Comparative Feature Sets
The 15 feature sets shown in Table 3 were compared to analyze the contribution of each feature set in estimating BigFive personality traits.

Table 3. Comparative feature sets for classifying BigFive personality traits

Unimodal
AFs: Acoustic Features
HMs: Head Motion Features
LFs: Linguistic Features
CSs: communication skill
Bimodal
AF_HM: Fusing AFs and HMs
AF_LF: Fusing AFs and LFs
AF_CS: Fusing AFs and CSs
HM_LF: Fusing HMs and LFs
HM_CS: Fusing HMs and CSs
LF_CS: Fusing LFs and CSs
Multimodal
AF_HM_LF: Fusing AFs, HMs, and LFs
AF_HM_CS: Fusing AFs, HMs, and CSs
AF_LF_CS: Fusing AFs, LFs, and CSs
HM_LF_CS: Fusing HMs, LFs, and CSs
All: Fusing all features (AFs, LFs, HMs, CSs)

5.1   Classification Techniques

In this study, several classification algorithms implemented in scikit-learn [23] were utilized to investigate the effectiveness of the identified features. Scikit-learn is an open-source machine learning library built in python programming language environment. We investigated the support vector machine (SVM), random forest, Naïve Bayes, and decision tree algorithms. The brief explanation for these algorithms is presented below.

– Support Vector Machine (SVM)
  SVM is considered one of a good classification algorithm in any kind of tasks, for example, text categorization problem and face detection [24]. This technique applies kernel trick (finding optimal hyperplane) for separating or classifying the data. In this experiment, we used the SVM classifier with radial basis function (RBF) kernel.

– Random Forest (RF)
   RF is also considered as an alternative to deal with a big number of
   features. Since it creates a set of decision trees from randomly selected
   training subset, it can reduce overfitting and produce a very robust, high-
   performing model [25]. In this experiment, we used RF classifier with maxi-
   mum depth $= 3$.
– Gaussian Naïve Bayes (GNB)
   Naïve Bayes classifier is well-known for its simpleness because it requires less
   training data to perform classification task [26]. The main disadvantage from
   this classifier is that since it holds NB conditional independence assumption, it
   cannot learn the relationship among the features (may causes oversensitivity
   to redundant features). Although it has some disadvantages, it was reported
   can achieve good performance for some domains. In this experiment, we used
   Gaussian NB.
– Decision Tree (ID3)
   This algorithm is also utilized in this experiment because it is also easy to
   use (no need a big effort on preprocessing the data). In this experiment, we
   use the CART algorithm of the decision tree with the default parameters.

5.2   Evaluation Criteria

F1-score is used to evaluate the performance of our estimators. F1-score conveys
the balance between precision and recall since it is calculated as the harmonic
mean of these two parameters. In this research, we performed k-fold cross-
validation with $K = 10$ to confirm the performance is not overfitting to the
testing data. In the result section, we refer the F1-score as the average of this
10-folds performance, defined as follows:

$$\overline{F_1} = \frac{1}{K} \sum_{k=1}^{K} F_1(k) \times 100\%. \tag{1}$$

6   Result

From Tables 4 and 5, the overall experimental results (for all tasks) show that
random forest technique achieved the best F1-score for estimating neuroticism
(68.07%), openness (63.84%), and agreeableness (73.75%) traits. On the other
hand, Gaussian NB could estimate well the extraversion (64.32%) and SVM
for estimating conscientiousness (65.84%). The best estimators for neuroticism,
extraversion, openness, agreeableness, and conscientiousness were obtained by
using AFs, HMs, HM_CS_LF (combination of HMs, CSs, and LFs), AFs, and
HMs, respectively. Our experimental results also showed that the agreeableness
is the most predictive trait which achieved the best F1-score in almost all cases.
Although AFs has the best contribution as unimodal feature set, CSs played
a slightly less important role compared to AFs in estimating the agreeableness
trait. Compared with the previous research (Okada et al. [6]), these results are

better in estimating neuroticism, openness, agreeableness, and conscientiousness traits. However, for estimating extraversion trait, the research of Okada et al. [6] by using co-ocurrent event discovery could obtain better accuracy (up to 69.61%).

Figure 1 shows the highest classification F1-score with regards to the task types. The average of F1-scores from each task shows that order from less predictive task to more predictive task is from All_Tasks (0.672), Task 1 (0.693),

Table 4. BigFive personality traits estimation using all tasks dataset with regards to machine learning technique

Technique	BigFive personality traits (%)				
	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
SVM	59.26	62.14	56.07	67.96	65.84
RF	68.07	62.81	63.84	73.75	64.32
GNB	59.94	64.32	56.30	64.94	65.60
ID3	63.34	63.96	61.20	65.62	62.88
Best	68.07	64.32	63.84	73.75	65.84
	GNB	RF	RF	SVM	HMs

Table 5. BigFive personality traits estimation using all tasks datasets with regards to feature sets

| Feature set	BigFive personality traits (%)				
	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
AFs	68.07	55.28	52.19	73.75	53.91
HMs	59.94	64.32	47.22	52.18	65.84
CSs	48.23	46.19	53.13	66.78	52.93
LFs	57.64	63.96	58.17	62.25	61.92
AF_HM	54.54	59.84	56.40	62.13	62.61
AF_CS	51.44	59.25	59.32	64.30	60.55
AF_LF	53.76	59.64	52.84	60.32	57.04
HM_CS	60.26	58.50	59.83	67.96	65.44
HM_LF	63.34	60.42	61.83	62.69	62.88
CS_LF	58.45	62.11	62.36	62.72	64.32
AF_HM_CS	55.33	59.77	57.32	63.45	59.47
AF_HM_LF	52.97	62.81	59.40	62.02	58.40
AF_CS_LF	56.08	56.70	61.20	61.63	63.69
HM_CS_LF	62.09	62.06	63.84	61.40	62.49
All	63.82	54.76	53.10	66.29	55.79
Best	68.07	64.32	63.84	73.75	65.84
	AFs	HMs	HM_CS_LF	AFs	HMs
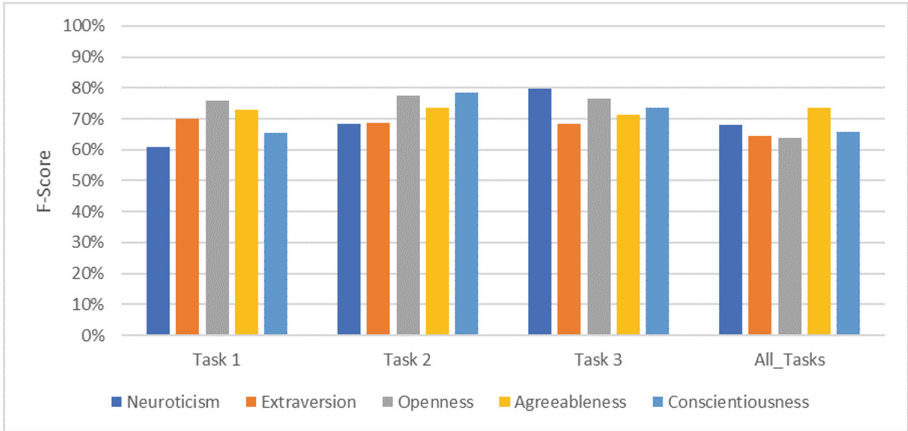
Fig. 1. BigFive personality classification results of depending upon task type (as described in Sect. 5) in term of best F1-score. The order from the less to the most predictive task is from All_Tasks, Task 1, Task 2, and Task 3.

Task 2 (0.733), and Task 3 (0.739), respectively. In addition, we figured out that there are similarities between Task 2 and Task 3. The performance estimators for several personality traits (extraversion, openness, and agreeableness) from Task 2 and Task 3 achieved almost the same F1-score.

7   Discussion

This section contains the discussion of our experimental results. The discussion mainly consists of three parts, i.e. analysis on the relationship between the BigFive personality traits with the multimodal feature sets and analysis on the relationship between the BigFive personality traits with discussion task type.

7.1   Analysis on Relationship Between BigFive and Multimodal Feature Sets

In this section, we discuss the effectiveness of the multimodal features. The overall experimental result in Table 5 shows that for unimodal feature sets, AFs and LFs are highly correlated with the BigFive personality traits (the average F-score for all traits estimation could reach around 60%). This result suggests that the way and the content of speaking play the most important role in estimating the speaker personality traits. Fusing the feature sets can also give a promising result, especially for estimating the openness trait (fusing HMs, CSs, and LFs).

Although the correlation between the personality traits and HMs and CSs feature sets are not as high as AFs and LFs, utilizing them implies better estimation for several traits. For instance, utilizing HMs is best in estimating extraversion and conscientiousness traits. This verified the previous finding that extraversion is positively associated with gesturing [27].

7.2   Analysis on Relationship Between BigFive and Discussion Task Type

Based on the experimental result (as shown in Fig. 1), every task type is highly associated with several traits. For instance, the prediction of extraversion trait is best by using in-basket task (Task 1). Hence, task 2 is highly associated with openness, agreeableness, and conscientiousness traits. On the other hand, the case study without prior information (task 3) can best predict the neuroticism trait. Although not as good as the prediction by task 2, task 3 can also give good performance for predicting openness and conscientiousness traits (the maximum F1-score is more than 60%).

From this result, it is hard to define which task is more predictable for all traits since the highest evaluation score for each trait estimation is not the same. However, based on the average maximum F1-score, task 1 reached a smaller number than task 2. Likewise, task 2 compared with task 3. The ANOVA test was also conducted to check the statistical significance of the result of predicting each trait. From the test, we obtained that only agreeableness trait estimation could reach p-value less than 0.05 (0.0006). Moreover, the estimation of extraversion trait could reach p-value = 0.0713 (weak statistically significant). From this result, we conclude that having free or non-strict conversation dataset may lead to a better automatic BigFive personality traits estimation, especially for estimating the agreeableness trait.

In the case of employing all tasks, the BigFive prediction became more difficult (the average maximum F1-score is the smallest). This was probably because of the characteristics of each task is different. Moreover, this may also be caused by the different manner of the target when the different discussion task was assigned (conclusion from Okada et al. [16]). In other words, we suggest that the homogeneous dataset (unvaried task type) is more predictive than the heterogeneous dataset (varied task type) for predicting BigFive personality traits.

7.3   Analysis on Relationship Between BigFive and Communication Skill Indices

The experimental result shows that taking the CSs indices is useful for estimating agreeableness trait (reached 66.78% F1-score). The high association between CSs and agreeableness may because a good communicator is usually a broad-minded and friendly person. Furthermore, the one who can express an opinion well on a decision affects how he/she can agree or disagree on something. Alternatively, the CSs indices did correlate with the openness and conscientiousness traits. However, we could not conclude that the CSs indices did correlate significantly with BigFive (since the F1-score was not higher than 60%). In addition, because the F1-score for neuroticism and extraversion traits estimation by using CSs indices did not even reach 50%, we conclude that these traits did not correlate with CSs indices.

8   Conclusion and Future Work

This paper presented multimodal analysis in multiple discussion types dataset to estimate the BigFive personality traits, which consists of neuroticism, extraversion, openness, agreeableness, and conscientiousness. This research aimed to clarify the effectiveness of multimodal (verbal and nonverbal) features and communication skill indices to predict the BigFive personality traits and to clarify the relationship among multimodal features, discussion type, and BigFive personality traits. Based on the results shown in Sect. 6, the best estimators for neuroticism, extraversion, openness, agreeableness, and conscientiousness were obtained by using AFs, HMs, HM_CS_LF, AFs, and HMs, respectively. The agreeableness was also reported as the most predictive trait. Although AFs has the best contribution, CSs played a slightly less important role compared to AFs in estimating the agreeableness trait. With regards to the task types, we figured out that the scope and freedom of conversations affected the performance of personality traits estimator. The experimental results suggested that having free or non-strict conversation dataset may lead to a better automatic BigFive personality traits estimation, especially for the agreeableness trait.

As the future work, we would like to investigate not only self-context features but also other-context features (the relationship between the multimodal features of other participants and the personality trait of the speaker). Another important future direction for this work is to consider the dynamics of the features. Zhu et al. [28] suggested that the temporal amplitude modulation played an important role in emotion perception. This implies that utilizing dynamic features instead of static features may lead to a better personality traits inference result. Furthermore, in the current result, we figured out that there may be a non-linearity effect. For instance, as unimodal feature set, AFs were relatively good features compared to HMs for predicting openness trait. However, fusing HMs, CSs, and LFs (HM_CS_LF) resulted in best prediction score. To deal with this issue, the non-linear model will be employed to account for BigFive as the future direction.

References

1. McCrae, R.R., John, O.P.: An Introduction to the five-factor model and its applications. J. Pers. **60**, 175–215 (1992)
2. Celiktutan, O., Eyben, F., Sariyanidi, E., Gunes, H., Schuller, B.: MAPTRAITS 2014 - the first audio/visual mapping personality traits challenge - an introduction: perceived personality and social dimensions. In: 16th International Conference on Multimodal Interaction, ICMI 2014, New York, NY, USA, pp. 3–9 (2014)
3. Metallinou, A., Katsamanis, A., Narayanan, S.: Tracking continuous emotional trends of participants during affective dyadic interactions using body language and speech information. J. Image Vis. Comput. **31**(2), 137–152 (2013)
4. Vinciarelli, A.: Speakers role recognition in multiparty audio recordings using social network analysis and duration distribution modeling. J. IEEE Trans. Multimed. **9**(6), 1215–1226 (2007)

5. Zancanaro, M., Lepri, B., Pianesi, F.: Automatic detection of group functional roles in face to face interactions. In: 8th International Conference on Multimodal Interfaces, ICMI 2006, Banff, Alberta, Canada, pp. 28–34 (2006)
6. Okada, S., Aran, O., Gatica-Perez, D.: Personality trait classification via co-occurrent multiparty multimodal event discovery. In: ACM on International Conference on Multimodal Interaction, ICMI 2015, New York, NY, USA, pp. 15–22 (2015)
7. Nihei, F., Nakano, Y.I., Hayashi, Y., Hung, H.H., Okada, S.: Predicting influential statements in group discussions using speech and head motion information. In: Proceedings of the ACM ICMI, pp. 136–143 (2014)
8. Pianesi, F., Mana, N., Cappelletti, A., Lepri, B., Zancanaro, M.: Multimodal recognition of personality traits in social interactions. In: Proceedings of ACM ICMI, pp. 53–60 (2008)
9. Aran, O., Gatica-Perez, D.: One of a kind: Inferring personality impressions in meetings. In: Proceedings of ACM ICMI, pp. 11–18 (2013)
10. Jayagopi, D.B., Sanchez-Cortes, D., Otsuka, K., Yamato, J., Gatica-Perez, D.: Linking speaking and looking behavior patterns with group composition, perception, and performance. In: Proceedings of ACM ICMI, pp. 433–440 (2012)
11. Batrinca, L., Mana, N., Lepri, B., Sebe, N., Pianesi, F.: Multimodal personality recognition in collaborative goal-oriented tasks. IEEE Trans. Multimed. **18**(4), 659–673 (2016)
12. Valente, F., Kim, S., Motlicek, P.: Annotation and recognition of personality traits in spoken conversations from the AMI meetings corpus. In: Proceedings of INTER-SPEECH, pp. 1183–1186 (2012)
13. Fang, S., Achard, C., Dubuisson, S.: Personality classification and behaviour interpretation: an approach based on feature categories. In: Proceedings of the 18th ACM International Conference on Multimodal Interaction, pp. 225–232. ACM (2016)
14. Lin, Y., Lee, C.: Using interlocutor-modulated attention BLSTM to predict personality traits in small group interaction. In: Proceedings of the 2018 on International Conference on Multimodal Interaction, pp. 163–169. ACM (2018)
15. Lepri, B., Pianesi, F.: Computational approaches for personality prediction. In: Burgoon, J., Magnenat-Thalmann, N., Pantic, M., Vinciarelli, A. (eds.) Social Signal Processing, pp. 168–182. Cambridge University Press, Cambridge (2017)
16. Okada, S., et al.: Estimating communication skill using dialogue acts and nonverbal features in multiple discussion datasets. In: ACM on International Conference on Multimodal Interaction, ICMI 2016, Tokyo, pp. 169–176 (2016)
17. Eyben, F., Wöllmer, M., Schuller. B.: OpenSMILE: the munich versatile and fast open-source audio feature extractor. In: Proceedings of the 18th ACM International Conference on Multimedia, MM 2010, pp. 1459–1462. ACM, New York (2010)
18. Schuller, B., et al.: The INTERSPEECH 2012 speaker trait challenge. In: INTER-SPEECH (2012)
19. Schuller, B., et al.: A Survey on perceived speaker traits: personality, likability, pathology, and the first challenge. Comput. Speech Lang. **29**(1), 100–131 (2015)
20. Core, M.G., Allen, J.: Coding dialogs with the DAMSL annotation scheme. In: AAAI Fall Symposium on Communicative Action in Humans and Machines, pp. 28–35 (1997)
21. Shriberg, E., Dhillon, R., Bhagat, S., Ang, J., Carvey, H.: The ICSI meeting recorder dialog act (MRDA) corpus. In: Proceedings of SIGDIAL, pp. 97–100 (2004)

22. Kudo, T., Yamamoto, K., Matsumoto, Y.: Applying conditional random fields to Japanese morphological analysis. In: EMNLP, vol. 4, pp. 230–237 (2004)
23. Pedregosa, F., et al.: Scikit-learn: machine learning in python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
24. Hearst, M.A.: Support vector machine. J. IEEE Intell. Syst. **13**(4), 18–28 (1998)
25. Breiman, L.: Random forest. J. Mach. Learn. **45**(1), 5–32 (2001)
26. Ratanamahatana, C.A., Gunopulos, D.: Scaling up the naive Bayesian classifier: using decision trees for feature selection (2002)
27. Ethier, N.A.: Paralinguistic and nonverbal behaviour in social interactions: a lens model perspective. UWSpac (2010)
28. Zhu, Z., Miyauchi, R., Araki, Y., Unoki, M.: Modulation spectral features for predicting vocal emotion recognition by simulated cochlear implants. In: INTER-SPEECH, pp. 262–266 (2016)