

Using A Probabilistic Topic Model to Link Observers' Perception Tendency to Personality

Shiro Kumano, Kazuhiro Otsuka, Masafumi Matsuda, Ryo Ishii, Junji Yamato
NTT Communication Science Laboratories
Atsugi, Kanagawa, Japan
Email: kumano@ieee.org

Abstract—Targeting multiparty conversations, the present study aims to elucidate how an observer will tend to perceive others' emotional states; develops a computational model that realizes the automatic inferencing of the observer's perception tendency. This paper proposes a probabilistic model that automatically discovers the correlation between perception tendency, gender, and personality traits of a target observer. Perception tendency, a probability distribution, explains how likely the observer is to perceive a certain state/level of a target emotion. Personality traits are measured by a variety of questionnaires. The proposed model links these three factors via a latent variable and explains observer's characteristics as a mixture of prototypical characters. An experiment is conducted with fifty observers. They watch 97 short conversation videos and give their impressions about the empathy between each interacting pair. The results demonstrate that the proposed method can find a reasonable framework that underlies the factors: e.g. 1) people who have high scores in Davis's empathy measures show empathy-biased response tendency, and 2) people who have strong sense of consideration for others tend to show an extreme response tendency, and such people are likely to be females. The proposed method shows promise in estimating an observer's perception tendency from his/her gender and personality traits, even when the target perception tendency is quite different from the average perception tendency among observers.

Index Terms—perception; tendency; personality; pLSA;

I. INTRODUCTION

Face-to-face conversation is the primary way of sharing information, understanding others' emotion, and making decisions in social life. Unfortunately, it's not so easy for people to fully understand what the others are feeling in a conversation, or reach full agreement about a controversial topic. The quality and efficiency of communication can be enhanced by applying information technologies to conversation support systems, such as in real-time computer-mediated visual telecommunication. This requires the automatic understanding of not only human behavior but also the interlocutors' emotions. Of interest, the main target of automatic meeting analysis is now shifting from behavior to emotion [1], [2], [3].

When considered in social situations, emotion has two distinct aspects: felt emotion, i.e. what the target person is actually feeling, and perceived emotion, i.e. the emotion that is perceived by an observer¹. The latter are vital to understand conversations; the emotion of an interlocutor is perceived by others via his/her behaviors, and the perception evolves

over the course of the interaction. A practical application of perceived emotions is to visualize emotional states of a group meeting for non-meeting members or social psychologists for deep understanding/analysis of the meeting and its atmosphere. Because even no meeting participant knows the real emotions of other participants, perceived-emotion-based descriptions, i.e. third-party objective descriptions, are a reasonable approach.

The affective computing research community is keen to infer perceived emotions. Most works so far try to make interpersonal perception dependent on a specific stimulus, i.e. the verbal/non-verbal behavior of the target [2]; e.g. when the target person is smiling, what type of emotion will observers perceive? To alleviate the subjectivity of perceivers, most previous works gathered the perceptions of multiple observers, and targeted their representative value, e.g. the majority/peak [5] or mean [6], or the distribution [7], [8]. To gather objective descriptions of emotions, observers unacquainted with the target people are often employed, like [7]. Such a collective perception includes a large amount of perception biases, as demonstrated in [9]. That is, different observer groups would produce different collective perceptions. However, most previous computational models fail to explain what emotions a specific perceiver will tend to favor.

We probabilistically model the relationship between the perception tendency of the perceiver and his/her gender and personality traits, as extracted from existing psychological questionnaires. Here, perception tendency means the type of emotion an observer is likely to perceive without considering the stimulus; e.g. a bias to positive/negative emotion, or a central tendency or extreme response tendency, i.e. tendency to prefer or avoid the center level of emotion. Perception tendency is expressed as a probability distribution.

Our model explains the characteristics of the target observer as a mixture of latent prototypical characters that are automatically discovered from a dataset. Each prototypical character is assumed to generate, probabilistically, a unique perception tendency. The proposed model can be considered as a family of probabilistic topic models, often called probabilistic latent semantic analysis (pLSA) [10]; which are widely used for information retrieval and natural language processing, where latent variables are often called topics. To the best of our knowledge, this is the first study to use pLSA to link the perception tendency of an observer to his/her gender and

¹Cowie [4] calls them cause- and effect-type descriptions, respectively.

personality traits. As key emotional states in conversation, this paper targets pair-wise empathy/antipathy (emotional contagion/conflict) between a pair of conversation interlocutors on a 5-point bipolar scale, like [7]. However, a variety of discretized perception targets, e.g. categorical and Likert-scale descriptions of perception, are also supported.

The proposed framework does not, per se, output the perception of a specific observer *for a specific stimulus*. However, its output, the stimulus-independent characteristics of a specific observer, can be input as prior knowledge to existing observer-independent but stimulus-dependent computational models mentioned above; this combination is not the focus of this paper. Although the proposed method requires a set of personality trait scores of target observers, these scores could also be used for other perception targets. It is superior to directly preparing perception data of individuals, especially when we target a variety of emotional states.

The remainder of this paper first introduces related works in II to position this study. Next, the proposed probabilistic model is described in III. Conversation data and empathy perception labels used in the present experiment are explained in IV. The proposed model is evaluated and discussed in V and VI. Finally, we summarize this study in VII.

II. RELATED WORKS

For over 50 years, personality and social psychology researchers have studied how humans judge another's internal states or characteristics e.g. emotion, personality and skills. One famous model of interpersonal perception is Kenny's social relations model (SRM) [9]. SRM introduces three effects to explain interpersonal perception: perceiver effect, target person effect, and their relationship effect. They demonstrated that the perceiver effect is dominant for unacquainted individuals.

For inferring the coefficients of the three factors of SRM, some researchers have demonstrated the effectiveness of the Bayesian framework [11], [12]. The present study differs from these studies in two main points: First, the present study focuses on the relationship of the interpersonal perception of perceivers with their gender and personality traits, while their targets are the mutual relationships of interpersonal perceptions between a pair. Second, we introduce a latent variable between these factors, while they directly link the factors.

A variety of probabilistic models have been introduced to estimate observer perceptions of emotional states [2]. Of particular note, some researchers tried to estimate the distribution of perception among observers, i.e. how emotions are differently perceived by observers for specific stimuli (human behavior) [7], [8]. However, none of them addressed the modeling of the perception tendency of a specific observer.

Jayagopi et al. [13] recently introduced latent Dirichlet allocation (LDA) to estimate the hidden topics of multi-party conversations; latent conversation topics are found only from group behavior. The present study, however, uses the latent characteristics of observers to link their perception tendency, gender and personality traits.

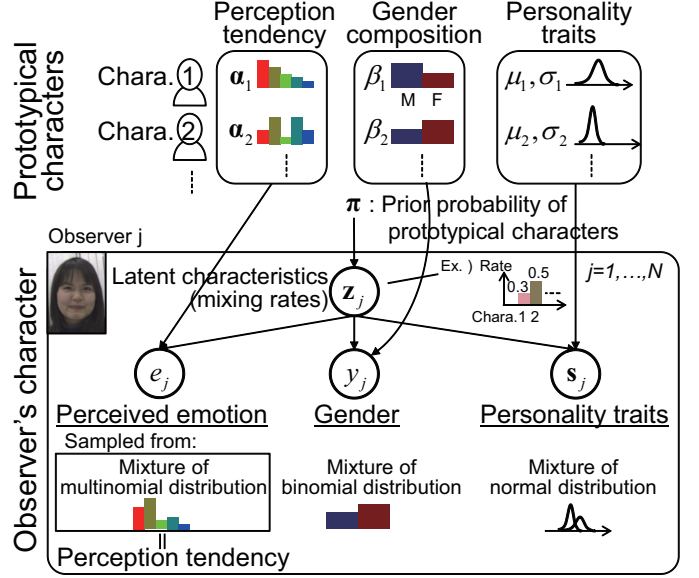


Fig. 1. Graphic representation of the proposed model

III. PROBABILISTIC TOPIC MODEL

This section proposes a probabilistic topic model for the perception tendency of empathy. Probabilistic modeling is used because the process of perception of other's internal states contains a mixture of various types of ambiguities in decision making. Bayesian theory well handles such ambiguities as probabilities.

A. Overview

The proposed model explains the observer's tendency as regards the perception of others' empathy. The tendency is linked to his/her gender and personality traits (found in existing psychological questionnaires) via his/her latent characteristics. These characteristics are described as a mixture of K prototypical human characters. The proposed model makes it possible to infer the empathy perception tendency of a target observer from his/her gender and psychological scale scores.

The proposed model is a so-called generative model. First, the latent characteristics of each observer are stochastically determined according to a prior probability distribution, or the component ratio of prototypical characters in a target population, $\pi = \{\pi_k\}_{k=1}^K$. Here, prior probability that observer j has the k -th prototypical character is expressed as $P(z_{j,k} = 1) = \pi_k$, where $z_j = \{z_{j,k}\}_{k=1}^K$, and $z_{j,k} \in \{0, 1\}$. Each prototypical character probabilistically determines the perception tendency, gender, and personality traits of the target observer. Fig. 1 is a graphic representation of the proposed model.

The perception tendency is expressed as a probability distribution, $\alpha = \{\alpha_e\}_{e=1}^{N_e}$; where $\sum_{e=1}^{N_e} \alpha_e = 1$. Probability α_e means how likely emotion of a category or level, $e \in \{1, \dots, N_e\}$, will be selected by a target observer. This paper models the process that the k -th prototypical character will perceive emotion e for a target with a multinomial

distribution, denoted by $\mathcal{M}(e|1, \alpha_k)$; where α_k is the perception tendency of the k -th prototypical character. The k -th prototypical character determines the observer's gender, $y \in \{'M', 'F'\}$, by following a binomial distribution, denoted by $\mathcal{B}(y|1, \beta_k)$; where β_k denotes the ratio of males in the k -th prototypical character. Personality trait scores, s , of the k -th prototypical character, are assumed to follow a Gaussian distribution, denoted by $\mathcal{N}(s|\mu_k, \Sigma_k)$, where μ_k and Σ_k are the mean vector and covariance matrix of the k -th character. In summary, model parameters are $\Theta = \{\pi_k, \alpha_k, \beta_k, \mu_k, \Sigma_k\}_{k=1}^K$.

This paper models the joint probability of the frequency distribution of an empathy perception of observer j , e_j , his/her gender, y_j , and personality traits, s_j , given by model parameters Θ ; $P(e_j, y_j, s_j|\Theta)$. The joint probability can be obtained by marginalizing latent variable z as:

$$\begin{aligned} P(e_j, y_j, s_j|\Theta) &= \sum_k P(z_{j,k} = 1|\Theta) P(e_j, y_j, s_j|z_{j,k} = 1, \Theta) \\ &= \sum_k \pi_k \mathcal{M}(e_j|1, \alpha_k) \mathcal{B}(y_j|1, \beta_k) \mathcal{N}(s_j|\mu_k, \Sigma_k). \end{aligned} \quad (1)$$

This paper assumes that the personality traits of observer j , s_j , are mutually independent given by his/her latent characteristics z_j , for mathematical simplicity. In this case, multivariate Gaussian distribution $\mathcal{N}(s_j|\mu_k, \Sigma_k)$ is decomposed into a product of univariate Gaussian distributions.

B. Model training

The proposed model is trained with gender, personality traits, and perception labels of observers given to target scenes in a training dataset. In the proposed model, if latent variable z is decided, the maximum likelihood estimators of model parameters Θ are analytically obtained. This problem can be effectively solved by using an expectation-maximization (EM) algorithm [14], which repeats E-step and M-step until convergence. In E-step, current mixing rate $\gamma(z_{j,k})$, or the posterior probability of $z_{j,k}$ given by all explanatory variables, is calculated as follows:

$$\begin{aligned} \gamma(z_{j,k})^{(t+1)} &= \frac{\pi_k^{(t)} \mathcal{M}(h_j|M, \alpha_k^{(t)}) \mathcal{B}(y_j|1, \beta_k^{(t)}) \mathcal{N}(s_j|\mu_k^{(t)}, \Sigma_k^{(t)})}{\sum_{k'} \pi_{k'}^{(t)} \mathcal{M}(h_j|M, \alpha_{k'}^{(t)}) \mathcal{B}(y_j|1, \beta_{k'}^{(t)}) \mathcal{N}(s_j|\mu_{k'}^{(t)}, \Sigma_{k'}^{(t)})} \end{aligned}$$

where h_j denotes the frequency distribution (histogram) of empathy perception of observer j in the training data including M scenes, and t denotes iteration step number. In M-step, parameters Θ are updated in a manner similar to the standard EM-algorithm for mixture of Gaussians or multinomial distributions, like [14].

C. Inference of observer perception tendency

The aim of the inference is to calculate the posterior probability of the perception tendency of observer j given by his/her gender y_j and personality traits s_j . By using the Bayes rule, it can be obtained as $P(e_j|y_j, s_j, \Theta) =$



Fig. 2. Snapshots of conversation (left) and labeling (right) scenes.

$P(e_j, y_j, s_j|\Theta)/P(y_j, s_j|\Theta) \propto P(e_j, y_j, s_j|\Theta)$, where the last relationship can be obtained by considering $P(y_j, s_j|\Theta)$ as a constant term because y_j, s_j , and Θ are given. The joint probability, i.e. (1), represents that an expected perception tendency of observer j , $\bar{\alpha}_j$, can be obtained as the expectation of the mixture of multinomial distributions. That is, $\bar{\alpha}_j = \sum_k \bar{\gamma}(z_{j,k}) \alpha_k$, where $\bar{\gamma}(z_{j,k})$ is the estimated mixing rates of the k -th character for observer j , i.e.

$$\bar{\gamma}(z_{j,k}) = \frac{\pi_k \mathcal{B}(y_j|1, \beta_k) \mathcal{N}(s_j|\mu_k, \Sigma_k)}{\sum_{k'} \pi_{k'} \mathcal{B}(y_j|1, \beta_{k'}) \mathcal{N}(s_j|\mu_{k'}, \Sigma_{k'})}. \quad (2)$$

IV. EXPERIMENTAL DATA

A. Stimulus video clips for observers

This paper assumes that a full set of targeted population stimuli yields well balanced empathy perceptions. As such stimuli, we prepare a series of short conversation video clips, where the total frequencies of the perception of several coders in preliminary experiments are well-balanced. The perception tendencies of target observers are obtained for the representative video set.

This paper uses multi-party face-to-face conversations [7], as shown in the left part of Fig. 2. The interlocutors were instructed to hold alternative-type discussions and to build consensus as a group, i.e. agree on a single answer, on each discussion topic within 6-8 minutes. Topics included “Who are more beneficial, men or women?”. The interlocutors were 20 Japanese women (five four-person groups) in their twenties or thirties. All conversations were captured at 30 fps by IEEE1394 color cameras.

Some of the conversations were annotated in advance with perceived empathy labels by 5 or 9 non-expert coders [7]. None of them acted as the observers in the present study. From this conversation data, the present study picked up 97 short scenes that yielded a variety of voting rates, e.g. empathy-dominant, empathy-inferior, and flat, to balance the frequency of empathy perception as much as possible. Each scene was six seconds long. The selected scenes showed only a pair of interlocutors sitting side by side, see right pane of Fig. 2; this made it easy to understand their mutual gaze behaviors.

We then generated a short video clip of each scene. Each clip showed a short scene twice without audio at normal speed. This forced the observers to focus on the emotions exchanged by visual nonverbal behaviors. The end of each clip was announced by a bell. Each clip was 36 sec long, and consisted

of: 1) scene ID, 5s; 2) viewing time ('first time' or 'second time'), 2s; 3) the first frame of the scene (the video has not started yet), 1.5s; 4) scene video, 6s (2)-4) are repeated twice); 5) remaining time to answer empathy perception, 8s; 6) bell ring, 0s; and 7) waiting time for next clip, 4s. These 97 clips were presented in random order. In addition, three clips were randomly selected from these 97 clips for a reproducibility test, and they were reshown at the end of the original 97 clips. Finally, these 100 clips in total were combined into two 50 clip videos. So, each video had fifty 36-sec clips, i.e. the length is 30 min. The present study ignores the last three scenes. So, the number of target scenes, M , in this study is 97.

B. Observers and their labeling of perceived empathy

Fifty observers participated ($N = 50$, 25 males and 25 females). They were Japanese university students in their early twenties. They had not met the interlocutors before the experiment. The observers were asked to watch the videos, separated by a 5-min interval, played at normal speed without stopping, and to assign one of the following bipolar labels, the one closest to their perception, to each scene: "Empathy" (+2), "Weak Empathy" (+1), "Neither Empathy nor Antipathy" (0), "Weak Antipathy" (-1), and "Antipathy" (-2). That is, the labeling was scene-by-scene, and the number of empathy perception states, N_e , was five. All observers completed the set task. All labeling was done in isolation. Each observer used a laptop computer with a 15.6-inch monitor. Headphones were used just to catch the bell ring indicating clip end. The right part of Fig. 2 shows a typical labeling scene.

C. Personality traits

To obtain observers' personality traits, we also asked the observers to answer three psychological questionnaires after the labeling task: Davis' Interpersonal Reactivity Index (IRI) [15], Emotional Skills and Competence Questionnaire (ESCQ) [16], and Tokyo University Egogram (TEG) [17]. IRI and ESCQ measure ability/tendency to understand others' emotions. TEG is a measure of basic personalities and tries to explain how people function and express their personality through their behavior.

IRI measures perspective taking (PT), fantasy (FA), empathic concern (EC), and personal distress (PD). ESCQ measures the following three subscales: a) ability to perceive and understand emotion (PU), b) ability to express and label emotion (EL), and c) ability to manage and regulate emotion (MR). TEG is summarized by the following five subscales and a lie scale: 1) CP (Critical Parent), indicating strong sense of responsibility and critical, 2) NP (Nurturing Parent), indicating sympathetic and consideration for others, 3) AD (Adult), indicating realistic and objective, 4) FC (Free Child), indicating freewheeling, bright and cheerful, and 5) AC (Adapted Child) indicating unassertive, tendency to anticipate other people's criticisms.

These three scales yield 13 (=4+3+6) subscales in total, i.e. the dimension of s is thirteen. The present study uses the Z-score of these subscales, i.e. scores in each scale were

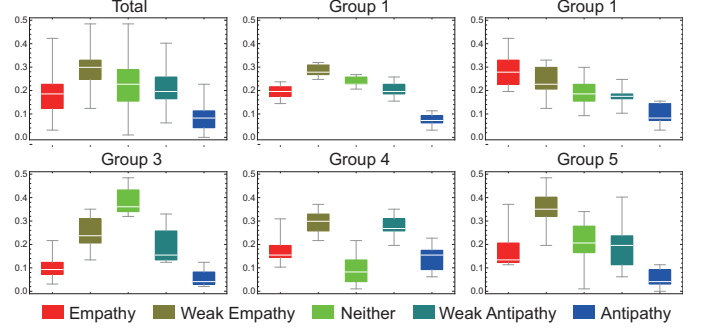


Fig. 3. Box plots of the histogram of empathy labels for the 97 scenes, i.e. $\{h_j\}_{j=1,\dots,N}$, of all fifty observers (upper left), and each observer group (others). Vertical axes denote probabilities (frequencies). The observers have a variety of perception tendencies.

standardized to zero-mean-unit-variance. Note that the main aim of this study is not to find/propose the best psychological questionnaires for directly measuring the perception tendency of an observer, but to propose a data-driven method to automatically find the relationship between the perception tendency and the psychological traits obtained from a variety of existing questionnaires.

V. EVALUATION

This paper quantitatively evaluates how accurately the proposed model can recreate an unseen perception tendency of observers from their gender and personality traits.

A. Settings

We first explore how many types of perception tendencies the observers have. A standard k-means clustering technique [18], applied to the perception tendencies of the fifty observers, more precisely their frequency distributions of empathy perception h_j , found five discriminant tendency types that achieve both small intraclass variations and large interclass variations; the number of observers in each group is 10. Fig. 3 represents the mean and variance of the perception tendencies of each group. Group 1 observers have average perception tendencies. Group 2 observers are biased to empathy. Group 3 observers have central tendencies, i.e. a preference for the center (Neither) label. Group 4 observers have extreme response tendencies, i.e. a tendency to avoid the center (Neither) label. Group 5 is a miscellaneous group, i.e. a group of other observers who have no remarkable tendencies.

Based on these results, this paper employs five-fold cross validation. That is, the model is trained with the data (personality traits, gender, and empathy perception labels for 97 scenes) of four groups (40 observers), and the perception tendencies of the remaining group (10 observers) are inferred from their gender and personality traits by using the trained model. Thus, the number of prototypical characters, K , is set to 4.

B. Training results: discovered prototypical characters

Fig. 4 shows the training results achieved with Groups 1 to 4. The upper part represents discovered empathy percep-

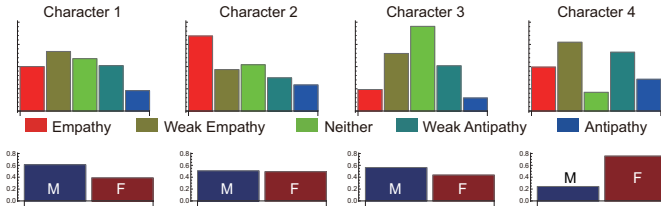


Fig. 4. (Upper) Perception tendencies of the prototypical characters that were discovered in observer Groups 1 to 4. Each figure describes expected perception tendency generated from prototypical character k , i.e. α_k . (Lower) Gender composition of each prototypical character, i.e. β_k .

TABLE I
PERSONALITY TRAIT SCORES FOR EACH PROTOTYPICAL CHARACTER

Scale		Chara. 1	Chara. 2	Chara. 3	Chara. 4
		μ_1	μ_2	μ_3	μ_4
IRI	PT	-0.15	0.43	0.07	0.19
	FA	-0.29	0.74	0.23	-0.16
	PD	-0.20	0.75	0.13	-0.34
ESCQ	MR	-0.15	-0.14	0.37	-0.09
TEG	CP	-0.06	-0.43	-0.01	0.13
	NP	0.00	-0.12	-0.08	0.31
	AD	0.39	-0.55	0.06	0.12
	FC	0.11	-0.43	0.00	-0.16
	AC	-0.09	0.39	0.08	-0.31

tion tendencies of K prototypical characters, $\{\alpha_k\}_{k=1}^K$. The perception tendencies of Groups 1-4 shown in Fig. 3 were successfully found. The lower part of Fig. 4 denotes the composition ratio of gender for each prototypical character, $\{\beta_k\}_{k=1}^K$. While the middle two characters do not depend on gender, the left-most and right-most characters exhibit male and female dominance, respectively. These results suggest that males more frequently show average perception tendencies than females, while females more often show extreme response tendencies than males.

Table I shows the trained distributions of personality trait scores for each prototypical character. Only the scales with $|\mu| \geq 0.3$ are listed. This is because a scale having at least one large absolute value of μ contributes to the identification of the prototypical character(s), while other scales fail to explain any feature of the characters. Compared to Fig. 4, each character indicates the following reasonable personality traits.

Character 1, associated with average tendency, yields high AD scores. That is, observers who are realistic and objective have average perception tendency. This further suggests that it is reasonable to determine an objective description of perceived emotions to be the average of collective perceptions, like [6]. Character 2, associated with empathy-biased tendency, yields remarkable personality traits. Such people basically show high IRI scores. This is reasonable, because IRI is a measure of empathy. This character also shows high AC score and low CP, AD, and FC scores. Such people are neither critical nor objective, but concerned about rules and other people's criticisms. So, such people might tend to give empathy labels by considering social desirability.

Character 3, associated with central tendency, yields high

TABLE II
ACCURACY OF THE INFERENCE OF PERCEPTION TENDENCY

Model	MAE($\downarrow$)	RMSE($\downarrow$)	BC($\uparrow$)	OA($\uparrow$)
Average among all groups				
The proposed model	.066 $\uparrow$	.079 $\uparrow$	.971 $\uparrow$	.835 $\uparrow$
Baseline	.068	.082	.969	.829
Group 1				
The proposed model	.023	.027	.997	.942
Baseline	.022	.026	.997	.944
Group 4				
The proposed model	.071***	.085***	.963***	.824***
Baseline	.084	.101	.952	.791

" $\uparrow$ " and " $\downarrow$ " mean higher and lower are higher performances. Symbols $\uparrow$ and *** mean $p < .1$ and $p < .001$ in two-tailed paired t-test, respectively.

MR score, i.e. high ability to manage and regulate emotion. This suggests that a cool person perceives less empathy/antipathy. This character showed no other distinctive personality traits. Character 4, associated with extreme response tendency, shows high NP score, and low AC and PD scores. The first two can well explain this character: Such people can often perceive empathy and antipathy because they have a strong sense of consideration for others. They are not concerned about their extreme responses, because they do not anticipate other people's criticisms.

C. Inference results

We here evaluate the proposed model in terms of its inference accuracy. This paper utilizes four similarity measures between two probability distributions p and q (N_e -dimensional vectors): mean absolute error (MAE), root mean square error (RMSE), Bhattacharyya coefficients (BC), and overlap area (OA). MAE and RMSE measure errors, while BC and OA measure similarity; MAE and RMSE become zero at best, i.e. perfect inference, while OA and BC become one (zero) at maximum (minimum, i.e. worst inference). These measures are selected from those used in [7], [8].

Table II summarizes the results of the average accuracies among five observer groups, as well as the accuracies of Groups 1 and 4 as typical cases. Table II also includes the inference results achieved by using a baseline model that always returns the mean label distribution in the training samples. Overall, the average accuracies of the proposed model are marginally significantly better than those of the baseline model (two-tailed paired t-test, $p < .1$). However, their accuracies for individual observer groups vary widely. Group 4 well demonstrates the characteristics of the proposed model, where the target perception tendency is quite different from the average perception tendency among observers. The proposed model statistically significantly outperforms the baseline model ($p < .001$). As to Group 1, though the baseline model achieves quite high inference performance, this simply means that Group 1 consists of observers who have a tendency close to the average perception tendency, as shown in Fig. 3.

Fig. 5 presents the estimated mixing rates of prototypical characters for each observer, i.e. $\bar{\gamma}(z_{j,k})$. For example, the perception tendencies of Group 1 observers are modeled by

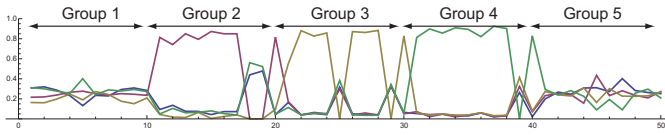


Fig. 5. Estimated character mixing rates, $\hat{\gamma}$. Different colors mean different prototypical characters. The observers in Groups 1 and 5 are modeled by equally mixing all prototypical characters, while the observers in Groups 2 to 4 are mostly explained by a single character.

equally mixing the trained perception tendencies of the prototypical characters. This is reasonable, because the observers in Group 1 have average perception tendencies, and their perception tendencies can be well represented as the average of other tendencies. On the other hand, the perception tendencies of the observers in Groups 2 to 4 are modeled by basically a single perception tendency. It is difficult to explain these tendencies by averaging/summing the other tendencies, unlike the average tendency.

VI. DISCUSSION

The results in V demonstrate the effectiveness of the proposed method, even when the target perception tendency is quite different from the average perception tendency among observers. This study used three psychological questionnaires, IRI, ESCQ and TEG in an initial study. Of course, a variety of psychological scales, including the Big Five personality traits, have been developed so far, and various combinations of such scales are possible; there may be sets of scales that better explain empathy perception, but this is not the focus of this paper. Instead, the proposed method can use scales that are expected to somewhat correlate to the target, to explain the perception tendency. Scales with little influence can be ignored by giving similar parameters to all prototypical characters. The proposed framework is general, and can be used for a variety of perception targets.

The proposed method has some difficulty in inferring an unseen perception tendency. A possible explanation is as follows: The proposed method tries to explain the perception tendency of a target observer by *summing* prototypical tendencies modeled with multinomial distributions (*non-negative* vectors). Thus, it requires a training data set that includes a set of perception tendencies, the summation of which is close to the target tendency, or a perception tendency that is close to the target tendency.

Furthermore, this paper prepared a representative set of conversation stimuli by assuming that the total frequencies of the perception are well balanced in target population stimuli. This could be replaced by normalizing the perception frequency of each observer against the total frequency of all observers, though it might over enhance the difference in perception tendencies. However, more importantly, the stimuli in daily life are not necessary balanced; in fact, in our previous study, the perception frequencies before scene selection were strongly unbalanced to positive perception (generally biased

to empathy), as explained in [7]. How to compensate the perception frequency is another issue.

VII. CONCLUSION

This paper proposed a probabilistic model that automatically discovers the correlation between perception tendency, gender, and personality traits of a target observer. In the proposed model, these three factors are linked via latent variables that explain the observer's characteristics as a mixture of prototypical characters. An experiment was conducted with fifty observers. They watched 97 short conversation videos and gave their impressions about empathy between each interacting pair. The results demonstrated that the proposed method can find a reasonable structure between the factors. The proposed method showed promising performance in estimating an observer's perception tendency from his/her gender and personality traits, even when the target perception tendency is quite different from the average perception tendency among observers.

REFERENCES

- [1] D. Gatica-Perez, "Analyzing group interactions in conversations: A review," in *Proc. IEEE Int'l Conf. MFI*, 2006, pp. 41–46.
- [2] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. PAMI*, vol. 31, no. 1, pp. 39–58, 2009.
- [3] K. Otsuka, "Conversation scene analysis," *IEEE Signal Proc. Mag.*, vol. 28, pp. 127–131, 2011.
- [4] R. Cowie, "Describing the emotional states expressed in speech," in *ITRW on Speech and Emotion*, 2000, pp. 11–18.
- [5] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan, "IEMOCAP: interactive emotional dyadic motion capture database," *Language Resources And Evaluation*, vol. 42, no. 4, pp. 335–359, 2008.
- [6] M. Nicolaou, H. Gunes, and M. Pantic, "Output-associative RVM regression for dimensional and continuous emotion prediction," in *Proc. IEEE Int'l Conf. FG'11*, 2011.
- [7] S. Kumano, K. Otsuka, D. Mikami, M. Matsuda, and J. Yamato, "Understanding communicative emotions from collective external observations," in *Proc. CHI EA '12*, 2012, pp. 2201–2206.
- [8] H. Meng, A. Kleinsmith, and N. Bianchi-Berthouze, "Multi-score learning for affect recognition: the case of body postures," in *Proc. ACII*, vol. 1, 2011, pp. 225–234.
- [9] D. A. Kenny and L. Albright, "Accuracy in interpersonal perception: A social relations analysis," *Psychol. Bull.*, vol. 102, no. 3, pp. 390–402, 1987.
- [10] T. Hofmann, "Probabilistic latent semantic indexing," in *Proc. Int'l ACM SIGIR Conf. Res. Dev. Inf. Retrieval*, 1999, pp. 50–57.
- [11] P. D. Hoff, "Bilinear mixed-effects models for dyadic data," *J. Amer. Statist. Assoc.*, vol. 100, no. 469, pp. 286–295, 2005.
- [12] P. S. Gill and T. B. Swartz, "Bayesian analysis of dyadic data," *Am. J. Math. Manage. Sci.*, vol. 27, no. 1–2, pp. 73–92, 2007.
- [13] D. Jayagopi, D. Sanchez-Cortes, K. Otsuka, J. Yamato, and D. Gatica-Perez, "Linking speaking and looking behavior patterns with group composition, perception, and performance," in *Proc. ICMI*, 2012, pp. 433–440.
- [14] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. R. Stat. Soc.*, vol. B 39, no. 1, pp. 1–38, 1977.
- [15] M. H. Davis, "Measuring individual differences in empathy: Evidence for a multidimensional approach," *J. Pers. Soc. Psychol.*, vol. 44, no. 1, pp. 113–126, 1983.
- [16] V. Taksic, "The importance of emotional intelligence(competence) in positive psychology," in *Proc. Int'l positive psychology summit*, 2002.
- [17] H. Suematsu, S. Nomura, and M. Wada, *Handbook of TEG. 2nd edition [in Japanese]*. Tokyo: Kaneko-shobo, 1993.
- [18] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 1967, pp. 281–297.