Analysis of Tourists' Nationality Effects on Behavior-based Emotion and Satisfaction Estimation

Yuki Matsuda*†, Dmitrii Fedotov‡§, Yutaka Arakawa¶‖, Hirohiko Suwa*†, Wolfgang Minker‡, Keiichi Yasumoto*†
* *Nara Institute of Science and Technology*, Nara, Japan, Email: {yukimat, h-suwa, yasumoto}@is.naist.jp.
† *RIKEN Center for Advanced Intelligence Project AIP*, Tokyo, Japan.
‡ *Ulm University*, Ulm, Germany, Email: {dmitrii.fedotov, wolfgang.minker}@uni-ulm.de.
§ *ITMO University*, Saint Petersburg, Russia.
¶ *Kyushu University*, Fukuoka, Japan, Email: arakawa@ait.kyushu-u.ac.jp.
‖ *JST Presto*, Tokyo, Japan.

***Abstract*—Smart tourism is attracting attention of researchers in recent years. Its technologies can be used by tourists in order to obtain useful information during sightseeing with smart devices etc. To provide suitable and personalized tourism information according to the situation of tourists, understanding psychological status during sightseeing, especially, emotional status and satisfaction level, is important. We assume that the psychological status of tourists is appearing and represented through unconscious behaviors during sightseeing such as head/body movements and facial/vocal expressions, and have proposed methods to estimate emotion and satisfaction statuses by sensing and analyzing tourists' behaviors. Through in-the-wild experiments with 22 participants, we found that the difference in tourists' attributes might give effects for the estimation. In this paper, we have statistically analyzed those effects, focusing on tourists' nationality. As a result of the two-way ANOVA, we found the interaction effect (disordinal interaction) between tourists' nationality and estimation performance, the main effect in differences of features, and the main effect in differences of tourists' nationality. The results imply that we need to take tourists' nationality into account for building estimation models. Contribution: We have statistically analyzed the nationality effects on tourist emotion and satisfaction estimation, and confirmed significant differences in feature contributions for estimation models.**

***Index Terms*—emotion recognition, satisfaction estimation, statistical analysis, ubiquitous computing, wearable computing, smart tourism**

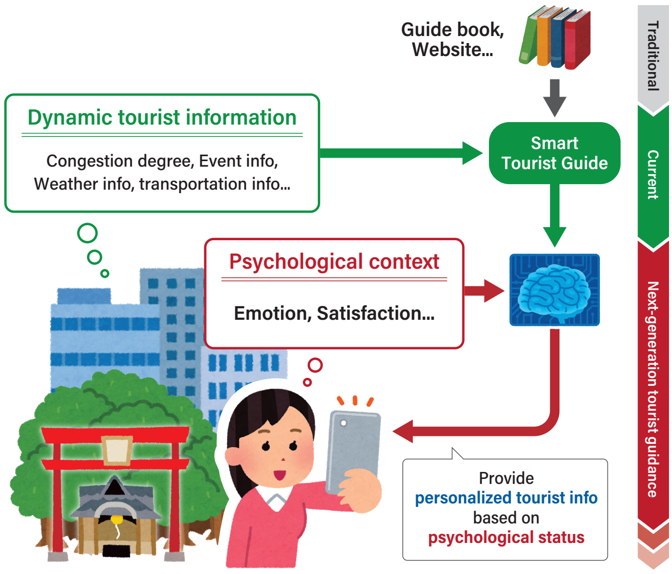

Fig. 1. The paradigm in tourist guidance systems, and the objective of our research project.

I. Introduction

Thanks to the progression of information technologies, we can get various information about our livelihood through smartphones and wearable devices in real-time, which is useful in our daily life. By providing dynamic tourist guidance in consideration of such environmental context, tourists can acquire useful information during sightseeing [1], [2]. However, current services such as navigation systems, recommender systems do not necessarily reflect the tourist's sensation (e.g., emotion, satisfaction level). To provide richer content, not only environmental information but also the real-time psychological perspective of tourists should be considered as shown in Fig. 1.

Therefore, we are studying estimation methods of the tourists' psychological status during sightseeing based on objective data collection. In this study, it is presumed that the psychological status of tourists appears in the form of unconscious behaviors during sightseeing, such as head/body movements, facial expressions, and vocal expressions. We hypothesize that the emotion and satisfaction estimation model might be built by using these clues. In our research so far [3], [4], we have proposed the estimation method of the psychological status for each tourist spot (session) as shown in Fig. 2, based on sensing of tourist's unconscious behavior using multiple wearable devices (eye tracker, motion sensor) and smartphone (camera). Such devices are not common at this moment yet, we assume more various wearable devices will be available as consumer devices in the near future

This research was funded by Japan Society for the Promotion of Science KAKENHI (grant number 19K24345) and the technology transfer project "Do it yourself, but not alone: Companion Technology for DIY support" of the Transregional Collaborative Research Centre SFB/TRR 62 "Companion Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG). 978-1-7281-9331-1/20/$31.00 ©2020 IEEE

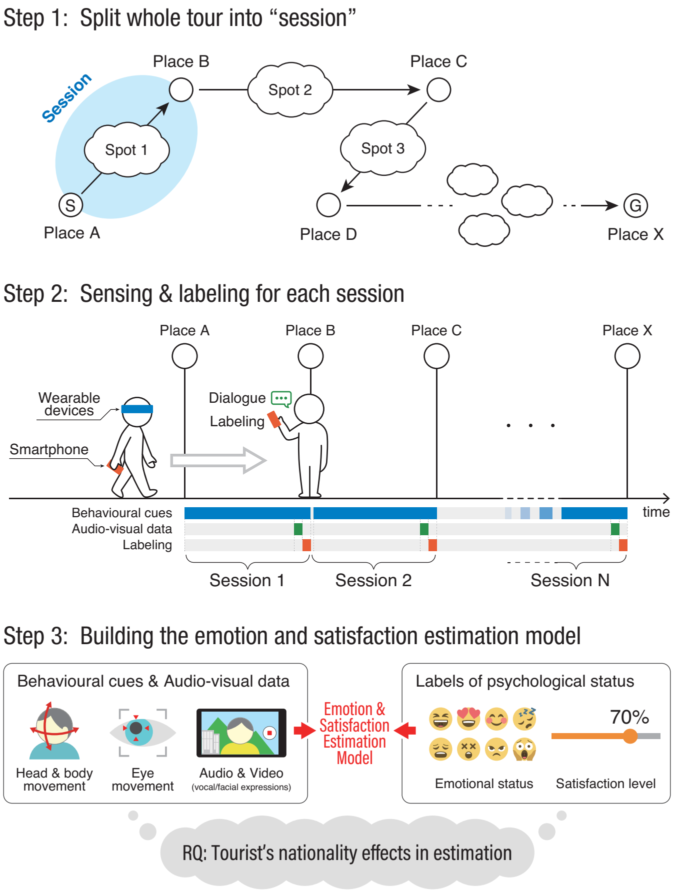

Fig. 2. Workflow of tourists emotion and satisfaction estimation. This figure includes the figure taken from our previous paper [3].

such as smart glasses, smartwatches, smart shoes, which have capabilities for sensing physical/physiological data used in this study.

Through the evaluation of our estimation model on in-the-wild experiments with 22 participants including Japanese and Russian on two sightseeing areas (Ulm, Germany and Nara, Japan), we found that the difference in tourists' attributes might give effects for the estimation. In this paper, we have statistically analyzed that effects, especially focusing on tourists' nationality. As a result of the two-way ANOVA, we found (1) the interaction effect (disordinal interaction) between tourists' nationality and estimation performance, and the main effect in differences of features, and (2) the main effect in differences of tourists' nationality.

The organization of this paper is as follows. First, we explain related work in Section II. Second, the collected data to be used for analysis is described in Section III. In the Section IV, we summarize the evaluation results as the reference for the performance of our estimation model. Then, we analyze the statistical significance between tourists' nationality and these results, and provide discussions. Finally, we conclude this paper with Section V.

II. Related Work

There are many research works focusing on emotion recognition and satisfaction estimation. To recognize such a psychological status of the user, audio and visual information are often used as popular modalities [5]–[8]. Also, physiological features [9], [10] and physical features [11]–[14] are also used for emotion recognition. In addition, the estimation performance might be improved by using multimodal features [15]–[18]. Although such systems work comparatively well, they are evaluated with the acted dataset collected in laboratory environments in many cases. Hence, they might have troubles when using real-condition data [19].

In addition, in the study of sensing-based emotion recognition, it has been revealed that the form of emotional expression may differ depending on nationality [20]–[22]. This study aims to estimate the psychological status, e.g., emotions, by focusing on the *unconscious behaviors* that tourists are taking during sightseeing [3]. Hence, it is assumed that tourists' attribute including nationality has an influence on estimation.

In the following section, we will statistically analyze the influence of nationality on the emotion and satisfaction estimation model during a touristic activity.

III. Tourist Behavior Data and Psychological Status Data During Sightseeing

This section provides an overview of a dataset including tourist behavior data and psychological status data which are collected during sightseeing.

To collect a dataset, we have conducted in-the-wild experiments in two sightseeing areas. Each area has a totally different condition and characteristic. The first area is the center of Ulm, Germany. The route through the center of the city surrounded by buildings, and is sometimes crowded depending on the time. There are several places with high tourist value, such as Schiefes Haus. The route has eight small sessions which include at least one tourist place. Its total length is approximately 1.5 km. The second area is around Nara Park which is the famous historic area of Nara, Japan. The sights in this area located in nature, and there are many religious and scenic buildings (e.g., shrines and temples). The route has seven small sessions, and its total length is approximately 2.0 km.

We have conducted this experiment with 22 participants. The age range is 22–31 years old (average: 24.3), gender is 17 males and 5 females, and nationalities are 12 Japanese and 10 Russian. We selected these nationalities as they comprise the two largest clusters of participants in our dataset. The full version of the dataset consists of recordings from people with other nationalities as well, but their number is insufficient for analysis. In total, we have 183 sessions (approx. 25 hours): 143 sessions with 17 participants (10 Russians, 7 Japanese) in Germany, 40 sessions with 5 participants (all Japanese) in Japan.

TABLE I
FEATURES DERIVED FROM BEHAVIOR DATA DURING THE SIGHTSEEING.

Feature	Description
Eye movement	Intensity of eye-ball movement (average)
	Statistical values of eye-ball movement (average, standard deviation) * Values calculated with time window of 1, 5, 10, 20, 60, 120, 180, 240 sec.
Head movement	Count of turning face toward upper direction (/sec)
	Time interval of turning face toward upper direction (average, standard deviation) * These values also calculated for right, left, lower direction.
	Count of turning face toward upper/lower direction (/sec)
	Time interval of turning face toward upper/lower direction (average, standard deviation)
	Intensity of turning face toward upper/lower direction (average, standard deviation) * These values also calculated for right/left direction.
	Count of turning face toward upper/lower/right/left direction (/sec)
Body movement	Footstep count (/sec)
	Time interval per one footstep (average, standard deviation)
	Intensity of footsteps (average, standard deviation)
Audio (vocal expression)	Low-level descriptors (LLDs) * 65 LLDs which can be extracted by using openSMILE [23]
Video (facial expression)	Action Units (AUs) [24], [25] * AUs (01, 02, 04, 05, 06, 07, 09, 10, 12, 14, 15, 17, 20, 23, 25, 26, 28, 45) which can be extracted by using OpenFace [26], [27].

TABLE II
THE EVALUATION RESULT OF EMOTION AND SATISFACTION ESTIMATION USING ALL PARTICIPANTS' DATA.

Feature used for building estimation model	Emotion (UAR)		Satisfaction (MAE)	
	Avg.	SD	Avg.	SD
Eye movement	0.432	0.073	1.124	0.178
Head/body movement	0.428	0.070	1.187	0.170
Behavioral cues (eye & head/body movement)	**0.496**	0.130	1.171	0.188
Audio (vocal expression)	0.410	0.069	1.124	0.154
Video (facial expression)	0.404	0.092	1.101	0.155
Audiovisual data (audio & video)	0.431	0.098	1.108	0.165
Feature-level fusion	0.465	0.097	1.204	0.195
Decision-level fusion	0.485	0.098	**1.076**	0.134

A. *Dataset of tourist behavior*

In this study, we employ three sensor data to collect tourists' behavior. A summary of collected data from tourists during sightseeing are described in following sections. Then, we derive features from this data as shown in Table I. These features are later used as an input to our models for emotion and satisfaction recognition. For more detail explanation of data, features, and derivation processes, see our previous paper [3].

1) Eye movement: Due to tourists mainly acquire information during sightseeing through the sense of sight, it is assumed that eye movements reflect his/her interests naturally. To collect eye movement data of tourists during sightseeing, we employ Pupil Labs Eye Tracker [28] with two infrared global shutter eye cameras. As features, we use statistical values calculated from *theta* and *phi* values of the eye-ball movement represented in spherical coordinates.

2) Head/body movement: After gaining interest, it is supposed that tourists will take some actions, e.g., looking up, and walking slower. To obtain such actions, we employ SenStick multi-sensor board [29] with inertial sensors. As features, we use head tilt derived by gyroscope data; and as a feature of body movement, we use a footstep count calculated by accelerometer data.

3) Selfie movie (audio, video): In general, many tourists might take photos and movies of tourist spots during sightseeing. Additionally, the *selfie* photo/movie is coming to be popular due to the widespread social networking services (SNS; e.g., Instagram, Twitter, and Facebook). Hence, audiovisual data can be used for tourist emotion and satisfaction estimation. To extract features for building the model, we use OpenSMILE [23] and OpenFace [27] which are open-source toolkits.

B. *Psychological status data*

As metrics of the psychological status which tourists felt during the sightseeing, we employed the emotional status and the satisfaction level. We asked tourists to manually enter the ratings of the session by using the smartphone application which we developed when each session ended. Each metric is described in the following sections:

1) Emotional status: As a metric of the tourists' emotional status, we have defined nine emotion categories and three emotion groups based on the Russell's two-dimensional (valence/arousal) circumplex model [30]. The emotion categories and groups are as follows:

Positive: 0 (Excited), 1 (Happy/Pleased), 2 (Calm/Relaxed)
Neutral: 3 (Neutral)
Negative: 4 (Sleepy/Tired), 5 (Bored/Depressed), 6 (Disappointed), 7 (Distressed/Frustrated), 8 (Afraid/Alarmed)

2) Satisfaction level: As a metric of the tourists' satisfaction level, we have employed the Seven-Point Likert scale which is officially used by the Japanese government (Ministry of Land, Infrastructure, Transport, and Tourism). Tourists can select their satisfaction level of each session between "6" (fully satisfied) and "0" (fully unsatisfied). We regard "3" represents the middle satisfaction level, and the psychological status of the tourist at the start of sightseeing.

IV. TOURISTS' NATIONALITY EFFECTS ANALYSIS

A. *Overview of estimation performance evaluation*

We have built the emotion and satisfaction estimation model using the tourist behavior data and the psychological status labels mentioned above. The model training scheme has been

TABLE III
THE EVALUATION RESULT OF EMOTION AND SATISFACTION ESTIMATION BY PARTICIPANTS' NATIONALITY.

Feature used for building estimation model	Emotion (UAR)				Satisfaction (MAE)			
	Japanese		Russian		Japanese		Russian	
	Avg.	SD	Avg.	SD	Avg.	SD	Avg.	SD
Eye movement	0.438	0.086	0.426	0.061	1.001	0.142	1.248	0.114
Head/body movement	0.417	0.082	0.438	0.056	1.198	0.228	1.176	0.093
Behavioral cues (eye & head/body movement)	0.415	0.067	**0.576**	0.129	1.093	0.237	1.249	0.071
Audio (vocal expression)	0.447	0.069	0.372	0.048	1.032	0.098	1.217	0.146
Video (facial expression)	0.463	0.098	0.346	0.027	1.019	0.110	1.184	0.152
Audiovisual data (audio & video)	0.445	0.092	0.417	0.106	1.014	0.106	1.201	0.164
Feature-level fusion	0.423	0.048	0.507	0.117	1.124	0.239	1.285	0.095
Decision-level fusion	**0.473**	0.064	0.496	0.125	**0.995**	0.103	**1.157**	0.112

proposed in our previous paper [3]. For the emotion estimation model, we have built a 3-class classification model of Positive, Neutral, and Negative emotions. For the satisfaction estimation, we have built a regression model for predicting values in the range of 0–6.

The evaluation results of each built model are shown in Table II. As an evaluation metric, unweighted average recall (UAR) was used for emotion estimation in consideration of the fact that the number of psychological status labels was not uniform, and mean absolute error (MAE) was used as an evaluation index for satisfaction estimation. A detailed explanation is provided in our previous paper [3]. Each row of Table II represents the evaluation result for each feature used when building the model. In addition, feature-level fusion is a fusion method that builds a single model using all features, and decision-level fusion is a fusion method to get the final result by a combination of estimates from models built using each feature. As a result of Table II shows that it is possible to estimate emotion with 49.6% of UAR and satisfaction with 1.1 of MAE.

This experiment has employed participants with two nationalities, Japanese and Russian. Table III shows the results of the evaluation of emotion and satisfaction estimation models according to the nationality of tourists. We found differences in the best performances of emotion estimation between different nationalities (for Japanese, the highest UAR of 47.3% has been obtained using decision-level fusion, and for Russian, the highest UAR of 57.6% has been obtained using behavioral cues). Regarding satisfaction estimation, the best performances of 0.995 (Japanese) and 1.157 (Russian) have been obtained using decision-level fusion. However, in most cases, the MAE for the Russian group tends to be larger than Japanese group. In the next section, we confirm these observations with statistical analysis.

B. Statistical analysis

Through the evaluation, we found differences in estimation performance between the Japanese and Russian groups. Here, we conduct statistical analysis of the results mentioned above (Table III) to confirm the effects on the accuracy of estimation by nationality of tourists. As an analysis method, we have employed the two-way ANOVA which is a statistical test method used to determine the effect of two factors, e.g., effects by an independent factor, synergy and disordinal interaction of two factors.

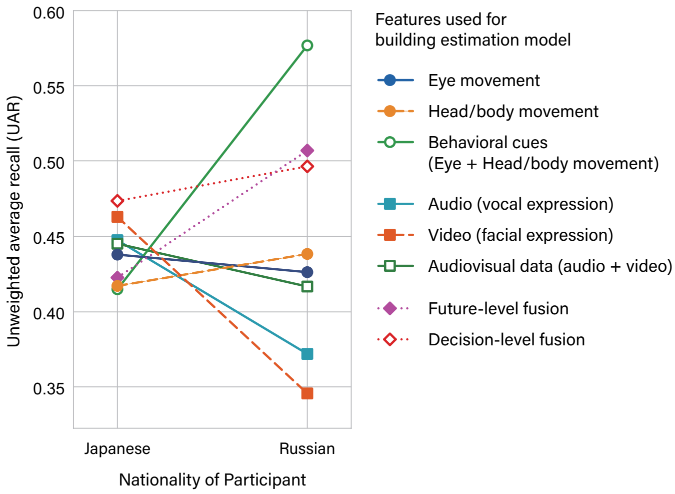

Fig. 3. Interaction plots for the nationality effect of emotion estimation results.

TABLE IV
RESULT OF TWO-WAY ANOVA IN EMOTION ESTIMATION.

	TSS	DF	F-value	p-value
Main effect (nationality)	0.002	1.0	0.286	0.594
Main effect (feature)	0.162	7.0	3.208	0.003 *
Interaction effect	0.270	7.0	5.355	0.000 *

* TSS: Total Sum of Squares, DF: Degree of Freedom.

1) Emotion estimation: The interaction plots for the nationality effect of emotion estimation results are shown at Fig. 3. Then, Table IV shows the result of the two-way ANOVA for emotion estimation results.

As a result of the analysis, the main effect of the tourist's nationality is not significant, but the main effect of the feature used for building the estimation model is significant. Furthermore, the interaction effect is significant. It suggests the main

TABLE V
RESULT OF TUKEY–KRAMER TEST IN EMOTION ESTIMATION.

Feature	MD	L	U	Sig.
Eye movement	-0.012	-0.082	0.059	False
Body movement	0.021	-0.045	0.088	False
Behavioral cues (eye & head/body movement)	0.161	0.065	0.258	True
Audio (vocal expression)	-0.075	-0.131	-0.020	True
Video (facial expression)	-0.117	-0.185	-0.049	True
Audiovisual data (audio & video)	-0.028	-0.122	0.065	False
Feature-level fusion	0.084	-0.000	0.168	False
Decision-level fusion	0.023	-0.071	0.117	False

* MD: Mean Difference, L/U: Lower/Upper bound, Sig.: Significant difference.

effect of the tourist's nationality is canceled by this interaction (disordinal interaction).

Then, due to the interaction significance, we analyze in detail for whole groups of used features. As an analysis method, we have employed Tukey-Kramer multiple comparison test, which is a statistical testing method focusing differences of the average value between every two groups of the multiple groups. Family Wise Error Rate (FWER) has been set as 5%. The multiple comparison result is shown in Table V. As a result, significant differences between nationality groups (Japanese and Russian) have been observed when the estimation model is built using the following features: behavioral cues (eye & head/body movement), audio (vocal expression), and video (facial expression).

These results suggest that the necessity of constructing the model by selecting the features to be used based on the tourist's nationality to improve the performance of the emotion estimation model. On the other hand, it has a possibility to reduce the data collection cost by changing the viewpoint. For example, for Russians, it is difficult to estimate emotion by using facial and vocal expressions (from selfie videos), but in contrast, eye movements and head/body movements are helpful. It suggests the emotion estimation model can be built without collecting videos.

2) Satisfaction estimation: The interaction plots for the nationality effect of satisfaction estimation results is shown in Fig. 4. Then, Table VI shows the result of the two-way ANOVA for satisfaction estimation results.

As a result of the analysis, different from the case of emotion, the main effect of the features used for building the estimation model is not significant, but the main effect of the tourist's nationality is significant. Also, the interaction effect is not significant. This tendency can be found in Fig. 4 except the case of head/body movement.

From this statistical analysis, we found the tendency that estimating the satisfaction level of Russian tourists is difficult in comparison to Japanese tourists. To improve estimation performance, we have to consider a better feature extraction method and/or additional modality.

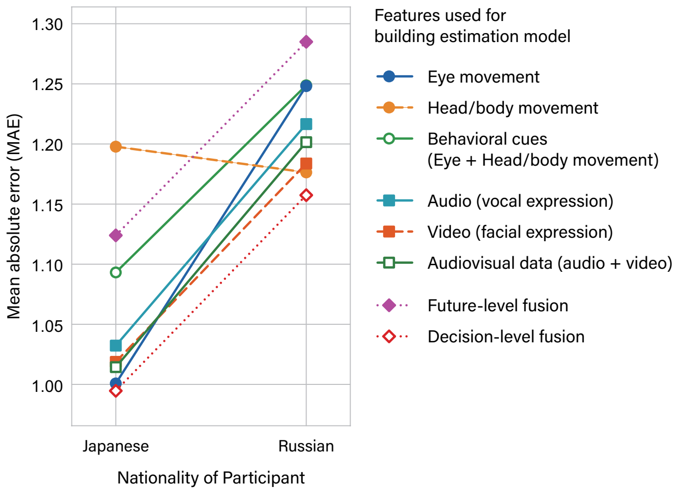

Fig. 4. Interaction plots for the nationality effect of satisfaction estimation results.

TABLE VI
RESULT OF TWO-WAY ANOVA IN SATISFACTION ESTIMATION.

	TSS	DF	F-value	p-value
Main effect (nationality)	0.964	1.0	44.204	0.000 *
Main effect (feature)	0.287	7.0	1.883	0.076
Interaction effect	0.209	7.0	1.367	0.223

* TSS: Total Sum of Squares, DF: Degree of Freedom.

We also confirmed that there are no significant differences in the importance of each feature to the estimation. It suggests a possibility that can omit the devices which require a high burden, e.g., Pupil Labs Eye Tracker [28] (participants is required to use wire-connected PC during the whole sightseeing). Such feature selection might help to realize the simple measurement.

C. Discussion

In this paper, we have confirmed that nationality affects the contribution of features in emotion/satisfaction estimation models using the dataset including only two nationalities, Japanese and Russian. To get more general insights regarding the effects of nationality, we need to expand varieties of nationalities as future work.

This paper provides statistical analysis with the nationality as a tourist attribute, but we consider that there are other tourist attributes that affect the estimation model. For example, general personal attributes (e.g., gender, age), personalities used in tourist spot recommendation systems (e.g., Travel Personality [31], Big Five Factor [32]). Tourist attributes such as preferences [33] will need to be investigated in further analysis.

In addition, tourists' behavior during sightseeing might be affected by the tourist sight itself. The experiments in this paper have been conducted in two different sightseeing areas, Germany and Japan. Hence, we will analyze the effects of sightseeing areas to estimation models as future work. Also, the combination of the touristic area and the tourist's nationality might give effects on the performance of the estimation model. The effects of location-nationality combination should be analyzed and discussed as future work.

The current performance of our proposed method is not high. It suggests the difficulty of making a "general" estimation model that can be applied to everyone around the world. However, if we find tourist attributes that affect estimation performance, there is a possibility that the estimation model can be improved, e.g., employing multiple models and model selection/combination algorithm based on tourists' attributes.

V. Conclusion

In this study, we aim to implement the tourists' emotion and satisfaction estimation methods based on measurement and analysis of the behaviors during sightseeing, assuming that the psychological status of tourists appears in the form of unconscious behavior. Through the evaluation of the built emotion and satisfaction estimation model, it was suggested that differences in tourist attributes might affect the accuracy of the estimation model.

According to the results, we focused on nationality among these tourist attributes, and statistically analyzed how it affects the emotion and satisfaction estimation model in this paper. As a result of the two-way ANOVA, we found the interaction effect (disordinal interaction) between tourists' nationality and estimation performance, the main effect in differences of features, and the main effect in differences of tourists' nationality. The results imply that we need to take tourists' nationality into account for building estimation models.

As future work, we will conduct further investigation of the effects of tourists' attributes, e.g., personality, preferences, gender, age, on the estimation model. Then, we will consider a simpler estimation method, and improve the emotion and satisfaction estimation performances.

References

[1] E. Balandina, S. Balandin, Y. Koucheryavy, and D. Mouromtsev, "Iot use cases in healthcare and tourism," in *2015 IEEE 17th Conference on Business Informatics*, vol. 2, July 2015, pp. 37–44.

[2] S. Morishita, S. Maenaka, N. Daichi, M. Tamai, K. Yasumoto, T. Fukukura, and K. Sato, "Sakurasensor: Quasi-realtime cherry-lined roads detection through participatory video sensing by cars," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UBICOMP '15)*, 2015, pp. 695–705.

[3] Y. Matsuda, D. Fedotov, Y. Takahashi, Y. Arakawa, K. Yasumoto, and W. Minker, "EmoTour: Estimating emotion and satisfaction of users based on behavioral cues and audiovisual data," *Sensors*, vol. 18, no. 11, 2018. [Online]. Available: http://www.mdpi.com/1424-8220/18/11/3978

[4] D. Fedotov, Y. Matsuda, Y. Takahashi, Y. Arakawa, K. Yasumoto, and W. Minker, "Towards real-time contextual touristic emotion and satisfaction estimation with wearable devices," *2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops '19)*, pp. 358–360, 2019.

[5] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, "End-to-end multimodal emotion recognition using deep neural networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 8, pp. 1301–1309, 2017.

[6] A. Subramaniam, V. Patel, A. Mishra, P. Balasubramanian, and A. Mittal, "Bi-modal first impressions recognition using temporally ordered deep audio and stochastic visual features," in *Computer Vision - ECCV 2016 Workshops*. Springer International Publishing, 2016, pp. 337–348.

[7] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, "Emotion recognition using facial expressions," *Procedia Computer Science*, vol. 108, pp. 1175–1184, 2017.

[8] W. Y. Quck, D.-Y. Huang, W. Lin, H. Li, and M. Dong, "Mobile acoustic emotion recognition," in *Region 10 Conference (TENCON), 2016 IEEE*. IEEE, 2016, pp. 170–174.

[9] P. J. Bota, C. Wang, A. L. N. Fred, and H. Plácido Da Silva, "A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals," *IEEE Access*, vol. 7, pp. 140 990–141 020, 2019.

[10] Y. Hsu, J. Wang, W. Chiang, and C. Hung, "Automatic ecg-based emotion recognition in music listening," *IEEE Transactions on Affective Computing*, vol. 11, no. 1, pp. 85–99, 2017.

[11] A. Kleinsmith and N. Bianchi-Berthouze, "Affective body expression perception and recognition: A survey," *IEEE Transactions on Affective Computing*, vol. 4, no. 1, pp. 15–33, 2013.

[12] B. Resch, A. Summa, G. Sagl, P. Zeile, and J.-P. Exner, "Urban emotions – geo-semantic emotion extraction from technical sensors, human sensors and crowdsourced data," in *Progress in Location-Based Services 2014*, 11 2014, pp. 199–212.

[13] J. Deng, C. Leung, and Y. Li, "Beyond big data of human behaviors: Modeling human behaviors and deep emotions," in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, April 2018, pp. 282–286.

[14] Z. Zhang, Y. Song, L. Cui, X. Liu, and T. Zhu, "Emotion recognition based on customized smart bracelet with built-in accelerometer," *PeerJ*, vol. 4, p. e2258, 2016.

[15] B. Zhao, Z. Wang, Z. Yu, and B. Guo, "Emotionsense: Emotion recognition based on wearable wristband," in *2018 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 2018, pp. 346–355.

[16] M. Soleymani, S. Asghari-Esfeden, Y. Fu, and M. Pantic, "Analysis of eeg signals and facial expressions for continuous emotion detection," *IEEE Transactions on Affective Computing*, vol. 7, no. 1, pp. 17–28, Jan 2016.

[17] E. Kanjo, E. M. Younis, and N. Sherkat, "Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach," *Information Fusion*, vol. 40, pp. 18–31, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253517303433

[18] T. W. AlHanai and M. M. Ghassemi, "Predicting latent narrative mood using audio and physiologic data," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI '17)*, 2017, pp. 948–954.

[19] A. Dhall, R. Goecke, S. Ghosh, J. Joshi, J. Hoey, and T. Gedeon, "From individual to group-level emotion recognition: Emotiw 5.0," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ACM, 2017, pp. 524–528.

[20] J. T. Stanley, X. Zhang, H. H. Fung, and D. M. Isaacowitz, "Cultural differences in gaze and emotion recognition: Americans contrast more than chinese," *Emotion*, vol. 13, no. 1, pp. 36–46, February 2013.

[21] L. Pragst, S. Ultes, M. Kraus, and W. Minker, "Adaptive dialogue management in the kristina project for multicultural health care applications," in *Proceedings of the 19th Workshop on the Semantics and Pragmatics of Dialogue (SEMDIAL)*, 2015, pp. 202–203.

[22] J. Miehle, W. Minker, and S. Ultes, "What causes the differences in communication styles? a multicultural study on directness and elaborateness," in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA), May 2018. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2018/pdf/167.pdf

[23] F. Eyben, M. Wöllmer, and B. W. Schuller, "Opensmile: the munich versatile and fast open-source audio feature extractor," in *Proceedings of the 18th ACM International Conference on Multimedia*, ser. MM '10. ACM, 2010, pp. 1459–1462.

[24] P. Ekman and E. L. Rosenberg, *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, USA, 1997.

[25] P. Ekman and W. V. Friesen, *Manual for the facial action coding system*. Consulting Psychologists Press, 1978.

[26] T. Baltrušaitis, "Openface," https://github.com/TadasBaltrusaitis/OpenFace, 2017, (accessed 10 May 2019).

[27] T. Baltrušaitis, P. Robinson, and L. P. Morency, "Openface: An open source facial behavior analysis toolkit," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, March 2016, pp. 1–10.

[28] M. Kassner, W. Patera, and A. Bulling, "Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, ser. UbiComp '14 Adjunct, 2014, pp. 1151–1160.

[29] Y. Nakamura, Y. Arakawa, T. Kanehira, M. Fujiwara, and K. Yasumoto, "Senstick: Comprehensive sensing platform with an ultra tiny all-in-one sensor board for iot research," *Journal of Sensors*, vol. 2017, 2017.

[30] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[31] D. Jani, "Relating travel personality to big five factors of personality," *Turizam: medjunarodni znanstveno-stručni časopis*, vol. 62, no. 4, pp. 347–359, 2014.

[32] S. D. Gosling, P. J. Rentfrow, and W. B. Swann Jr, "A very brief measure of the big-five personality domains," *Journal of Research in personality*, vol. 37, no. 6, pp. 504–528, 2003.

[33] M. Hidaka, Y. Matsuda, S. Kawanaka, Y. Nakamura, M. Fujimoto, Y. Arakawa, and K. Yasumoto, "A system for collecting and curating sightseeing information toward satisfactory tour plan creation," *The Second International Workshop on Smart Sensing Systems (IWSSS '17)*, 2017.