Multimodal Tourists’ Emotion and Satisfaction Estimation Considering Weather Conditions and Analysis of Feature Importance

Ryoya Hayashi1,2, Yuki Matsuda1,2,3, Manato Fujimoto1,2, Hirohiko Suwa1,2, and Keiichi Yasumoto1,2,
1 Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan
Email: {hayashi.ryoya.ho3, yukimat, manato, h-suwa, yasumoto}@is.naist.jp.
2 RIKEN Center for Advanced Intelligence Project AIP, Tokyo, Japan.
3 JST Presto, Tokyo, Japan.

Abstract—Smart tourism, which provides rich tourism support to tourists, is becoming increasingly common, but does not reflect the experiences of individual tourists. To provide richer tourism support, we have to recognize the psychological state of each tourist, such as emotions and satisfaction level. We previously reported that we could estimate the psychological states of tourists by measuring and analyzing their unconscious behaviors (head movements, body movements, facial expressions, and vocal expressions) during sightseeing. In this paper, we propose a new method using principal component analysis and support vector machine, which achieves higher performance than the previous method in estimating emotion and satisfaction. We evaluated the performance of the proposed method using the data of 46 participants. The result of the emotion estimation showed that the proposed method achieved unweighted average recall (UAR) of 69% compared to 51.3% for the previous method for the three-class classification task. The satisfaction level estimation achieved mean absolute error (MAE) of 1.008 for the proposed method compared to 1.033 for the previous method for a seven-level regression task. In addition, we developed a new model for estimating emotion and satisfaction considering weather conditions during sightseeing, assuming that the emotions and satisfaction levels of tourists are affected by the weather conditions. For the emotion estimation, UAR increased from 53.8% to 58.0% and for the satisfaction levels estimation, MAE decreased from 1.244 to 1.238 compared to the previous model for the data of 24 participants. We believe that considering weather conditions is effective in estimating tourists’ emotion and satisfaction.

Index Terms—smart tourism, emotion estimation, satisfaction estimation, mobile sensing, wearable device, weather

I. INTRODUCTION

Real-time urban environment information (e.g., congestion levels, event information) is becoming more easily available with the spread of various smart devices such as smartphones and the advancement of sensing technology. Smart tourism, which provides rich tourism support to tourists by utilizing these technologies, is attracting attention in the field of tourism. To make these tourism support systems more useful, we need to consider the psychological states of tourists, such as their emotions and satisfaction, which mean the feedback from users. For example, even if a tourist visits the same sightseeing spot, his/her emotions and satisfaction may be different from the average ones depending on his/her personality and preferences. If we estimate the psychological state of tourists at any time, we can utilize this information in the on-site tourism support system, which dynamically recommends sightseeing spots and routes during sightseeing [1].

The objective of our study is to investigate a method for estimating the psychological state of tourists based on quantitative data. The most common approaches to collecting tourist emotions and satisfaction levels are review website postings and surveys [2]–[5]. As is well known, these approaches have difficulties in motivating people to post reviews or participate in surveys. Also, there are issues regarding the comprehensiveness of the information and the effects of psychological bias also cannot be ignored. We assume that the psychological state of a tourist appears in various unconscious gestures such as head movement, body movement, facial expression, and vocal expression during sightseeing. Thus, we estimate the psychological state of tourists by multimodally sensing them with wearable devices and smartphones without using questionnaire surveys that were often used in existing approach [6].

In our previous study [6], we applied a neural network as the machine learning algorithm for estimating tourists’ emotions and satisfaction levels (hereinafter called the previous method). However, we believe that there is room for improvement in the estimation accuracy and computation time of the model. In this paper, we propose a new method for estimating tourists’ emotions and satisfaction levels. In the proposed method, First, we conduct principal component analysis (PCA) for dimension reduction of the features used for the model. The previous method used 188-dimensional features as input, but the proposed method reduces it to 60 dimensional features by PCA. Next, we apply support vector machine (SVM) as a machine learning algorithm. We evaluated the performance of the proposed method using the data of 46 participants. The results of the emotions estimation showed that the proposed method achieved unweighted average recall (UAR) of 69% compared to 51.3% for the previous method for the three-class classification task (positive, negative and neutral). The satisfaction level estimation achieved mean absolute error (MAE) of 1.008 for the proposed method compared to 1.033 for the previous method for a seven-level regression task.

In addition, we developed a new model for estimating emo-
tions and satisfaction levels considering weather conditions during sightseeing, assuming that the emotions and satisfaction levels of tourists are affected by the weather conditions. For the emotion estimation, UAR increased from 53.8% to 58.0% and for the satisfaction levels estimation, MAE increased from 1.244 to 1.238 compared to the previous model. On the other hand, we confirmed that if we construct a model for dataset that contains information on various seasons and different tourist destinations, the model does not improve its accuracy. We believe that it is necessary to consider the season or weather conditions before building the model to improve the estimation of tourists’ emotion and satisfaction levels.

The contributions of this paper are summarized as the following two points. First, we confirmed that the estimation performance of tourists’ emotion and satisfaction can be improved by using dimensional reduction of features and devising machine learning algorithms. Second, we revealed that tourists’ emotion and satisfaction could be affected by seasons and weather conditions by constructing the model considering weather conditions.

The rest of this paper is organized as follows. Section II explains the existing work related to our study. Section III describes the method for tourist psychological state estimation. Section IV provides the results of modeling methods described in Section III. Section V discusses the implications of considering weather conditions in the model. Finally, Section VI concludes this paper.

II. RELATED WORK

In general, the individual psychological state data is collected by questionnaire surveys. However, questionnaires are not always accurate, and the reliability of the results is problematic. Moreover, when the number of questions is large, it becomes a burden for the people who answer it. It is not desirable to conduct the questionnaire survey during sightseeing, because it will disturb the sightseeing.

In this study, we are investigating a method for automatically estimating an individual’s psychological state by sensing his/her biological information, without placing a burden on the respondent to answer. Actually, there are many researches on methods of estimating psychological states by sensing. We believe that these methods can apply to the estimation of psychological states during sightseeing.

Resch et al. [7] proposed a system (Urban Emotions) that collects physical motion data of individuals and estimates their emotions using it. They use a wristband-type wearable device to collect the data. A method for estimating emotions by sensing eye gaze is also proposed [8], [9]. A method for estimating emotion by collecting people’s voices is also proposed [10], [11]. These methods are based on laboratory environment, and their accuracy is not high when applied in noisy outdoor environment. However, Tzirakis et al. [12] reported that the combination of audio and video data has the possibility to estimate emotions even in outdoor environments.

Howarth et al. [13] investigated the relationship between the weather and the mood of college students by asking them to report their mood every day, and found that students tended to concentrate less when the humidity was higher. Klimstra et al. [14] reported that there are individual differences in the relationship between weather and mood. Denissen et al. [15] reported that mood changes with the season. In winter, high temperatures positively affect people’s mood, while in summer, it negatively affects people’s mood. However, these studies use psychological state data collected by questionnaire surveys.

III. METHOD

In this section, first, we explain the method to estimate the psychological state of tourists we have proposed previously [6]. Second, we explain collection of weather conditions during sightseeing, which is additionally included in this paper. Third, we explain the datasets that are used to construct the model for estimating the psychological state of tourists. Finally, we explain the construction method of the new model proposed in this paper.

A. Method Overview

Fig. 1 shows the workflow to estimate tourists’ emotion and satisfaction. For sensing of tourist behaviors, the tourists wear some wearable devices and have a smartphone. We defined one sightseeing spot as one session, and labeled each session with one emotion state and one satisfaction level.

1) Collection of Sensing Data: Our method can collect the following tourist behavior data. First, we collect eye movements and eye gaze during sightseeing by using Pupil Labs Eye Tracker [16], assuming that these data express the tourists’ interest in sightseeing spots because tourists can obtain a lot of visual information during sightseeing. Second, we collect the head and body movements during sightseeing.
by using the SenStick multi-sensor board [17], assuming that these also express the tourist’s psychological state because tourists are usually looking at various landmarks and moving around during sightseeing. Third, we collect facial expressions and vocal expressions by asking tourists to take selfies using the smartphone, assuming that photos and movies that are taken during sightseeing express their interest in sightseeing. For more details about the way each feature is calculated, please refer to our previous work [6].

2) Collection of Psychological State Labels: We collect the emotional state and satisfaction levels of tourists as a ground truth using smartphone application that tourists can answer on their own emotional state and satisfaction levels at the end of each session. For emotional states, we use the spatial model that expresses emotional states on two axes of Valence (Positive/Negative) and Arousal (Active/Passive), defined by Russell et al. [18]. Based on this model, emotional states are divided into three groups and nine categories: Positive group (Excited, Happy/Pleased, Calm/Relaxed), Neutral group (Neutral), and Negative group (Sleepy/Tired, Bored/Depressed, Disappointed, Distressed/Frustrated, Afraid/Alarmed). For satisfaction levels, we use 7-point Likert scale. Tourists can choose their own satisfaction level from 0 (unsatisfied) to 6 (satisfied). The initial level (before the start of sightseeing) is set to 3.

B. Collection of Weather Information

We collect weather information during sightseeing as the environmental information. It is assumed that the psychological state during sightseeing is affected by the weather conditions at that time. For example if it rains, tourists may not enjoy sightseeing fully due to the limited visibility by using an umbrella during sightseeing. Also, if it is sunny and the temperature is very high, tourists may feel uncomfortable due to sweating. In this study, we use temperature, humidity, and atmospheric pressure as the weather data that can be obtained from SenStick [17] which is used to measure head and body movements during sightseeing.

C. Dataset

The dataset we use in this study is collected from experiments in three sightseeing areas: Ulm in Germany, Nara and Kyoto in Japan. For details of sightseeing routes and information, please refer to our previous papers [6], [19]. The dataset of Ulm and Nara consists of 22 people, however, the period when we conducted the experiment is widely different for each tourist. In fact, the period we used in this study is when the weather conditions are similar for each tourist. In December 2017, January 2018, April 2018, May 2018, June 2018, August 2018, and September 2018. Therefore, the weather information such as temperature or humidity during sightseeing is different for each tourist. On the other hand, the dataset of Kyoto consists of 24 participants, its collection period is three consecutive days, from March 25 to 27, 2019. Therefore, the weather conditions are similar for each tourist during sightseeing in the dataset of Kyoto.

For these reasons, we report the results of two cases of estimating emotion and satisfaction to investigate the effect of weather conditions - one is when the weather conditions are similar; only using the dataset collected in Kyoto, the other is when the weather conditions are dissimilar: using all the dataset collected in Ulm, Nara, and Kyoto.

D. Modeling

The features used in this study are 188 dimensions (194 dimensions if we include weather data). In this study, we used PCA to reduce the dimensions to approximately one-third, which is 60 dimensions, and then applied SVM. We conducted grid search with radial basis function (RBF) as kernel, C parameter in the range of $10^0$ to $10^6$, and gamma parameter in the range of $10^{-1}$ to $10^{-6}$, and selected the best performing model among them.

IV. RESULTS

Based on the dataset we collected in the experiments described in the previous section, we constructed a machine learning model for estimating tourists’ emotions and satisfaction levels. In this paper, we explored two new approaches in comparison with the previous studies. First, we propose a new method for model construction that is different from the previous methods. We applied neural Network as the machine learning algorithm in [6], however, we newly applied PCA and SVM to improve the estimation performance of the model and to reduce the computational complexity. Secondly, we constructed a model considering weather conditions during sightseeing. In previous study, we used data on tourists’ unconscious behavior to construct models. In this paper, we newly constructed models considering weather conditions in addition to tourists’ unconscious behavior data, assuming that tourists’ emotions and satisfaction are affected by weather conditions during sightseeing. In this section, we report the results of exploring these approaches. The results are summarized in Table I and Table II.

A. Comparison of Methods

In this subsection, we report the results of the differences in the methods used to construct the model for estimating tourist emotion and satisfaction. We compared the results of the case of using neural network that we applied in [6], and the results of the case of using PCA and SVM that we applied in this study. First, we report the results of emotion estimation. Second, we report the results of satisfaction estimation.

1) Results of Emotion Estimation: Table I shows the results of emotion estimation. In this study, we treat the emotion estimation as a three-class classification task (positive, negative, and neutral). We use unweighted average recall (UAR) as the evaluation metric. Note that the previous method uses 10-fold cross validation for the evaluation. Also, the proposed method uses 5-fold cross validation to consider the small of minority data. The results of constructing the model by using all datasets collected in Ulm, Nara, and Kyoto, showed that the proposed method using PCA and SVM achieves UAR of 0.69
TABLE I
RESULTS OF EMOTION ESTIMATION

<table>
<thead>
<tr>
<th></th>
<th>Kyoto</th>
<th>Ulm &amp; Nara &amp; Kyoto</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o weather w/</td>
<td>w/o weather w/</td>
</tr>
<tr>
<td>Neural Network</td>
<td>0.423</td>
<td>0.456</td>
</tr>
<tr>
<td>Proposed Model</td>
<td>0.538</td>
<td>0.580</td>
</tr>
</tbody>
</table>

*The result is based on the model that includes the object detection and the Empatica features. For more details, please refer to [19].

TABLE II
RESULTS OF SATISFACTION ESTIMATION

<table>
<thead>
<tr>
<th></th>
<th>Kyoto</th>
<th>Ulm &amp; Nara &amp; Kyoto</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o weather w/</td>
<td>w/o weather w/</td>
</tr>
<tr>
<td>Neural Network</td>
<td>1.226</td>
<td>1.192</td>
</tr>
<tr>
<td>Proposed Model</td>
<td>1.109</td>
<td>1.092</td>
</tr>
</tbody>
</table>

*The result is based on the model that includes the object detection and the Empatica features. For more details, please refer to [19].

whereas the previous method using neural network achieves UAR of 0.513. When using only the dataset collected in Kyoto, the proposed method achieved UAR of 0.538, compared to UAR of 0.423 for the previous method.

2) Results of Satisfaction Estimation: Table II shows the results of satisfaction estimation. In this study, we treat the satisfaction level estimation as a regression task (seven levels). We use mean absolute error (MAE) as evaluation metric. Note that the previous method uses 10-fold cross validation for the evaluation. Also, the proposed method uses Leave one person out cross validation to validate more appropriate generalization performance. The results of constructing the model by using all datasets collected in Ulm, Nara, and Kyoto, showed that the proposed method using PCA and SVM achieves MAE of 1.008 whereas the previous method using neural network achieves MAE of 1.033. When using only the dataset collected in Kyoto, the proposed method achieved MAE of 1.109, compared to MAE of 1.226 for the previous method.

B. Considering Weather Information

In this subsection, we report the results of the model construction with and without weather data for the estimation of tourists’ emotion and satisfaction. First, we report the results of emotion estimation. Second, we report the results of satisfaction estimation.

1) Results of Emotion Estimation: Table I shows the results of emotion estimation. The results of constructing the model using only the dataset collected in Kyoto showed UAR of 0.456 when considering the weather conditions, whereas UAR of 0.423 when not considering the weather conditions. On the other hand, the results of constructing the model by using all the datasets collected in Ulm, Nara, and Kyoto, showed that UAR are lower when considering the weather conditions than when not considering them. These results are similar for both the previous method (using neural network) and the proposed method (using PCA and SVM).

2) Results of Satisfaction Estimation: Table II shows the results of satisfaction estimation. The results of constructing the model using only the dataset collected in Kyoto showed MAE of 1.192 when considering the weather conditions, whereas MAE of 1.226 when not considering the weather conditions. On the other hand, the results of constructing the model by using all the datasets collected in Ulm, Nara, and Kyoto, showed that MAE are lower when considering the weather conditions than when not considering them. These results are similar for both previous method using neural network and proposed method using PCA and SVM.

V. DISCUSSION

In this section, we discuss the findings based on the results of the previous section. First, we discuss the influence of considering weather conditions in constructing emotion and satisfaction estimation models. Second, we analyze the feature importance in the emotion and satisfaction estimation model constructed by considering the weather conditions.
A. Considering Weather Information

As explained before, the dataset of Kyoto was collected for three consecutive days from March 25 to 27, 2019, thus the weather conditions are almost same. Since the performance of the model is higher when considering the weather conditions for this dataset, we believe that tourists’ emotions and satisfaction during sightseeing are affected by the weather conditions during sightseeing. On the other hand, if we constructed the model considering the weather conditions for all datasets of Ulm, Nara, and Kyoto, the performance of the model did not improve. As previously mentioned, the datasets for Ulm and Nara were collected in December 2017, January 2018, April 2018, May 2018, June 2018, August 2018, and September 2018, which are different for each tourist, therefore, the weather conditions during sightseeing are also different for each tourist. In other words, it may not be appropriate to use directly the weather information obtained for model constructing. For example, if the temperature during sightseeing is the same, but the season is different, the tourists may feel the different in the feeling temperature. They may feel that the temperature is comfortable, or they may feel that the temperature is uncomfortable. Therefore, one of the effective methods to construct a model considering the weather conditions is to construct each model for each season independently. Another effective method is to construct one model for all seasons by setting the explanatory parameters that consider the effects of the seasons in advance.

B. Feature Importance

We used Permutation Importance to evaluate the feature importance. Permutation Importance is a technique to evaluate the importance of a feature based on the difference in the performance of the model when a feature is shuffled so that it no longer contributes to the model. If the performance of the model is lowered when shuffling one kind of features, we can consider that the feature is highly important. However, if the performance of the model is not changed when shuffling one kind of features, we can consider that the feature is not important because it does not affect the model.

Fig. 2 shows the feature importance of the emotion estimation model with the highest UAR score in the evaluation. The horizontal axis represents each feature. Note that the proposed method applied dimensionality reduction to the original feature vectors in advance, thus it represents the principal component (PC) No. The vertical axis represents the weight of feature importance. The higher weight of the feature importance means that the feature is important in the construction of the model. Fig. 3 shows the feature importance of the satisfaction estimation model with the lowest MAE score in the evaluation. The horizontal and vertical axes represent the same information as in Fig. 2. We can see that PC No. 1, No. 6, No. 5, No. 12, and No. 44 are the most important features in order in emotion estimation. Table III lists the 30 original features ranked in order of the contribution of each of the five PCs mentioned above. For example, PC_1 represents a high contribution of audio features, PC_4 represents a high

---

TABLE III
Principal Component

<table>
<thead>
<tr>
<th>Rank</th>
<th>PC_1</th>
<th>PC_2</th>
<th>PC_3</th>
<th>PC_4</th>
<th>PC_5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F0emitoneFrom27.5Hz_sma3_mean</td>
<td>slope500-1500_sma3_mean</td>
<td>alphaRatio_sma3_std</td>
<td>Loudness_sma3_mean</td>
<td>AU26_f_std</td>
</tr>
<tr>
<td>2</td>
<td>F0emitoneFrom27.5Hz_sma3_mean</td>
<td>HammondIndex_sma3_mean</td>
<td>mfc2_sma3_mean</td>
<td>Loudness_sma3_mean</td>
<td>AU26_f_std</td>
</tr>
<tr>
<td>3</td>
<td>F0emitoneFrom27.5Hz_sma3_mean</td>
<td>up_down_span_mean</td>
<td>mfc3_sma3_mean</td>
<td>Loudness_sma3_mean</td>
<td>AU26_f_std</td>
</tr>
<tr>
<td>4</td>
<td>shimerIdioh_sma3_mean</td>
<td>AU14_f_std</td>
<td>mfc3_sma3_mean</td>
<td>Loudness_sma3_mean</td>
<td>AU26_f_std</td>
</tr>
<tr>
<td>5</td>
<td>shimerIdioh_sma3_mean</td>
<td>AU14_f_std</td>
<td>mfc1_sma3_std</td>
<td>Loudness_sma3_mean</td>
<td>AU26_f_std</td>
</tr>
</tbody>
</table>

The suffix _sma3 means that the data is filtered with a moving average filter for time window 3 (sma3 is non-zero conditional).

Audio features are defined in GeMAPS [20].
that are more contextual for tourists’ feelings, to improve we need to explore ways of considering weather conditions the model considering weather conditions performs higher than the previous method. Second, we constructed the model considering weather conditions, as- suming that tourists’ emotions and satisfaction are affected computing complexity. The result of the emotion and satis- faction estimation showed that the proposed method achieved higher performance than the previous method. Second, we constructed the model considering weather conditions, as- suming that tourists’ emotions and satisfaction are affected by weather conditions during sightseeing. We confirmed that the model considering weather conditions performs higher than the model not considering weather conditions when the collected datasets are in similar conditions. As future work, we need to explore ways of considering weather conditions that are more contextual for tourists’ feelings, to improve the performance of the estimation of tourists’ emotion and satisfaction.

VI. CONCLUSIONS

Aiming to estimate tourists’ psychological states from their unconscious behaviors, in this paper, we explored two new approaches in comparison with our previous study. First, we constructed a machine learning model using PCA and SVM to improve the estimation performance and reduce the computational complexity. The result of the emotion and satisfaction estimation showed that the proposed method achieved higher performance than the previous method. Second, we constructed the model considering weather conditions, assuming that tourists’ emotions and satisfaction are affected by weather conditions during sightseeing. We confirmed that the model considering weather conditions performs higher than the model not considering weather conditions when the collected datasets are in similar conditions. As future work, we need to explore ways of considering weather conditions that are more contextual for tourists’ feelings, to improve the performance of the estimation of tourists’ emotion and satisfaction.

ACKNOWLEDGEMENT

This study was supported in part by JSPS KAKENHI under Grant No. JP19K24345 and JST PRESTO under Grant No. JPMPJR2039.

REFERENCES


