Exploring Travel Photo Albums as Behavioral Signals for Inferring Tourist Preferences and Personality Traits

Minaho Yamasaki
minaho.yamasaki@cocolab.jp
Okayama University
Okayama, Okayama, Japan

Yuki Matsuda
yukimat@okayama-u.ac.jp
Okayama University
Okayama, Okayama, Japan

Abstract

Personalized tourism recommendation systems often rely on explicit user inputs, such as ratings or manually specified preferences. This study explores whether user-created travel photo albums can serve as a passive behavioral data source for inferring tourist characteristics without requiring users to manually select photos for analysis. We investigate an album-level approach that estimates personality traits and tourist preferences from visual features and photo-taking behavior features derived from EXIF metadata, shooting intervals, and redundant photo-taking patterns. As an initial empirical study, we constructed a dataset of 69 travel albums from 23 participants and evaluated regression models using Leave-One-Album-Out (LOAO) cross-validation. The results showed moderate performance for selected tourism-related traits, with the best result for Independence & History in the Seven-Factor Model ($R^2 = 0.50$, correlation coefficient = 0.71). Although the results mainly reflect within-participant regularities, they suggest that travel photo albums may provide useful behavioral signals for future personalized tourism recommendation systems.

CCS Concepts

• **Human-centered computing** → *User models.*

Keywords

Tourism Recommendation, Personality Estimation, Preference Estimation, Image Analysis, Personalization

ACM Reference Format:
Minaho Yamasaki and Yuki Matsuda. 2026. Exploring Travel Photo Albums as Behavioral Signals for Inferring Tourist Preferences and Personality Traits. In *Companion of the 2026 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '26), October 11–15, 2026, Shanghai, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3798063.3837194

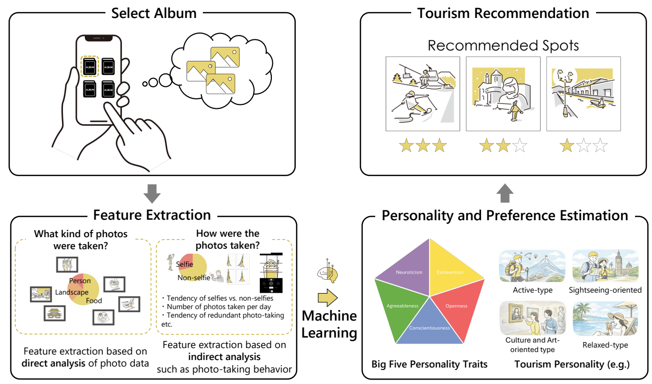


Figure 1: Overview of this study

1 Introduction

The demand for personalized tourism recommendation systems has been growing. However, traditional systems rely heavily on explicit user inputs, such as ratings or search histories, which are burdensome and fail to capture latent behavioral tendencies [4, 5, 7].

Several studies have explored travel photos or social media posts to infer user interests and personality traits [2, 3, 10, 12]. They often require manual photo selection, rely on interaction logs, or lack optimization for tourism behavior. Furthermore, most existing approaches focus exclusively on visual content, overlooking behavioral patterns embedded in the photo-taking process itself.

In this study, we explore travel photo albums as a lightweight input source for tourist profiling. Unlike approaches that require users to manually select representative photos, our approach uses participant-created albums as they are provided, without additional researcher-side photo selection. This reduces explicit user effort and preserves album-level temporal patterns. Our key idea is to analyze not only what users photographed but also how they photographed it. Specifically, we extract visual features, such as colors and image categories, together with photo-taking behavior features, such as shooting intervals, shooting volume, and redundant photo-taking patterns derived from EXIF metadata and image similarity.

Using an empirical dataset of 69 albums from 23 participants, we examine whether these album-level signals are associated with personality traits and tourist preferences. The results suggest that travel photo albums may contain useful signals related to certain tourist characteristics, while generalization to unseen users remains challenging.

2 Proposed Method

Figure 1 shows the workflow. The user only selects an existing travel album, and the system extracts album-level features from the photos and metadata without requiring manual photo selection for analysis. The features consist of direct visual features, which describe what appears in the photos, and indirect photo-taking behavior features, which describe how the photos were taken. The extracted features are input into a regression model to estimate Big Five personality traits and tourist preferences.

ACM ISBN 979-8-4007-2533-3/2026/10
https://doi.org/10.1145/3798063.3837194

Table 1: Features based on direct analysis

Tool Used	Analysis Content	Original Data Item	Calculated Features
OpenCV	Pixel ratio per color interval	Ratio of each color interval	Mean, Variance
	Ratio of warm to cool colors	Ratio of cool/warm colors	Mean, Variance
	Saturation distribution calculation	Saturation	Mean, Variance
	Saturation composition	Saturation	Ratio of Low/Medium/High (Mean, Variance)
	Value distribution calculation	Value	Mean, Variance
	Value composition	Value	Ratio of Low/Medium/High (Mean, Variance)
	Emotional dimension	PAD model [11]	Mean, Variance
LAVIS [8] (VQA)	Object attribute determination	Natural or artificial object	Number of types, Entropy
	Shooting format determination	Selfie or not	Number of types, Entropy
	Shooting format determination	Group photo or not	Number of types, Entropy
	Person containment rate	Person included or not	Number of types, Entropy
	Shooting environment	Indoor or outdoor	Number of types, Entropy
	Crowdedness of environment	Crowded or not	Number of types, Entropy
LAVIS [8] (Classification)	Statistics of category probabilities	Probability of each category	Mean, Variance
	Appearance rate of high-ranked categories	Probability of each category	Ratio in Top-n ($n = 1, 3, 5$)
	Appearance rate of high-probability categories	Probability of each category	Ratio above threshold (0.3, 0.6)
	Classification confidence metric	Probability difference of top 2	Mean, Variance
	Category diversity metric	Highest-ranked category	Ratio of top category, Entropy

For data collection, the album data used as input was obtained through the Picker API[1] (Google Photos API), from which we retrieved both the image data and associated EXIF data. Preliminary verification confirmed that image data and available EXIF metadata could be retrieved, although GPS information was not accessible and some EXIF fields were missing depending on the device or camera settings. As ground truth data for users' personality traits and tourist preferences, we employed three established scales. To measure personality traits, the Japanese version of the Ten Item Personality Inventory[2] (TIPI-J) was utilized to briefly assess the Big Five personality traits. To capture tourist preferences, we introduced two metrics: the scale developed by Jiang et al. [6] based on the International Tourist Role Scale (ITR) to evaluate familiarity/novelty orientation (Familiarity), and the Seven-Factor Model proposed by Neidhardt and Werthner [9] and adopted in PicTouRe [10] to score integrated personality and tourism-related characteristics on a scale from 0 to 100.

The features extracted from the time-series photo data are broadly classified into two categories: direct visual analysis (*"what kind of photos were taken?"*) and indirect behavioral analysis (*"how were they taken?"*).

Features based on direct analysis were extracted using OpenCV[3] and LAVIS [8], as listed in Table 1. Using OpenCV, we extracted album-level color features from the HSV color space based on prior studies [2, 3]. These features include the ratios of six hue intervals and warm/cool colors, as well as the mean and variance of Saturation and Value and their low/medium/high compositions. We also calculated PAD-model metrics [11] derived from saturation and brightness values. Using LAVIS, we extracted semantic features from each image. First, with its Visual Question Answering (VQA) function, we queried subjects and shooting environments, such as selfies, group photos, presence of people, indoor/outdoor scenes, and crowdedness, and then calculated the number of answer types and entropy $H(p)$ for each album. Second, with its classification function, we classified images into 12 tourism-related categories based on Cho et al. [1], such as Food & Beverage, Shopping, Activities, and Natural Landscape. Here, "Korean Traditional Architecture" in the original category set was adapted as "Traditional Architecture." Based on these results, album-level direct visual features were calculated as summarized in Table 1.

Features based on indirect analysis were extracted from LAVIS [8] outputs and EXIF metadata, as summarized in Table 2. Using LAVIS, we modeled redundant photo-taking behavior by combining image captioning and image-to-text similarity functions. Specifically, images and their generated captions were embedded into vectors, and image groups whose cosine similarity exceeded a predefined threshold were identified as redundant photos. Album-level statistics of these groups, such as the number of groups, redundancy rate, group size, and temporal intervals, were then calculated. From EXIF metadata, we extracted photo-taking behavior features related to camera settings, devices, shooting time, shooting volume, and image orientation. For numerical fields such as ISO sensitivity and exposure time, we calculated basic statistics, while for categorical fields such as device information and flash usage, we calculated the number of unique values and entropy. We also computed time-series features, including the average number of photos per day, ratios of photos across six time slots, shooting intervals, and burstiness across multiple time windows. Image aspect ratios and orientations were also extracted. Missing EXIF values were excluded during statistical calculations, and missing feature elements were imputed with 0 or NaN.

Considering the scale of the collected data, LightGBM was adopted for the prediction model. The input data consists of features derived from both direct and indirect analyses, while the ground truth data comprises numerical scores from the personality and preference questionnaire. Since the ground truth consists of continuous values, a regression model was employed. For model evaluation and validation, Leave-One-Album-Out (LOAO) cross-validation was used. In this method, a single arbitrary album is used as test data, and all remaining albums, including other albums from the same user, are used for training. This method was employed to evaluate how user-specific photo-taking habits impact the estimation performance. As evaluation metrics, the Mean Absolute Error (MAE), coefficient of

[1] https://developers.google.com/photos/picker/reference/rest
[2] https://doi.org/10.2132/personality.21.40
[3] https://opencv.org/

Table 2: Features based on indirect analysis

Tool Used	Analysis Content	Original Data Item	Calculated Features
LAVIS [8]	Scale of redundant shooting	Similar image groups	Number of groups, Number of redundant photos, Redundancy rate ($\frac{\text{Number of redundant photos}}{\text{Total number of photos}}$)
	Density of redundant shooting	Similar image groups	Group density ($\frac{\text{Number of groups}}{\text{Total number of photos}}$)
	Group size tendency	Similar image groups	Group size statistics (Mean, Variance)
	Inter-group interval (Time)	Shooting time	Inter-group temporal interval (Mean, Variance, Min, Max, Median)
	Inter-group interval (Photo ratio)	Number of photos between groups	Inter-group photo interval (Mean, Variance, Min, Max, Median)
	Inter-group interval type tendency	Time difference between groups	Interval type ratio (Negative, Zero, Positive)
	Intra-group shooting interval	Consecutive photos within a group	Intra-group shooting interval (Mean, Variance)
EXIF	ISO sensitivity	ISO	Mean, Variance, Min, Max, Median
	Exposure time	ExposureTime	Mean, Variance, Min, Max, Median
	Brightness	BrightnessValue	Mean, Variance, Min, Max, Median
	Shutter speed	ShutterSpeedValue	Mean, Variance, Min, Max, Median
	Exposure compensation	ExposureCompensation	Mean, Variance, Min, Max, Median, Number of types, Entropy
	Device/Lens information	LensMake, Make, etc.	Number of types, Entropy
	Flash usage rate	Flash	Number of types, Entropy
	Focal length tendency	FocalLength, FocalLengthIn35mmFilm	Number of types, Entropy
	Camera orientation tendency	Orientation	Number of types, Entropy
	Aperture tendency	ApertureValue	Mean, Variance, Min, Max, Median
	White balance tendency	WhiteBalance	Number of types, Entropy
	Shooting volume tendency	Number of photos per day	Mean
	Shooting tendency by time of day	Ratio of photos per time slot	Ratio of each time slot, Entropy
	Shooting interval tendency	Interval between consecutive photos	Mean, Variance, Min, Max, Median
	Burstiness	Number of photos per unit time	Ratio
	Aspect ratio tendency	Photo aspect ratio	Mean, Variance
	Device orientation tendency	Portrait or Landscape photo	Ratio

Table 3: Estimation results of personality traits and tourist preferences (LOAO)

Category	Item	MAE (*Mean ± SD*)	Coefficient of Determination R^2	Correlation Coefficient
Big Five[2]	Extraversion	2.34 ± 1.87	0.34	0.59
	Agreeableness	1.49 ± 1.34	0.11	0.37
	Conscientiousness	1.96 ± 1.38	0.27	0.54
	Openness	2.31 ± 1.94	0.13	0.45
	Neuroticism	2.29 ± 1.94	-0.34	0.03
ITR [6]	Familiarity	7.18 ± 6.67	0.49	0.71
Seven-Factor Model [9, 10]	Sun & Chill-Out	14.48 ± 14.26	-0.003	0.34
	Knowledge & Travel	20.91 ± 15.15	0.01	0.35
	Independence & History	14.02 ± 13.06	0.50	0.71
	Culture & Indulgence	13.04 ± 14.74	0.07	0.39
	Social & Sports	18.95 ± 13.14	0.15	0.42
	Action & Fun	15.35 ± 12.01	0.24	0.49
	Nature & Recreation	13.90 ± 12.80	-0.14	0.19

determination (R^2), and correlation coefficient were calculated. We also performed Leave-One-Person-Out (LOPO) cross-validation as a stricter evaluation setting for unseen users.

3 Experiment and Evaluation

Experimental Setup. We evaluated regression models for estimating user characteristics from travel albums, using questionnaire-based ground truth data on personality traits and tourist preferences.

The data for this experiment were collected from a total of 23 participants who each possessed three or more travel albums containing 30 or more photos. A total of 69 albums were collected, with an average of 89 photos per album (minimum: 30, maximum: 356, total: 6,158). Furthermore, a questionnaire survey regarding personality traits and tourist preferences was administered to each participant, and their responses were collected as ground truth data. The datasets exhibited a diverse distribution across all personality scales, including the five subscales of TIPI-J[2], ITR [6], and the Seven-Factor Model of PicTouRe [10]. All participants provided informed consent, and the collected photos and metadata were used only for feature extraction; derived features were stored in anonymized form for analysis.

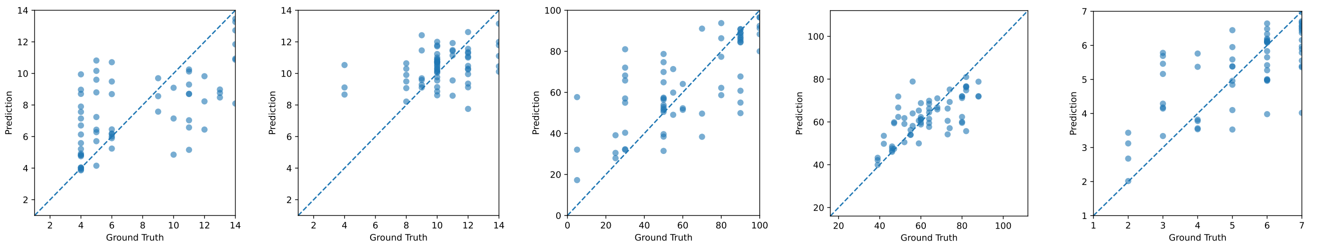
Table 4: Ablation study of direct and indirect feature groups for Openness estimation

Features	MAE (*Mean ± SD*)	Coefficient of Determination R^2	Correlation Coefficient
All	2.31 ± 1.94	0.13	0.45
Direct	2.76 ± 1.93	-0.16	0.13
Indirect	2.14 ± 2.00	0.13	0.45

Evaluation Results. The results of LOAO cross-validation are shown in Table 3 and Fig. 2. Overall, the proposed method achieved moderate prediction performance for several tourism-related dimensions, while the results varied substantially across traits. In Table 3, Independence & History and Familiarity achieved the highest correlation coefficient of 0.71, with $R^2 \geq 0.49$.

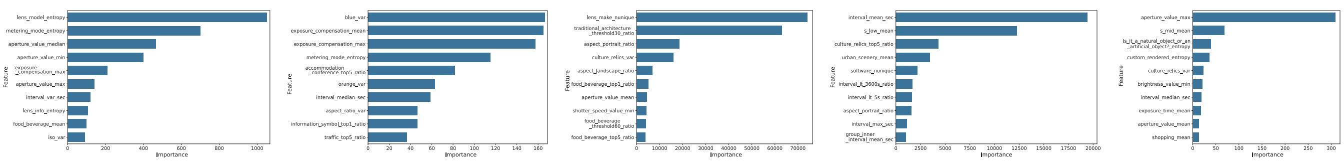
As shown in Fig. 2, Familiarity (Fig. 2(d)), Independence & History (Fig. 2(c)), and the ITR item Q6 (Fig. 2(e)) exhibited relatively clear positive trends. In contrast, several other traits showed larger deviations from the 45-degree line and limited predictability. These results suggest that album-level visual and photo-taking behavior features may capture within-participant regularities related to certain tourist preferences.

Ablation Study. To clarify feature contributions, we evaluated models trained with all features, direct features only, and indirect features only for Openness as a representative case. As shown in



(a) Extraversion (Big Five[2]) **(b) Agreeableness (Big Five[2])** **(c) Independence & History (Seven-Factor Model [9, 10])** **(d) Familiarity (ITR [6])** **(e) Q6 (ITR [6]) "I prefer to travel to countries with well-developed tourism industries."**

Figure 2: Part of the scatter plots of the estimation results in LOAO cross-validation



(a) Extraversion (Big Five[2]) **(b) Agreeableness (Big Five[2])** **(c) Independence & History (Seven-Factor Model [9, 10])** **(d) Familiarity (ITR [6])** **(e) Q6 (ITR [6]) "I prefer to travel to countries with well-developed tourism industries."**

Figure 3: Feature importances

Table 4, indirect features alone achieved performance comparable to all features, whereas direct features alone degraded performance. This result suggests that, at least for Openness, photo-taking behavior features may provide stronger signals than direct visual features. Further ablation across all target variables is needed to determine whether this tendency generalizes to other traits and preferences.

Discussion. We investigated the feasibility of estimating tourism-related user characteristics using visual and photo-taking behavior features extracted from travel photo albums. Although some traits showed moderate estimation performance, several limitations were also identified.

First, the dataset consisted of only 23 participants, which may not sufficiently capture the diversity of user characteristics. Incorporating additional features, such as photographic composition, may also improve estimation performance. The minimum photo requirement and participants' album curation practices may have affected behavioral features such as shooting intervals and redundant photo-taking patterns. Moreover, device-dependent variations and missing values in EXIF metadata may have reduced feature stability. HSV-related features are also strongly influenced by destinations and seasons, potentially introducing biases when estimation is based on a single album. Integrating multiple albums may help mitigate such effects.

Figure 3 shows that EXIF-derived and category-related features frequently appeared among the top-ranked features, suggesting that semantic and photo-taking behavior features may serve as useful signals. However, some EXIF-derived features, such as lens- and software-related metadata, may partially reflect device- or user-specific patterns; future work should evaluate models excluding such features.

To examine generalization to unseen users, we additionally performed Leave-One-Person-Out (LOPO) cross-validation, where all albums from one participant were held out for testing. Compared with LOAO validation, LOPO showed substantially lower performance, with most traits yielding low or negative R^2 values and weak correlations. Only a few traits retained moderate performance, such as Agreeableness (R^2 = 0.36, correlation coefficient = 0.63) and Independence & History (R^2 = 0.25, correlation coefficient = 0.51). These results suggest that the stronger LOAO performance mainly reflects within-participant regularities in photographic behavior, while generalization to unseen users remains challenging.

4 Conclusion

We explored the feasibility of estimating tourism-related user characteristics from travel photo albums for future personalized tourism recommendation systems. The proposed approach uses visual and photo-taking behavior features derived from EXIF metadata, shooting intervals, and redundant photo-taking patterns, reducing explicit user effort while capturing album-level behavioral signals.

Experimental results using Leave-One-Album-Out (LOAO) cross-validation showed moderate performance for selected traits, with the best result for Independence & History in the Seven-Factor Model (R^2 = 0.50, correlation coefficient = 0.71). The results suggest that travel photo albums may contain useful signals for understanding tourist characteristics, while generalization to unseen users remains challenging. Future work includes expanding the dataset, evaluating models under stricter user-independent settings, integrating multiple albums, and improving robustness across diverse shooting devices.

Acknowledgments

This work was supported in part by the SCAT Research Grant and JSPS KAKENHI Grant Number JP24K20763.

References

[1] Nahye Cho, Youngok Kang, Jiyoung Yoon, Soyeon Park, and Jiyeon Kim. 2022. Classifying Tourists' Photos and Exploring Tourism Destination Image Using a Deep Learning Model. *Journal of Quality Assurance in Hospitality & Tourism* 23, 6 (2022), 1480–1508. doi:10.1080/1528008X.2021.1995567

[2] Siham El Bahy, Noureddine Aboutabit, and Imad Hafidi. 2024. Analysis and prediction of personality traits using a self-generated database of Moroccan instagram users: impact of gender on image content and quantity on prediction accuracy. *Multimedia Tools and Applications* 83, 29 (2024), 73583–73606. doi:10.1007/s11042-024-19101-2

[3] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. 2016. Using Instagram Picture Features to Predict Users' Personality. In *Proceedings, Part I, of the 22nd International Conference on MultiMedia Modeling (MMM 2016, Vol. 9516)*. 850–861. doi:10.1007/978-3-319-27671-7_71

[4] Masato Hidaka, Yuki Kanaya, Shogo Kawanaka, Yuki Matsuda, Yugo Nakamura, Hirohiko Suwa, Manato Fujimoto, Yutaka Arakawa, and Keiichi Yasumoto. 2020. On-Site Trip Planning Support System Based on Dynamic Information on Tourism Spots. *Smart Cities* 3, 2 (2020), 212–231. doi:10.3390/smartcities3020013

[5] Masato Hidaka, Yuki Matsuda, Shogo Kawanaka, Yugo Nakamura, Manato Fujimoto, Yutaka Arakawa, and Keiichi Yasumoto. 2017. A System for Collecting and Curating Sightseeing Information toward Satisfactory Tour Plan Creation. In *The Second International Workshop on Smart Sensing Systems (IWSSS '17)*. 1–6.

[6] Jeff Jiang, Mark E. Havitz, and Robert M. O'Brien. 2000. Validating the international tourist role scale. *Annals of Tourism Research* 27, 4 (2000), 964–981. doi:10.1016/S0160-7383(99)00111-5

[7] Yohei Kurata, Yasutaka Shinagawa, and Tatsunori Hara. 2015. CT-Planner5: a computer-aided tour planning service which profits both tourists and destinations. In *Workshop on Tourism Recommender Systems, RecSys*, Vol. 15. 35–42.

[8] Dongxu Li, Junnan Li, Hung Le, Guangsen Wang, Silvio Savarese, and Steven C.H. Hoi. 2023. LAVIS: A One-stop Library for Language-Vision Intelligence. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*. 31–41. doi:10.18653/v1/2023.acl-demo.3

[9] Julia Neidhardt and Hannes Werthner. 2017. Travellers and Their Joint Characteristics Within the Seven-Factor Model. In *Information and Communication Technologies in Tourism 2017*. 503–515. doi:10.1007/978-3-319-51168-9_36

[10] Mete Sertkan, Julia Neidhardt, and Hannes Werthner. 2020. PicTouRe - A Picture-Based Tourism Recommender. In *Proceedings of the 14th ACM Conference on Recommender Systems (RecSys '20)*. 597–599. doi:10.1145/3383313.3411526

[11] Patricia Valdez and Albert Mehrabian. 1994. Effects of Color on Emotions. *Journal of Experimental Psychology: General* 123 (12 1994), 394–409. doi:10.1037/0096-3445.123.4.394

[12] Pengpeng Zhao, Xiefeng Xu, Yanchi Liu, Victor S. Sheng, Kai Zheng, and Hui Xiong. 2017. Photo2Trip: Exploiting Visual Contents in Geo-tagged Photos for Personalized Tour Recommendation. In *Proceedings of the 25th ACM International Conference on Multimedia (MM '17)*. 916–924. doi:10.1145/3123266.3123336