Vision-Based Happiness Inference: A Feasibility Case-Study

Zhanna Sarsenbayeva
The University of Melbourne
z.sarsenbayeva@student.unimelb.edu.au

Denzil Ferreira
University of Oulu
denzil.ferreira@ee.oulu.fi

Niels van Berkel
The University of Melbourne
n.vanberkel@student.unimelb.edu.au

Chu Luo
The University of Melbourne
c.luo@student.unimelb.edu.au

Mikko Vaisanen
University of Oulu
mjvaisanen@gmail.com

Vassilis Kostakos
The University of Melbourne
vassilis.kostakos@unimelb.edu.au

Jorge Goncalves
The University of Melbourne
jorge.goncalves@unimelb.edu.au

UbiComp/ISWC'17 Adjunct, September 11–15, 2017, Maui, HI, USA

ACM ISBN 978-1-4503-5190-4/17/09.
https://doi.org/10.1145/3123024.3124438

Abstract

To humanize interaction between users and computers, one needs the ability to infer the users' mood. One approach is to use a vision-based approach. We quantify the 'preview effect' bias in visual mood assessment. We demonstrate that automated tools which infer user mood from photographs or video may be affected by the presentation methodology used while performing image capture. Specifically, we demonstrate that showing a "preview" of oneself, *i.e.*, a mirror, increases the accuracy of the visual mood inference algorithms present in Google's Mobile Vision API. Our findings show that studies that incorporate visual mood assessment should include "preview" images to reduce bias and increase the reliability of vision-based happiness inference.

Author Keywords

Mood tracking; mobile instrumentation; smartphones; front-facing camera; MoodTracker; Google Mobile Vision.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

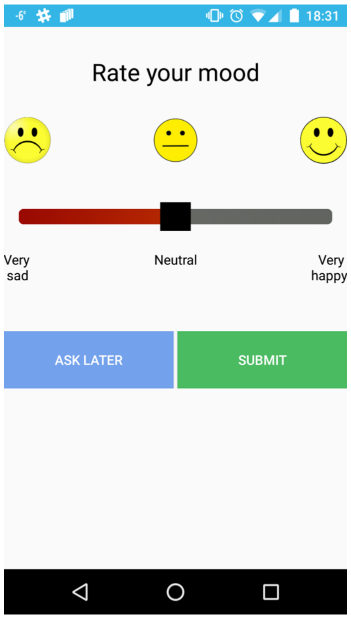

Figure 1. Self-report Mood ESM.

Introduction

Mood tracking using mobile devices is an active area of research, leveraging emerging mobile sensing technologies. Previous work shows that mood influences our behaviour and plays a significant role in our daily lives [9], and that one's facial expressions may be used to infer mood [11,12]. A core challenge is how to *accurately* and *reliably* collect mood data from mobile devices. Google recently released a lightweight Mobile Vision API (Google Vision API [14]) including a facial expression recognition algorithm that quantifies users' expression of happiness (*i.e.*, a positive mood). In psychology and medical fields, an increasing number of studies rely on participants/patients submitting self-reports of mood on a regular basis. However, it remains unknown whether collecting such mood data on smartphones can be reliable, and especially whether the self-report protocol itself may introduce bias. In this paper, we demonstrate one potential source of bias that we refer to as the 'preview effect.' Specifically, we show that visual mood inference is affected by allowing users to see themselves (*i.e.*, presented with an image preview) when assessment occurs. In our feedback sessions, we find that by seeing themselves, we affect the participant's willingness to express their most truthful mood via a facial expression.

Related Work

Mood Measure

Mental wellbeing plays a profound role in people's health and their quality of life [9,10]. Mood is considered to be a compartment of cognitive aspect of human nature [16] and can therefore be identified as a situational impairment which potentially influences cognitive aspects of interaction with the mobile device [18]. In our MoodTracker app, we focus on measuring happiness. Among the multiple moods, "feeling happy" correlates with positive valence and high activation. Happiness is perceived as one of the most beneficial and precious emotions for wellbeing [3,6].

Smartphone-based Mood Tracking

Today's smartphone capabilities go well beyond a communication device, being intertwined in todays' daily life [1,2,6]. Leveraging smartphones' omnipresence, the Photographic Affect Meter (PAM) [17] measures mood by allowing users to select 1-out-of-16 photos which best describes how they feel. Each photo represents different mood states, arranged in a 4x4 grid, along with valence for x-axis and arousal for y-axis. PAM is based on Russell's circumplex model for its simplicity, quick administration, and various definitions of mood states. PAM is considered effective and is validated against PANAS [17].

Another popular method to capture mood is using the Experience Sampling Method (ESM) via brief questionnaires displayed on the smartphones. They have been shown to measure dynamic precompetitive emotions without causing priming effects [4]. Mappiness [3] investigates the relations between individual's random momentary wellbeing states (*e.g.*, happy, relaxed, awake) and their experiences to paid work.

Most closely related to our work is EmoSnaps [15], an application that captured pictures of user's facial expressions throughout the day and used them for the later recall of momentary emotions, in which users reported via Experience Sampling Method (ESM) questionnaires. Noteworthy, EmoSnaps allowed users to recall their emotions based on captured facial expressions with considerable accuracy, even a week

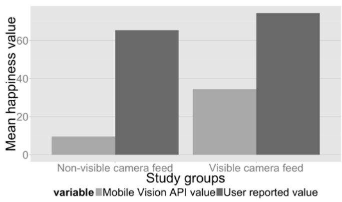

Figure 2. Mean happiness values computed by Google's Mobile Vision API and user reports for each study group.

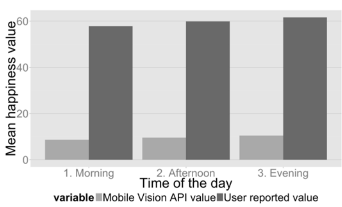

Figure 3. Mean happiness values computed by the mobile vision API and user reports for each study group during different time of the day.

after the sampling process. MoodScope [12] predicts user's mood based on their current phone usage.

Our application, MoodTracker, differs from EmoSnaps as it does not focus on emotion recall or recognition, but instead investigates if user's captured facial expressions correspond with their self-reported mood in run-time. MoodTracker extends MoodScope by capturing and validating user's mood semi-autonomously through participant self-reports.

MoodTracker

We created MoodTracker, a plugin for AWARE [7], that uses Google Mobile Vision API [14] to record users' happiness levels as inferred by analysing a photo taken using the front-facing camera. In addition, our application explicitly collects self-reported mood data via a small ESM questionnaire (**Figure 1**). Both mood capture approaches (vision-based and self-report) are simultaneously collected when the participant unlocks his phone. Following the recommendation by Conner *et al*. [5] to avoid emotion data degradation, we delimit mood reporting to once in the morning (8-12h), once in the afternoon (13-18h) and again in the evening (19-22h). In case the participant did not unlock the device in a designated slot, we would discard the request. Participants were enrolled for a 2-3 week deployment.

To ensure users' privacy and compliance to our ethical guidelines, only the inferred happiness scores were stored remotely on our server. The image captures were locally stored on the participants' device and later discarded at the end of the study. MoodTracker collected:

- The vision-based happiness score (at 0.01 steps between 0-1), provided by Google Mobile Vision API and inferred locally, *i.e.*, the image is not sent remotely.
- Self-report Mood Score (0 to 7 at intervals of 1). Similar to [13], mood is scaled between 'very sad' and 'very happy' with a neutral as the default selected option (**Figure 1**).

Study and Method

We utilise a between-subjects study design with two experimental conditions: "preview hidden" (A), and "preview shown" (B). In condition A, we captured both the vision-based happiness inference and the self-report, but hide the camera preview. In condition B, we capture the same information and include the front-facing camera preview.

Vision-based happiness values are captured as the self-report is submitted in both conditions. We test the performance of the inferred user happiness (Google Mobile Vision API), and for ground-truth, we rely on users' self-reported happiness scores. Our hypothesis is that the correlation between happiness scores measured by the vision-based approach and self-preview is visible.

Participants

Participants (N=15) were recruited through mailing lists in our campus. Participants were aged 21 to 30 years (M=26.47, SD=2.13) (4 female). Participants were rewarded with a movie voucher for their participation. Participants used their own devices during the study. The data collection occurred during the Spring semester, 2016.

Procedure

Participants arrived to a scheduled meeting with a researcher and were briefed about the purpose of the study. Then, we recorded their demographic information (*e.g.*, age, gender) and asked them to sign the consent form. We then installed the AWARE client and our MoodTracker plugin on their mobile device and randomly allocated them to one of the two conditions. We demonstrated the application functionality, given the condition assigned.

Participants were instructed to be as truthful as possible when answering the self-reports, and reassured that all the data is anonymised and the data is transferred to our server using encryption.

MoodTracker logged users' mood three times a day upon a phone unlock (*e.g.*, morning between 8h and 12h, afternoon between 13h and 18h, and evening between 19h and 22h). As users explicitly indicated their mood (self-report), MoodTracker assessed the vision-based happiness value using the front-facing camera when submitting the self-report. At this stage, participants in condition B could observe a live image of themselves, otherwise hidden for participants in condition A.

Results

From an optimal number of 945 mood records, in the end we obtained from all participants 637 entries of mood self-reports, and 368 entries of vision-based scores. This discrepancy is due to several factors: battery running out, participants not responding to the prompt for the mood within the designated time slot, or due to the inability to detect and infer the mood due to sudden motion, camera performance, and light conditions. In such cases, the API reports a value of -1, which we discarded.

A Pearson product-moment correlation coefficient was computed to assess the relationship between the self-reported happiness scores and vision-based ones. We found no correlation between the two variables ($r = 0.01$, $df = 98$, $p = 0.93$) for condition A (*i.e.*, preview hidden). However, for condition B (*i.e.*, preview shown) there is a significant positive correlation between self-reported and vision-based happiness scores ($r = 0.55$, $df = 254$, $p < 0.01$).

The mean happiness values reported by Google's Mobile Vision API and self-reports are summarised in **Figure 2**. **Figure 2** suggests that the mean happiness provided by the vision-based approach is generally low when compared to the self-reports. This difference is substantially higher for participants in condition A (preview hidden) (M=65.40, SD=22.54 self-report vs. M=9.58, SD=12.59 vision-based) when compared to condition B (preview visible) (M=74.38, SD=23.76 vs. M=34.47, SD=39.80, respectively).

We assessed if there exist diurnal effects of the reported mood (**Figure 3**). The vision-based approach, regardless of preview condition, recorded a higher mean happiness score in the evenings (M = 10.50, SD = 13.92) than in the mornings (M = 8.66, SD = 12.16) or afternoons (M = 9.62, SD = 13.35). This trend aligns well with the self-reported happiness scores, with the highest mean happiness scores in the evenings (M = 61.62, SD = 20.86), followed by afternoons (M = 59.87, SD = 19.23), and with the lowest happiness score in the mornings (M = 57.81, SD = 18.38). This indicates that both types of report vary similarly throughout the day in both preview conditions.

Discussion

From our feasibility case study, we found there is a high positive correlation between self-reported happiness scores and vision-based happiness scores for participants who could see the video preview ($r = 0.55$, $p < 0.01$).

However, there is no correlation between self-reported and vision-based happiness scores for participants in condition A (preview hidden) (*$r = 0.01$, $p = 0.93$*).

The presence of a camera preview made participants aware of their happiness state and they wanted/tried to match their facial expression to their current happiness affect state. Our findings are supported by a related study where Kleinke *et al*. [10] discovered that the participants in their experiment reported improvements in their positive mood when they simulated positive facial expressions, and deterioration in positive mood if negative facial expressions were mimicked. These changes in mood were stronger when participants could see themselves in a mirror.

We observed a temporal effect in self-reports that correlate significantly with the vision-based mood. Participants had higher happiness scores towards the evening when compared to other times of the day. Similar to [8], the trend for low happiness scores at 10am, and increased happiness values towards the late afternoon and the evening is also reflected in our study, albeit to a lesser degree. This finding suggests that to measure mood it is important to take into account the temporal effects of self-reports, and therefore request affect state measurements at different times of the day.

We acknowledge our population sample (both students, and size) are a limitation of our study and our results may apply to the atypical user.

Conclusion

We demonstrate that self-reported happiness values positively correlate with happiness values calculated by Google's Mobile Vision API when there is a visible preview. There is no significant correlation between the two variables for a group with a hidden preview. Our findings provide insight to future mood tracking experiments: a visible preview results in more consistent and reliable happiness values. In addition, we validate that users' happiness varies throughout the day, and therefore one should consider diurnal effects in the self-report contingency, *i.e.*, scheduling strategy.

Additional research is needed to investigate if other contextual factors, *e.g.*, location, ambient light, ambient noise and other situational impairments may influence happiness scores, which were not covered in our experiment. These features might prove helpful to build a model to predict user's mood based on various external predictors, such as facial expression, applications usage, physical location, and others. Then one day, self-reports can be less of a potential burden to study participants.

Acknowledgements

This work is partially funded by the Academy of Finland (Grants 276786-AWARE, 286386-CPDSS, 285459-iSCIENCE, 304925-CARE), the European Commission (Grant 6AIKA-A71143-AKAI), and Marie Skłodowska-Curie Actions (645706-GRAGE).

References

1. Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger and Gernot Bauer. 2011. Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, ACM, 47-56.
2. Barry Brown, Moira McGregor and Donald McMillan. 2014. 100 Days of iPhone Use: Understanding the Details of Mobile Device Use. In *Proceedings of the 16th International Conference on Human-computer*

Interaction with Mobile Devices & Services, ACM, 223-232.

3. Alex Bryson and George MacKerron. 2015. Are You Happy While You Work? *The Economic Journal*. http://dx.doi.org/10.1111/ecoj.12269.
4. Ester Cerin, Attila Szabo and Clive Williams. 2001. Is the Experience Sampling Method (ESM) appropriate for studying pre-competitive emotions? *Psychology of Sport and Exercise* 2, 1: 27-45.
5. Tamlin S. Conner and Katie A. Reid. 2012. Effects of Intensive Mobile Happiness Reporting in Daily Life. *Social Psychological and Personality Science* 3, 3: 315-323.
6. Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus and Anind K. Dey. 2014. Contextual Experience Sampling of Mobile Application Micro-Usage. In *International Conference on Human-Computer Interaction with Mobile Devices and Services*, ACM, 91-100.
7. Denzil Ferreira, Vassilis Kostakos and Anind K. Dey. 2015. AWARE: mobile context instrumentation framework. Frontiers in ICT 2, 6: 1-9.
8. Jorge Goncalves, Pratyush Pandab, Denzil Ferreira, Mohammad Ghahramani, Guoying Zhao and Vassilis Kostakos. 2014. Projective Testing of Diurnal Collective Emotion. In *International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 487-497.
9. Aron Halfin. 2007. Depression: the benefits of early and appropriate treatment. *The American journal of managed care* 13, 4 Suppl: S92-7.
10. Chris L. Kleinke, Thomas R. Peterson and Thomas R. Rutledge. 1998. Effects of self-generated facial expressions on mood. *Journal of Personality and Social Psychology* 74, 1: 272.
11. Peter Leijdekkers, Valerie Gay and Frederick Wong. 2013. CaptureMyEmotion: A mobile app to improve emotion learning for autistic children using sensors. In *International Symposium on Computer-Based Medical Systems*, IEEE, 381-384. http://dx.doi.org/10.1109/CBMS.2013.6627821.
12. Robert LiKamWa, Yunxin Liu, Nicholas D. Lane and Lin Zhong. 2013. MoodScope: Building a Mood Sensor from Smartphone Usage Patterns. In *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*, ACM, 389-402.
13. Alexander Meschtscherjakov, Astrid Weiss and Thomas Scherndl. 2009. Utilizing Emoticons on mobile devices within ESM studies to measure emotions in the field. *Proc. MME in conjunction with MobileHCI* 9.
14. Mobile Vision. Retrieved 10/09/2016 from https://developers.google.com/vision/
15. Evangelos Niforatos and Evangelos Karapanos. 2014. EmoSnaps: a mobile application for emotion recall from facial expressions. *Personal and Ubiquitous Computing* 19, 2: 425-444. http://dx.doi.org/10.1007/s00779-014-0777-0.
16. L. Palinkas. 2001. Circannual pattern of hypothalamic– pituitary–thyroid (HPT) function and mood during extended antarctic residence. *Psychoneuroendocrinology* 26, 4: 421-431. http://dx.doi.org/10.1016/s0306-4530(00)00064-0
17. John P. Pollak, Phil Adams and Geri Gay. 2011. PAM: a photographic affect meter for frequent, in situ measurement of affect. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 725-734.
18. Zhanna Sarsenbayeva, Jorge Goncalves, Juan García, Simon Klakegg, Sirkka Rissanen, Hannu Rintamäki, Jari Hannu and Vassilis Kostakos. 2016. Situational impairments to mobile interaction in cold environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 85-96.