

# A Closer Look at Photographers' Intentions: a Test Dataset

Mathias Lux  
Klagenfurt University  
Unverisitätsstrasse 65-67  
Klagenfurt, Austria  
mlux@itec.aau.at

Mario Taschwer  
Klagenfurt University  
Unverisitätsstrasse 65-67  
Klagenfurt, Austria  
mt@itec.aau.at

Oge Marques  
Florida Atlantic University  
777 Glades Road  
Boca Raton, FL, USA  
omarques@fau.edu

## ABSTRACT

Taking a photo is a process typically triggered by an intention. Some people want to document the progress of a task, others just want to capture the moment to re-visit the situation later on. In this contribution we present a novel, openly available dataset with 1,309 photos and annotations specifying the intentions of the photographers, which were eventually validated using Amazon Mechanical Turk.

## Categories and Subject Descriptors

H.4 [Information Systems Applications]: Miscellaneous

## General Terms

Experimentation, Human Factors

## Keywords

User intentions, digital photos

## 1. INTRODUCTION

The process of visual data production, e.g., taking a photo, is guided by the user's goals or intentions<sup>1</sup>. Similar resulting images may reflect different intentions and relate to very different contexts, e.g., a picture of a car's engine could be used to document a repair process or simply because the photographer just purchased the car and is really happy about it. Photographers' intentions may be implicit or explicit; in either case, they are usually not documented (or encoded) with the visual asset, which makes it very hard to determine, after the fact, "why was this photo taken?". In this paper we present a dataset with 1,309 photos and associated annotations, specifying the intentions of the photographers when taking those photos.

There has been a modest amount of research in the general direction of "user intentions in multimedia", partly because currently there is no commonly agreed-upon model for user intentions and in part due to the fact that only a few datasets are currently available. This paper contributes to alleviate the latter limitation, by introducing a new dataset geared towards research in the field of image retrieval and analysis based on users' intentions.

<sup>1</sup> *a thing intended; an aim or plan* as defined by the Oxford Dictionary.

## 2. THE DATASET

We collected an initial dataset from Flickr, employing an RSS feed featuring "recent additions to Flickr", between June and September 2011. For each of the selected photos we posted a comment to invite the photographer and owner of the published photo to support a scientific project by participating in a survey. We received a positive response from photographers of 1,309 photos. After obtaining their permission to use the data for scientific and non-commercial purposes, we asked the photographer the question: "What was your intention of taking the photo?". Moreover, we requested that they rate the statements listed below (whose selection was based on the results of [1]) on a five-point Likert scale, from 1 (*strongly disagree*) to 5 (*strongly agree*):

- I took the photo to support a task of mine (e.g., archive a document, communicate work progress, etc.).
- I took the photo to capture a bad feeling (e.g., sadness, anger, depression, etc.).
- I took the photo to preserve a good feeling (e.g., luck, joy, happiness, etc.).
- I took the photo to capture the moment or recall a specific situation later on.
- I took the photo to publish it online.
- I took the photo to show it to friends and family.

Fig. 1 shows the distribution of answers to the six questions in the survey. Four of the six possible intentions – *preserve good feeling* (3.34 on average), *publish online* (3.35 on average), *recall a situation* (3.81 on average) and *show to friends & family* (3.59 on average) – show a trend towards agreement. The *preserve bad feeling* intention, on the other hand, leans heavily towards disagreement (1.4 on average), whereas the remaining intention (*support task*) shows a slight tendency towards agreement (2.72 on average). Table 1 shows the Pearson correlation coefficients  $\rho$  between answers. Most of the answers are not correlated ( $|\rho|$  is very small). The highest correlation (with  $\rho = 0.45$ ) can be found between *recall a situation* and *preserve a good feeling*. The second largest observed correlation (with  $\rho = 0.29$ ) is between *show to friends and family* with both *recall a situation* and *preserve a good feeling*. In addition to the photos and the results of the survey we captured the EXIF metadata, tags, and title of the photo as given on Flickr.

Due to the open nature and the process of collecting answers of the survey, the integrity and quality of the survey's

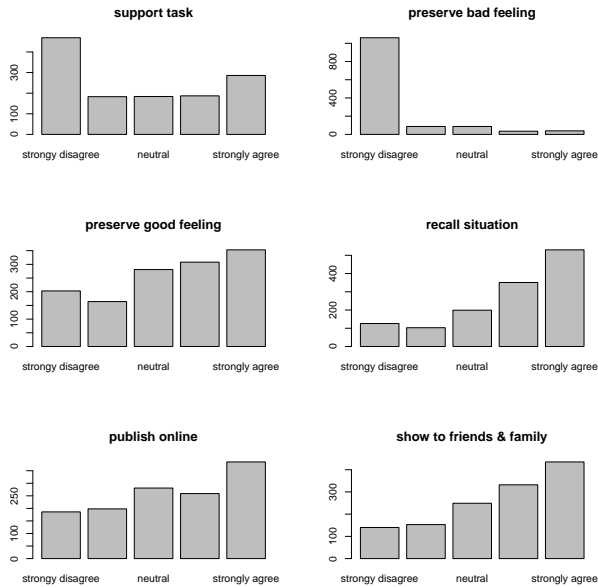


Figure 1: Survey results in absolute numbers according to the Likert scale from *strongly disagree* (left) to *strongly agree* (right).

	bad	good	recall	pub.	f&f
task	0.14	-0.05	-0.08	0.21	-0.08
bad		-0.01	-0.05	0.01	-0.08
good			<b>0.45</b>	0.04	<b>0.29</b>
recall				0.01	<b>0.29</b>
pub.					0.19

Table 1: Pearson correlation coefficient  $\rho$  of answers given by the survey participants.

answers is questionable. Hence we decided to validate the results of this first step by undertaking a second study on Amazon Mechanical Turk (mturk)<sup>2</sup>, whose goal was to validate the participants’ input and the integrity of their answers. The basic idea behind Amazon Mechanical Turk is that small, well defined, and simple tasks – called HITs (Human Intelligence Tasks) – can be offered on a marketplace-like system. Mechanical Turk workers then accept and submit the HITs by solving the tasks (for a modest pay). For each of the 1,309 photos and associated data we created five HITs to be completed by five different workers. Within each of the HITs we presented basic instructions on how to fill out the form, the actual photo, the tags assigned to the photo, and the free text description of the intention given by the photographers. The ratings of the six intentions-related questions were **not** shown within the HITs. The tasks assigned to the worker were: (i) to read and review the given information; (ii) to rate the level of digital manipulation that the image might have been subjected to (with three options: *unaltered digital photo*, *heavily manipulated picture* or *artificial image*); (iii) to rate the readability and expressiveness (i.e., whether they can infer an intention from it) of the free

<sup>2</sup><http://www.mturk.com>



Figure 2: Example photo from the dataset.

text answer; and (iv) to give a rating for each of the above six statements in the same manner as the photographers did.

## 2.1 Example

Fig. 2 gives an example of a photo from the dataset titled “packed up for the day”. The photographer states in the free text answer: “I took this photo because there was something relaxing about it. What was relaxing about this photo was that the water was very still, very calm, and because this boat’s covers were on [...]”. Tags assigned to the image by the photographer were: *lake, boat, mountain, building, calm, still, water, cover, nature, photography, ship, sea, and strings*. The photographer’s ratings of this image are 5, 1, 4, 5, 4, 4 (in the order of the questions stated above, with 1 being *strongly disagree* and 5 being *strongly agree*).

## 3. CONCLUSIONS

In this paper we introduced a new dataset geared towards research in the field of “user intentions in multimedia”. The dataset and the associated verification data, based on mturk HITs, present a novel way to approach the topic of user intentions in image production. We invite other researchers to download our test dataset and use it to advance the state of the research in this field: <http://www.itec.uni-klu.ac.at/~mlux/intentions-data-set.htm>.

## 4. ACKNOWLEDGMENTS

We’d like to thank A. Kopper and D. Khura for collecting and preparing the data. This work was supported by Lakeside Labs GmbH, Klagenfurt, Austria, and funding from the European Regional Development Fund and the Carinthian Economic Promotion Fund (KWF) under grant 20214|22573|33955.

## 5. REFERENCES

- [1] M. Lux, M. Kogler, and M. del Fabro. Why did you take this photo: a study on user intentions in digital photo productions. In *ACM SAPMIA '10*, pages 41–44, New York, NY., 2010.