

Tags You Don't Forget: Gamified Tagging of Personal Images

Nina Runge, Dirk Wenig, Danny Zitzmann, and Rainer Malaka

Digital Media Lab, TZI, University of Bremen, Germany
{nr,dwenig,malaka}@tzi.de, zitzmann.danny@gmail.com

Abstract. Mobile multi-purpose devices such as smartphones are progressively replacing digital cameras; people use their smartphones as everyday companions and increasingly take pictures in their daily life. Tagging is a way to organize huge collections of photos but raises two challenges. First, tagging (especially on mobile devices) is a boring task. Second, remembering the assigned tags is important to find images with tags. We propose gamification for more entertaining tagging. Most gamification approaches use crowd-based assessments of good or bad tags, which is a good way to prevent cheating and to not assign improper tags. However, it is not appropriate for personal images because users don't want to share every image with the crowd. We developed and evaluated two mobile apps with gamification elements to tag images, a single-player and a multiplayer app. While both variants were more entertaining than a simple tagging app, the single-player app helps users to remember significant more tags.

Keywords: gamification, image tagging, mobile devices

1 Introduction

Consumer photography has fundamentally changed twice in the last 15 years. First, moving from analog to digital photography allowed users not only to easily edit photographs but also to take as many pictures as they want without considerable additional costs for photographic films. Second, multi-purpose devices (first and foremost smartphones) are more and more replacing digital cameras specialized in taking pictures. This allows users to extensively take spontaneous pictures in their daily life.

Furthermore, smartphones are often the only device for saving, sharing and presenting images. Even though there are many programs for sorting or tagging images on a PC, the possibilities for tagging images directly on the mobile device are very limited. Most mobile image applications do only tag images based on simple computer vision algorithms, e.g. Google and Apple detect people in images. Some of them tag images when they are uploaded into the cloud with computational expensive computer vision solutions. For some users this is a serious privacy issue. To prevent massive amounts of unorganized data, it would be beneficial to have apps for tagging images directly on the mobile device.

© IFIP International Federation for Information Processing 2015

K. Choriantopoulos et al. (Eds.): ICEC 2015, LNCS 9353, pp. 301–314, 2015.

DOI: 10.1007/978-3-319-24589-8_23

Tags are well suited for organizing large collections of data. Especially for images, manual tagging is better than using simple computer vision algorithms [11]. Good tags for personal images should describe the image content well, but most important, tags should help the users to organize images and to easily find them after a while. Therefore, it is not possible to judge the suitability of tags through an expert or other people. Because the quality of produced tags cannot be determined by majority vote, crowd-based tagging is inappropriate for private photos. The users need to tag their images on their own or with friends and family.

The focus of this work is on the playful interaction that helps to encourage users to contribute user-generated personal semantic data for mobile image tagging. New object recognition algorithms [19] work quite well for detecting objects in images. This is a good starting point for an automatic tagging system, but tags are not only about the image content, but also about the context. An image from a wedding cake for example could be automatically tagged with object tags like “cake” or “knife”, but to find it later again it may be more important whose wedding the user attended or how she/he liked it. Such more personal tags cannot be solved by computer vision only or by verifying precompiled tags. While there are other approaches to learn [16] tags from the users, it is always necessary that the users add new tags from which the system could learn. Tagging images by crowd workers is also not a solution for private images.

Gamification is a good method to enforce people to solve boring tasks and has been used to tag images but, in most cases, only for public and not for private image collections [21]. Imagine you return from a wonderful holiday with your family, where you took hundreds of images. Even if an app helps you to tag your images, you probably won't be motivated to tag all the images. Wouldn't it be great to have a game you could play with fun alone or together with your children that tags the images alongside? Such a gamified app has to be entertaining and necessarily needs to produce tags of good quality. The user has to remember them to be able to find the images again later.

We developed two apps with gamification elements for mobile devices: a single-player and a multiplayer app. In a user study we evaluated both against each other and a simple non-gamified tagging app. The goal was to find out whether the gamified approaches are more entertaining than the simple tagging app and how game elements affect the quality of the image tags. Therefore, we conducted a post questionnaire after one week to determine whether the users can remember the tags they assigned to the images.

The rest of the paper is structured as follows. First, we introduce the related work. Afterwards, we present the gamified apps we developed. We conducted an evaluation with these two apps and tested them also against a non-gamified approach, as a baseline. Finally, we conclude with a discussion and present our plans for future work.

2 Related Work

Tags are a good way to organize large data collections, and are used of many large companies to visualize their data or in social media to enhance the visibility. Especially for multimedia data, tags are a good way to find them or to organize them. Users are more motivated to tag images for sharing/visibility purposes, but not for organizing their own data [1]. Therefore, many approaches try to tag data automatically. Image tagging is a large research field in the area of computer vision techniques. Images can be tagged automatically using a computer vision algorithm [3]. Many approaches try to detect objects in images; a good overview is given by Zhang [25]. New results [19] are very promising for detecting and tagging objects in images. All these approaches only use the image data to tag images.

In a collection of images, the co-occurrence of image features could be used to tag images [18]. Some approaches also use context information from the mobile device, like the location or the date, to recommend new tags [15]. But context is more than date and location [17]. Qin et al. [24] tried to recommend tags based on the information from all friends and the status of their mobile devices but this does not seem to be a very practical solution: it would require that all nearby users share their mobile device's sensor data.

Automatically computed tags are not perfect and lack contextual information. Humans are very good in detecting and identifying objects or text in images (a crucial ability to tag images), but in addition they also might know information about the context of the image. This is a reason why we should bring humans into the loop [14]. A way might be to recommend new tags to the users and let them add additional ones. Recent approaches try to learn new tags from user input [16]. They require extensive training and in particular personalization needs to adapt the mapping of features to tags for each and every single user. For tagging of private images, a combination of different approaches might be a solution. Apart from that, it is also possible to improve computer vision algorithms with user tags [9].

Human computation tries to motivate people to solve tasks that could not be solved by computers. Gamification is often used to better motivate people to solve these tasks in a gamified environment or in a game. Games with a purpose (GWAPs) are often used to tag images because tagging is a task which could not be fully solved computer-based and is very boring and repetitive [6, 23]. The ESP game for example is used to tag images for the Google search engine. Von Ahn showed that people do not play these games because they want to help solving tasks but want to play entertaining games; e.g. Peekaboom [22], KissKissBan [8].

Krause et al. [12] argued for GWAPs that encapsulate the task as much as possible in the game because it motivates users to play the game. However, hiding a task like inserting text is not easy to be included e.g. in a first-person shooter. Takhtamysheva and Smeddinck [20] showed that it is often sufficient enough to add only some playful elements like sound and graphics to motivate the users. Mekler et al. [13] found out that even simple game elements increase the amount of tags users assign to paintings. Especially points, levels and leaderboards should be used to boost user

performance. They also showed that the intrinsic motivation is not effected by the game elements they used.

For personal images there are only a few games or gamified approaches. One example is a kind of memory game [2]. A problem of this approach is that only one expert rated the quality. Until now, it has also not been evaluated how good users can remember their own tags. Some approaches also use new interaction techniques to motivate the user [7]. But such games are only designed for selecting tags and not for adding tags, which is crucial for a good tagging system.

3 Gamified Tagging for Personal Images

Our goal was to develop apps with simple gamification elements that allow users to tag images in an entertaining way and help them to remember the tags they have created. The system should run on mobile devices because many people only use these devices to organize, present and share images. Because we believe that users can remember tags best when they create their own individual tags (not limited to a predefined list), the apps should allow free entries. Moreover, the time required for creating tags should be similar to simple tagging apps. This requires efficient (soft-) keyboard input for new text and excludes gamification of text-entry (e.g. hitting appearing characters in the manner of a Whac-A-Mole game).



Fig. 1. Simple tagging app without gamification

We aimed for an intuitively usable gamified application. A shooter game or very complex role-play would be problematic for users who are not familiar with these genres. Furthermore, it is complicated to integrate them with an un-gamified input

like the keyboard. Therefore, we decided to develop a casual game that is easy to understand and to easy to play for everybody. The decision for the game genre was also influenced by the requirements that we did not want to use the crowd to verify or recommend tags. To develop something like a guessing game, wrong tags need to be created. When a user plays a game just after creating an image, nothing is known about these images and no wrong tags can be established.

Furthermore, we were interested whether a competitive task leads to more cheating or different tags and developed a single- and a multiplayer application. Both apps were realized for the Android platform using the libGDX framework. For the multiplayer system we used a server to establish the communication between the players. We tested these two apps against one simple app as a baseline. It allows adding tags to an image but without any gamification elements. Figure 1 shows the app: the user looks at an image, can add tags and edit or remove them afterwards. We call this app the simple app. The simple tagging app and the single-player app are self-explanatory for the users, for the multiplayer app we developed a small tutorial directly embedded in the app.

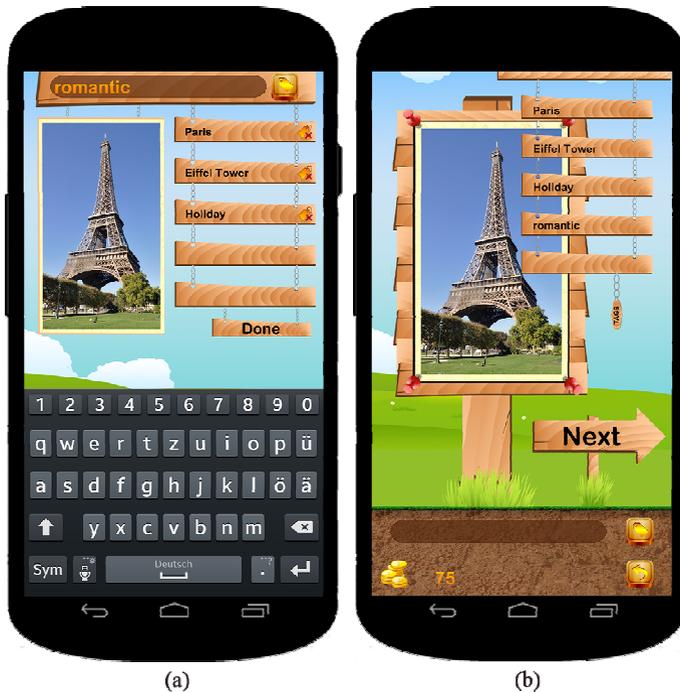


Fig. 2. Gamified single-player tagging app: (a) text entry and (b) tagged image

3.1 Single-Player App

The requirements for the gamified app are similar to the simple tagging app. A user should see an image, add tags and edit tags. This prototype is based on the assumption

that even very simple game elements can motivate people to solve tasks [13]. Gamification of tagging with text input is not a trivial task, especially if nothing is known about the images. Only the numbers or the lengths of the tags are candidate measures for points and highscores. As users realize such mechanisms, they might be enforced to just add tags based on these measures. Therefore, we carefully balanced the gamification elements and the main task. In particular, we only used small and minimalistic but very typical game elements: background graphics that give the impression of a casual game, background music and sound for every input, graphics and points for every tag users add to an image.

The gamified variant is shown in Figure 2. In this prototype, an image is presented to the user in gamified graphics. The user has to add tags for which she/he receives points. Adding the same tag multiple times is excluded, but there are no semantic evaluations of the tags. People may add tags, which are semantically not correct (e.g. “petersparty” instead of “peter” and “party”), but they might be able to remember these tags very well.

Every point a player gets is accentuated with sound. These are the only game elements. We decided that the length of the tags do not score, because ‘the longer, the better’ is not true for tags and we do not want to bias the users or set them on a wrong track how to tag images. The only differences to the simple tagging app are:

- Background graphics
- Background sound and sound for adding tags
- Five points for every added tag

In the evaluation we investigated how these simple elements entertain the user and influence the recall of tags compared to the multiplayer app.

3.2 Multiplayer App

The second prototype uses the game elements of the single-player app and additional elements that enhance the game-character of the app. Competitive image tagging in a game or gamified app might be fun but leads to some problems: people might be more interested in the game and winning aspects and thus neglect the serious tagging task. This is the reason why many multiplayer apps for image tagging use a tag matching procedure for distinct users who tagged an image. Those mechanisms, like used in the ESP Game, can prevent the users from cheating. However, in our approach we use personal images which excludes crowd-based mechanisms. We also used no experts or a gold standard to evaluate the tags.

Von Ahn and Dabbish [23] distinguish three types of Human computation games:

- Input-agreement: One player gets an input and has to assign an output to it. The second player also receives an input and the output from the first player. Output is added if the user decides input and output fit.
- Output-agreement: Both players have to produce an output for a given, equal input. If the output is equal, it is added.
- Inversion-problem: One player creates the output and the second one has to decide if this output is correct.

We decided to implement a game which uses an inversion-problem because only using this technique we get new and verified tags from every round played by the users. Furthermore, not both players have to add tags with the keyboard, which is probably the most boring part. In our app, a user starts the competition by tagging an image. Then a second player joins the session and has to guess which tags are added by the first one. Alternately, the players are either tagger or guesser. Tags are shown for a few seconds on the screen and the guesser has to click on the tags that seemed to be correct for her/him. To realize the multiplayer app, we used a client-server implementation.

Each round contains six images. The tagger assigns three tags and one category (from the categories home, animals and people, landmark, event, or on the road) to an image. An overview of this approach is shown in Figure 3. The inversion-problem approach requires wrong tags: the guesser achieves points for correct tags while s/he loses points for incorrect ones. Wrong tags should not be fully improper (e.g. person for a landmark). Therefore, we decided to learn wrong tags directly from the user and introduced the *evil tag*. One evil tag is chosen by the tagger and can be used to set the guesser on a wrong track (e.g. a wrong spelled name, or a wrong event name). If a family returns from holidays and wants to tag the taken images together, this is a good way to challenge each other. Furthermore, we also used the non-selected categories as wrong tags. Users get points for a correct category and even more for a correct tag, but also discount for wrong selected tags (categories and the evil tag). For the correct category a user receives five points and for a tag, which was typed in by the other user, 25 points. We used the same discount if a user selects the evil tag or a wrong category.

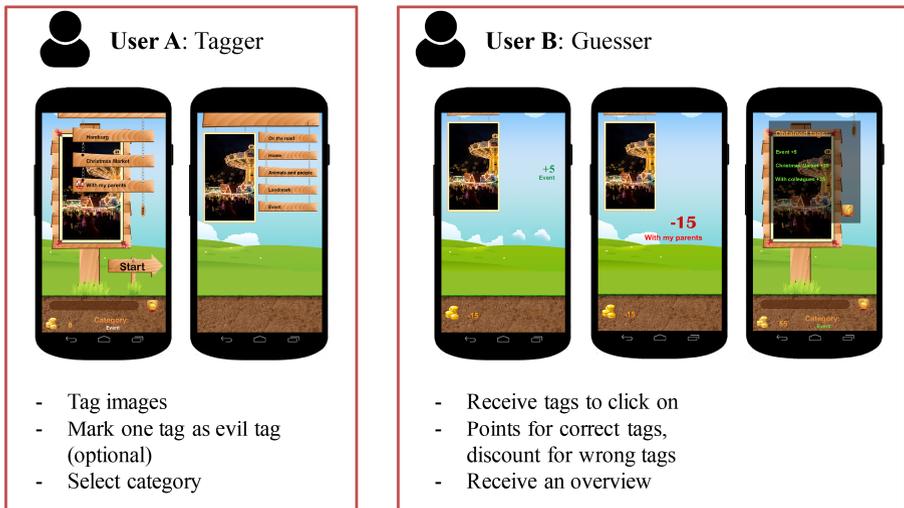


Fig. 3. Multiplayer app: Tagger on the left, guesser on the right

4 Study Design

We recruited 27 participants (20 male and 7 female) aged between 18 and 32 years ($M=26$). Most participants (26) own a smartphone and use its camera to take pictures. Nine of them use the camera at least once a month, 13 once a week and four of them even daily. None of our participants has ever tagged images. Twenty-two participants have games installed on their mobile device. The majority favor strategy games, followed by quiz games. But most of the participants do not play games very often.

For the user study we decided to use our own images because not every participant may comply with sharing her/his private images. Even though this has some disadvantages such as users do not know the whole story behind the images, most important is the advantage that the results are comparable. We selected 18 pictures which are good to tag, representative for mobile photos (like events, people, objects) and not displeasing. In the start questionnaire we asked the participants for the tags they would add to most of their images on their mobile devices. Most prominent tags are: family (added by 18 participants), holiday (16), friends (14), party (6), me (5), which corresponds to our image collection.

We evaluated the apps in a controlled lab setting using a within-subject study design. Every participant had to tag 18 images; six with every prototype. We shuffled the order of the prototypes based on Latin square. Additionally, we varied the order of the images so that every image was equally often tagged with every prototype on every position. We tested the prototypes with a Nexus 7 tablet, so the participants could see every detail of the image and all game elements.

The participants had to answer one general questionnaire at the beginning and different questionnaires after each prototype. After using the simple tagging app, they filled out the System Usability Scale (SUS [4]), which allows us to exclude that usability issues influence the tagging procedure. After the two gamified approaches, the users had to answer the Post Game Experience Questionnaire [10] to verify the motivation and emotions during the gameplay. We also used the Intrinsic Motivation Inventory [5], which is very good for gamified tasks, because we want to know if the users play the game only to solve the task or if they are intrinsically motivated to play the game.

All test runs were recorded on video. We did not enforce the participants to think aloud because this could influence the game play. After the participants have tested every prototype, they were interviewed and asked how they liked the prototypes in general, what they think about their own performances and about the quality of the tags they have created with them. We also conducted an interview at the end, asking what they liked about the apps and what might be improvements.

Because we wanted to know if the game elements help the participants to remember the tags they have created, we contacted them one week later. We asked them directly or via video chat to tag all the images again. Using this information we evaluated how good the tags, assigned with the different prototypes, are. Image tags are used to find images again later, so the recall of the tags is the crucial for the tag quality.

5 Results and Analysis

The simple tagging app had no fundamental usability problems and achieved a SUS score of $M=89.8$ ($SD=10.85$). So the differences in the following results are not caused by usability problems in the simple app, which we used as a baseline.

5.1 Gamification

For the two gamified apps, the game experience questionnaire gives quite good results but no major differences for single- and multiplayer approach. The participants rated the use of the app as a positive experience (single-player: $M=3.59$, $SD=.71$; multiplayer: $M=3.87$, $SD=.56$; 1=strong reject, 5=strong accept). They also rated the app as not exhausting (single-player: $M=1.09$, $SD=.28$; multiplayer: $M=1.17$, $SD=.34$). Furthermore, the evaluation showed that it is easy to return to reality in both apps (single-player: $M=1.47$, $SD=.51$; multiplayer: $M=1.58$, $SD=.56$) probably because of the game genre. But this is typical for mobile and also for casual games. Because there are no statistical differences between the single- and multiplayer app, we can conclude that no further results are based on different game experiences.

Using the IMI-Questionnaire we tested how motivated the people are to play with the apps and not only to solve the tagging task. These results are good and very similar for the single- and multiplayer app. The IMI results are summarized in four dimensions (scale: 1=total disagree, 5=total agree; single-player=sp; multiplayer=mp):

- *Interest* (sp: $M=4.72$, $SD=.31$; mp: $M=4.89$, $SD=.17$)
- *Competence* (sp: $M=4.58$, $SD=.27$; mp: $M=4.42$, $SD=.12$)
- *Perceived Choice* (sp: $M=3.87$, $SD=.24$; mp: $M=3.74$, $SD=.17$)
- *Perceived Pressure* (sp: $M=2.61$, $SD=.20$; mp: $M=2.85$, $SD=.16$)

Because statistical tests showed that the dimensions are not normally distributed we used a Wilcoxon test. The dimension *interest* is significantly higher for the multiplayer app ($T=23.5$; $p<0.001$; $r=-0.74$), *perceived choice* ($T=35.5$; $p<0.001$; $r=-0.71$) and *perceived pressure* ($T=14.5$; $p<0.01$; $r=-0.81$) are also highly significant, but not the dimension *competence*. We can conclude that the multiplayer app leads to more intrinsic motivation, which is interesting because they also have to use the keyboard and repetitive add tags.

5.2 Fun and Perceived Quality of Tags

After the participants tested all three prototypes we asked them to score every prototype with respect to the fun they had (see Figure 4) and the tag quality they think they have achieved. Every prototype was scored on a 5-point scale (1=worst, 5=best).

The simple tagging application was not rated as entertaining by most of the participants ($M=3.07$, $SD=1.37$), the single-player app was more fun ($M=3.81$, $SD=1.00$) and the multiplayer approach was rated best ($M=4.52$, $SD=.80$).

The multiplayer app was significantly better rated than the simple tagging app ($t(26)=1.99, p<.001$). All other relations showed no significant differences.

The participants also rated the perceived quality of the tags they added while using the different applications. All results are very good without major differences for the prototypes (simple app: $M=3.93, SD=.96$; single-player: $M=3.89, SD=1.05$; multiplayer: $M=4.04, SD=.90$). This corresponds with the results from the IMI questionnaire stating that there were no major differences in the perceived competence. As the participants did not have the feeling that the game part distracted the quality of the tagging task, this is quite astonishing.

In the last question we asked the participants to rate the three prototypes. The results are very clear: 88% of our participants favored the multiplayer and 83% would start tagging images with this prototype.

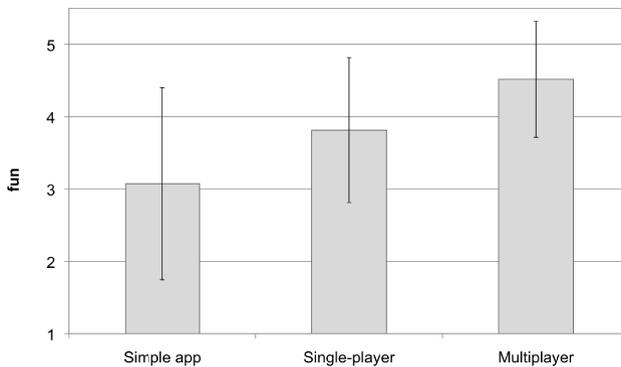


Fig. 4. Mean fun during tagging, error bars indicate standard deviation

We conclude that especially a multiplayer tagging app for playing with family and friends (and not strange people in the cloud) would be beneficial, because it motivates the users to tag their own images with fun.

5.3 Tag Recall

While one expert evaluated all tags to be sure that they all fit the images, for private images it is more important whether the users can remember their own assigned tags. All participants tagged the images one week after the initial trial again. We computed a score for each participant how good they produce the same tags again. Therefore we computed the f1-score, based on the precision

$$P = \text{TruePositiv} / (\text{TruePositiv} + \text{FalsePositiv})$$

and the recall:

$$R = \text{TruePositiv} / (\text{TruePositiv} + \text{FalseNegativ})$$

The f1-score is a measurement for classification tasks. We used it to measure whether tags assigned with the prototypes are equal to those in the post questionnaire.

The score describes how good people can remember the tags with each prototype (0=worst, 1=best).

$$f1 = 2 \cdot \frac{P \cdot R}{P + R}$$

Figure 5 shows the mean f1 scores and standard deviation for each prototype. The single-player ($M=.72$, $SD=.11$) application produced significant better results than the simple tagging application ($M=.60$, $SD=.15$), $t(26)=1.99$, $p<.001$. Because we did not implement any differences in these two apps despite the graphics, music and the points for a tag, we can conclude, that just these gamification elements lead to better image tags the users can remember.

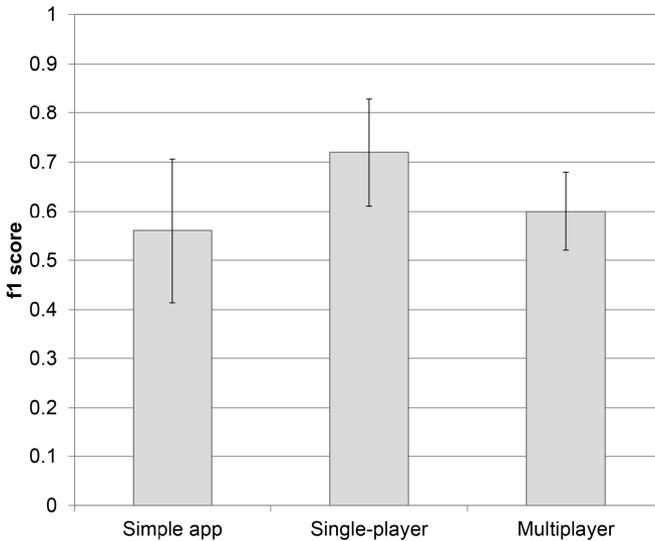


Fig. 5. Mean f1-scores for all participants, error bars indicate standard deviation

The f1-Score for the single-player application is also significantly higher than the score for the multiplayer gamified approach ($M=.56$, $SD=.08$), $t(26)=1.99$, $p<.001$ which means that our multiplayer app produced less valid tags. This might have different reasons. For instance, some elements distracted the users and they forgot some tags during the game play. Another reason might be the tag comparison between the users. Tags were only counted as correct tags, if and only if the first player has assigned the tag to the image and the second player guessed it correctly.

6 Discussion

Tagging images requires time and is a tedious repetitive task. Especially on mobile devices it is not very motivating to tag all images. Gamifying the task is a promising approach to improve the motivation. A problem is that many users are more

motivated to win the game than to create good tags. Most approaches use many different people (such as *crowd workers*) to prevent the users from cheating and use the majority vote to evaluate tags. But using the crowd is not appropriate for personal images. Therefore we created a gamified approach which does not need any other users or experts: one for a single-player and one for multiple players. Both are evaluated against each other and against a non-gamified application.

In the evaluation, we provided evidence that gamification does not only turn the task into an entertaining activity but also helps remembering the assigned tags which in turn facilitates human recall. In the single-player condition, the participants remembered significantly more tags than in the non-gamified one: a few game elements helped to improve the recall significantly. The multiplayer variant was most entertaining and the participants had significantly more fun than in the simple app condition. Most of them favored multiplayer and would start tagging images because of it. This leads to the conclusion that there is a need for gamified multiplayer apps that are not only entertaining but also help the user to remember the tags. In this paper we have shown that gamified approaches do not need other experts or the crowd to create good tags. Even with a multiplayer approach we were able to create mechanisms to force the users to add good tags.

While multiplayer gamification approaches should be competitive, realizing such mechanics is a challenge. People focus on winning the game and might begin to cheat and to enter inappropriate tags. On the other side, multiplayer approaches can be used to prevent users from cheating (e.g. different users suggest the same tag). One expert evaluated all the image tags from the prototypes and all of them could be rated as correct (no cheatings or totally wrong tags). Most important, it is astonishing, that users could remember the assigned tags significantly better if they use a gamified approach.

7 Conclusion and Future Work

The approach of gamification is promising for tagging personal images. It does not only provide an entertaining alternative but also helps users to remember the assigned tags better. Furthermore, it seems that the users are not interested in cheating when tagging personal images. This allows creating apps with gamification elements that not necessarily include verification through an expert or the crowd.

In the study, the users were not interested in cheating the system at all. This might have very different reasons:

- They are focused on the tagging aspect.
- The results are influenced by the innovativeness.
- The lab setting influenced the results.

A long-term study is required to find out more about the reasons. Future work should investigate, if cheating in the long run is not a problem for personal images.

Currently we are planning to develop a gamified app for multiple players that ideally does not only make more fun but also helps to remember tags. This is

interesting whenever a family or a circle of friends tag images together. A multiplayer approach with more than two players might be a solution to evaluate the tags based on the majority vote of the other players. Integrating a recommender system might allow new variegated game mechanics. We are planning to integrate computer vision algorithms and machine learning mechanisms to recommend tags or to integrate them into the system.

Furthermore, this work does only investigate the part of creating tags. Although the presented approach helps the users to remember the tags, systems and user interfaces that allow her/him to easily find the images s/he is looking for are still needed. Based on our insights we want to develop more gamified apps, not only for creating new tags, but also for the verification of tags, and combine these with a recommender system.

References

1. Ames, M., Naaman, M.: Why We Tag: Motivations for Annotation in Mobile and Online Media. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2007, pp. 971–980. ACM (2007)
2. Bieliková, M.: Personal Image Tagging: a Game-based Approach. In: Proceedings of the 8th International Conference on Semantic Systems, I-SEMANTICS 2012, pp. 88–93. ACM (2012)
3. Binti Zakariaa, L.Q., Lewisb, P., Hallc, W.: Automatic Image Tagging by Using Image Content Analysis. In: Proceedings of the International Conference on Informatics and Applications, pp. 91–101. The Society of Digital Information and Wireless Communication (2012)
4. Brooke, J.: SUS - A Quick and Dirty Usability Scale. In: Usability Evaluation in Industry, Taylor & Francis (1996)
5. Deci, E.L., Ryan, R.M.: Self-Determination. John Wiley & Sons (2010)
6. Goh, D.H., Lee, C.S.: Perceptions, quality and motivational needs in image tagging human computation games. *Journal of Information Science* 34(5), 515–531 (2011)
7. Gonçalves, D., Jesus, R., Correia, N.: A Gesture Based Game for Image Tagging. In: Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems, CHI EA 2008, pp. 2685–2690. ACM (2008)
8. Ho, C.J., Chang, T.H., Lee, J.C., Hsu, J.Y.J., Chen, K.T.: KissKissBan: A Competitive Human Computation Game for Image Annotation. In: Proceedings of the ACM SIGKDD Workshop on Human Computation, pp. 11–14. ACM (2009)
9. Hwang, S.J., Grauman, K.: Reading between the lines: Object Localization Using Implicit Cues from Image Tags. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34(6), 1145–1158 (2012)
10. IJsselsteijn, W., De Kort, Y., Poels, K., Jurgelionis, A., Bellotti, F.: Characterising and Measuring User Experiences in Digital Games. In: Proceedings of the International Conference on Advances in Computer Entertainment Technology, ACE Conference 2007, pp. 27–31. ACM (2007)
11. Kennedy, L.S., Chang, S.F., Kozintsev, I.V.: To Search or To Label? Predicting the Performance of Search-Based Automatic Image Classifiers. In: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval, pp. 249–258. ACM (2006)

12. Krause, M., Takhtamysheva, A., Wittstock, M., Malaka, R.: Frontiers of a Paradigm: Exploring Human Computation with Digital Games. In: Proceedings of the ACM SIGKDD Workshop on Human Computation, pp. 22–25. ACM (2010)
13. Mekler, E.D., Brühlmann, F., Opwis, K., Tuch, A.N.: Do Points, Levels and Leaderboards Harm Intrinsic Motivation? An Empirical Analysis of Common Gamification Elements. In: Proceedings of the First International Conference on Gameful Design, Research, and Applications, pp. 66–73. ACM (2013)
14. Motoyama, M., Levchenko, K., Kanich, C., McCoy, D., Voelker, G., Savage, S.: Re: CAPTCHAs – Understanding CAPTCHA-Solving Services in an Economic Context. In: USENIX Security Symposium, vol. 10, pp. 3–21 (2010)
15. Proß, B., Schöning, J., Krüger, A.: iPiccer: Automatically retrieving and inferring tagged location information from web repositories. In: Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices & Services, MobileHCI 2009, pp. 69–71. ACM (2009)
16. Runge, N., Wenig, D., Malaka, R.: Keep an Eye on Your Photos: Automatic Image Tagging on Mobile Devices. In: Proceedings of the 16th international Conference on Human-Computer Interaction with Mobile Devices & Services, pp. 513–518. ACM (2014)
17. Schmidt, A., Beigl, M., Gellersen, H.W.: There is more to context than location. *Computers & Graphics* 23(6), 893–901 (1999)
18. Sigurbjörnsson, B., Van Zwol, R.: Flickr Tag Recommendation based on Collective Knowledge. In: Proceedings of the 17th International Conference on World Wide Web, WWW 2008, pp. 327–336. ACM (2008)
19. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. In: Proceedings of the International Conference on Learning Representations, ICLR 2015 (2015)
20. Takhtamysheva, A., Smeddinck, J.: Serious questions in playful questionnaires. In: Herrlich, M., Malaka, R., Masuch, M. (eds.) ICEC 2012. LNCS, vol. 7522, pp. 449–452. Springer, Heidelberg (2012)
21. Von Ahn, L.: Games with a Purpose. *Computer* 39(6), 92–94 (2006)
22. Von Ahn, L., Liu, R., Blum, M.: Peekaboom: A Game for Locating Objects in Images. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2006, pp. 55–64. ACM (2006)
23. Von Ahn, L., Dabbish, L.: Designing Games With A Purpose. *Communications of the ACM* 51(8), 58–67 (2008)
24. Qin, C., Bao, X., Roy Choudhury, R., Nelakuditi, S.: TagSense: A Smartphone-based Approach to Automatic Image Tagging. In: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, MobiSys 2011, pp. 1–14. ACM (2011)
25. Zhang, D., Islam, M.M., Lu, G.: A review on automatic image annotation techniques. *Pattern Recognition* 45(1), 346–362 (2012)