Using Microtasks to Enhance the Recognition of Receipt Entities to Keep Track Of Expenses

Maximilian Altmeyer
DFKI GmbH,
Saarland University
Saarbrücken, Germany
maximilian.altmeyer@dfki.de

Pascal Lessel
DFKI GmbH,
Saarbrücken Graduate School
of Computer Science
Saarbrücken, Germany
pascal.lessel@dfki.de

CHI 2016. Workshop: Productivity Decomposed: Getting Big Things Done with Little Microtasks

Introduction

Optical character recognition (OCR) has improved significantly in recent years [6] and is used in various domains ([2, 5, 8]. However, OCR results are still error-prone and heavily depend on the quality of both the picture taken and the printed text [4]. This makes it hard to achieve reliable OCR results in domains that are confronted with e.g. text fonts that are hard to recognize, pale ink, or crumpled paper [9] and gets even worse when taking pictures with a smartphone camera [4]. One approach to fix spelling errors is using crowdsourcing which relies on the concept of the wisdom of crowds [7]. This concept states that a group of people is able to come to a better decision than an individual [7]. We investigated crowdsourcing and gamification in the domain of receipt capturing to keep track of expenses as we found evidence that users are interested in this but shy away from the huge effort. We developed a budgeting app for smartphones that allows for tracking expenses by taking pictures of receipts to extract relevant entities such as the total sum, store name, single articles and their corresponding prices, and a categorization of each article. To enhance the recognition algorithm, we used the outcome of different microtasks that were solved by users of our app. Solving these microtasks is not motivated by monetary rewards. Instead, we use gamification - the use of game elements in non-game contexts [3].

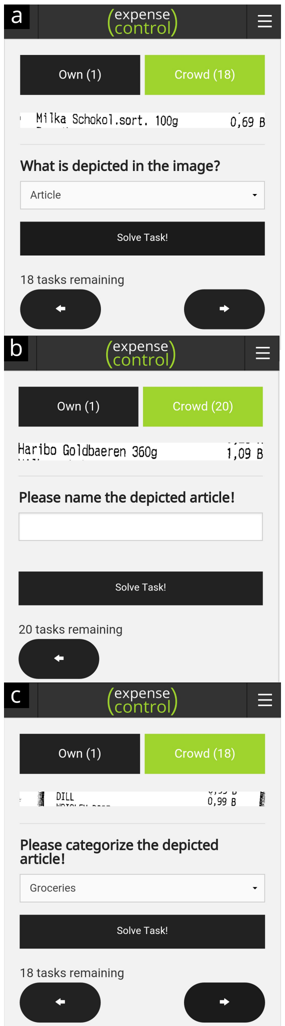

Figure 1: Different microtasks that were solved by users of the app

Microtasks used in the System

By using microtasks we aimed to provide information for our extraction algorithm to solve the following challenges:

C1 Identification of articles and extraction of their prices

C2 Categorization of single articles and the whole expense

C3 Identification and extraction of the total sum

In the app, each microtask is shown isolated and consist of an image of the unknown receipt line and a short task description. We decided to use three different task types, which we assume to be helpful to match articles and prices, to extract the total sum, and to categorize articles and the overall expense (cf. Figure 1):

Classification Microtasks
We identified three different types for entities: **article names**, **additional information** (e.g. article numbers or quantity indications) and **total sum**. Given the classification of a line (whether it contains an article, additional information or a total sum), we are able to match articles and prices and furthermore extract an overall sum solving **C1** as well as **C3**. The user is asked to identify the entity to be an article, additional information or total sum, as depicted in Figure 1a).

Article Correction Microtasks
This microtask is created once a line is identified as an article by the crowd. The user is asked to name a depicted article as shown in Figure 1b). The outcome of this microtask is used to correct OCR errors (spelling errors), identify articles and provide an article name, since the articles are often abbreviated on the receipt.

Article Categorization Microtasks
Again, this microtask is generated after an entity was classified as an article, to obtain a category for it, thus solving **C2** (cf. Figure 1c).

Discussion

The outcomes of the used microtasks, that were derived by majority votings, reduced the error rate significantly when extracting receipt entities (The error rate was reduced by nearly two thirds, cf. [1]). However there were still errors (10.36%) after the changes of the crowd were applied, posing the question for reasons. First, there were errors in the algorithm cutting out every line from the receipt picture leading to microtask pictures that were ambiguous, incomplete or broken. However, there were also wrong classifications and corrections made by the crowd, even when the picture was perfectly extracted. Reasons might be uncertain microtask descriptions or the will to collect points and unlock achievements to compete other players on the leaderboard, as these elements were offered due to the gamification concept. Furthermore the aggregation method could have been altered: Instead of using a majority voting to obtain outcomes of the microtasks, other options could lead to fewer errors (such as an expert voting in which votes of trustfully users are more weighted). We moreover found that users subjectively did not consider solving microtasks fun or engaging, although the use of gamification significantly lead to a higher amount of solved microtasks. It is therefore necessary to investigate further incentive methods to motivate and retain users to solve more microtasks over a longer timespan.

Biography

Maximilian Altmeyer holds a bachelor's degree in media computer science from Saarland University and is currently enrolled as a master student. In 2015 he has started to work as Junior Researcher at the German Research Center for Artificial Intelligence focusing on crowdsourcing, gamification and persuasive technologies.

References

[1] Maximilian Altmeyer, Pascal Lessel, and Antonio Krüger. 2016. Expense Control : A Gamified , Semi-Automated , Crowd-Based Approach For Receipt Capturing. *Proceedings of the 21th International Conference on Intelligent User Interfaces. ACM, 2016.* (2016), (to appear). DOI: http://dx.doi.org/10.1145/2856767.2856790

[2] Erin Brady, Meredith Morris, Yu Zhong, Samuel White, and Jeffrey Bigham. 2013. Visual Challenges in the Everyday Lives of Blind People. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2013.* (2013), 2117–2126. DOI: http://dx.doi.org/10.1145/2470654.2481291

[3] Sebastian Deterding and Dan Dixon. 2011. From Game Design Elements to Gamefulness: Defining "Gamification". *Proceedings of the 15th International Academic MindTrek Conference. ACM, 2011.* (2011), 9–15. DOI: http://dx.doi.org/10.1145/2181037.2181040

[4] Daniel Esser, Klemens Muthmann, and Daniel Schuster. 2013. Information Extraction Efficiency of Business Documents Captured with Smartphones and Tablets. *Proceedings of the 2013 ACM Symposium on Document Engineering. ACM, 2013.* (2013), 111–114. DOI:http://dx.doi.org/10.1145/2494266.2494302

[5] Rose Holley. 2009. Many Hands Make Light Work : Public Collaborative OCR Text Correction in Australian Historic Newspapers. *National Library of Australia Staff Papers (2009)* March (2009), 1–28.

[6] M.P. Nevetha and A. Baskar. 2015. Applications of Text Detection and its Challenges : A Review. *Proceedings of the Third International Symposium on Women in Computing and Informatics. ACM, 2015.* (2015), 712–721.

[7] James Surowiecki. 2005. *The Wisdom of Crowds*. Anchor.

[8] Takumi Toyama, Daniel Sonntag, Andreas Dengel, Takahiro Matsuda, Masakazu Iwamura, and Koichi Kise. 2014. A Mixed Reality Head-Mounted Text Translation System Using Eye Gaze Input. *Proceedings of the 19th International Conference on Intelligent User Interfaces. ACM, 2014.* (2014), 329–334. DOI: http://dx.doi.org/10.1145/2557500.2557528

[9] Luis Von Ahn, Benjamin Maurer, Colin McMillen, David Abraham, and Manuel Blum. 2008. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. *Science 321.5895* 321, 5895 (2008), 1465–1468. DOI: http://dx.doi.org/10.1126/science.1160379