

MobiMed: Comparing Object Identification Techniques on Smartphones

Andreas Möller,
Stefan Diewald, Luis Roalter
Technische Universität München
Munich, Germany
andreas.moeller@tum.de,
stefan.diewald@tum.de, roalter@tum.de

Matthias Kranz
Luleå University of Technology
Department of Computer Science,
Electrical and Space Engineering
Luleå, Sweden
matthias.kranz@ltu.se

ABSTRACT

With physical mobile interaction techniques, digital devices can make use of real-world objects in order to interact with them. In this paper, we evaluate and compare state-of-the-art interaction methods in an extensive survey with 149 participants and in a lab study with 16 participants regarding efficiency, utility and usability. Besides radio communication and fiducial markers, we consider visual feature recognition, reflecting the latest technical expertise in object identification. We conceived MobiMed, a medication package identifier implementing four interaction paradigms: pointing, scanning, touching and text search.

We identified both measured and perceived advantages and disadvantages of the individual methods and gained fruitful feedback from participants regarding possible use cases for MobiMed. Touching and scanning were evaluated as fastest in the lab study and ranked first in user satisfaction. The strength of visual search is that objects need not be augmented, opening up physical mobile interaction as demonstrated in MobiMed for further fields of application.

Author Keywords

Physical mobile interaction, object identification, touching, pointing, scanning

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

General Terms

Experimentation, Human Factors, Performance

INTRODUCTION

More than a decade ago, the idea of bridging the gap between the virtual and physical world has been coined [22], and later been referred to by the term of *physical mobile interaction*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NordiCHI'12, October 14-17, 2012 Copenhagen, Denmark
Copyright ©2012 ACM 978-1-4503-1482-4/12/10... \$15.00

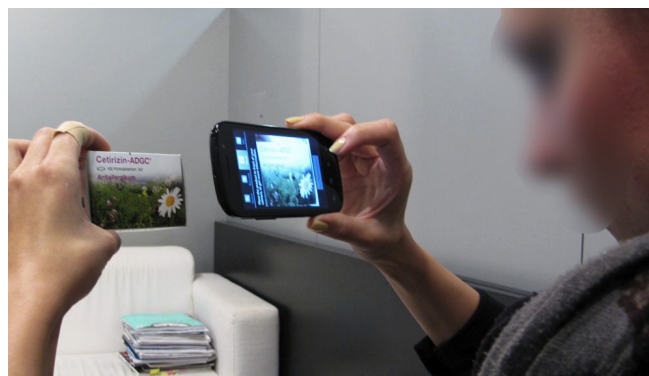


Figure 1. MobiMed identifies drug packages using different methods of physical mobile interaction: pointing, touching, scanning, or text search. For a distinctive definition of the interaction techniques, see the Related Work section. The subject's face was blurred for privacy reasons.

[16]. This concept subsumes any system where digital devices interact with physical objects in their environment [9]. Typical interaction paradigms such as *touching*, *pointing* and *scanning* [21, 17] have been presented, as well as numerous fields of applications in which they have been used and evaluated, such as identifying and remote-controlling devices at home or at the workplace [8, 23], and interacting with displays or posters in public spaces [7, 20, 2].

Today, both quantitative and qualitative results of prior comparisons are valid only with restrictions for several reasons. Applications have been realized with communication technologies which were popular at that earlier time, but are not any more today (such as infrared). Interaction methods were restricted by hardware limitations, e.g. VGA resolution cameras [12] or were entirely custom-made, e.g. with a laser pointer and photo diodes [17] and have not been further pursued. A more detailed discussion of previously presented approaches is provided in the Related Work section.

Meanwhile, smartphones offer high-resolution cameras and plenty of processing power, which opens up new technological possibilities and enables approaches formerly not possible. Earlier approaches had in common that real-world artifacts needed to be augmented with markers or similar in order to be interacted with. Today's fast multi-core smartphones enable vision-based feature recognition, making object recognition in real-time applicable for the *pointing* technique and

function with any object, so that items need not to be augmented any more.

In addition, people have now prior experience with some types of physical interaction, unlike some years ago where researchers reported on subjects without prior knowledge in their studies [10]. Scanning visual codes has e.g. become widely popular, e.g. for product price comparisons. This must be taken into account, since users might prefer techniques they know a-priori, or which seem at first glance more suitable, faster, or more intuitive. This evolution justifies an updated comparative view on object identification techniques.

The contribution of our paper is twofold. First, we present an updated view and comparison of physical mobile interaction techniques, in light of prevalent technologies today and advances of mobile phones in the recent years. Second, we incorporated these techniques into an example application, MobiMed, and evaluate them in a medically motivated use case. We conducted a large online study to get an up-to-date view on real users' opinions on the different techniques and their perceived advantages and disadvantages today. With a smaller user group, we conducted a lab study with a prototype in which we confirmed the large-scale results and compared four techniques regarding their efficiency, utility and perceived usability.

The remainder of the paper is structured as follows. We start with an overview of physical interaction techniques and related work. Subsequently, we introduce the MobiMed application and use case, and explain the implemented object identification methods. After that, the two studies (online survey and lab study) are described and their results are presented and discussed. We conclude with a summary of our findings and a brief outlook.

PHYSICAL INTERACTION TECHNIQUES

Approaches for identifying objects with smartphones and interacting with them are manifold. In accordance with previous research, we distinguish three general paradigms: *pointing*, *touching* and *scanning* [21, 17].

Pointing

Pointing denotes the process of aiming at an object with the smartphone and is considered as a natural technique, since we are used to point at things with our fingers [17]. The pointing paradigm can be implemented by an optical beam (e.g. infrared or laser) or visual codes [12, 17, 21]. Both techniques require the target object to be augmented and a line of sight between mobile device and target object. Wu et al. detected the pointing direction with the phone's accelerometer [23]; however, the position of the target objects in the room must be known. Due to the availability of high-quality cameras in smartphones, the popularity of 1-dimensional (e.g. bar codes) or 2-dimensional visual codes (e.g. quick response (QR) codes) steadily increases. Infrared is meanwhile found rarely in smartphones and thus a less attractive interaction technology to use. With increasing processing power, content-based image retrieval (CBIR, based on visual feature recognition) has become an alternative to identify objects. Föckler et al. [6] present a mobile phone

museum guide that visually recognizes exhibits and provides additional information on them. A system by Nistér et al. [11] finds similar-looking CD covers based on a reference image in real time. However, feature recognition is often more error-prone than explicit visual codes.

Touching

Touching is a proximity-based approach that allows to identify an object by bringing the phone close to it. Objects are augmented with electronic tags based on radio communication [22], such as RFID (radio frequency identification) or NFC (near field communication), which can be read with a capable smartphone in a distance of few centimeters. The touching modality has e.g. been used for interaction with posters [20] and public displays [7] with a grid of NFC tags. Benelli et al. [3] presented scenarios for the medical area with usage of mobile devices. They propose NFC-based access control for work hour tracking and access control in hospitals, rapid exchange of patient information and case histories.

Scanning

In accordance with a definition by O'Neill et al. [12], scanning is a proximity-based approach where the user points with a camera at a visual tag (cf. 'to scan a bar code'). While the touching technique requires the target to be in direct proximity, scanning works for objects in close proximity (as far as the visual code is readable).

Scanning can also be understood as searching for available services in an environment by wireless techniques such as Bluetooth or WLAN [21, 17]. In our work, however, we use the term of *scanning* in the sense of targeting a fiducial marker.

Comparative Work

Rukzio et al. [17] compared three mobile physical interaction techniques regarding performance, cognitive load, physical effort and other factors: Objects could be touched (using NFC tags and a NFC-enabled phone), scanned (using Bluetooth access points and a corresponding mapping of nearby objects), or pointed at with a light beam (using a laser pointer attached to a smartphone and a light sensor attached to the object). A further factor of distinction was the distance to the objects to interact with: scanning was preferred for more distant objects, pointing for objects in intermediate range, and touching for close objects. The authors could not give universal recommendations, since questionnaires and experiments showed that the preferred technique depends on location, activity and motivation.

In a trial by O'Neill et al. [12] where untrained users interacted with posters using two-dimensional fiducial markers and RFID tags, the visual QR codes were read faster. Another study, conducted by Mäkelä et al. [10], investigated usability of RFID and 2D bar codes and revealed that users have a limited understanding of both interaction techniques. Visual codes vary in their dimension, layout, size and data density, which influences reading speed [13]. Higher data density also places higher requirements on camera and processing power. Ailisto et al. [1] reviewed visual codes, infrared, RFID and

Bluetooth as base technologies for physical selection. They pointed out that infrared has the advantage of bidirectional communication and mentioned the wide support in devices as benefit of this method, compared to others. Due to the progress in mobile hardware, these results do not apply today any more.

Reischach et al. [14] proposed the usage of physical mobile interaction for product identification and demonstrated an advantage of NFC and barcode identification over conventional product search in dimensions of task completion time and perceived ease of use.

Physical Mobile Interaction vs. Object Identification

Our comparison of techniques is motivated by physical mobile interaction as understood by e.g. Valkynnen et al. [21] or Rukzio et al [17]. However, in order to interact with a physical object, we need to *identify* it in the first place. This includes the aspect of recognizing its presence as well as distinguishing it from other, similar objects. Hence, we generalize the concept of interaction techniques and speak of *object identification techniques* in the remainder of this paper. We investigate not universal, but *domain-specific* object identification applicable for our scenario described in the following section. Our ambition is to recognize objects, such as drug packages, with state-of-the-art smartphones in reasonable time (i.e., an order of magnitude of few seconds).

MOBIMED: IDENTIFYING DRUGS WITH A SMARTPHONE

In times of food supplements and vitamin compounds, and with regard to the aging society, managing multiple drugs is an issue people are struggling with. We observe an increase of health-related apps and services, such as drug reference guides¹, pill reminders or identifiers² and food ingredient databases³. This development motivated us to combine personal medication support with physical mobile interaction as use case for our evaluation. We conceived MobiMed, a system for identifying medication packages as example application to evaluate object identification techniques on the smartphone. MobiMed can be imagined as digital package insert replacement, applicable e.g. when the original package insert is lost, or for quickly comparing medications. It provides detailed information on drugs, such as active ingredients, application areas, intake instructions and side effects.

Use Case

In order to evaluate MobiMed, it is important for subjects to understand how the application is used in a specific scenario. We conceived the following use case, which we presented to study participants, to motivate MobiMed.

John, 45, architect, is an active golfer and likes cycling tours in his holidays. He regularly takes food supplements (carotenes, Vitamin E) and anticoagulants since he is a cardiac patient. He needs to take up to four different medications a day in different intervals, why he is sometimes not sure about the correct dosage. Since he

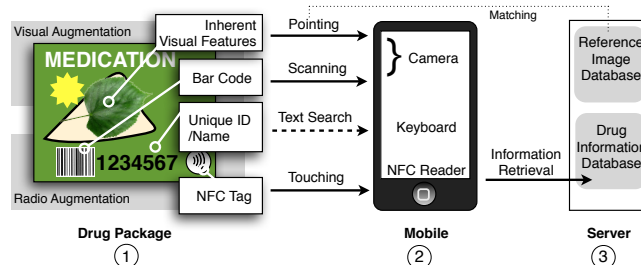


Figure 2. Principle of drug identification with MobiMed. Drug packages have bar codes, a unique identification number, inherent visual features and are augmented with NFC tags (Step 1). By pointing at a package or scanning the bar code (using the portable device's camera), the package is identified visually. By touching, the package's NFC tag is read. Additionally, the drug can be searched by its name or unique ID (Step 2). Once the package has been identified, drug information is retrieved from a database and displayed on the phone (Step 3).

had trouble pulling the blister packages out of the box, he removed the package inserts and cannot refer to the instructions. John points at the drug package with MobiMed which identifies the medication by the appearance of the box. He gets detailed information on the product and scrolls to the correct dosage instruction.

The other day, John wants to get an influenza medication at the pharmacy. Since he is allergic to acetaminophen, he scans the package with MobiMed and checks whether the product contains the substance in order to know if he can safely buy it.

Prototype Implementation

The interplay between drug packages, smartphone and the MobiMed server is illustrated in Fig. 2. The package's bar code or NFC tag and inherent visual features allow its identification by four methods: *pointing*, *scanning*, *touching* and *text search*. For the first three methods, the phone's camera and NFC reader are used. Text input is performed on the phone's soft keyboard (see Fig. 3, left image). The user can switch between these identification methods in the MobiMed application by four tabs at the top of the screen (see Fig. 3).

MobiMed was implemented in Android 2.3 using the default NFC and Camera APIs. For the *scanning* modality, the ZXing Barcode reader library⁴ was used.

The drug information database and ~100,000 reference images for feature recognition are stored on a server, where also the feature extraction is performed. Our database currently contains entries on more than 47,000 drugs, making MobiMed's product search an extensive real-world scenario. Drug information as well as reference images were retrieved from online pharmacy websites. The individual identification methods work as described in the following.

Scanning

In many countries, medical products have a unique identification number, such as the NDC (National Drug Code) in the US, the PZN (Pharmazentralnummer) in Germany, or the

¹WebMD. <http://www.webmd.com/mobile>

²Drugs.com. <http://www.drugs.com/apps>

³FDDB. <http://fddb.info>

⁴ZXing. <http://code.google.com/p/zxing/>

PPN (Pharmacy Product Number) which will be introduced in the European Union and be encoded in a two-dimensional DataMatrix code on every product package. Currently, the bar code on each drug package holds a 7-digit unique PZN and can hence be used to unambiguously recognize a product. The bar code is targeted (scanned) with the phone's camera at a distance of few centimeters. Upon recognition of the code, the name of the medication appears in a popup and the user is asked whether it is the searched one. After confirmation, the details page for the medication is displayed (see Fig. 3, right image). The recognition starts immediately with the camera preview screen, it is not necessary to take a photo.

Touching

For this method, medication packages were enhanced with NFC tags on which we stored the same unique PZN as in the bar code. We used MIFARE Ultralight tags operating at 13.56 MHz (NFC Forum Type 2) with 48 bytes of memory, complying with the ISO 14443A standard. Touching a drug package with a NFC-capable phone reads the PZN stored in the tag and shows the drug's information page.

Pointing

Each drug package has inherent visual features, such as logos, imagery, colors, package shape and imprinted text. These characteristics can serve for identification, which is referred to as content-based image retrieval (CBIR). We use an extension to the vocabulary tree approach [19] with a training set of 100,000+ images. At a distance of 20–50 centimeters, the user points at the drug package with the camera, which continuously captures several query images. A visual feature tracker identifies characteristic features (maximally stable extremal regions, MSER) and matches them with reference images in the database. Moving the camera while pointing is unproblematic; varying perspectives in the query images even improve matching. The result is a candidate list of potential matches from which the user selects the desired medication. The correct hit is usually amongst the first page of results.

Text Search

In this modality, a free text search is performed on all database fields, so that drugs can be found based on their PZN as well as by their name, active ingredients, side effects and so on. A list of search results that match the specified search term is presented to the user.

Omission of Speech Recognition

Speech recognition on mobile devices is experiencing a revival with Apple's Siri and Google Now. We did not include voice-based interaction as modality in our example scenario for various reasons. Medication names are often hard to pronounce or similar-sounding, which is still problematic for speech recognition engines. Often, there exist variants with different dosage but similar name, which entails a high risk of inappropriate results. For these reasons, we considered speech input not as appropriate for identifying drugs.

ONLINE SURVEY

At an early stage of the development of MobiMed, we conducted a large-scale survey with 149 participants to investigate the following research questions (RQ):

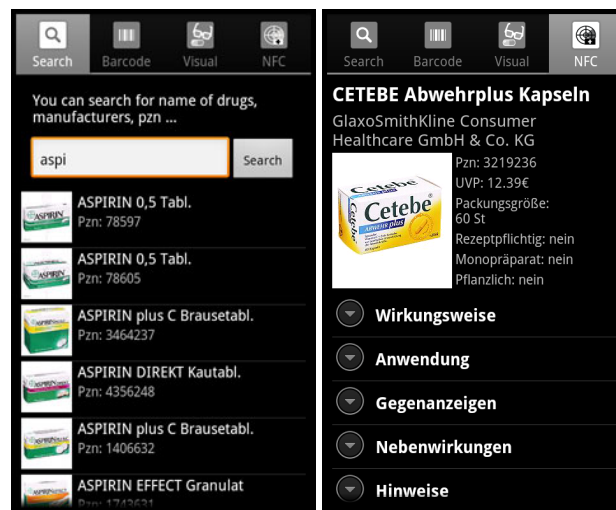


Figure 3. Left: Text search for medications in MobiMed. Right: The result screen with detailed drug information. The top of the screen shows tabs for switching between the identification methods.

- RQ1: What advantages and disadvantages of identification techniques as presented in MobiMed do people see?
- RQ2: Which method is preferred by users a priori?
- RQ3: What potential do people see for MobiMed as a whole?

The gathered responses are *a priori* feedback, i.e. user's opinions are reflected based on textual descriptions of MobiMed and the incorporated methods, without having used it by themselves.

Questionnaire

Since our goal here was broad feedback, and a lab study with a number of 100+ participants would not have been feasible, the survey was conducted as online questionnaire. At the beginning, the concept of MobiMed, the use case and the four identification methods were introduced and illustrated with descriptive screenshots. By explaining in detail for each method how the process of identifying a package works, we tried to give subjects a best possible idea of the system without actually using it.

Then followed two sets of questions, addressing the individual research questions. The first set of questions addressed the advantages and disadvantages of MobiMed's identification techniques, and of MobiMed as a whole. In the second block, we asked whether subjects would use the app themselves, how much money they would spend for it, and which other functions they would like to see in it. All questions were asked in a neutral manner (e.g. explicitly asking for both advantages and disadvantages) to exclude a confirmation bias.

Participants

The questionnaire was uploaded as Human Intelligence Task (HIT) to Amazon Mechanical Turk⁵. Completion of a questionnaire was compensated with \$0.30. Excellent answers in

⁵Mechanical Turk. <https://www.mturk.com>

length and quality were rewarded with a bonus of an additional \$0.50. 149 participants worldwide (thereof 100 living in the US) aged between 17 and 79 (average age: 31 years, SD = 11) took part in the survey; 74 were females, 75 were males. 110 participants owned a smartphone.

Results and Discussion of the Online Survey

RQ1: Advantages/Disadvantages of Identification Techniques

We collected a variety of statements on the presented methods, giving an impression what people did and did not like. Items were formulated as open questions with free text answers. Hence, results reflect main tendencies and opinions, but cannot expose the spectrum of answers in its entirety.

Scanning

People considered this method as ‘quick and convenient’ way to identify packages. Further attributes mentioned repeatedly were ‘precise’, ‘easy’, ‘specific’ and ‘cool’. Respondents particularly liked that ‘you can know exactly that it is the right product’, since the bar code uniquely identifies it, even if brand names or packages look similar. There was much familiarity with this technique due to market penetration of bar code scanner apps. Some people reported to ‘scan products all the time’ with their phones when shopping for price comparisons. This a priori knowledge might have biased participants’ responses towards ‘scanning’.

As drawbacks, people mentioned the necessity for a camera and a readable bar code. Furthermore, it takes time to find the bar code on the product. It could happen that it is too small to be recognized (due to lightning conditions, a damaged package, or a weak camera). Usability problems were reported more frequently for bar code scanning than for other methods, perhaps because people have tried this method already for themselves. Statements such as ‘doesn’t always work for me’ and ‘sometimes hard to focus the bar code’ indicate that users have experienced problems on their own, which makes them perceive scanning more difficult than other methods.

Touching

NFC tags were affirmed to allow fast and precise identification, combined with good usability. It was highlighted that touching the object with the phone is the only necessary user interaction, which makes the method, according to respondents, ‘hassle-free’, ‘fool-proof’ and suitable for ‘old people and non-expert persons’. Other adjectives used were ‘modern’, ‘cool’, and ‘satisfying to [...] get so much information so quickly’.

Downsides mentioned were the ‘extensive’ requirements: a NFC-capable phone, augmented medication packages, and an increased energy consumption on the phone. People mentioned potential error-proneness of NFC, being a novel technology, which indicates that they were less familiar with NFC and thus more skeptical, compared to bar codes. Other disadvantages addressed costs (NFC augmentation would raise drug prices⁶) and privacy concerns due to the radio technology. People here seemed to overestimate the proximity range

of NFC (which is actually only few centimeters), since also ‘interferences with other packages around’ were listed as potential problem.

Pointing

Respondents found feature recognition simple, convenient and easy to use. In accordance with earlier findings [21], pointing at objects was considered as very intuitive. One person said that it is ‘the most human form of scanning’. Respondents appreciated that searching for details on the package, such as tags or codes, is not necessary with visual recognition: ‘No need to fumble about looking for the bar code on the product’. Instead, you ‘could scan from any angle’. People recognized that also images from websites could be identified; one would not have to hold the product in their hands. It was mentioned several times that damaged packages could be recognized as well (but this was not investigated in this research). A person added that it would be ‘excellent [...] for persons with sight or motor skill disabilities’.

On the negative side, people mentioned high processing demand and potentially slower recognition, compared to the other methods. Respondents recognized that visual search, unlike explicit (fiducial or radio) tags, cannot provide unambiguous identification. They supposed that drugs could be confused ‘by slight deviations from standard packaging, e.g. a pack with a sticker saying 2 for 1’, or ‘if companies produce packaging designs that are really similar’.

However, this method also works with ‘untagged’ objects. Its inherent ambiguity can be a strength for searches for similar products, such as different package sizes of the same brand, other dosage forms (powder instead of pills), etc. A person said: ‘Although this may not get your specific product, it can identify similar products. That’s incredibly helpful.’ By presenting a list of possible results, the ambiguity is made explicit and the responsibility is given to the user.

Text Search

Text search was attested the highest familiarity due to daily use like e.g. in search engines. People liked its inherent accurateness and its multi-functionality, allowing to search for other keywords than the product name. A respondent accentuated that ‘general terms’ could be used for search, and a ‘broad range of results’ would show up. Another said that you can ‘find products of the same category, and [...] make a comparison among them’ (although this was not the goal of the app). In particular, if subjects do not know what drug they are looking for exactly, they consider text search as a good method. An important point mentioned was that the method is independent of sensors and drug package (‘you don’t need to be near the product’), as long as the name is known. This was considered as advantage not only in case of recognition failures, but also in light of the fact that in some countries (including the US), pills are often handed to patients without original packaging. Respondents also came up with other usage scenarios, e.g.: ‘Say you’re allergic to acetaminophen you can see what drugs contain it to know what to avoid’. While this goes beyond the original task of identifying drugs, they are nevertheless an interesting example of what MobiMed could additionally be used for.

⁶actually, the mass market price of RFID tags is <0.15\$/piece, see e.g. the RFID Tag Pricing Guide at www.odinrfid.com

Drawbacks mentioned were the high amount of interaction with the device (i.e., the necessity to type), longer search time and potential misspelling (which is likely for complicated drug names and medical terms). Not only the problem of no results due to typos was seen, but also that ‘incorrect wording [...] could end up giving people information on the wrong medicine’. Also, similar items in the result list could lead to confusion. Text search was considered more difficult than the other methods especially due to small on-screen keyboards on smartphones.

RQ2: Preferences

When asked which method respondents would prefer, 48.3% chose *scanning*, 25.3% *text search*, 13.6% *touching*, and 12.2% *pointing* (see Fig. 4).

User Preferences for Methods in Online Survey

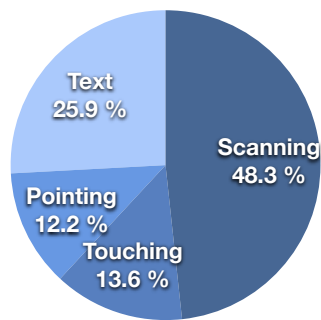


Figure 4. Preferences for identification methods in the online survey. Respondents liked scanning most (probably due to higher familiarity than NFC or visual search) followed by text search, pointing and touching.

The high preference for scanning and text search could be explained with their level of familiarity. Most smartphone users are experienced with text search and bar code scanning, while they are less familiar with NFC and visual feature recognition. ‘I have experience using other software using barcodes, and have liked their ease.’, one person stated. A respondent who chose text search said: ‘This is a tried and true way of researching information’. Previous knowledge and positive experience may have attracted respondents to choose a ‘familiar’ method as their favorite.

It seemed more difficult for people to evaluate pointing (visual search) and touching (NFC) without hands-on experience. In particular, recognizing objects just by pointing was partly seen as ‘science-fiction’, and the range of NFC was overestimated, leading to the assumption that closely placed products could interfere with each other and make targeting the desired one difficult. Some users even worried about being inundated with information when passing by the shelves in a store without having a hand in the matter.

RQ3: Potential of MobiMed

Respondents highly appreciated MobiMed. 81.8% answered ‘yes’ to the question ‘Would you use such a system as described above?’. They liked the idea of finding drug information fast and easily, and envisaged various focus groups who could benefit from MobiMed, such as pharmacists, doctors,

or people who take multiple drugs. A person said it’s ‘perfect for if you have something at home that you want to find somewhere so you can pick more up or learn more about it’.

In average, interviewees would spend \$8.40 for MobiMed (with a standard deviation of \$17.12). The high variance is rooted in the difference between older and younger respondents: Those under 25 would averagely pay \$6.34, those older than 25 in average \$14.01. There are two possible reasons: First, older people might have a higher need for medical applications, so that the personal value is higher for them. Second, they have a different idea of product pricing than younger people who are used to get software in mobile app stores for small amounts of money.

LAB STUDY (PROTOTYPE EVALUATION)

After having collected a priori large-scale feedback on identification methods in the MobiMed use case, we wanted to verify our findings with a smaller number of participants in a hands-on lab study. We investigated the following research questions (RQ):

- RQ1: Which object identification method is superior in terms of efficiency?
- RQ2: Which method is preferred by users in practical use of the MobiMed application?
- RQ3: What potential do people see for MobiMed as a whole after having used the application?

RQ1 was evaluated in an experiment, RQ2 and RQ3 with a questionnaire after the experiment. The lab study’s research questions corresponded with those of the online survey (see previous section) and confirmed the initial results. While the online survey revealed a priori findings (i.e. before subjects actually used the application) of a large user group, this time the research questions were investigated quantitatively and qualitatively in an experiment in a smaller scale.

Participants

16 people (6 females, 10 males), aged between 22 and 69 years (average age: 31, SD = 12) took part in the evaluation. All of them had a mobile phone and used it regularly, 9 were smartphone owners. No subject had physical disabilities that could have hindered the execution of the demanded tasks (such as difficulties with holding the smartphone steadily). Participants were recruited among acquaintances of some of the researchers; none were involved in the project. They were rewarded for their participation with a small compensation in the form of sweets.

Experimental Task

Subjects identified medication packages using the four methods described above: *scanning*, *touching*, *pointing* and *text search* (conditions). Each participant ran through all conditions (within-subjects design) in one of four orders, which were counterbalanced using a 4×4 Latin Square [5].

The experiments were conducted at a table in a separated, brightly lit room at a medical office. 13 NFC-augmented medication packages (see Fig. 5) were placed in a box on



Figure 5. Drugs packages like they were used in the experiment. All packages contain bar codes; our sample packages for the study were additionally augmented with NFC tags for the *touching* method.

the table. Subjects were handed a Samsung Nexus S smartphone with MobiMed installed. They were introduced to the device and the task, and could make themselves familiar with the phone and the MobiMed app prior to the experiment.

For each of the four conditions, 10 packages had to be identified. Participants were asked to fetch one package at a time out of the box (blind draws; the order was randomized) and to identify it with MobiMed. After successful identification of 10 packages, they were put back in the box and the condition was changed. In the *text search* condition, users were free to either type drug name or identification number (PZN) which was printed below the bar code on each package.

During the experiment, subjects were encouraged to express any thoughts that came to their mind ('think aloud'). Fig. 6 shows subjects performing different identification methods. The experiment took about 30 minutes per participant. A researcher was present during the entire time.

Data Collection

Quantitative Data

A timer built in to the application (in millisecond resolution) measured the time needed for each single identification. The measurements were saved to a log file on the smartphone. Subjects were asked to tap a button on the screen when they were ready, which started the recognition process. At this moment, the timer was initialized, and it was stopped when the identification process was complete and the drug's description page was shown.

For touching and scanning, the measured time was equivalent to the recognition time of the NFC tag or the bar code. For pointing and text search, recognition led to a result list, since these methods can return ambiguous hits. In these cases, the total time added up of the recognition time and the user selection time of the correct list item. The timer was always stopped upon the first appearance of the drug's description screen, i.e. it was assumed that no corrections of the user's choice from the list were required.



Figure 6. Subjects (faces were blurred for privacy reasons) identifying drug packages with MobiMed in the lab study: Bar code scanning (1), touching (2), pointing (3) and checking results after text search (4).

Qualitative Data

Qualitative feedback was gathered in a questionnaire after the experimental task (addressing RQ2 and RQ3). Subjects were left alone while they filled out the form to prevent any influence of a researcher present in the room.

In the questionnaire, people were asked how they liked each identification method and whether they would continue using the app with each particular method. Moreover, utility and usability of the MobiMed app were evaluated.

Results of the Lab Study

RQ1: Experimental Comparison of Methods

A summary of the measurements is shown in Fig. 7. Subjects needed 1.8 seconds (all figures are mean values) to identify a single drug by touching the NFC tag. This was significantly faster than scanning the bar code (13.5 s), pointing at the object (16.4 s) and using text search (20.5 s).

The standard deviations (SD) were 3.7 s for *touching*, 9.6 s for *scanning*, 6.1 s for *pointing* and 22.9 s for *text search*. Although scanning was performed slightly faster than pointing, the higher variance for scanning and some large single values (above 40 s, maximum at 61.7 s) indicate that subjects sometimes struggled with this identification method. In the experiments, this became especially evident when the camera had focusing problems with small bar codes.

While the upper quartile of the NFC identification time ends at 3.9 s, some subjects took more than 10 s for the task, with a maximum of 18.1 s. We explain this variation with different proceedings we observed in the experiment. While most participants focused the NFC tag on the package prior to pressing the start button on the phone, few others pressed the button and began afterwards to turn the package in their hands to search where to touch it with the NFC reader.

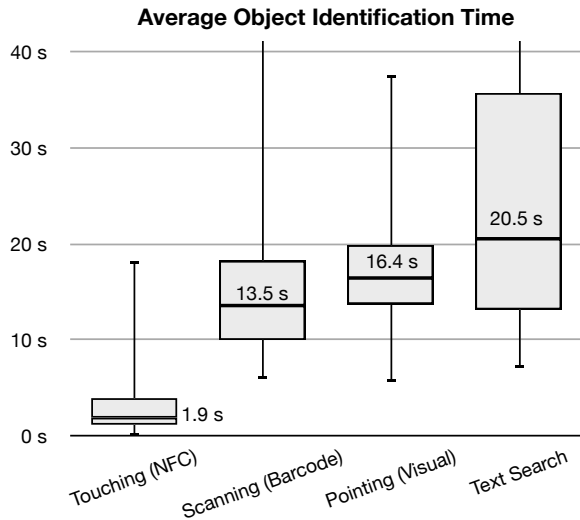


Figure 7. Box plots indicating speed measures of object identification with MobiMed (figures are means). Boxes represent the interquartile range (IQR), whiskers represent extrema. Maxima for scanning (61.7 s) and text search (102.3 s) were cut off for better readability.

The high variance for text search reflects the diverging typing capabilities of participants. With a maximum of 102.3 s, text search took more than five times longer than the longest NFC identification (18.1 s). It could also be explained by the length of some drug names, providing no upper bound for text input length. It is worth to mention that these results were still obtained under ‘ideal’ conditions, i.e. under the assumption that subjects selected the correct item from the result list. In practice, the need to correct accidental choices might entail even longer total times for the *text search* modality.

RQ2: Qualitative Comparison of Methods

Fig. 8 visualizes users’ preferences for the individual methods in the experiment. The ratings reflect the measured identification times, i.e. faster methods were rated better. On a 7-step Likert scale ($-3 =$ strongly disagree, $3 =$ strongly agree), participants responded with 3.0 that they liked touching ($SD = 0.0$), which was also the fastest method. Scanning, the second fastest method, was rated with 2.0 ($SD = 1.2$). Pointing received a rating of 1.4 ($SD = 1.9$) and text input was rated averagely with 0.2 ($SD = 2.0$). When asked whether users would use the methods *in future*, the pattern was generally the same, but at a lower agreement level. Subjects agreed with 2.8 ($SD = 0.4$) that they would like to use the touching method. The scores for scanning and pointing were 1.9 ($SD = 1.0$) and 0.6 ($SD = 2.0$), respectively. The wish to use text search was expressed with only -1.1 ($SD = 1.8$).

RQ3: Utility and Usability of MobiMed as a Whole

In order to evaluate the perceived usability and general acceptance of MobiMed, we used the System Usability Scale (SUS) [4] which contains a 10-item set using a 5-step Likert scale ($1 =$ strongly disagree, $5 =$ strongly agree). The set is a mixture of positive and negative statements, such as ‘I thought the system was easy to use’ (positive), ‘I felt very confident using the system’ (positive) or ‘I thought there was too much inconsistency in this system’ (negative). A SUS

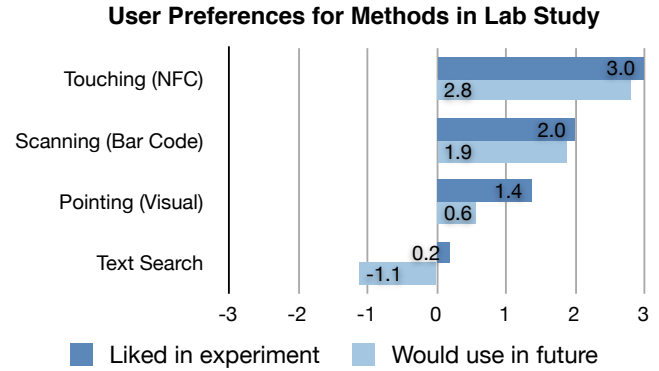


Figure 8. Likert scale averages on whether users liked a specific identification method and whether they would use this method in the future ($-3 =$ strongly disagree, $3 =$ strongly agree).

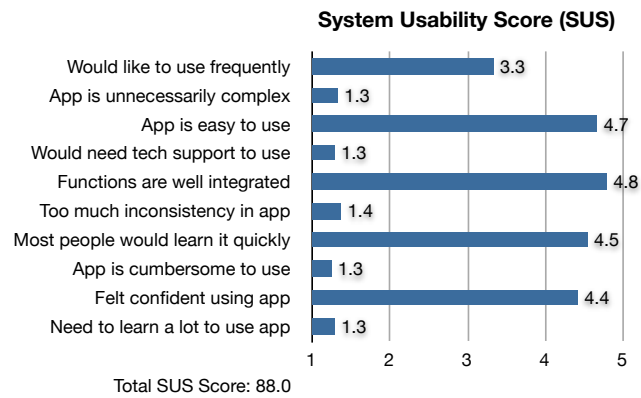


Figure 9. Individual aspects of the System Usability Score for MobiMed on a 5-step Likert scale ($1 =$ lowest, $5 =$ highest agreement with statements). Statements are abbreviated; for exact wording, see [4].

score ranging from 0 to 100 is then calculated from the individual items (negative statements are inversely incorporated).

Fig. 9 illustrates the SUS values for the individual items. Positive items were usually rated with a score of more than 4, except the statement ‘I think that I would like to use this system frequently’, which was rated with 3.3. Subjects averagely disagreed with negative items with a score of less than 1.4. Only one participant stated to ‘need support of a technical person’ to use the system. The resulting SUS score is 88.0 points out of 100 possible points, which can be considered as a good result [18] so that we can assume that subjects appreciated the MobiMed app and did not find major usability concerns.

In order to evaluate MobiMed’s utility in everyday life, we asked people which information sources they use to learn about medications, their active ingredients, dosage, side effects etc. 75% ask their doctor or pharmacist, 69% read the package insert. 56% stated to consult books or the internet, 13% use other sources. 75% of subjects use more than one single source to get information on drugs. After the study, 14 out of 16 participants (88%) declared that they were interested to use MobiMed as alternative source to inform themselves on drugs. Not only is this an indicator that subjects

appreciated the prototype; MobiMed was also the most popular individual source of information of all other ones.

Additional Feature Suggestions

With an open question, we asked subjects for desired additional features in MobiMed. They came up first and foremost with shopping-related functions: lookup of prices, finding cheaper generic products, providing a list of suppliers, and the possibility for direct order. Subjects were also interested in active ingredient analysis: MobiMed could suggest products that show fewer cross-correlated side effects for a specific combination of components. One participant suggested a tool that helps diagnosing based on symptoms a user enters. Several subjects were interested in a personalized medication management tool that allows to manage drug intake, creates medication lists and reminds users to take their pills.

DISCUSSION AND SUMMARY

We investigated advantages and disadvantages, efficiency and user acceptance of four object identification techniques at the example of a drug information system (RQ1 and RQ2), and evaluated utility and usability of MobiMed, a prototype implementation thereof (RQ3). We examined the research questions in two steps: qualitatively with a 149 users in an online survey, and both quantitatively and qualitatively with 16 subjects in a lab study.

We summarize our findings and observations in five main points. Results from both studies coincide in large parts, although we found some divergences which we also try to explain below.

1. Physical Mobile Interaction is Popular and Efficient

In the lab study, subjects preferred all physical mobile interaction techniques over conventional text search, and would also prefer them in future use. Quantitative measures showed that physical mobile interaction techniques were faster for identifying drug packages than text search, and that the standard deviation was significantly lower. Today, physical mobile interaction is widely still neglected compared to traditional input approaches, although all requirements to the hardware are now met (which they have not even few years ago). Our findings motivate to foster the use of object identification techniques in everyday applications, making the interaction with the physical world more effective and intuitive.

2. Touching and Scanning Evaluated Best

In the comparison of the individual techniques, *touching* and *scanning* were evaluated as fastest methods, and also preferred by study participants. *Touching* was significantly faster than scanning, and also adopted more positively by subjects in the lab study.

In the online survey people responded differently: almost half of respondents stated that they liked *scanning* most, followed by *text search*. A reason for this high agreement might simply be that they knew these due to their wide popularity, while they could not imagine so well how touching using NFC or pointing using visual search would work. Lab study participants who tried all modalities might have been guided by their actual experience rather than by their previous knowledge.

It remains to be seen how the further distribution of NFC on the market, e.g. through digital payment services such as Google Wallet, will influence people's attitude towards this interaction technique. The practical use for identifying everyday objects with NFC is currently not yet given since few products are augmented with tags yet. However, a widespread use of NFC to improve customer services has been suggested previously [15], and could be promoted by cheap tags, as they are already integrated in clothing.

3. Visual Search as Interesting Alternative for Future Systems

The *pointing* paradigm, which we realized by visual feature-based search, was in our study less popular and slightly slower than touching and scanning. However, pointing almost reached the performance of *scanning*, and had no outliers like *scanning* due to unreadable bar codes. We believe that future hardware and improved implementations could further increase recognition speed and reliability of visual search, and outperform scanning of visual codes in terms of both speed and usability. The major advantage of visual search is that it can work for any object, which opens up this method for a variety of applications beyond identifying marker-augmented products. In the online survey, subjects recognized this potential, mentioning that there is no need to search for the bar code and that it is the most natural way of physical mobile interaction.

4. Best Method Depends on Scenario

We had to confine to a selected task in the lab study in order to compare the performance of four identification methods, but subjects were creative about alternative use cases of MobiMed. Mentioned were, among others, product comparisons in the pharmacy, online price checks, a medication diary and pill reminder systems that incorporate side effects of the combinations of active ingredients. It became clear that the 'best' interaction method depends on the selected scenario. A particular example is the question whether a method should return unambiguous results or a result list. For product comparisons, multiple results as produced by visual search are desired, while reliable information for drug intake at home requires a method that provides a unique result, e.g. scanning the bar code. Visual search can be an interesting alternative when the bar code is not readable or invisible, e.g. when drug packages are placed behind glass in the pharmacy, or information should be retrieved from a picture, e.g. on a website. NFC tags are nowadays not in widespread use, but were included in the comparison as they are a state-of-the-art technology on the rise. Both the short identification time and the compelling user feedback of *pointing* using NFC imply that this method could be an interesting alternative in cases where objects can be augmented with respective tags.

It also turned out that *performance* is only one factor for user preferences, which suggests future research on factors that influence the likability of physical mobile interaction techniques for a specific scenario.

5. General Demand for Medical Apps

Responses of both the online survey and the lab study a high level of interest in medical apps. This might be related to an

increased awareness for a healthy lifestyle (which includes interest in food supplements and ingredients), but also to a rising need for medication support in light of the aging society. Our studies indicate a high demand for support in managing and retrieving information on drugs, their dosage, effects and interactions, which can be complicated for lay people. MobiMed was evaluated by subjects as helpful complement to other information sources the pharmacist's advice.

OUTLOOK AND FUTURE WORK

In future work, we will obtain more detailed data w.r.t. human and system performance. The results of this work showed that performance and user preference not necessarily coincide. They bring up the new research question whether a more 'likable' interaction technique is probably an argument against a performant, but unpopular method, and what actually accounts for pleasant physical mobile interaction. We will further investigate other scenarios of physical mobile interaction in order to gain insights on adequate identification methods for individual use cases. In particular, we will refine visual feature recognition for the use of the *pointing* modality to evaluate in which cases it could possibly replace scanning of bar codes.

ACKNOWLEDGMENTS

We would like to thank Hans Heid for his valuable work with the MedScan prototype. Furthermore, we thank all online and lab study subjects for their participation.

REFERENCES

1. Ailisto, H., Korhonen, I., Plomp, J., Pohjanheimo, L., and Strömmer, E. Realising physical selection for mobile devices. In *Mobile HCI 2003 Conference, Physical Interaction (PI03) Workshop on Real World User Interfaces, Udine, Italy* (2003).
2. Ballagas, R., Borchers, J., Rohs, M., and Sheridan, J. G. The smart phone: A ubiquitous input device. *IEEE Pervasive Computing* 5, 1 (Jan. 2006), 70–77.
3. Benelli, G., and Pozzebon, A. Near field communication and health: Turning a mobile phone into an interactive multipurpose assistant in healthcare scenarios. *Biomedical Engineering Systems and Technologies* (2010), 356–368.
4. Brooke, J. SUS—a quick and dirty usability scale. *Usability evaluation in industry* 189 (1996), 194.
5. Cochran, W., and Cox, G. *Experimental designs*. John Wiley & Sons, 1957.
6. Föckler, P., Zeidler, T., Brombach, B., Bruns, E., and Bimber, O. Phoneguide: museum guidance supported by on-device object recognition on mobile phones. In *Proceedings of the 4th international conference on Mobile and ubiquitous multimedia, MUM '05*, ACM (New York, NY, USA, 2005), 3–10.
7. Hardy, R., and Rukzio, E. Touch & interact: touch-based interaction of mobile phones with displays. In *Proc. of the 10th Intl. Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '08*, ACM (NY, USA, 2008), 245–254.
8. Kranz, M., Möller, A., Hammerla, N., Diewald, S., Plötz, T., Olivier, P., and Roalter, L. The mobile fitness coach: Towards individualized skill assessment using personalized mobile devices. *Pervasive and Mobile Computing, Special Issue on Mobile Interaction with the Real World* (2012), DOI=10.1016/j.pmcj.2012.06.002.
9. Kranz, M., and Schmidt, A. Prototyping Smart Objects for Ubiquitous Computing. In *Proceedings of the International Workshop on Smart Object Systems in Conjunction with the Seventh International Conference on Ubiquitous Computing* (Sept. 2005).
10. Mäkelä, K., Belt, S., Greenblatt, D., and Häkkinen, J. Mobile interaction with visual and rfid tags: a field study on user perceptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '07*, ACM (New York, NY, USA, 2007), 991–994.
11. Nistér, D., and Stewenius, H. Scalable recognition with a vocabulary tree. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, IEEE (2006), 2161–2168.
12. O'Neill, E., Thompson, P., Garzonis, S., and Warr, A. Reach out and touch: using NFC and 2D barcodes for service discovery and interaction with mobile devices. *Pervasive Computing* (2007), 19–36.
13. Querini, M., Grillo, A., Italiano, G. F., and Lentini, A. 2D Color Barcodes for Mobile Phones. In *International Journal of Computer Science & Applications, Special issue on Multimedia Applications*, vol. 8 (2011), 136–155.
14. Reischach, F., Michahelles, F., Guinard, D., Adelman, R., Fleisch, E., and Schmidt, A. An evaluation of product identification techniques for mobile phones. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I, INTERACT '09*, Springer-Verlag (Berlin, Heidelberg, 2009), 804–816.
15. Resatsch, F., Karpischek, S., Sandner, U., and Hamacher, S. Mobile sales assistant: NFC for retailers. In *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services, MobileHCI '07*, ACM (New York, NY, USA, 2007), 313–316.
16. Rukzio, E. *Physical mobile interactions: Mobile devices as pervasive mediators for interactions with the real world*. PhD thesis, 2006.
17. Rukzio, E., Leichtenstern, K., Callaghan, V., Holleis, P., Schmidt, A., and Chin, J. An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. *UbiComp* (2006), 87–104.
18. Sauro, J. Measuring usability with the system usability scale (SUS). <http://www.measuringusability.com/sus.php>, 2011.
19. Schroth, G., Al-Nuaimi, A., Huitl, R., Schweiger, F., and Steinbach, E. Rapid image retrieval for mobile location recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE (2011), 2320–2323.
20. Siorpaes, S., Broll, G., Paolucci, M., Rukzio, E., Hamard, J., Wagner, M., and Schmidt, A. Mobile interaction with the internet of things. In *Adjunct Proc. of Pervasive*, Citeseer (2006).
21. Välikynen, P., Korhonen, I., Plomp, J., Tuomisto, T., Cluitmans, L., Ailisto, H., and Seppä, H. A user interaction paradigm for physical browsing and near-object control based on tags. In *Mobile HCI, Physical Interaction Workshop on Real World User Interfaces* (Udine, Italy, 2003), 31–34.
22. Want, R., Fishkin, K., Gujar, A., and Harrison, B. Bridging physical and virtual worlds with electronic tags. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, ACM (1999), 370–377.
23. Wu, J., Pan, G., Zhang, D., Li, S., and Wu, Z. MagicPhone: pointing & interacting. In *Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing, Ubicomp '10*, ACM (New York, NY, USA, 2010), 451–452.