Towards a Natural User Interface to Support People with Visual Impairments in Detecting Colors.

Sergio Mascetti[1,2], Chiara Rossetti[1], Andrea Gerino[1,2],
Cristian Bernareggi[1,2], Lorenzo Picinali[3], and Alessandro Rizzi[1]

[1] Università degli Studi di Milano, Dept. of Computer Science, EveryWare Lab
[2] EveryWare Technologies, Milano, Italy
[3] Dyson School of Design Engineering, Imperial College London, UK

Abstract. A mobile application that detects an item's color is potentially very useful for visually impaired people. However, users could run into difficulties when centering the target item in the mobile device camera field of view. To address this problem, in this contribution we propose a mobile application that detects the color of the item pointed by the user with one finger. In its current version, the application requires the user to wear a marker on the finger used for pointing. A preliminary evaluation conducted with blind users confirms the usefulness of the application, and encourages further development.

Keywords: Blind people, visual impairments, mobile device, smartphones, color detection, computer vision.

1 Introduction

Research on assistive technologies is currently facing a shift in its main focus; from computer accessibility to real world accessibility through computer. This phenomenon is particularly evident in the field of assistive technologies for people with visual impairment. Nowadays most devices (including mobile ones) are accessible to blind people and people with low vision through well engineered and generally stable commercial applications. There are still some open issues, for example with image exploration, but overall well established interaction paradigms exist to access most interface elements, like buttons, lists, etc... Consider for example accessibility to mobile devices, where VoiceOver and TalkBack (for iOS and Android devices, respectively) make most applications accessible, even if not specifically designed for users with visual impairments.

This situation allows researchers to focus on applications, specifically designed for people with special needs, that ease the access to the real world and provide support in completing everyday activities. Clearly, mobile devices are convenient platforms to support such applications, as they are easier to carry, generally ready for use and they also have sensors (e.g., accelerometer) that are unavailable on traditional devices. There are indeed many commercial applications, as well as research projects, aimed at supporting visually impaired people

in their everyday tasks. Among others applications, some support outdoor orientation and independent mobility (commercial applications like *AriadneGPS*[4], *iMove*[5] and many research projects [1–6]), others help recognising objects (like *AudioLabels*[6]) or sources of lights (like the commercial application *LightDetector*[7]).

In this paper we consider the problem of detecting the color of items (e.g. objects, surfaces, etc) in the real world. There are many applications, already on the market, aimed at supporting this task; the user points the device camera towards the desired item and then the app "reads aloud" the color in the image center or in the point where the user taps[8]. This interaction is suitable for color-blind users and/or for users with residual vision, who can still see shadows and distinguish between objects. However it is ineffective for blind users, who in many cases have problems in pointing the target item.

This contribution presents our ongoing research to address the above problem with a natural user interface aimed at supporting people with visual impairments in discovering the color of real world items. The ultimate goal is to have a software that recognizes a user's finger, detects the color being pointed by the finger and then describes it to the users, either with text to speech or with a sonification approach (i.e., with non-speech audio). The preliminary prototype *JustPoint* presented in this contribution is aimed at evaluating this idea with one simplification: instead of recognizing the actual finger, *JustPoint* recognizes a marker, a sort of ring that the user wears on the distal phalanx (see Fig. 1). The reason for this simplification is that this contribution focuses on the evaluation of system usability and ergonomics. Finger detection will be added as a future work: it is a non-trivial but doable task, as proved by many papers addressing this problem [7, 8].

This contribution first describes *JustPoint* (Sec. 3) and then presents the experimental evaluation we conducted to evaluate the prototype effectiveness (Sec. 4). Finally (Sec. 5) we discuss how we intend to address the issues that emerges from the evaluation and to improve current prototype.

2   Related work

In recent years, different solutions have been investigated to enable people with visual impairments to recognize colors. These solutions can be roughly divided into two categories: those addressed to people with color deficiency and those addressed to people who are not able to see colors at all (i.e., totally blind or severely sight impaired people[9]). In the following, the works concerning both categories are examined.

---

[4] https://itunes.apple.com/en/app/ariadne-gps/id441063072?mt=8
[5] https://itunes.apple.com/en/app/imove/id593874954?mt=8
[6] https://itunes.apple.com/en/app/audiolabels/id496513473
[7] https://itunes.apple.com/en/app/light-detector/id420929143?mt=8
[8] Examples are the "See Color" and "ColorVisor" apps, available on the AppStore.
[9] Henceforth we refer to these people as "blind people" or "blind users".

In the first category (i.e., solutions for people with color deficiency), Tanuwidjaja et al. and Popleteev et al. independently propose solutions on wearable smart glasses [9, 10]. These solutions automatically adapt the colors in the scene according to the type of color blindness and show the result on the smart glass display with an augmented reality approach. Both solutions cannot be directly applied to support blind people.

Manaf et al. [11] propose a system to support color blind people to detecting colors. One of the system's feature is similar to the one we propose in this contribution: to recognize the color pointed by the fingertip, however the interaction with blind users is not discussed.

Different solutions have been specifically proposed for blind people. Abboud et al. [12] propose EyeMusic, an application that enables the sonification of images captured through a camera. Colors are represented through different musical instruments and color combinations are mapped on a pentatonic music scale. Bologna et al. [13] propose SeeColor, an application aimed at supporting independent mobility. The scene is captured through the camera and most relevant objects are sonified. Colors are mapped into different musical instruments. The focus of these two contribution is on the sonification method to convey image information (including color) rather than on the user-interaction. Both solutions suffer from this limitation: a blind person can run into difficulties when he/she needs to point the camera towards a small target.

Findlater et al. [14] propose a technique for conveying visual information (e.g., color, labels, etc.) on objects that can be touched, capturing images through micro-camers mounted on the fingers. While this solution is promising, it also incurs into two issues. First, it requires specialized hardware, second, since the camera is mounted on the finger it can prevent the finger mobility, if not properly designed.

3   Color Identification With *JustPoint*

JustPoint is a fully functional Android prototype that (i) acquires images from the device camera, (ii) detects the marker, (iii) computes the "target area" pointed by the user, (iv) identifies colors in the target area and (v) provides the result to the user. Figure 1 shows an example image for each of these 5 steps.

(i) The image acquisition phase captures images from the camera video stream, using automatic exposure and focus. In our experiments, conducted with a Nexus 5, about 30 frames per seconds can be acquired, while fewer (i.e., about 3) can be processed. Since, at each instant, we want to process the most recent frame, a new frame is acquired only when the previous one has been processed.

(ii) In its current version, the marker is a QR-code. To exclude that other QR-codes that appear in the image are erroneously detected as the marker, *JustPoint* verifies that the string encoded in the actual marker matches a predefied string. When detecting the QR code, the position of its vertexes in the image is computed (see Figure 1(b)).
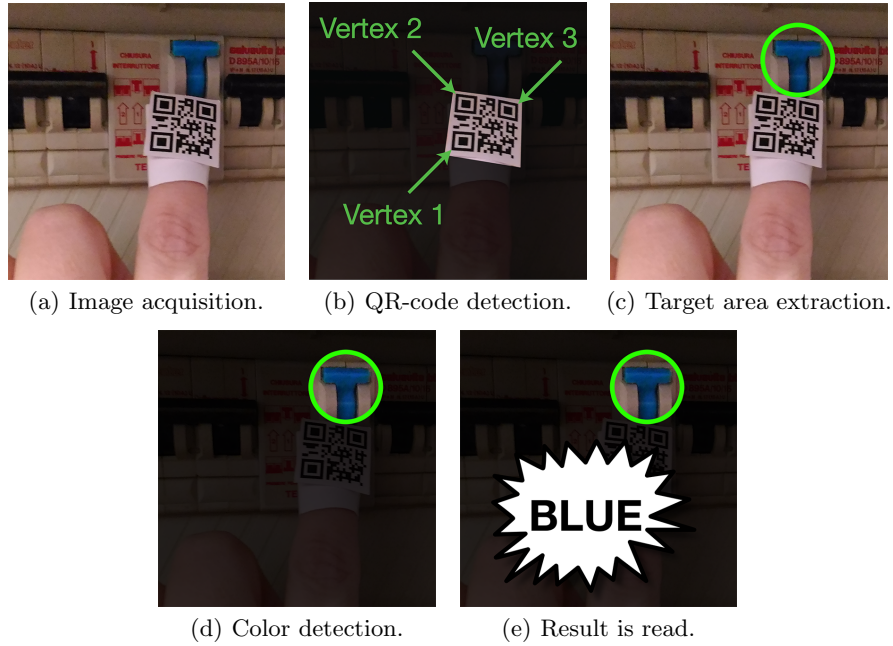
(a) Image acquisition.   (b) QR-code detection.   (c) Target area extraction.

(d) Color detection.   (e) Result is read.

Fig. 1. Five steps of the color recognition procedure.

(iii) In the third step *JustPoint* extracts the "target area", i.e., the image portion the user is pointing to. Thanks to some evaluations with the final users, we defined this as a circular region, approximately the same size as the finger width (i.e., circle diameter is the same as QR code edge), positioned right after the QR code, along the finger direction (see Fig. 1(c)).

(iv) The idea to identify colors in the target area is to label each pixel with a color name and then to count the number of pixels for each label. Adopting this approach each color name is associated with its "frequency" i.e., the number of pixels with that color divided by the total number of pixels in the target area. *JustPoint* considers 11 color names, as defined by Berlin and Kay [15]: black, white, grey, red, yellow, green, blue, violet, brown, orange. To assign a color name to each pixel, *JustPoint* adopts a simple and effective solution: the target area is converted in its HSV representation (from RGB) and then threshold values on S and V channels are used to distinguish "black", "white" and "grey" colors. Other colors are distinguished though threshold values on the H channel with the only exception of "orange" and "brown" that are distinguished also considering the V channel. Note that this coarseness of the color bin division increases the method robustness with respect to device camera gamut shifts.

(v) User is informed with a Text To Speech (TTS) synthesizer that reads the name of the colors identified in the target area. With a standard TalkBack-based interaction, the user can set the "number of colors" preference that represents

how many colors, among those with frequency higher than the "min-frequency" threshold, should be read to the user. Consider the example shown in Fig. 1(e). If "number of colors" is set to 1, then "blue" is read, as this color has the highest frequency ($\sim 40\%$). If the parameter is set to 2, "blue and grey" is read, as grey is the second color with highest frequency ($\sim 30\%$). The same text is read if "number of colors" is set to 3. Indeed, the third color with highest frequency is "brown" ($\sim 15\%$), but its frequency is below the "min-frequency" threshold (i.e., below 25%).

4   System Evaluation

JustPoint was designed with a user centered approach. The development stage was continuously monitored and evaluated by one member of the design team, who is congenitally blind. In addition, we conducted three evaluation sessions with blind subjects; one before the prototype was developed, one with a preliminary prototype and, finally, one with the final version of *JustPoint*.

In the preliminary evaluation we described the idea of the application to two congenitally blind subjects (one 37 years old male, one 37 years old female) and we asked them to point with one hand an object whose color they want to discover. At the same time subjects had to hold the device in the other hand and point the camera towards the object. While doing that, instead of the actual application (which at the time had not been developed yet), subjects were asked to record a video. The result is therefore a set of recordings representing what the camera was pointing at during the simulated use of the application. The test was conducted in a non-supervised environment.

Two main considerations emerged from this evaluation. First, for a blind person it is not trivial to hold the device with one hand and point its camera towards the other hand that is pointing an object. In most cases the two subjects were holding the device very close to the pointing hand, which was out of the camera field of view. The second consideration is that a blind person may not have a clear idea of the environmental light conditions. Indeed, one video was recorded in an almost totally dark room where colors were not perceivable.

Based on the experience of the first evaluation session, we added two functions to the application; first, when the application runs for the first time, it suggests the user to put the device camera right above the pointing finger and than move the camera away. Second, if light is too low *JustPoint* warns the user and turns the flash light on.

The second evaluation session was conducted during the app development stage. It was performed with the two subjects involved in the previous evaluation and by one late-blind subjects (28 years old male) in a supervised environment. Subjects were asked to use the application to detect the color of some objects provided to them. Two main observations arose. Firstly, while the user is trying to point the device in the correct direction, the app does not give any feedback, hence the user does not know if the app is working properly. To fix this problem we added a feature that plays a pure tone sound every 2 seconds when no QR code

is detected. The second observation was that the color detection technique, based on RGB representation at that time, was not accurate enough. Consequently we modified it as described in Section 3.

The third evaluation session was conducted with an improved version of *JustPoint* and was aimed at evaluating the usability of the application while identifying the color of different objects. We asked the same blind subjects of the second evaluation session to recognize the color of three sets of objects: books, small toys and liquid in bottles. The three subjects were able to accomplish all tasks successfully. They did not run into difficulty in pointing the camera correctly and this operation generally took about one second, or even less. All subjects independently commented that the application is very useful but that wearing the ring with the QR code is not convenient. In particular one subject noted that she would not always have the ring with her when she would need to recognize a color.

5   Future Work

While the current version of *JustPoint* is still preliminary, its user-centred design highlighted a number of features and its evaluation suggests that it can be practical and useful.

The main limit of *JustPoint*, in its current version, is that it requires the user to wear a marker on the finger. We intend to address this by implementing a finger detection technique.

Considering the user interaction, we intend to investigate the use of wearable devices (like *Google Glass*[10] or *Sony SmartEyeglass*[11]). One research question is whether it is easier to point the target object if the camera is placed on the user's head.

Another important improvement for *JustPoint* is to implement sonification for delivering information about the colors. Sonification could be used as the sole audio feedback method, or in conjunction with speech information. A similar approach has been followed by the same team when working with a navigational task [2]. This will allow the user to distinguish between a continuous range of colors, rather than a discrete one. Furthermore, sonification will allow the transfer of additional information to the user (e.g. colors of the areas surrounding the finger, indications about how to orient the phone in order for the camera to be able to catch the pointing finger, etc.).

Finally, the performance of *JustPoint* can be improved by adopting parallel computation (i.e., process more than one frame a given time) and possibly by using the highly parallel GPU architecture to speed up computation.

---

[10] https://developers.google.com/glass/
[11] https://developer.sony.com/develop/wearables/smarteyeglass-sdk/

References

1. Mascetti, S., Ahmetovic, D., Gerino, A., Bernareggi, C., Busso, M., Rizzi, A.: Robust traffic lights detection on mobile devices for pedestrians with visual impairment. Computer Vision and Image Understanding (In Press)

2. Mascetti, S., Picinali, L., Gerino, A., Ahmetovic, D., Bernareggi, C.: Sonification of guidance data during road crossing for people with visual impairments or blindness. International Journal of Human-Computer Studies (2015)

3. Ahmetovic, D., Manduchi, R., Coughlan, J., Mascetti, S.: Zebra crossing spotter: Automatic population of spatial databases for increased safety of blind travelers. In: Proc. of the 17th Int. ACM SIGACCESS Conf. on Computers and Accessibility. (2015)

4. Manduchi, R., Coughlan, J., Miesenberger, K., Coughlan, J.M., Shen, H.: Crosswatch: a system for providing guidance to visually impaired travelers at traffic intersection. Journal of Assistive Technologies **7**(2) (2013) 131–142

5. Hara, K., Azenkot, S., Campbell, M., Bennett, C.L., Le, V., Pannella, S., Moore, R., Minckler, K., Ng, R.H., Froehlich, J.E.: Improving public transit accessibility for blind riders by crowdsourcing bus stop landmark locations with google street view: An extended analysis. ACM Transactions on Accessible Computing (TACCESS) **6**(2) (2015) 5

6. Brady, E.L., Sato, D., Ruan, C., Takagi, H., Asakawa, C.: Exploring interface design for independent navigation by people with visual impairments. In: Proc. of the 17th Int. ACM SIGACCESS Conf. on Computers & Accessibility, ACM (2015) 387–388

7. Baldauf, M., Zambanini, S., Fröhlich, P., Reichl, P.: Markerless visual fingertip detection for natural mobile device interaction. In: Proc. of the 13th Int. Conf. on Human Computer Interaction with Mobile Devices and Services, ACM (2011) 539–544

8. Chun, W.H., Höllerer, T.: Real-time hand interaction for augmented reality on mobile phones. In: Proceedings of the 2013 international conference on Intelligent user interfaces, ACM (2013) 307–314

9. Tanuwidjaja, E., Huynh, D., Koa, K., Nguyen, C., Shao, C., Torbett, P., Emmenegger, C., Weibel, N.: Chroma: A wearable augmented-reality solution for color blindness. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. UbiComp '14, New York, NY, USA, ACM (2014) 799–810

10. Popleteev, A., Louveton, N., McCall, R.: Colorizer: Smart glasses aid for the color-blind. In: Proceedings of the 2015 Workshop on Wearable Systems and Applications. WearSys '15, New York, NY, USA, ACM (2015) 7–8

11. Manaf, A.S., Sari, R.F.: Color recognition system with augmented reality concept and finger interaction: Case study for color blind aid system. In: ICT and Knowledge Engineering (ICT & Knowledge Engineering), 2011 9th International Conference on, IEEE (2012) 118–123

12. Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S., Amedi, A.: Eyemusic: Introducing a visual colorful experience for the blind using auditory sensory substitution. Restorative neurology and neuroscience **32**(2) (2014) 247–257

13. Bologna, G., Gomez, J.D., Pun, T.: Vision substitution experiments with see color. In: Natural and Artificial Models in Computation and Biology. Springer (2013) 83–93
14. Findlater, L., Stearns, L., Du, R., Oh, U., Ross, D., Chellappa, R., Froehlich, J.: Supporting everyday activities for persons with visual impairments through computer vision-augmented touch. In: Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. ASSETS '15, New York, NY, USA, ACM (2015) 383–384
15. Berlin, B., Kay, P.: Basic color terms: Their universality and evolution. Univ of California Press (1991)