Accessible On-Line Floor Plans

Cagatay Goncu
Monash University, Australia
cagatay.goncu@monash.edu

Anuradha Madugalla
Monash University & NICTA
Victoria, Australia
amad29@student.monash.edu

Simone Marinai
Universita degli Studi di
Firenze, Italy
simone.marinai@unifi.it

Kim Marriott
Monash University & NICTA
Victoria, Australia
Kim.Marriott@monash.edu

ABSTRACT

Better access to on-line information graphics is a pressing need for people who are blind or have severe vision impairment. We present a new model for accessible presentation of on-line information graphics and demonstrate its use for presenting floor plans. While floor plans are increasingly provided on-line, people who are blind are at best provided with only a high-level textual description. This makes it difficult for them to understand the spatial arrangement of the objects on the floor plan. Our new approach provides users with significantly better access to such plans. The users can automatically generate an accessible version of a floor plan from an on-line floor plan image quickly and independently by using a web service. This generates a simplified graphic showing the rooms, walls, doors and windows in the original floor plan as well as a textual overview. The accessible floor plan is presented on an iPad using audio feedback. As the users touch graphic elements on the screen, the element they are touching is described by speech and non-speech audio in order to help them navigate the graphic.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces; I.7.5 [**Document And Text Processing**]: Document Capture—*Graphics recognition and interpretation*; K.4.2 [**Computers And Society**]: Social Issues—*Assistive technologies for persons with disabilities*

General Terms

Human Factors, Design

Keywords

graphics; accessibility; vision impairment; transcription; graphics recognition

1. INTRODUCTION

One of the most disabling consequences of being blind (by whom we mean any person whose level of vision does not allow them to easily read printed text or graphics even when magnified) is a lack of access to information presented graphically, such as maps, plans, diagrams and illustrations in instruction manuals. Such graphics are increasingly accessed on the web. However, on-line graphics are typically not accessible to users who are blind because despite accessibility guidelines [1] the vast majority either do not include an alternative text description of the graphic or the text is very general and not helpful [16]. While in theory a blind person might request an audio description or a tactile graphic version of an on-line graphic from a support organisation such as Vision Australia this is expensive and may take weeks for the accessible version to be produced.

We present a new model for providing access to on-line information graphics: automatic transcription from an on-line image to an accessible graphic designed for presentation on a portable touch screen device such as an iPad. The accessible graphic can be viewed immediately by using audio feedback, and can also be improved by a transcriber with semi-automatic transcription. We believe this model has the potential to dramatically improve access to a wide range of on-line information graphics.

We describe and evaluate the use of the model for provision of accessible on-line floor plans. The reason for focusing on floor plans is that the fear of leaving the safety of home and having to find one's way in a new environment is another great disabling consequence of blindness. Navigation in buildings and public spaces such as hotels, conference venues, airports or shopping centres is particularly challenging. When visiting a building for the first time it is difficult to find doors, elevators, stairs or the location of the desired destination. Accessible floor plans help with this task by allowing a person who is blind to plan their visit beforehand and to identify if they will need help to reach their location. Floor plans are not only useful for navigation. For most sighted people they are an important part of the decision making process when finding an apartment or house for renting or buying. However, if you are blind it is currently virtually impossible to access floor plans that are available on most on-line real estate sites, making renting or buying that much more difficult.

Nowadays floor plans of buildings and public spaces are usually accessed on the web. Thus, there is a real practical
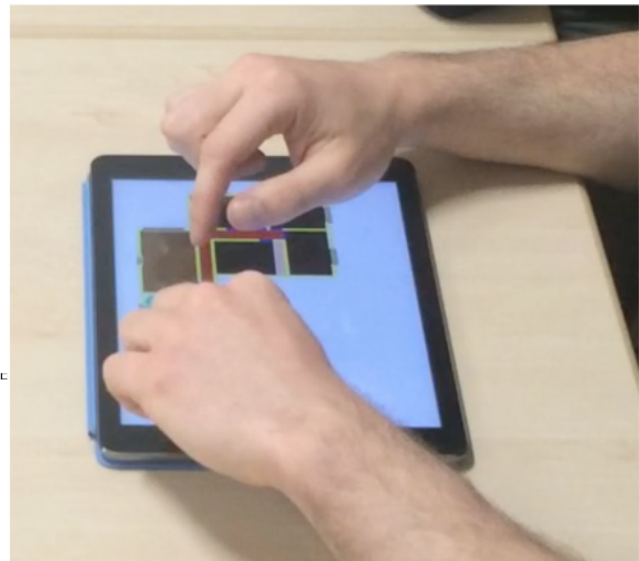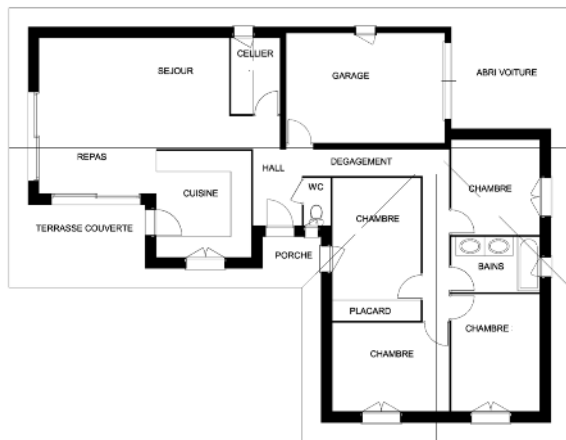
Figure 1: The web service can be used to convert an on-line image of an apartment or house floor plan (for instance the one the left) to an accessible version that is being read by a blind person using the iPad app (shown on the right).

need to provide better access to such on-line floor plans. Our model, illustrated in Figure 1, addresses both presentation and automatic production of accessible floor plans from on-line images. It allows a blind person to immediately view building and floor plans that they encounter on-line and does not require printing expensive tactile graphics. This potentially provides much greater access to such information and removes a significant barrier to travel by blind people. Our approach relies on two innovations.

The first innovation is that we present an accessible version of the floor plan on an iPad using audio feedback. This is based on prior research into audio and audio-haptic presentation of street maps on mobile devices. We have implemented a new application that we call GraViewer. The presentation is not a simple VoiceOver labelling of the objects in the floor plan. As the users move their fingers over the accessible floor plan they receive carefully chosen feedback, e.g sharp clicks as a finger moves through a wall. Based on transcription guidelines used for tactile floor plans we also include a high level description of the floor plan that allows the user to obtain an overview when they first open the floor plan.

Our approach is in contrast to research on producing accessible versions of on-line bar charts and line graphs which has focused on producing a textual description of the on-line graphic [6, 7]. While a textual description can provide an overview of the building's layout, such a description necessarily loses information and it is also difficult to use it to build a mental model of the layout. For this reason accessibility guidelines recommend the use of tactile graphics for presenting graphics such as floor plans in which spatial relationships are important, e.g. [24]. Unfortunately tactile graphics are relatively expensive and require specialised printing devices and/or paper: touch screen mediated presentation potentially provides the spatial benefits of a tactile graphic but at a fraction of the cost.

Currently accessible versions of a graphic such as a floor plan are created manually, typically by trained transcribers. This is expensive and time consuming. It is not practical to manually transcribe floor plans in order to support travel by blind people and completely infeasible to provide manual transcription of the dozens of floor plans encountered when searching on an on-line real estate site. Thus, some sort of automatic transcription is required. In the case of maps, accessible graphics can be generated from GIS data. Unfortunately, equivalent geospatial data is not available for most buildings or public spaces. What is required is automatic transcription from on-line floor plan images.

Our second innovation is to have developed a web service, called GraFloor, that automatically creates an accessible graphic in SVG format from a floor plan image. This uses sophisticated graphics recognition algorithms and OCR to recognise the rooms, walls, windows, doors and furniture. From this GraFloor generates a simplified graphic for display on the GraViewer which shows the rooms, walls, doors and windows. This graphic also includes an automatically generated textual overview.

Robust floor plan recognition is a very difficult problem. While GraFloor is very effective, in applications where manual transcription is not cost prohibitive (such as for the web pages of a hotel or convention center) manual transcription will typically produce a better result. In order to support manual transcription we have also integrated GraFloor with GraAuthor, a web application which is designed to create content for the GraViewer. GraAuthor takes the output of GraFloor web service and allow the transcriber to correct/edit the accessible version. Such semi-automatic transcription is much faster than traditional manual transcription which can take many hours.

We provide two preliminary empirical evaluations of our approach. The first is a user-study with 8 blind participants to show that GraViewer's audio-only presentation allows blind users to understand an accessible floor plan and

use it in common tasks such as planning how to get from one room to another or identifying the number of bedrooms. The second is an evaluation of the quality of the automatic transcription provided by GraFloor.

In the next section we discuss related work. In Section 3 we describe the presentation app GraViewer, in Section 4 we describe the authoring tools GraFloor and GraAuthor, and in Section 5 we provide the evaluations.

2. RELATED WORK

Related work falls into three categories: presentation technologies, graphics recognition and automatic transcription.

Presentation Technologies for Accessible Graphics

Presentation technologies fall into four main categories: tactile graphics, audio only presentation which is either based on a textual description of the graphic or sonification, e.g. [14, 4]; tactile overlay on top of a pressure-sensitive screen which provides audio feedback when a graphic element is pressed, e.g. IVEO or TTT [9, 15]; haptic presentation using devices like the phantom, e.g. TeDUB [21] or the VTPlayer tactile mouse with a graphics tablet [28]; or presentation on a touchscreen device using a mixture of haptic or audio feedback as the user explores the graphic, e.g. GraVVITAS [11] or [10]. The advantage of the last approach is that it does not rely on expensive tactile graphics or overlays but still allows the user to use both hands to explore the graphic and build up a mental model of it. This is the approach we use.

Our implementation of GraViewer is based on that of GraVVITAS [11]. The main difference is that GraVVITAS provides haptic feedback using a custom data glove that used small vibrators attached to fingers to indicate when an element was touched. In this project we focussed on audio-only feedback because most users do not have access to a data glove and we wanted to create an application that could be easily downloaded and deployed on any iPad.[1] Our work extends earlier work on GraVVITAS by demonstrating that an audio only presentation of a complex graphic such as a floor plan can be understood and by the use of automatic transcription to create the graphic.

Our approach is also similar to audio-only and haptic-audio presentation of on-line street maps on mobile devices, e.g. [23, 13]. Here the main difference is that we are focussing on floor plans which have different characteristics and tasks to street maps (which are primarily used for way-finding) and the use of automatic transcription from an image rather than an on-line GIS. We also note a mobile app being developed for generating textual directions for blind users when navigating large public spaces [25]. This does not provide a graphical representation or consider automatic transcription from an image.

Graphics Recognition

The second field of research is graphic recognition. Graphics recognition has proven to be extremely difficult and most previous research has focussed on recognising particular visual notations such as tables, charts, music, CAD drawings, floor plans etc [17]. Traditionally graphics recognition has dealt with scanned images, but some recent work has dealt with recognition of born-digital, i.e. computer generated, images from PDF documents [8]. We also only consider born-digital images and not scanned images as information graphics are increasingly born-digital. This means that recognition is easier as the images are clean and so few ambiguities arise in low-level image pre-processing and state-of-the-art OCR tools recognize text with few errors.

Our automatic transcription tool is based on a state-of-the-art application for vectorizing and recognising images of floor-plans [2]. The original application integrated opensource image processing tools with Markov Logic Networks to achieve robust identification of walls, doors, stairs, windows and furniture as well as identification of the room type and function based on furniture. We have extended the recognition software to handle text and to use this to aid in room identification and also added functions to determine the geometry of rooms including multi-use open plan areas.

Automatic Transcription

Researchers at the University of Washington have built an image processing tool (TGA) that can be used by transcribers to extract and replace text from scanned graphics by the equivalent braille [12]. ViewPlus have a vectorization and OCR-to-speech tool for constructing presentations for IVEO [9], while TMAP and Talking TMAP are tools for creating tactile maps from a geographic database [19]. However, all of these tools employ relatively simple image processing, OCR and vectorization technologies and cannot perform anything like the sort of complex transcription provided in our tool.

There has been some previous research into automatic generation of accessible graphics from on-line images. This has focussed on automatically generating a textual summary or audio presentation of on-line bar charts (Interactive SIGHT) [6] and (iGraph-LITE) line graphs [7]. Our work differs from this because of its focus on floor plans rather than bar or line charts and because we generate a simplified spatial representation as well as a textual summary. Because of the more open ended tasks performed with floor plans and the fact that they are inherently spatial in nature we feel that it is important to provide a graphical representation.

3. FLOOR PLAN PRESENTATION

The most common ways to provide an accessible version of a floor plan are either to provide a textual description or a tactile graphic. While a textual description can provide an overview of the building's layout such a description necessarily loses information and it is also difficult to use it to build a mental model of the layout. Research has shown that touch allows blind people to build internal spatial representations that are functionally equivalent to those obtained from visual input [3] and tactile maps have been shown to help blind children and adults build a survey-like representation of their environment, thus facilitating navigation [27]. For this reason a tactile graphic together with a textual overview is the preferred presentation approach [24].

The disadvantage of tactile graphics as a presentation medium is that they are relatively expensive to produce and require specialised printing devices and/or paper. Thus, it

---

[1]However, we note that it is likely that in the near future touch screen devices will provide haptic feedback natively through technologies such as Tactus (`http://www.tactustechnology.com/technology.html`), in which case it would be straightforward to modify GraViewer to take advantage of this.

is not practical to use them to present accessible versions of most on-line floor plans. Instead we decided to use touch screen with audio feedback. As we have noted, this approach has previously been used to present street maps to blind users.

We have built an iPad application called GraViewer for accessible presentation of information graphics. When the blind user touches a graphic element on the display, audio feedback describes the element and is also provided in order to help with navigation. The great advantage of this approach is that it does not require the use of expensive tactile overlays or additional haptic presentation devices like the phantom, yet allows the user to explore the graphic with both hands using strategies similar to those used with tactile graphics.

The user interface of GraViewer and the presentation model for floor plans was developed using a participatory design methodology with blind end-users. The GraViewer user interface is designed to be consistent with that of iOS accessibility framework: it utilises standard iOS accessibility gestures for menu navigation and application control as well as using VoiceOver for speech.

When the accessible floor plan is first displayed an overview of the graphic is read to the user. This describes the rooms in the floor plan. If required again the overview can be accessed through the menu items that appear when the user double taps with two fingers. This menu also allows the user to open files and change application settings.

Components in the floor-plan–rooms, walls, windows and doors–have both a textual description and non-speech audio associated with them. The colour of the object determines which non-speech audio is associated with it. Initially we experimented with high-affordance non-speech audio with different rooms (such as the sound of cutlery in the dining room) but feedback from blind-users was that they preferred more neutral non-speech audio and so we now simply use different frequencies for different colours.

GraViewer displays the accessible floor plan on the iPad. While people who are totally blind obviously cannot see this image, it is useful to provide a visible graphic with distinct colours and thick simple lines for users who have low-vision. Furthermore it is useful to have a visible graphic in a classroom like environment where a sighted instructor might wish to see what the blind person is doing.

When a floor plan component is first touched its associated textual description is read aloud using VoiceOver. This is read again if the user double taps the component to "query" it or if the user's fingers leave the component and then reenter it. One subtlety in floor plans are open plan areas which may contain multiple functionally different spaces that do not have clear boundaries. Our convention is to associate a description of form "Open plan: kitchen" and "Open plan: dining" to the different areas in the open plan room.

One of the most difficult tasks for blind-users is following the boundary of a graphic element or following thin-lines such as those indicating walls in a floor plan. To help this we use a sharp click to indicate when a graphic element is entered or left and use a volume gradient along the boundary to reduce the non-speech audio as the user's fingers are about to leave the room or wall. Junction points such as where two walls meet are also indicated with a different sharp click sound to the user.

Users can use multiple fingers at once. However our earlier studies found that using more than two fingers simultaneously was confusing because of the multiple speech and non-speech feedback. Furthermore it was quite common for the user to inadvertently touch the screen with their thumb or another fingers and then to become quite confused by the associated feedback. Therefore, GraViewer now lets the users know when they are touching the screen with more than two fingers by using a short non-speech audio warning. It also warns the user when one of their fingers leaves the display area and enters the surrounding bezel area using another distinct non-speech audio. If the users use two fingers, they simultaneously receive separate audio feedback for each finger if they are touching different objects but if both fingers are touching the same object the non-speech audio feedback is amplified.

GraViewer displays graphic content specified in SVG (the W3C standard for Scalable Vector Graphics) on a graphic canvas which is part of the iOS framework. The canvas loads a SVG file and uses metadata associated with the shapes to control the interaction. The metadata associated with a shape is: its ID, audio volume level for the interior of the shape and for its boundary, the text string to be read out when the shape is queried, and the name of a (non-speech) audio file and/or the color code for generating the sound associated with the shape during navigation. Figure 2 shows an example floor plan and the associated meta data. The metadata also includes the title, category, and the overview summary of the graphic.

While we have focussed on the presentation of floor plans, GraViewer is a generic technology that can be used to display accessible web graphics. Because it displays an SVG file it fits well with HTML5 and current web technologies. Web authors can choose to link or embed SVG files in their web pages that are suitable for display using GraViewer and so provide an accessible alternative to the default web page graphic. This is most suitable for use in situations where a tactile graphic would be preferred to a textual description, i.e. when the spatial layout of graphic elements is important. This is true for floor plans but also maps and many scientific and technical drawings in on-line educational material and instruction manuals.

4. FLOOR PLAN CREATION

While GraViewer provides an exciting new technology for displaying accessible graphics it is unfortunately not realistic to expect that most web page authors will have the time or recognise the need to create an accessible version of an on-line floor plan or other graphic. We know that it is still rare for web authors to provide even the most rudimentary accessibility information for a graphic such as a detailed alternative text description. Thus, in practice we need an automated method for creating an accessible floor plan from an on-line image that can be used by both blind and non-blind web authors when creating a web page or blind people when they encounter a floor plan image on the web. We have created a web service called GraFloor to do this.

GraFloor has two main components: the *graphics recognition module* and the *transcription module*. The graphics recognition module takes an floor plan image in common formats such as png and jpg, and produces a text file with a high level specification of the rooms, walls, windows, doors and furniture in the floor plan including textual labels and

Overview: "This graphic is the floor plan of a house which has six rooms. The rooms are located in clockwise direction as follows: cellar, toilet, hall, study, living room, dining room, and kitchen. Hall, living room, dining room and kitchen are in an open plan area."

"open plan: kitchen" "cellar"

"door" "toilet"

"open plan: dining room" "open plan: hall"

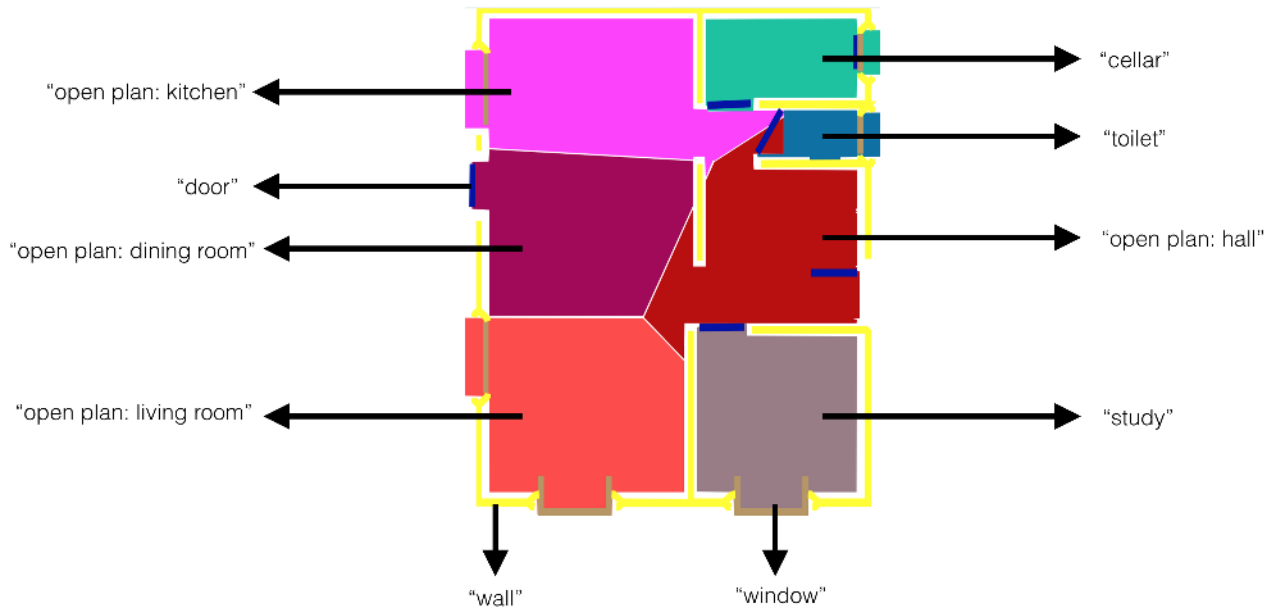"open plan: living room" "study"

"wall" "window"

Figure 2: A floor plan annotated with the associated textual descriptions and overview. Different coloured regions have different associated non-audio feedback.

their geometry. The transcription module gets the output of the recognition module in JSON format and generates an SVG file with additional meta data information that is suitable for display on the GraViewer.

4.1 Graphics Recognition Module

The application developed to recognize the floor plan images is based on the following main steps. They are illustrated in Figure 3.

1. Preprocessing. Thin dashed and solid lines are removed as they usually describe items on other levels of the house (e.g. the shape of the roof).

2. Text recognition. Text regions are identified and recognized by `tesseract`, a state-of-the-art Optical Character Recognition (OCR) tool (`https://code.google.com/p/tesseract-ocr/`). These regions are also removed from the input image.

3. Image representation. The floor plan image is described by means of a *region adjacency graph* whose nodes correspond to connected components in the image while edges connect touching components.

4. Object recognition is performed by means of an inexact graph matching algorithm on the region adjacency graph. Objects include furniture and windows.

5. Wall identification and vectorization. This includes determining the location of stairs.

6. Door recognition and removal from the floor plan image.

7. Room identification and sub-division of open-plan areas into separate "virtual" rooms.

Our implementation extends that described in [2]. We use the image preprocessing and the floor plan representation and object retrieval provided in the first release of the tool and have extended wall, door, and room identification. While the focus of [2] was identifying the purpose of a room from its furniture and without considering textual labels, the main aim of the floor plan analysis system in GraFloor is the delineation of the geometry and function of rooms as well as their mutual arrangement. In the rest of this section we describe the above mentioned steps more important for the extraction of the floor plan structure.

Text Localization and Recognition

Text strings are identified in the image by looking for connected components (CC) with "character-like" features. The latter are CCs having dimensions (width and height) close to a square and with neighbouring CCs with similar features. A text string is hypothesized if there are at least three character-like objects aligned. Additional smaller components in the neighbourhood of the string can be added to the string in order to recover punctuation marks that could be skipped in the first check.

All the potential text strings (Figure 3(b))are then processed with the OCR tool in order to identify the room function and the position of the corresponding text string.

Wall Identification

In the preprocessing step the use of adaptive thresholding allows us to obtain a black-and-white image with empty walls
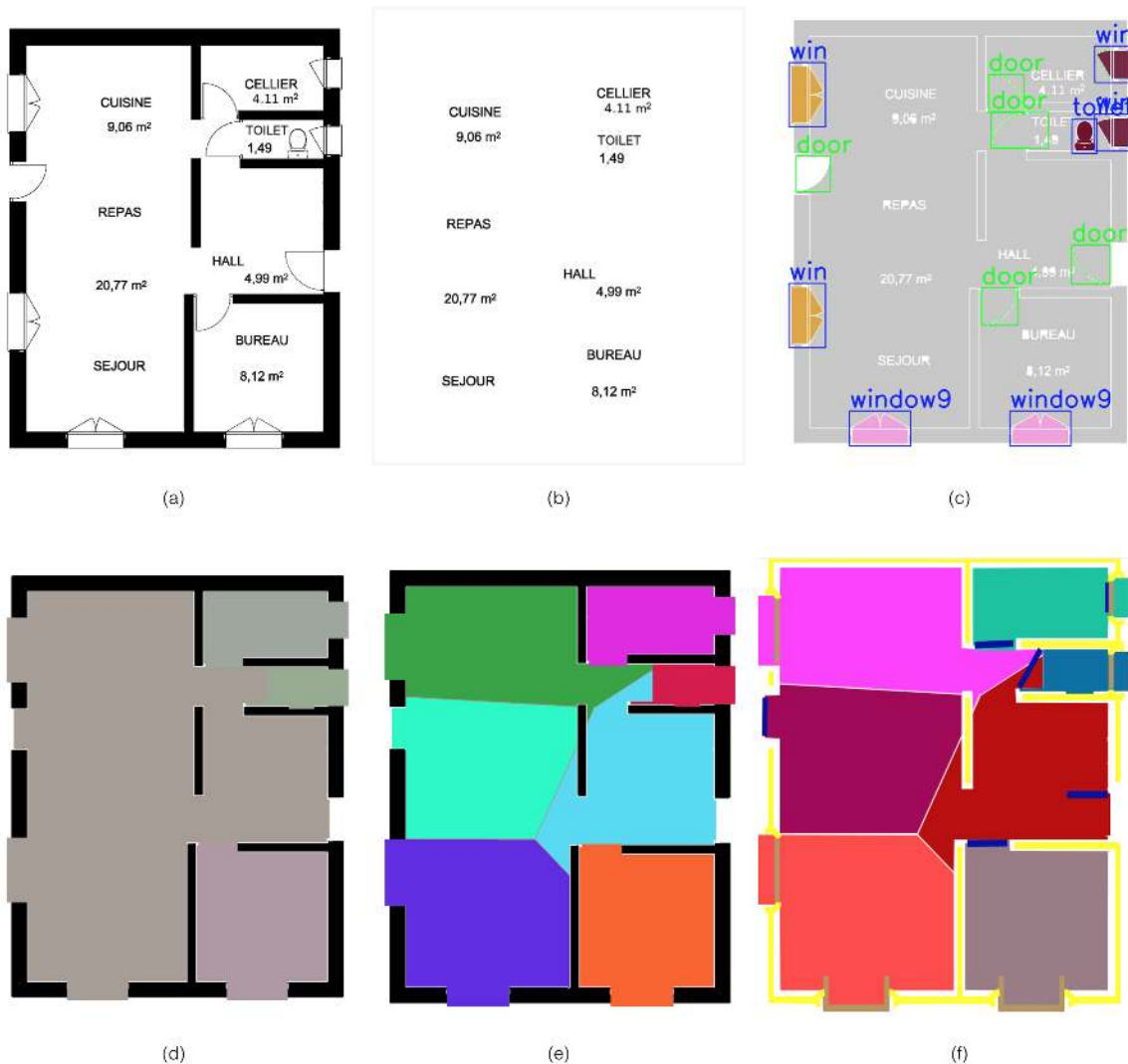
392

Figure 3: Main processing steps in automatic floor plan transcription: (a) input floor plan; (b) text extraction using OCR; (c) object recognition; (d) wall recognition; (e) room identification and sub-division of open plan area; (f) accessible graphic produced for display on GraViewer.

even when the input image has black walls. This is achieved by smoothing the image with a 3×3 Gaussian Filter and then thresholding the resulting image with a fixed-size moving window where the binarization threshold is adapted in each position on the basis of the average pixel intensity in the window. operations are then performed to fill black components opened by the previous operations. After this step the objects are identified with an inexact graph matching algorithm [2]. Among objects, we consider also windows that are recognized with the same algorithm. For example, Figure 3(c) represents the objects found in Figure 3(a) including doors and windows.

Walls are then searched among components not recognized as objects. To this purpose we first identify the skeleton of the component and then look for straight lines formed by pixels of the skeleton with the Hough transform (HT). Potential walls correspond to segments in the output of HT and are verified by means of the standard deviation of the

local thickness of the component along the segment. Segments are stored as walls If the standard deviation is low and the component is dense.

After recognizing the walls, the corresponding connected component is represented by means of a poligonalization of its skeleton that is computed with the Douglas-Peucker algorithm. The description of walls by means of their skeleton allows us to adapt the wall thickness to the needs of the final visualization so that different users can modify the wall thickness.

Door Identification

In floor plans, doors are often represented with one segment that denotes the door in the opened position and one circular arc that represents the doorway. In our system we analyze the contour of large connected components (Figure 4(a)) and search potential doors (short segments, Figure 4(b)). To verify whether one potential door is a real one we search the
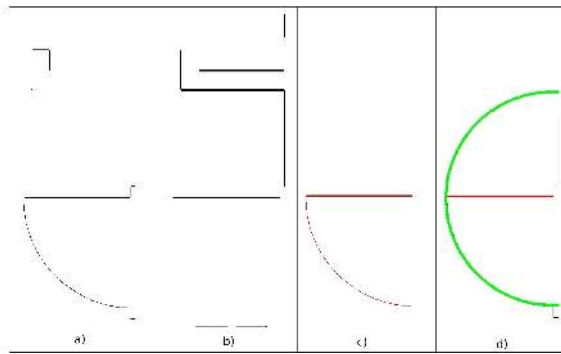
Figure 4: Steps in door identification.

Figure 5: GraAuthor: The web application which is designed for creating content for GraViewer.

doorway by first extracting the pixels of the contour closer to the segment (Figure 4(c)) and then verifying whether these points belong to a circle by using the Hough transform for circles (Figure 4(d)).

With this procedure it is possible to identify also the opening direction of the door. This can be represented in the accessible floor plan.

Room Identification

Room identification is made more difficult because modern houses and flats frequently contain multi-purpose rooms where open-plan living spaces share different functions that are sometimes identified by suitable text strings. In this case, in Figure 3(a) one room has three labels: "CUISINE" (kitchen), "REPAS" (meal room), and "SEJOUR" (living room). For a sighted reader this makes it clear that while this is a single large room it is split into three different regions whose approximate location is given by the textual label.

In order to extract the shape of rooms we need to remove the objects and the doors identified in each room. A room is recognised as a region completely enclosed by walls, windows and doors. In Figure 3(d) initially four rooms are identified since "CUISINE", "REPAS", "SEJOUR" and "HALL" are not separated by doors and so GraFloor recognises them as a single open space room. In GraFloor open-plan rooms with multiple functions are split in virtual sub-rooms by considering the Voronoi partition of the original open-plan room by using the barycenters of the text strings as the central points. In Figure 3(e) we show the partition of the larger room on the basis of the four text strings.

While the Voronoi partition provides a reasonable first approximation to how to split the rooms we plan to explore this further by better taking into account natural boundaries between regions induced by stub walls and kitchen benches.

At the end of the room extraction, the polygonal approximation of its contour is stored, together with the furniture objects and structural items (doors and windows) extracted from the floor plan input image.

4.2 Semi-Automatic Transcription

Robust floor plan recognition is a difficult problem and while evaluation in Section 5 shows that GraFloor is quite effective the generated images typically contain small errors, such as not recognising small internal walls or mis-recognized doors or windows. In situations where manual transcription is not cost prohibitive (such as for a hotel or conven-

tion center) this will typically produce a better result. To support such manual transcription we have created another web-service, GraAuthor, for manual production of accessible versions of floor plans designed for display on GraViewer. Importantly, GraAuthor can take the output of GraFloor and allow the transcriber to correct/edit the accessible version. Such semi-automatic transcription is much faster than fully automatic transcription. An example of its use is shown in Figure 5.

The use of a JSON intermediate floor plan description also allows other ways of generating accessible floor plans. It is possible to specify a floor plan textually say by hand or to generate it from CAD or some other digital representation of the building and then use the transcription module to generate an accessible version.

5. EVALUATION

We have conducted two preliminary evaluations of our approach. The first is a user-study with blind participants to investigate whether GraViewer's audio-only presentation allows blind users to understand an accessible floor plan and to use it in common tasks. The second is an evaluation of the quality of the automatic transcription using GraFloor.

5.1 User Study

As discussed previously GraViewer has been developed using a participatory design methodology with blind students and adults. We believe participatory design with blind participants is vital for any project of this kind since our experiences, and those of other researchers in the field, suggest that people who have been blind from an early age may have quite different strategies for understanding graphics to people who are sighted. It has been reported that congenitally blind children perform significantly poorer than other blindfolded and late-blind children in unfamiliar spatial tasks [20].

The main aim of the user study was to evaluate whether blind people could understand a floor plan of a house or apartment generated using GraFloor and displayed using GraViewer and use the plan to help them perform a number of simple tasks.

Participants

Eight participants, all legally blind (2 born blind and 6 late blind), between the ages of 25 to 55 participated in the user study. It is worth pointing out that since there are relatively few blind people and it is often hard for them to travel, it

394

Figure 6: Test graphics: Input images (a, d), uncorrected graphics (b, e), and corrected graphics (c, f).

is quite difficult to find blind participants (also pointed out in [26, 22]). Hence the number of participants in any such study is necessarily quite small.

The participants had previously been recruited to participate in the design of GraViewer through an email list for print disabled people in Australia. They had prior experience with tactile graphics and most had some prior training in the accessibility features of iOS and had used GraViewer to view simply geometric shapes: polygons and lines.

Procedure and Materials

The experiments were performed in a private room at Monash University or at one of the offices of Vision Australia, with a single participant at a time. Each study took about one hour.

Participants were asked to sign a consent form which had previously been sent by email. This also provided a short explanation of the user study and what type of information would be collected.

Participants were then asked some background questions to clarify:

- Their prior experience in understanding tactile floor plans,

- Whether they wanted to be reminded how the GraViewer system works.

Next they were trained in the use of GraViewer for viewing floor plans. This had two parts: If they wanted, first they reviewed the common iOS accessibility gestures and interface, e.g. accessing menus and turning VoiceOver on and off, and the use of GraViewer for viewing an image containing a simple geometric shape. In the second stage of training participants were shown a simple floor plan on GraViewer and asked to explore it and answer questions similar to those used in the subsequent study. They were encouraged to use a rapid sweeping gesture with one finger across the screen to obtain an initial overview of the spatial layout while using the other finger as an anchor.

In the evaluation proper, the participants were shown two floor plans on GraViewer one after the other. For each floor plan they were asked to explore it and to let the experi-

Participant	Training 1	Training 2	Corrected Floor Plan A					Uncorrected Floor Plan B				
	Exp.	Exp.	Exp.	Q1	Q2	Q3	Q4	Exp.	Q1	Q2	Q3	Q4
P1	40	210	112	49	24	125	66	446	50 (X)	44	77	33
P2	60	120	127	5	5	61	30	127	5 (X)	5	22	61
P3	32	225	363	4	13	62	26	342	3	21	40	37
P4	23	305	164	1	7	45	97	167	1	2	36	103
			Uncorrected Floor Plan A					Corrected Floor Plan B				
P5	94	321	215	51	5	140	89	110	10	41	184	30
P6	45	170	210	10	13	89	170	424	7	12	38	152
P7	70	160	95	1	30	210	65	420	20	11	80	40
P8	67	83	46	38	35	71	37	219	27	28	29	65

Table 1: User study results: Task completion times for training and test graphics are given in seconds (X: incorrect answer, Exp: Exploration, Qn: Question n).

menter know when they felt that they knew the basic layout. They were then asked to answer the following questions:

Q1 How many bedrooms are there in the house?

Q2 Point to where the kitchen is on the floor plan?

Q3 What is the shortest route from the kitchen to the bathroom and which rooms do you pass through?

Q4 Can you describe the layout of the house?

The time taken for the initial exploration and to answer each question was recorded as well as the answers to the question.

The two floor plans (A and B) used in the evaluation are shown in Figure 6. Both were generated by GraFloor from a floor plan in the corpus from [18] with English room names replacing the original French names. Two versions of each floor plan were created. The first version was corrected manually in GraAuthor to remove any errors arising from mis-recognition while the second was the actual output from GraFloor and contained several transcription errors. The reason for using the results from both semi-automatic and fully automatic transcription was to better understand whether any difficulties in understanding were the result of the presentation on GraViewer or whether they were the result of imperfect transcription by GraFloor. Participants were either shown the the corrected version of A and then the uncorrected version of B or else the corrected version of B then the uncorrected version of A.

Finally the participants were asked how the tools and presentation could be improved and whether they thought automatic transcription of on-line floor plans would be useful.

Results

For the first two participants the fully automatic transcription of Floor Plan B did not include the overview and different bedrooms where all labelled as "Bedroom" while the manually corrected transcription of Floor Plan A included a textual overview and also bedrooms were labelled "Bedroom 1", "Bedroom 2", etc. As a result participants P1 and P2 did not correctly determine the number of bedrooms in the uncorrected Floor Plan B. In response to this and their suggestions we modified the transcription module to automatically generate a textual overview and also to label bedrooms "Bedroom 1", "Bedroom 2", etc. This modified version was used in the subsequent experiments.

We see that apart from this problem all participants could correctly answer the 4 questions for both the fully automatic

and manually corrected floor plans. The task completion times are given in Table 1.

Feedback from the participants included suggestions to help navigation: "It will be good to query where the rooms are located relative to the finger" and "There might be different sound themes which can represent the rooms better" and also "I like the overview summary. You can not picture the image in your mind, but it gives you an idea of what to expect."

All participants felt that access to on-line floor plans would be useful and that the tool provided a good approach to provide this. In fact one participant described how they had recently been searching for a house to buy and had paid Vision Australia to provide tactile versions of the on-line floor plans that he/she was interested in. They felt that this tool would have been very useful.

5.2 Quality of Automatic Transcription

Recognition of floor plans is a difficult task and so it is not realistic to expect 100% accuracy for even well-structured images. One of the great difficulties in robust floor plan recognition is that there are many different graphical conventions: for instance walls can be drawn as thick black lines or as two thin parallel lines.

In order to verify the effectiveness of GraFloor we evaluated the errors made in the automatic transcription of one benchmark dataset of floor plan images. The dataset [18] is composed of 90 images that contain floor plans with some basic furniture objects and rooms labelled with textual descriptions. These images have been provided by an architectural office and cover a period of more than ten years.

In our first experiment we evaluated how well GraFloor recognised rooms including open plan rooms. This was the primary purpose of the tool, to provide the blind user with an understanding of the rooms in the floor plan and their spatial arrangement. The results are summarized in Table 2. For each room (including regions in an open-plan area) in each floor plan in the corpus we manually determined if the room/region had been correctly recognised and included in the generated image. Full recognition meant that the room and its regions (if an open-plan room) was correctly recognised and displayed at the correct location in the generated image; partial recognition meant that though the room's extent was determined the kind of room was not (usually because the associated text was too short to be recognised). We see that GraFloor is quite effective and recognises nearly 95% of rooms at least partially. This is

Recognition	Open Plan Room	Non-Open Plan Room
Full	85% (82)	77%(718)
Partial	8%(8)	15%(138)
Failed	7%(6)	8%(77)

Table 2: Percentage and total number of items correctly recognized by the floor plan analysis system for the entire dataset.

Recognition	Door	Window	Sink	Tub	Cupboard	Toilet	Table	Shower	Stairs
Full	78%(162)	50%(90)	29%(9)	50%(5)	67%(37)	60%(12)	86%(6)	0%(0)	0%(0)
Partial	16%(33)	17%(31)	13%(4)	20%(2)	0%(0)	15%(3)	14%(1)	75%(3)	90%(9)
Failed	6%(12)	33%(60)	58%(18)	30%(3)	33%(18)	25%(5)	0%(0)	25%(1)	10%(1)

Table 3: Percentage and total number of items correctly recognized by the floor plan analysis system for 20 randomly chosen images.

comparable with other state of-the-art floor plan recognition software [5] where a detection rate of 94.76 % and a recognition accuracy of 94.29 % are reported. It is important to notice that [5] is not aimed at recognizing the type of room and does not address the subdivision of open-plan areas.

In a second experiment we evaluated how well other objects in the floor plan were recognised. In this case because of the time required to manually check all components of a floor plan we only considered 20 randomly chosen floor plans. The results are summarized in Table 3. Full recognition meant that the object was correctly recognised and displayed at the correct location in the generated image; partial recognition meant that it was either recognised as another kind of object, e.g. window as a door, or in the case of stairs its extent was not fully determined. While the results are not outstanding recognition rates are reasonable for most objects apart from showers (which are confused with sinks) and stairs because only some of the stairs in a staircase were recognised. We are currently working on improving this: however for the purpose of providing an overview, accurate object recognition is not required.

6. CONCLUSION AND FUTURE WORK

We have presented a new model for providing blind users with accessible on-line graphics: automatic transcription from an on-line image to an accessible graphic designed for presentation on a touch screen device such as an iPad that can be viewed using audio feedback with semi-automatic transcription as a back-up. We believe this approach can be used to handle a wide variety of information graphics including statistical charts, tables, simple maps, chemical structural diagrams etc. Thus it has the potential to dramatically improve accessibility of a wide range of on-line information graphics.

We have described and evaluated the use of this model for on-line floor plans. An empirical evaluation of our approach shows that it can provide an accurate overview of the layout of the rooms in a floor plan and that this presentation can be used by a blind person to understand the layout. The great advantage of our approach is that it allows a blind person to immediately view building and floor plans that they encounter on-line and does not require printing expensive tactile graphics. This potentially provides much greater access to such information and removes a significant barrier to travel by blind people.

However, while the initial evaluations presented in this paper are very promising there is some additional work required before we have a full-fledged system ready for mainstream use:

- We need to investigate how to include panning and zooming so that we can present large floor plans of buildings and public spaces.

- We would like to provide more advanced interaction searching and filtering features so that users can, for instance, be presented with a route to their desired destination.

- The recognition system has been tailored to work with a particular corpus of examples and the graphic conventions used in that corpus. We are now modifying the system so that given some initial training it can robustly identify floor plans using other conventions.

Acknowledgements

7. REFERENCES

[1] Web content accessibility guidelines (WCAG) 2.0. 2008.

[2] A. Barducci and S. Marinai. Object recognition in floor plans by graphs of white connected components. In *21st International Conference on Pattern Recognition (ICPR)*, pages 298–301. IEEE, 2012.

[3] Z. Cattaneo et al. Imagery and spatial processes in blindness and visual impairment. *Neuroscience and Biobehavioral Reviews*, 32(8):1346–1360, 2008.

[4] G. Daunys and V. Lauruska. Sonification system of maps for blind. In *Universal Access in Human-Computer Interaction. Ambient Interaction*, pages 349–352. Springer, 2007.

[5] L.-P. de las Heras, S. Ahmed, M. Liwicki, E. Valveny, and G. Sánchez. Statistical segmentation and structural recognition for floor plan interpretation. *International Journal on Document Analysis and Recognition (IJDAR)*, 17:221–237, 2014.

[6] S. Demir, D. Oliver, E. Schwartz, S. Elzer, S. Carberry, K. F. Mccoy, and D. Chester. Interactive sight: textual access to simple bar charts. *New Review of Hypermedia and Multimedia*, 16(3):245–279, 2010.

[7] L. Ferres, G. Lindgaard, L. Sumegi, and B. Tsuji. Evaluating a tool for improving accessibility to charts and graphs. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(5):28, 2013.

[8] R. P. Futrelle, M. Shao, C. Cieslik, and A. E. Grimes. Extraction, layout analysis and classification of diagrams in pdf documents. In *2013 12th International Conference on Document Analysis and Recognition*, volume 2, pages 1007–1007. IEEE Computer Society, 2003.

[9] J. Gardner and V. Bulatov. Scientific Diagrams Made Easy with IVEO. *Computers Helping People with Special Needs*, 4061:1243–1250, 2006.

[10] N. A. Giudice, H. P. Palani, E. Brenner, and K. M. Kramer. Learning non-visual graphical information using a touch-based vibro-audio interface. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, pages 103–110. ACM, 2012.

[11] C. Goncu and K. Marriott. Gravvitas: Generic multi-touch presentation of accessible graphics. In *Proc. INTERACT'11*, pages 30–48. Springer, 2011.

[12] C. Jayant, M. Renzelmann, D. Wen, S. Krisnandi, R. Ladner, and D. Comden. Automated tactile graphics translation: in the field. In *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, pages 75–82. ACM, 2007.

[13] N. Kaklanis, K. Votis, and D. Tzovaras. Touching openstreetmap data in mobile context for the visually impaired. In *Proceedings of the 3rd Workshop on Mobile Accessibility-ACM SIGCHI Conference on Human Factors in Computing Systems*, 2013.

[14] A. R. Kennel. Audiograf: a diagram-reader for the blind. In *Proceedings of the second annual ACM conference on Assistive technologies*, pages 51–56. ACM, 1996.

[15] S. Landau and K. Gourgey. Development of a Talking Tactile Tablet. *Information Technology and Disabilities*, 7(2), 2001.

[16] J. Lazar, A. Allen, J. Kleinman, and C. Malarkey. What frustrates screen reader users on the web: A study of 100 blind users. *International Journal of human-computer interaction*, 22(3):247–269, 2007.

[17] Llados, Josep and Rusinol, Marcal. Graphics recognition techniques. In D. Doermann and K. Tombre, editors, *Handbook of Document Image Processing and Recognition*, pages 489–521. Springer London, 2014.

[18] S. Mace, H. Locteau, E. Valveny, and S. Tabbone. A system to detect rooms in architectural floor plan images. In D. S. Doermann, V. Govindaraju, D. P. Lopresti, and P. Natarajan, editors, *Document Analysis Systems*, ACM International Conference Proceeding Series, pages 167–174. ACM, 2010.

[19] J. A. Miele, S. Landau, and D. Gilden. Talking TMAP: Automated generation of audio-tactile maps using Smith-Kettlewell's TMAP software. *British Journal of Visual Impairment*, 24(2):93–100, 2006.

[20] S. Millar. *Space and sense*. Psychology Press, 2008.

[21] H. Petrie et al. TeDUB: A system for presenting and exploring technical drawings for blind people. *Computers helping people with special needs*, pages 47–67, 2002.

[22] H. Petrie, F. Hamilton, N. King, and P. Pavan. Remote usability evaluations with disabled people. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1133–1141. ACM, 2006.

[23] B. Poppinga, C. Magnusson, M. Pielot, and K. Rassmus-Gröhn. Touchover map: audio-tactile exploration of interactive maps. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pages 545–550. ACM, 2011.

[24] Round Table on Information Access for People with Print Disabilities Inc., Australia & New Zealand. *Guidelines on Conveying Visual Information (2005)*, 2005.

[25] M. Spindler, M. Weber, D. Prescher, M. Miao, G. Weber, and G. Ioannidis. Translating floor plans into directions. In *Computers Helping People with Special Needs (ICCHP)*, pages 59–66. Springer, 2012.

[26] H. Takagi, S. Saito, K. Fukuda, and C. Asakawa. Analysis of navigability of Web applications for improving blind usability. *ACM Trans. on Computer-Human Interaction*, 14(3):13, 2007.

[27] S. Ungar. Cognitive mapping without visual experience. *Cognitive mapping: Past, Present and Future*, pages 221–248, 2000.

[28] S. Wall and S. Brewster. Feeling what you hear: tactile feedback for navigation of audio graphs. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1123–1132. ACM, 2006.