Assisting Visually Impaired People with Mobility through Technology in the Age of Context

A Survey

Martin Goldberg
CUNY Graduate Center
New York, NY. USA
mgoldberg@gradcenter.cuny.edu
2015

Committee Members:

Professor Zhigang Zhu, The City College of New York, CUNY (Advisor)
Professor Zhanyang Zhang, College of Staten Island, CUNY
Professor Ioannis Stamos, Hunter College, CUNY
Dr. William H. Seiple, Lighthouse Guild

A Literature Review Submitted to
the Graduate Faculty of Computer Science in Partial Fulfillment of
the Requirements of the Degree of Doctor of Philosophy,
The Graduate Center of City University of New York
Acknowledgements

This work is supported in part by the National Science Foundation (Award # EFRI-1137172), a grant from Provost’s Office at the City College of New York, and a CUNY Graduate Center Science Fellowship (to Goldberg).
Abstract

The objective of this survey is to understand how new technologies in the “Age of Context” can best help individuals with disabilities – in particular, the visually impaired – with mobility issues. Technical writers, Robert Scoble and Saul Israel, recently have defined the five forces of the Age of Context – Mobile Devices, Sensors, Location Awareness, Social Media, and Data. Through the prism of their framework, we would like to discover what has been achieved thus far with assistive technologies that aid with mobility. Each of these five forces will be considered first on their own, then an integration of them will be considered. When looking at each ‘force’, we will particularly consider how it helps or hinders the blind community and, in some cases, consider how they are used indirectly and may benefit the community with mobility issues. From this, we would like to establish where the research is at present, and then we would like to consider how we might take this further.
# Table of Contents

**Acknowledgements** ii  
**Abstract** iii  

## Introduction 1  
- Background and Motivation ........................................... 1  
- Definitions ............................................................ 2  
- Overview and Organization of the Survey .......................... 3  

### 1 Mobile Devices 4  
1.1 Mobile v. Portable Devices ........................................... 6  
1.2 Tablets ................................................................. 8  
1.3 Smartphones ........................................................... 9  
1.4 Wearable Technologies ............................................... 13  
  1.4.1 Google Glass ....................................................... 15  
  1.4.2 Smart Watches ................................................... 17  
1.5 Discussion ............................................................. 18  

### 2 Sensors 19  
2.1 Sound Sensors .......................................................... 19  
2.2 Image Sensors .......................................................... 20  
2.3 Depth and Distance Sensors .......................................... 21  
2.4 Localization Sensors ................................................... 22  
2.5 Force Sensors .......................................................... 23  
2.6 Some Additional Sensors .............................................. 24  
  2.6.1 Accelerometers .................................................... 24  
  2.6.2 Gyroscopes ......................................................... 24  
  2.6.3 Magnetometers .................................................... 24
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6.4 Inertial Measurement Unit (IMU)</td>
<td>24</td>
</tr>
<tr>
<td>2.6.5 Antennas</td>
<td>25</td>
</tr>
<tr>
<td>2.7 RGB-D Cameras</td>
<td>26</td>
</tr>
<tr>
<td>2.8 Using the Body as a Sensor</td>
<td>27</td>
</tr>
<tr>
<td>2.8.1 Argus II</td>
<td>27</td>
</tr>
<tr>
<td>2.8.2 Brainport</td>
<td>28</td>
</tr>
<tr>
<td>2.8.3 Haptic Output and Tactile Graphics</td>
<td>28</td>
</tr>
<tr>
<td>2.9 Discussion</td>
<td>30</td>
</tr>
<tr>
<td>3 Location Awareness</td>
<td>30</td>
</tr>
<tr>
<td>3.1 GPS based Localization</td>
<td>31</td>
</tr>
<tr>
<td>3.1.1 When GPS is not Available</td>
<td>32</td>
</tr>
<tr>
<td>3.2 Augmented Environments</td>
<td>32</td>
</tr>
<tr>
<td>3.3 Computer Vision Algorithms</td>
<td>33</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>35</td>
</tr>
<tr>
<td>4 Social Media</td>
<td>35</td>
</tr>
<tr>
<td>4.1 Social Media v. Data</td>
<td>36</td>
</tr>
<tr>
<td>4.2 VIP Usage of Social Media</td>
<td>37</td>
</tr>
<tr>
<td>4.3 Leveraging Social Media: BlindSquare</td>
<td>37</td>
</tr>
<tr>
<td>4.4 Expanding on the Social Media Paradigm: Viz-Wiz and Beyond</td>
<td>38</td>
</tr>
<tr>
<td>4.5 Discussion</td>
<td>38</td>
</tr>
<tr>
<td>5 Data</td>
<td>39</td>
</tr>
<tr>
<td>5.1 Big Data</td>
<td>39</td>
</tr>
<tr>
<td>5.2 “The Miracle of Little Data”</td>
<td>40</td>
</tr>
<tr>
<td>5.3 Audio and Braille Enhanced Devices and Maps</td>
<td>40</td>
</tr>
<tr>
<td>5.4 Enhanced GPS locators</td>
<td>41</td>
</tr>
</tbody>
</table>
5.5 Image Data

5.5.1 Geo-Tagging

5.5.2 Vision Algorithms for Location

5.5.3 3D Data Used for Orientation

5.6 Discussion

6 Integration of the Five Forces

6.1 Beyond Tactile Maps

6.2 An Indoor Navigation System

6.3 The BlindNavi Project

6.4 The Wayfindr Project

6.5 Sensor Enhanced Mobile Devices

6.6 A Interactive Approach

6.7 Discussion

7 Conclusions and Discussions

7.1 The Impact of the Five Forces on VIPs

7.2 Looking Forward

References
List of Figures

1. Manduchi/Coughlan use of a mobile camera [52] ........................................ 6
2. Smartphones displaying Personal Assistants - essential for VIPs – Left to Right: Microsoft: Cortana; Apple: Siri; Google (Android): Google Now – (September 2014) ................................................................. 8
3. PERCEPT-II as an example of the interactive nature of a good navigation system ................................................................. 10
4. Lin’s head worn prototype; An example of the need to better design wearable technology ................................................................. 13
5. Google Glass Explorer edition (September 2014) ................................................................. 15
6. The Sekiguchi team laser sensor ................................................................. 22
7. Overview of RGB-D Mapping. The algorithm uses both sparse visual features and dense point clouds for frame-to-frame alignment and loop closure detection. The surfel representation is updated incrementally. [31] ................................................................. 26
8. Subject using the Kinect-Enhanced Cane [47] ................................................................. 27
9. Tactile Map use in Brock’s user study ................................................................. 29
10. An example of the Bag of Words concept (Source: [20]) ................................................................. 42

List of Tables

1. Prominent Intelligent Personal Assistants ................................................................. 10
2. GPS solutions for VIPs ............................................................................................. 31
Introduction

They’re everywhere. The five forces of context are at your fingertips when you touch a screen. They know where you are and in what direction you are headed when you carry a smartphone. They are in your car to warn you when you are too close to something else. They are in traffic lights, stores and even pills.

... All five of these forces – mobile, social media, data, sensors and location – are enjoying an economic sweet spot.


Scoble and Israel have defined these ‘five forces of the Age of Context’ and we are to consider how these ‘forces’ have been, and might be, used to help with developing assistive technologies with respect to mobility. Mobility is a key issue in allowing the visually impaired to become active and productive members of society.

Background and Motivation

Scoble and Israel hail the ‘perfect technological storm’ of the Age of Context, citing five forces: mobile devices, sensors, location-based services, social media, and data. (Note: We have rearranged the order of the ‘forces’ as we will later refer to them.) The authors are primarily concerned with what this ‘Age of Context’ means for privacy. Thus in the later sections of the book, they show how these five ‘forces’ come together in a way that may undermine our sense of privacy; therefore, what is described in the book is that these forces acting together are quite a storm. However, it is our intention to look at these forces both individually and together and ask how we might harness this ‘storm’ in our efforts to better the mobility of visually impaired people (VIPs).
VIPs constitute approximately 2.5% of the population\(^1\) are of all races and genders and are thus a considerable enough part of society. They experience needs particular to their specific disabilities; however, it appears that when assistive technologies are developed, they are adopted by the general public for use as well. A case in point is that of DragonSpeak – designed to aid blind writers with ease of producing documents – it has been quite widely adopted in the medical profession as a means of transcribing notes taken during patient sessions\(^2\). Thus studying and understanding how these forces affect us, and how they might be harnessed to aid with assistive technologies, has much to offer.

**Definitions**

In order to make things clear, we think it wise to consider a few definitions as they may have different meanings to researchers of other fields.

- **VIP**: visually impaired person - either with limited sight, but still having some vision; or, having so little or no sight as to be considered unable to use natural vision in any meaningful way.

- **Mobility**: the area of dealing with a VIP’s ability to ‘get out and about’, preferably without needing a helper’s assistance.

- **Wayfinding**: deciding a correct or optimal path to get from one location to another.

- **Navigation**: the process of managing the traversal decided by wayfinding, being able to avoid hindrances or other dangers - such as holes or edges.

- **Orientation**:


\(^2\) This has been witnessed and discussed with at least four doctors who talk about it as common practice
– defining the correct direction to follow;

– or, defining the direction the VIP is facing.

**Overview and Organization of the Survey**

In this review, it is in our interest to look at these forces as they apply to our work in assistive technology research, in particular as it applies to helping visually impaired persons (VIPs) to manage with mobility.

Let us first explain the five forces as described in *the Age of Context*, and meanwhile briefly point out our intentions to apply them to help VIPs:

- **Mobile Devices**
  
  Here the authors talk about the ever increasing number of smartphones and tablets that are everywhere today, then go on to discuss wearables, in particular Google Glass. Written in 2013, this book does not refer to the smartwatches that are becoming more available as we write this review.

- **Sensors**
  
  The authors portray sensors as small devices that can keep track of most things, they are primarily concerned with the collection of personal data. In our case, we will consider the use of sensors to make people’s life easier.

- **Location Based Services**
  
  Here the authors concern themselves with Apple and Google Maps, noting that Google wants to know everything about you, whereas we are considering aiding VIPs with mobility issues, knowing where the subject is and having a reference to his/her surroundings can be very helpful.
• Social Media

“Social media is essential to the new Age of Context. It is in our online conversations that we make it clear what we like, where we are and what we are looking for.”

We would like to see what impact the social media could have to VIPs - good or bad.

• Data

The authors talk about how much data there is, and how online, searchable data is increasing in ways we could not have conceived of before. Noticeably, they concern themselves with the way in which so much data is handled and searched through in what they call the miracle of ‘little data’. As we will see, finding the proverbial needle in the haystack is more doable than previously believed. How we could make use of them for serving the VIPs?

We will look at each force individually, concentrating on the tangible forces in the first place, looking at mobile devices in Section 1 and, location awareness in Section 3. Then we will look at the two forces that drive this ‘Age of Context’: social media in Section 4 and data in Section 5. In the later part of this review we will consider how these forces have been integrated in Section 6 summing up our understanding in the last section, Section 7.

1 Mobile Devices

Mobile devices are a key aspect of assistive technologies. This was recognized back in 2000 when the author described the impact mobile devices would have on VIPs; ostensibly, freeing them of the ‘tether’ to their homes. This allows them to go about while still having the security of being able to contact others if needed. Indeed, mobile devices were far more ubiquitous in the communities they ‘untethered’; early mobile devices such as
laptops and cell phones freed both VIPs and non-VIPs alike. Scoble & Israel, however, do not consider laptops and other typical hand-held devices to be relevant to their ‘Age of Context’ paradigm unless they are connected seamlessly to the Internet and have sensors. However they are also part of the mobile devices that benefit VIPs.

In their 2008 presentation [53] Mandudchi and Coughlan already preempt our discussion as they note:

*Computing power, communications and Internet access are becoming increasingly untethered from the desktop and moving to the realm of portable, wireless technology.*

*Devices such as mobile (cell) phones and PDAs (personal digital assistants) have become affordable and ubiquitous, and offer not only substantial computational power but also telephone and Internet access, as well as a variety of sensors such as cameras, GPS and RFID readers.*

*... there is enormous potential to harness their capabilities for use in assistive technology. As of yet [2008], however, this potential remains largely untapped, with very few commercially available systems of this type.*

In their article [52] published the same year, the authors present a use of the cell phone with camera as a device held on a necklace for detecting the way in front of the user (Fig 1). This is but one, early, example of the use of a cell phone with a camera as a vision processor.

This discussion was further extended in 2012 [54] where the authors argue that computer vision is the optimal solution to aiding VIPs with mobility. They cover a range of relevant topics, such as the variance in VIP needs and abilities, a few algorithms and -
in the main - the use of the camera enabled mobile phone as the Electronic Travel Aid (ETA) of choice.

Much has changed in the proceeding years and one can find a veritable plethora of articles regarding applications one can use for guidance and monitoring on mobile and/or portable devices – including cell phones. In the following paragraphs we will consider some of those most pertinent to aiding VIPs.

1.1 Mobile v. Portable Devices

Both mobile and portable devices are small enough to be carried around, as the adjectives ‘mobile’ and ‘portable’ imply. A portable device can include a cane, a calculator and a wristwatch. Mobile devices are a subset of portable devices that specifically have both a modicum of computation power and can communicate with other devices. Thus, the most prolific mobile device of the current moment – i.e. Summer 2015 – is the smartphone. Also, in this group are tablet and laptop computers. Traditionally, these devices have been called hand-held devices – defining their portability by use of hands. However, as has become apparent over the past few years, non-hand-held portable computation
devices with connectivity are entering the mainstream – most noticeably is the Google Glass\(^3\) which is worn as a pair of glasses, rather than carried; the Argus II [2] project uses glasses-mounted cameras which communicate with a device worn on the body before communicating with the eye itself; and, most recently, Apple has showcased it’s Apple-Watch - released April 2015\(^4\) – which is a small smartphone-like device worn like a wristwatch and synced to the iPhone, presumably allowing the user the same access as to the iPhone i.e. activating the accessibility functions of the device through the Personal Assistant (on the Apple this is Siri, on Google Android devices - Google Now, and Microsoft has Cortana) (Fig 2).

In recent years, smartphones, equipped with cameras, accelerometers and accessibility software have been adopted by blind users - how exactly needs to be researched - and we should also ask: “How can we help here?” and, “Is this the same for low-visioned individuals?”

Before we review a number of research papers, I would like to say a few words about my personal experience. As a VIP, with low-vision caused by central scotoma, I have found it difficult to use smartphones, as they have very small screens; text is often too small to read reliably; and, many features hinder rather than aid me. However, Android phones now have much improved magnification, activated by triple tapping the phone to toggle in and out of full-screen magnified mode – this has greatly improved my personal ability to use smartphones.

Tablets, also mobile – only larger and so more cumbersome due to having to use both hands to utilize – are more useful to me. Having stretch screen technology helps, especially when I am able to stretch to a usable size, keep it there while navigating the screen (scrolling, drag and drop, etc.), then be able to quickly resize. Desktop applications

\(^3\) On January 15\(^{th}\) 2015, Google announced that the Google Glass project was to be shelved, see the BBC announcement: [http://www.bbc.com/news/technology-30831128](http://www.bbc.com/news/technology-30831128)

for VIPs such as Zoom-Text and J.A.W.S. are not readily available on mobile devices.

VIPs with little to no vision tend to adopt the iPhone – ostensibly a plain glass surface with a single easy-to-find button – as their device of choice. They can operate ‘Siri’ by pressing the button; then, with this personal assistant they can utilize VoiceOver to access all the applications they have.

Figure 2: Smartphones displaying Personal Assistants - essential for VIPs – Left to Right: Microsoft:Cortana; Apple: Siri; Google (Android): Google Now – (September 2014)

1.2 Tablets

A recent paper in CHI (2015) has explored use of tablet computers by VIPs. I have – as a low-visioned individual – expressed a preference for tablets over smartphones because of ‘real-estate’ i.e. the amount of screen space to work with. The paper deals with VIPs reliant on haptic interfaces, speculating that, given a better access to a full QUERTY keyboard, the VIP could better use the tablet. They have developed a product called SpacialTouch, which purports to help speed up use of a touch-screen keyboard. It is an

\[5\] acronym for **Job Access With Speech**
interesting way to go, but for mobility issues where hands-free is the by-word, this only highlights where this technology fails the VIP.

### 1.3 Smartphones

In an ERCIM workshop paper [1] written in 2000, predating the advent of the iPhone, Abascal and Civit looked at how mobile phones were changing the world, noting that even blind users would adopt them as they provide a sense of autonomy and the implicit security of having an always present device to make calls on when and if the need arose. Thus, even 15 years ago, it was evident that mobile devices - in particular mobile telephones - were going to enable VIPs to get out and about. The iPhone reached the market in 2007, and the personal assistant, ‘Siri’ [6] which gives the blind such good access to the iPhone and its Accessibility was released to the public in 2011. In April, 2014 Microsoft released Cortana [7]. As voice activated and voice propelled interaction with smart phones becomes more prevalent, one can see the blind community becoming most used to it, and one may presume that this means, with the right kind of software and hardware add-ons, that this will greatly improve blind mobility.

The PERCEPT [24] and later PERCET-II [25] projects are a good demonstration of how mobile technology can be utilized for aiding VIPs with navigation and general mobility (Fig. 3). The PERCEPT-II system is built to run on a mobile device - while they talk specifically about smartphones and tablets, this type of system is conceivably transferable to other devices. The local system - *i.e.* the app on the mobile device - interfaces with the environment - presuming the environment can communicate, and then with the back-end databases that hold the essential data for the process.

---

7 [https://en.wikipedia.org/wiki/Microsoft_Cortana](https://en.wikipedia.org/wiki/Microsoft_Cortana)
What has made the smartphone the premiere assistive technology in mobility thus far is the rise of the Intelligent Personal Assistant. Major players in this arena and the related information are summarized in Table 1.

<table>
<thead>
<tr>
<th>Company</th>
<th>Product</th>
<th>OS</th>
<th>Revealed</th>
<th>Marketed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google</td>
<td>Google Now</td>
<td>Android</td>
<td>2011</td>
<td>2012</td>
</tr>
<tr>
<td>Apple</td>
<td>Siri</td>
<td>iOS</td>
<td>2012</td>
<td>2013</td>
</tr>
<tr>
<td>Microsoft</td>
<td>Cortana</td>
<td>Windows Phone 8.1</td>
<td>2014</td>
<td>2015</td>
</tr>
</tbody>
</table>

For VIPs, it would seem that it was Siri that was the real breakthrough, as it could be activated by a press of the only button on the interface; then having it activated, the user is able to activate the other apps they can use.

The DUB Group\footnote{DUB = Design; Use; Build} at Washington State in Seattle has done work on understanding how computers are – and can be – utilized by VIPs and persons of other disabilities. In their 2009 paper \cite{40}, they undertook a comprehensive survey of 20 individuals with visual or...
motor impairments and had them keep dairy records of how they used mobile devices – in particular their smartphones. 19 of their participants kept diaries and the study intended to discover needs of VIPs and others discovered:

- Their subjects actively researched new devices;
- People with disabilities typically rely on multiple devices “in the wild”;
- Found use of specific devices to be situational
- Had difficulties with crowded spaces
- Were subject to weather and lighting related problems

A number of individual diary entries highlight the fact that this technology is not easy to use in all circumstances; and – in particular – when using the phone as an assistive device, using it as a phone was impossible. While this survey does not specifically relate to mobility issues, it does highlight issues when developing solutions aimed at addressing this issue. Later work \[41, 4\] has been aimed at getting more out of smartphones, including useful map reading and usable touch-screen keyboards. Later work by this group includes the Virtual Braille project \[36\] which intends to create braille readers for mobile phone touch screens through haptic output.

A pilot program \[35\] for navigating the Science Museum in New Delhi highlights both the positive and problematic aspects to smartphone use in navigation:

- **Positive aspects:** a simple ubiquitous device - the cell phone needs a simple app that then can read data provided by the environment and pass this information on to a VIP.

- **Problematic aspects:** In order for this to work, the environmental infrastructure needs to be
1. built: implementation in any truly usable fashion requires investment on the part of the stake-holders; this includes monetary investment, planning and logistics; and, underlying it all, the buy-in of the institutions affected by the implementation.

2. maintained: once built, corrections, updates and general maintenance is required. There is nothing worse than a redundant system that no-one has responsibility for. Ultimately no-one can predict the long-term, but at least for the medium-term, we should expect a commitment to keeping up a system. This requires money, man-power and the continued interest of those who put the plan into place in the first instance.

This encapsulates the underlying issue with this approach. The paper uses a museum as the environment - which typically builds environments such as are required here, and maintenance is also almost assured. However, this is not the case in most need scenarios, and - ironically - navigating the science museum is not a typical activity for a VIP.

It has become clear that the power of a smartphone is such that it is a fully capable computer, with access to data and services through Internet capabilities built into the systems, whose main problem is the amount of power it can use before it needs to be recharged.

The smartphone is the basic device, easy for a VIP to use in that simple activation of the personal assistant gives the user voice control over the device. In his thesis defence - working under Bigham\textsuperscript{9} - Zhong shows what can be done on the smartphone – both through photography and through computer vision; his message, however, is clear, the smartphone, while having some use, is still limited in how it services the VIP community. However, it is mainly an issue of software that then communicates externally and reports back to the user, so it is not necessarily a smart cellphone that is required. In the next

\textsuperscript{9} The lead on the VizWiz project, a crowd utilizing mobile application for blind users
subsection we will consider wearable technologies, which proport to make the smart device interaction hands-free – a major advantage for VIPs.

1.4 Wearable Technologies

In a way, smartphones - the prominent assistive technology at this time - are old news. They have been around for almost a decade. The new technologies that are appearing on the scene are wearables - in particular, Google Glass, smart watches and others - are appearing and must be considered. If mobile technologies untethered the user from the home, and smartphones gave them access to all kinds of assistive technology, wearables are the natural next step being that they provide the user a hands-free environment - effectively allowing VIPs use of the hands for other things, such as canes.

Figure 4: Lin’s head worn prototype; An example of the need to better design wearable technology

Lin [49] proposes a wearable technology to be worn on the head; the idea is to deal with ‘hanging’ object avoidance. The prototype image is of a box with cameras facing
out (Fig 4) - here we are reminded of Brabyn’s caution\textsuperscript{10} and realize that in order for this technology to be adopted, it should be downsized and produced as an easily wearable device. This is no small issue, as if the VIP is unable or unwilling to wear the device, then the solution becomes mute.

At the MIT Media Lab \textsuperscript{77} work has been done on a finger reader; ostensibly, this is a camera worn on the knuckles and pointed to the tip of the finger. As such, the idea is to give a VIP\textsuperscript{11} the ability to point at some text needed while “on the go”\textsuperscript{12} and thus a mobility aid that is hands-free.

Knighten et. al. \textsuperscript{45} have considered using a wristband to recognize social gestures. Here they place accelerometers onto the wristband to gauge how the hand is moving. It appears that they wish to understand the user, not someone they are looking at, perhaps to then convey this information by another method?

In their article \textsuperscript{89}, a wristband is presented. The wristband proports to achieve the hands-free promise of wearable technologies. An admirable start, unfortunately it could only be used by connecting it to an Arduino single-board computer that in turn was tethered to a wall socket for power. We are consider a similar – if somewhat more complex – approach utilizing the Raspberry Pi v. \textsuperscript{13} To solve the issue of power, we are contemplating a battery pack to power the device with the mobility expectancy of battery-life to exceed an hour.

Work in the City College Visual Computing Lab (CCVCL) - to which we belong - has been done with regard to wearable technologies as aids for VIPs. In particular work done by Palmer, et al \textsuperscript{63} deals with wearables that utilize range sensors which detect how far they are from solid objects. They then relay data to vibrating cells worn on the body and

\textsuperscript{10} In his presentation “A Lifetime of Mistakes”, the author cautions us against creating tools without keeping the user in mind - especially as they are not usually the young energetic people used in trial experimentation; and as pertains to this particular issue - design must consider how things look

\textsuperscript{11} It would seem that some vision is required for this idea as control of what the finger is pointing at seems essential.

\textsuperscript{12} As the paper’s title suggests

\textsuperscript{13} Released February 15, 2015
produce a vibration increasing in intensity as the object gets closer. Work done by Khoo, et al. [44] is able to test in a virtual environment the measurements of these wearables and other sensor based technologies that may be developed to aid VIPs.

1.4.1 Google Glass

Google Glass (Fig. 5) is a wearable device, as the name suggests, it is worn as a pair of glasses, only instead of lenses, it has a glass prism with sensors hanging above one eye, typically the right eye.

Wikipedia states its functionality as [15]:

- Touchpad: A touchpad is located on the side of Google Glass, allowing users to control the device by swiping through a timeline-like interface displayed on the screen. Sliding backward shows current events, such as weather, and sliding forward shows past events, such as phone calls, photos, circle updates, etc.

- Camera: Google Glass has the ability to take photos and record 720p HD video.

- Display: The Explorer version of Google Glass uses a Liquid Crystal on Silicon (LCoS), field-sequential color, LED illuminated display. The display’s LED illumination is first P-polarized and then shines through the in-coupling polarizing beam splitter (PBS) to the LCoS panel. The panel reflects the light and alters it to S-polarization at active pixel sites. The in-coupling PBS then reflects the S-polarized areas of light at 45° through the out-coupling beam splitter to a collimation reflector at the other end. Finally, the out-coupling beam splitter (which is a partially reflecting mirror, not a polarizing beam splitter) reflects the collimated light another 45° and into the wearer’s eye.

A recent article [3] is of interest, though not strictly related to mobility, it nevertheless deals with the Google glass device as a wearable assistive technology. They use it here to help with non-verbal cues derived from a face-to-face conversation. What is therefore of interest is that, given how they use the Glass, it is very likely possible to adapt their basic algorithm and introduce other methods more conducive to the issue of mobility. In particular, face recognition comes to mind as a way to aid both low-visioned and blind VIPs to identify who is in their vicinity which tangentially aids with the mobility issue.

At Nevada U. [22] a product named Headlock embraces the new wearable technology; designed for use with Google Glass and other head worn devices, Headlock helps extend the range of a cane for distant object finding - such as doorways - the VIP still uses the cane to direct him/herself, only now the head-worn device can help discover elements of interest that are beyond the cane’s reach, guiding the VIP to within cane detection range.

In a demonstration paper [29] the researchers propose adding fluorescent optical markers to help their blind navigator for Google Glass to be able to call their ego-motion function to be able to calculate a user’s rotation and translation in a dynamic environment. Poleg [67] derives a similar idea using the video stream capture with temporal segmentation to
calculate the motion of the user.

1.4.2 Smart Watches

In his May, 2014 literature review [38], Johnson explores the advantages and disadvantages of smart watch technology; He studies the *Pebble* and draws conclusions about possible use of this kind of device. Although no specific mention is made with regard to VIPs and mobility, this review sheds light on the personal area network (PAN) idea, which - when combining with the other ‘forces’ may prove to be a good approach to enabling VIPs with mobility. In his proposal paper [37] written at the same time, Johnson outlines his project for communicating between smart watch and smartphone.

The authors and designers of ‘Duet’ [14] - an application for coordinating interaction between smart wear (watches, glasses, bracelets, etc.) - have studied the benefits and problems found in use of smart wear and smartphones, mainly the issue of line-of-sight, but it certainly bodes well for VIPs in that the technology is here and with some improvement should be able to be turned into a very usable system with much potential going forward.

Twyman et. al. [82] have considered how best to adapt smart watches; in this instance, that use the smooth glass face of the device as a haptic output device. As smart watches use touch-screens, they are using speech technology [4] on a new device. Thus they are attempting to make the output device hands-free.

In a recent survey of smart watches and their potential, [83] reveals how the Apple Watch can help VIPs, much in the same way one could get from an iPhone, only now it is hands-free.

---

16 using Siri
1.5 Discussion

Wright and Keith’s article proclaims: “If the Tech Fits, Wear it!”. Their focus is on medical uses of wearables, but the value to us here is their description of what is anticipated in terms of wearable technology in the next decade. When discussing sensors, they are primarily concerned with bio-scanning technologies. What we need to consider, therefore, is how we can utilize what is out there and reconstitute their application to our needs - that being assistive technology.

Ultimately, when considering the environment, we must have intelligent systems which can listen for and communicate with a random user when that user comes into the vicinity. Scoble and Israel mention the Tom Cruise movie “Minority Report” while futuristic, this concept is what we have in mind - an interactive environment that can guide the user safely through whatever is required.

From some recent workshops, presents there understanding of using wearables for VIP mobility issues; while somewhat inconclusive in their findings, noting that more research on how VIPs would work with this technology, they nevertheless condone use of wearables as the next step in assistive technologies for mobility.

In recent work we have been looking at the communication issues with combining intelligent environments with wearable personal area networks. The literature definitely points towards wearable technology; we want to expand on this creating more robust solutions that will tailor to a wider population.

---

17 2002, Starring Tom Cruise. There is a scene where Cruise’s character is running through a mall – as he does so, the mall is scanning his retina to identify him and are sending

18 We would hope to have something that would appeal to the VIP and thus hopefully get them to adopt the approach.
2 Sensors

To read Scoble and Israel, they think that sensors are ubiquitous. In a conversation with a certain Google person\textsuperscript{19} on assisting the visually impaired people with sensors, I was asked, why not just throw a million [or was it a billion] dollars at the problem? Indeed, science fiction is all to ready to presume that computers - through the use of sensors - will be able to do almost anything. This presumption is – to say the least – presumptuous. Understanding what sensors are, how they might be applied, and what is required to have them perform as required is a study within itself, we will consider them in brief here.

A good definition of a sensor is:

\ldots a transducer whose purpose is to sense (that is, to detect) some characteristic of its environs. It detects events or changes in quantities and provides a corresponding output, generally as an electrical or optical signal.\textsuperscript{20}

Many sensors are used in our everyday world: We can break this up into several categories\textsuperscript{21} sound, images, depth and distances, localization, force, etc. In the following subsections, we will discuss each of the sensors including the above, particularly for assisting VIPs. When we discuss sensors, we may include those sensors that obtain sensory information from the environment, and those sensors of users to accept information.

2.1 Sound Sensors

Devices used to transduce sound to and from electric current are pervasive, and quite central to devices created to assist VIPs. Microphones\textsuperscript{15} for audio input, and speakers for audio output are the most typical of these. When sight is limited, one of most efficient means of delivering instructions, and

\footnotesize
\textsuperscript{19} Casual conversation with a Google salesperson
\textsuperscript{20} from http://en.wikipedia.org/wiki/Sensor
\textsuperscript{21} Based on slides taken from: Lecture on Sensors from Stanford U.
of receiving response is audibly. Desktop computers have had applications built to aid in this respect, and with the profusion of smaller digital sensors, these are becoming readily available on even the smallest devices, and smart phone technology has certainly adapted these into their firmware and software. As far as the totally blind are concerned, audio is often the optimum source of reference available to them; however, audible output – in particular – can be problematic, as when multiple instances of audio are being broadcast simultaneously, there can be confusion, not to mention, disruption; and, since hearing is the most important sense to VIPs, impeding one’s hearing by covering the ear with earphones can be as much a problem. Thus when considering audio as a means of output, we must be careful to consider the environment in which it is to be used. Bone-conducting phones would be a partial solution, as well as haptic feedback.

2.2 Image Sensors

Image sensing is done mainly through cameras [81], the digital versions of which being devices that capture light in various degrees of resolution and can transduce the resultant matrix of captured light into sets of numerical values. Typically these “images” are stored such that we can access individual pixels and get their color intensities with the ability to relate a given pixel to the matrix in which it exists. This said, the sub-disciplines of image processing, computer graphics and computer vision exist to best decide how to utilize this input. Cameras are used in most solutions to aid VIPs; audio non-withstanding, when attempting to understand a given situation, images are most often the best descriptors of what is in the immediate vicinity, the great challenge is, then, to interpret this data in a usable way. In their work [84], Vázquez and Steinfeld have pointed out a serious issue with VIPs and camera sensing, that being that – due to the lack of good vision – both of humans and machines, accurate camera pointing is not assured. They note many instances in which VIPs might utilize a camera, we are chiefly interested in the mobility related issues. Their solution utilizes a process known as focalization which is handled
between a user and a system which processes input data and rectifies the camera angle so as to get an accurate image.

2.3 Depth and Distance Sensors

Depth and distances are typically measures utilizing either light or sound. In both cases the sensor emits some energy and collects it on the rebound, calibrating the time taken for the signal to return and, thus, based on either the speed of light, or of sound, the distance (or depth) can be calculated. Long range sensors such as Sonar and RADAR are used to locate objects in the vicinity; for shorter, more exacting distances, infra-red, laser and ultrasound are used with varying degrees of accuracy over different distances. Lasers, for instance can measure quite well and accurately over long distances (e.g. 50 meters), typically accurate in terms of millimeters, but are relatively more expensive. Ultrasound is useful over medium distances (within 10 meters), but, accuracy in both distances and angles is limited to knowing something is close or far within a few centimeters. Infra-red is optimal for measuring accurate depth, but is severely limited in range (within a few meters) and effectiveness under certain lighting conditions.

Some work has been done utilizing sound producing devices used by totally sightless VIPs, they tend to be expensive and are not readily adopted. Some devices utilizing sound as a feedback approach proved to be sensually overwhelming with prolonged use, and as a rule they are not readily utilized by the blind community. A laser range finder typically needs to be coupled with other feedback approaches to be properly utilized, but as a data source are most likely the best distance measures, utilizing time-of-flight to calculate depth to a very acute measurement, with millimeter accuracy.

Sekiguchi et. al [76] propose a wearable system to aid VIPs with navigating T-junctions and forks while traversing indoor passages. In their system, they apply depth data collected by a laser sensor (Fig. 6) which utilizes a time-of-flight scan on a 2D plane; the calibration is modified by an on-board gyroscope on the printed circuit inside the sensing
device.

Figure 6: The Sekiguchi team laser sensor

LIDAR, and other laser scanning techniques have been around for sometime (since the 1970s), and 3D point cloud generation is well researched [79], we will discuss 3D models in §5 below.

2.4 Localization Sensors

We will consider the larger topic of Localization in the subsequent section relating to Location Awareness in Section §3. From a sensors perspective, we concern ourselves with GPS and Dead Reckoning.

GPS - Global Position System - is a means by which electromagnetic signals are received by a decoder that can calibrate when and from how far away the signal was sent. By triangulation of 3 such signals, a 2-dimensional location can be calibrated. With 4 signals, a 3-dimensional location can be calculated. These measurements depend on both direct line-of-sight and very precise time measurements (using atomic clocks); as the signals come from Earth’s orbit, minor discrepancies in the data can cause serious error in the readings. Brady [9] comments on how dependant VIPs are on GPS for not getting lost, so much so that they will carry a GPS device used in cars in addition to the GPS app they have on their smart phones. As we will discuss later, location awareness is a major issue with VIP mobility issues.
Another form of locating oneself is through “Dead Reckoning”: this being where the sensors are able to pinpoint (even in 3-D space) where a device is, and how it is moving and oriented, we will look at this in the following subsections.

2.5 Force Sensors

Here we relate to transducers that receive pressure and respond accordingly. Most typically, with electronic systems, these are switches and buttons, which then, by way of programming, activate some functionality. So, keyboards and keypads are good examples of pressure based sensing. Another much used pressure sensing technique is the touch screen: pressure is measured either by surface acoustic waves which when disturbed calibrate where the touch happened; this is similar to how, when clicking on a mouse, the place on the screen is calibrated and a response accorded.

Another form of this is the use of micro dielectric cables which form a mesh across a surface and can respond to touch without the internal calibration; this is a much more expensive option and not usable with glass surfaces, but for non-glass surfaces, this is what is utilized. Recent advances in flexible cables hold the promise of tactile clothing.

One of the best means of device accessibility among sightless VIPs is through touch. Braille, used for reading and writing was developed based on this; today, electronic braille is available. As a means of input and output to and from electronic systems, the haptic approach is at once the best for VIPs to adopt: unless also touch impaired, the totally blind use their fingers to explore and control, so having buttons, or small impressions and embossed artifacts, are usable, if proper training is done and a good understanding of where each button is in relation to each other. One reason the Apple iPhone has become so popular amongst the VIP community is that it only requires a press of an easily found lone button on the panel to activate the voice control, and thus is easy to use.
2.6 Some Additional Sensors

Here we will consider devices that have been created to sense more complex data than humans can perceive, which have – over time – been digitized and miniaturized such that they can be implanted into mobile devices as small as smart phones.

2.6.1 Accelerometers

An accelerometer is a device used to measure proper acceleration – in other words, the g-force being applied to an object. For human-computer interaction (HCI) applications, \[15\] uses these sensors to keep track of how the hand is moving.

2.6.2 Gyroscopes

Gyroscopes were created to deal with moving environments, such as on board a sea-going vessel, where pendulums and other such devices suffered from error due to constant motion. Originally rather large devices, they have been modified into computer chips and much miniaturized. These, along with accelerometers, are used in smart phones to aid with rectifying orientation.

2.6.3 Magnetometers

As their name suggests, they are used to measure either the strength of a magnetic field or to how fast the field is moving. In smartphones, these are used to aid with compass based applications.

2.6.4 Inertial Measurement Unit (IMU)

IMUs are combinations of accelerometers, gyroscopes and magnetometers. They have been used to help collect data on devices use for SLAM (Simultaneous Location And Mapping) in robotic applications, where the IMU helps calibrate the roll and yaw of a device in motion\[19\]. Kang \[42\] made an interesting use of IMUs (see Section \[3.1.1\]).
In general, none of these sensor devices are of particular help with current assistive technologies; however, as we consider how we might integrate a network of sensors into a workable system; the IMU may prove a useful tool for keeping track of individual device position and orientation.

2.6.5 Antennas

Antennas are used to transduce electric signals into radio waves and vice-versa. They are necessary for any wireless communication. Antenna size is dependant on wavelength, which in turn is inverse to frequency. As a result of the FCC allowing the MHz and GHz spectrum for WiFi, Bluetooth and other such technologies, this has allowed for very small antennas that can be inserted into dongles and imprinted on circuit boards.

Bluetooth technology was originally built to facilitate data transfer, but it is also capable of reporting the distance between transmitter and receiver. Many companies are producing beacon-based indoor navigation systems: for example, Nokia and IBM. An American-Israeli company named SPREO, founded in 2013, also provides a Bluetooth-beacon-signal-based indoor navigation service, and claims an accuracy of 1.5m. The system is used in many industrial areas: as in hospitals, for events and trade shows, in museums, in retail stores, for transportation and on university campuses.

Radio Frequency Identification (RFID) uses the electromagnetic field wirelessly to transfer data for automatic identification, and to track tags attached to objects. As opposed to barcodes, RFID tags do not necessarily have to be within the line of sight of the reader, and may be embedded into the tracked objects.

WiFi\textsuperscript{22} is a local area technology that allows an electronic device to exchange data or to connect to the Internet using 2.4GHz UHF or 5GHz SHF radio waves. Besides the data exchange function – similar to Bluetooth – WiFi is also used by many organizations and companies (such as Google) in localization and navigation devices.

\textsuperscript{22} also spelled ‘Wi–Fi’ or ‘Wifi’
Antennas are in use for all types of devices, primarily – where we concerned – in enabling mobile devices to communicate with the outside world. When considering use of RFIDs and Beacons for localization, antenna size and available wavelength become important considerations.

### 2.7 RGB-D Cameras

RGB-D cameras utilize both what we would consider a standard full color digital camera, – the RGB referring to the three color bands for which intensities are captured (Red, Green, Blue), and a depth detector, usually a combination of an infra-red projector and a sensor that captures the returning light waves as they bounce off the surface being measured. A more detailed discussion of this will be handled in Section §5.

![Figure 7: Overview of RGB-D Mapping. The algorithm uses both sparse visual features and dense point clouds for frame-to-frame alignment and loop closure detection. The surfel representation is updated incrementally.][31]

Here the challenge is to understand what can be achieved by RGB-D cameras, in our recent paper [48], we explored the possibility of using depth data to augment the regular image of a face to better perform face recognition. This is one of many such RGB-D applications released in the past few years.

An interesting application of the RGB-D device – the Kinect – [47] has been shown to be partially useful in detecting an elevator in an indoor environment. The setup (see Fig. 8) shows at once the viability of the project, and – recalling Brabyn [8] – demonstrates the inherent awkwardness of this approach, suggesting that we need to consider a less cumbersome approach to assistive technology.
2.8 Using the Body as a Sensor

Nature uses sensors all the time, and the human body is no different in this respect. What we now know, though, is that it is not the sense organs (eyes, ears, etc.) that process the information into comprehensible data, rather it is the brain itself that does the processing, compiling input from the organs. So work on utilizing this fact has been done with projects such as Argus II and Brainport. It should be noted here, that for totally blind persons, even the barest minimum of sight restoration is tremendous, and we are still in the pioneering stage of such experiments, so the hope that this can be realized is not without merit.

2.8.1 Argus II

Argus II is an example of how sensors are utilized to aid the totally blind: it is a system that allows for a prosthetic eye by allowing a camera to take a image in front of the user; through a computer the image is processed, being transmitted to an array of 6x10 electrodes that then are attached to the retina where the signal is attempted to be
transduced from the wires to the cells of the retina. This has met with some success and is now commercially available. [2] [18]

2.8.2 Brainport

The Brainport system tries to utilize the tongue as a sensing device, contemplating that since the visual cortex is not being fully utilized due to the lack of data being streamed from the eyes, some form of data recognition might be achieved by having the body use the tongue to “see”. This utilizes camera fed data to a pad of electrodes (between 100 and 600 depending on the testing unit) that stimulate the tongue in a 3cm by 3cm area at 30 frames per second. This has yet to prove itself [17].

2.8.3 Haptic Output and Tactile Graphics

Another approach to aiding the totally blind is to provide haptic output - by this we refer to output that can be sensed through touch. Most commonly known to us is braille which is a form of writing that can be interpreted by a user running their fingers over a preset encoded alphabet that is embossed onto a surface. This non-digital technology has found its digital equivalence in electronic braille keyboards and output devices [66].

Tactile output has always been a consideration for VIPs as it is a means by which they can learn to get outputted information. Prior to computerized tactile solutions, embossed material representing some output – for instance braille – was the medium of choice. Electronically generated tactile output is useful in that, unlike its physical counterpart, it can be readily redrawn; thus, where physical tactile graphics have the problem of taking time to produce and being incapable of updating, the electronic option can – in much the same way electronic documents best physical documents in this respect in everyday usage.

23 We had the opportunity in our lab to have a student who was a subject of a clinical trial of this product, from him we learned that many issues subsist using this device, including extremely limited data and heavy fatigue after relatively short usage
VIPs, however, rely on haptic interaction for much of their ‘classic’ interfacing with text and graphics. It has been commented\(^\text{24}\) that electronic tactile output is far less legible than physical tactile output as it relies on electronic pulses being emitted and are harder to distinguish easily. It is possible that electronically generated graphics may be rendered on tactile – non-electronic – screens\(^\text{25}\) and thus provide a solution. A decade ago, Miele et al. \(^\text{58}\) presented an idea for electronic tactile interfaces. They cover the early development of physical tactile maps and details how they envision using a product named TMAI\(^\text{26}\) which would render audible as well as tactile output. In a more recent paper \(^\text{10}\) a similar idea\(^\text{27}\) is presented – Fig. 9 is used for Brock’s user study\(^\text{28}\).

Work at Arizona State University \(^\text{56}\) and at CUNY City College \(^\text{44}\) has been done to create vibrating wearable technology that, again, uses the body as sensor in response to vibrations, usually generated by the data collected by another sensor, such as an infra-red depth detector to pass signals to the wearer giving them contextual information about where they are and how close they might be to certain objects. Similar work by Roberto Manduchi’s Lab \(^\text{23}\) has been done.

\(^{24}\) Dr. Giudice of U. Maine at a presentation to CUNY, July 22, 2015  
\(^{25}\) As mentioned in \(^\text{1}\), some research into electronically driven tactile screens where the reader has an actual physical tactile output is being considered.  
\(^{26}\) Acronym for Tactile Maps Automated Production  
\(^{27}\) References on this paper do not cite Miele’s paper  
\(^{28}\) See Brock’s homepage at http://people.bordeaux.inria.fr/abrock/
2.9 Discussion

Sensors are very much a part of everyone’s daily life, and digital sensors are indeed becoming ubiquitous; VIPs typically are people who either lack or seriously lose use of one of the body’s five most basic sensors – the eyes – and as far as mobility is concerned, vision plays a central part to being able to find one’s way, to orient oneself, and to assure oneself of the safety of going places. Thus, computer science in particular, and more specifically computer vision, is the optimal place to research the use of sensors and understand how they should be utilized to better aid VIPs with mobility issues.

3 Location Awareness

When dealing with location awareness, there are three situations which concern the VIP:

- *Where I am and where I need to go*
  This entails planning the route and following it from beginning to end.

- *Where I am in reference to where I need to be*
  Consider the situation where the VIP got off-track and needs to get back to the planned path.

- *Where I am*
  Dealing with the situation where the VIP has become totally lost

Scoble and Israel \[75\] name this ‘Location Based Services’, and concentrate on how advertisers are able to customize their message (especially political messages) to location specific demographics. Underlying their thesis is the complaint that Google et al. can ‘know’ where you are all the time.

Where VIPs are concerned, this is of paramount and positive importance. We look at mobility from two perspectives that concern the average VIP. The first of these is
navigational directions: i.e., making sure that they can get along, correctly going where they need to go and safely – avoiding objects, people etc. The other big concern is the spatial context: i.e. knowing where they are in reference to the greater world. This is no limited to VIPs, everyone needs to know where they are and use visual clues where possible to get their bearings. For a VIP who gets lost, this is quite traumatic, not being able to orient oneself by way of visual clues, this is evidenced in the Kane survey [40].

3.1 GPS based Localization

Research in making the GPS capability of smartphones available to VIPs has been done back in 2007 [33]. Thus we can ascertain that anything that applies GPS - either in directing or following a VIP, is already in use; or, at the very least, can be applied if required.

Table 2: GPS solutions for VIPs

<table>
<thead>
<tr>
<th>Product</th>
<th>Software</th>
<th>Device</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loadstone GPS</td>
<td>Symbian OS</td>
<td>Symbian (Nokia)</td>
<td>Since 2006</td>
</tr>
<tr>
<td>LoroDux</td>
<td>Java ME</td>
<td>non-specific</td>
<td>utilizes OpenStreetMaps, under development</td>
</tr>
<tr>
<td>Mobile Geo</td>
<td>Windows Mobile</td>
<td>Windows enabled</td>
<td>Seamless integrated with Mobile Speak</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Devices</td>
<td></td>
</tr>
<tr>
<td>BlindSquare</td>
<td>iOS</td>
<td>iPhones</td>
<td>Based on FourSquare combined with OpenStreetMaps</td>
</tr>
<tr>
<td>Trekker Breeze</td>
<td>Firmware</td>
<td>HumanWare</td>
<td>Device not app for smartphone previously Trekker (2003) requires preset routes to work</td>
</tr>
<tr>
<td>BrailleNote GPS</td>
<td>Firmware</td>
<td>HumanWare</td>
<td>PDA, allows note taking and recording of points of interest</td>
</tr>
</tbody>
</table>

Table 3.1 lists the available devices and applications for VIPs using GPS technology.

GPS technology is based on geo-synchronized satellites orbiting the Earth. For a two-

---

29 Global Positioning System
30 Compiled based on https://en.wikipedia.org/wiki/GPS_for_the_visually_impaired

31
dimensional location, 3 signals must be received, for a three-dimensional reading, 4 signals are required. As GPS is based on direct line-of-sight signals, it is not totally reliable. In urban environments where VIPs need assistance the most, signal deflection can be an issue and any interference – such as cloud cover – can affect true readings. As a general area localization instrument it is still useful to VIPs.

3.1.1 When GPS is not Available

When GPS is not available or its signals are bad, mostly in an indoor environment or an urban setting, technologies based antennas are used for localization, such as Bluetooth beacons, RFID and WiFi. These have been discussed in Section 2.6.5.

Kang introduces SmartPDR [42] in which he adopts robot location methods to humans, reckoning that the human issue is more complex than with robots, as the movement made is more dynamic and arbitrary. Pedestrian Dead Reckoning (PDR) - as Kang shows - has been much discussed. GPS, typically can only be used outside where a satellite ‘line of sight’ can be achieved, and WiFi triangulation requires the necessary infrastructure to allow for calibration of the subject’s location. Instead, the proposal is to utilize the sensors inside the latest smartphones to essentially track and recalibrate the movement of the subject, thus keeping track of the subject. In lieu of an augmented environment, this may be useful in aiding VIPs with helping know where they are in relation to where we need them to be. They also use an Inertia Navigation System (INS) which employs IMUs.

3.2 Augmented Environments

Another approach to dealing with navigation of VIPs is to augment the environment to assist them. A non electronic example of this is placing embossed placards on walls – typically with braille which can be read: the New York MTA transit system has some of these in several of their major transfer hubs. Use of such implementations requires, first
of all, knowledge of their being there, and then, training in their use so as to be effective. Oftentimes this requires that the VIP memorizes their position, the path needed to be followed and points along the way to assist in the journey.

In the past two decades, smaller antennas – made available due to higher frequencies being allowed by government authorities – have allowed for wireless communication as a solution.

As a practitioner who interacts with academics, Mr Joe Cioffi has related to us how he has created his tactile maps for institutions. He is currently promoting the idea of utilizing the iBeacon\(^{31}\) as a means to enhance the navigation of the prescribed paths by providing inter-journey destinations, the iBeacon giving both guidance and distance information – thus enhancing the aid provided by his maps.

As another example, [46] has researched the use of RFIDs and highlights some products, including Step-Hear\(^{32}\) which allows a user to ‘discover’ up to five preconfigured locations and once selected, a location will provide an audio signal to guide the user to the required destination.

Daly et al. [16] have taken this a little further, experimenting with RFIDs embedded into concrete paving stones, thus creating a traceable path.

In similar fashion, [62] has looked at the various systems that can compensate for no GPS indoors, considering NFC\(^{33}\) as a potential solution.

### 3.3 Computer Vision Algorithms

There is no direct connection between VIP adaptation of computational vision based solutions and the basic assistive solutions that may help VIPs with location issues. As we see with GPS (see above), that VIPs are able to utilize GPS has little – if any –

---

\(^{31}\) iBeacon is an Apple\(^{TM}\) product that applies Bluetooth technology to allow for the transmission of minimal data of a calculable distance.

\(^{32}\) http://www.step-hear.com/step-hear-overview.htm

\(^{33}\) Near Field Communication
bearing on the development of GPS for general use.

In Section 6 we will see how certain systems utilize computer vision algorithms (or similar such algorithms) to achieve their intentions. In his recent review, fellow researcher Feng Hu [32] summarizes computer vision algorithms that may be of use in aiding VIPs; while his focus is on indoor localization, these algorithms may lend themselves to outdoor navigation as well:

- 2D solutions - those techniques that do not explicitly infer 3D information of a scene;

- 3D solutions - those reconstruct or use 3D information of a scene and perform 3D to 3D matching;

- 2D-to-3D matching - those hybrid methods that use 2D images to match 3D models

- Emerging technologies - such as mobile computing, cloud computing and their impact to vision based localization approaches.

We will look at the details of some of these methods in Section 5. Of particular interest is the technology developed by [20, 21, 73] in which the ‘bag of words’ concept – ostensibly a way of searching images based on feature generated ‘textons’; similar to advanced textual searching in documents, only utilizing image feature generated base search terms. Here, it has been suggested that when accessing tagged data sets generated by Google images, we are able to glean sufficient context from our image to ‘know’ where we are.

Thus, what is proposed, is a computational intensive approach to guiding VIPs (or robots) which demand access to large data sets. This has the intrinsic problem of power capability in a mobile solution; in our research, we agree with [32] that a distributed approach to applying these solutions needs to be found.
3.4 Discussion

Localization is an extremely important aspect of the mobility needs for VIPs. Most VIPs are able to ‘get around’ with their white canes and, if fortunate enough, guide dogs. However, it is a quality of life issue as well as a employment related issue that VIPs are able to go beyond their comfort zones. Discussion with Mr. Joe Cioffi has highlighted some issues with localization, such as the problem with spacial awareness amongst the congenitally blind, and deaf-blind; and, then there is the issue of ‘hallway traversal, and similar issues. Motivated VIPs will adopt new technologies and figure them out, but for the greater population of VIPs that are not as motivated, more utilizable systems may be the way to proceed. Augmentation seems an easy enough solution, but how to go about it effectively, maintain it over long periods and similar problems need to be addressed.

4 Social Media

To Scoble and Israel — social media is what makes the Age of Context the thing it is. They portray inordinate amounts of data being passed around and then through giants like Google and Facebook, having all our information out there for everyone to see. We suggest that this is overly alarmist, and we will see if and how this might help in our research.

According to Dictionary.com Social Media is defined as:

noun, ( usually used with a plural verb) Digital Technology

websites and other online means of communication that are used by large groups of people to share information and to develop social and professional

---

34 Click and Go website: [http://www.clickandgomaps.com/about-us/](http://www.clickandgomaps.com/about-us/)
35 We use the term Hallway, but in truth, any long stretch that needs to be traversed without landmarks to aid with keeping track of where one is constitutes a problem
contacts:

Many businesses are utilizing social media to generate sales.

4.1 Social Media v. Data

At this point it may be necessary to clearly define the difference between social media and data as we tend to confuse the two definitions.

A little over a decade ago, the new dynamic way of programming web pages to allow for the input of client-side data to the server and having the server code behave according to what was input at the client, was coined Web 2.0\footnote{The term was popularized by Tim O’Reilly and Dale Dougherty at the O’Reilly Media Web 2.0 Conference in late 2004 – source: \url{https://en.wikipedia.org/wiki/Web_2.0}}. Since the early days and with the advent of Facebook, YouTube, Twitter and the like, where more robust interactions between client and server opened up the possibility of connectivity between separate individual clients to each other through the server has extended Web 2.0 into what we now call social media\footnote{The actual person who coined the term is disputed}. Indeed, email, originally developed for a small community of academics and professionals to pass messages between one another is the earliest form of social media, and as web-enabled email has proliferated, it has become a more inherent part of the fabric of web interfaces, known now as social media.

Data, on the other hand, is what can be stored on some device. As far as electronic data is concerned, any set of on and off signals that can be interpreted by a machine to represent something more than an on/off signal is data. We will look into what this means to us in Section \S 5.

In essence, we could say, data is the information that can be passed between individuals and social media is the means by which this data is shared among many users over the internet.
4.2 VIP Usage of Social Media

Two Facebook writers [87] conducted an online survey of some 50,000 VIPs. The main takeaway from their paper is that “Visually impaired people engage with online social networks just as everyone else does”. They concentrate on how Facebook usage is conducted, and do not deal with mobility issues. Their mention of WAI-ARIA[39] is of interest, and this article provides a great basis for development of social media to aid VIPs with mobility issues.

Some papers [69, 90] gave looked specifically at social media use on mobile phones, with Ye’s paper concentrating on news reading, arguing that navigation can be improved with news feeds in ad-hoc situations.

4.3 Leveraging Social Media: BlindSquare

Blindsquare, with Apps available for iPhone and iPad, has been talked about on sites aimed at VIP community[40] The product page[41] gives us the following explanation of how Blindsquare works:

> When BlindSquare has determined your location using your iOS-devices GPS capabilities, it will look up information about your surroundings on FourSquare and Open Street Map. Employing unique algorithms, it will then ascertain the information most useful to you and speak it in a clear synthetic voice.

So, we see that Blindsquare is built upon two social media systems that already exist. Foursquare [60] which is useful for saving destinations, and Open Street Map [27, 28] which are generated from crowd-sourced inputs. Then with technologies including Text-

---

[41] http://blindsquare.com/about/
To-Speech (TTS), BlindSquare aims to provide information that blind and visually impaired people need to travel independently.

4.4 Expanding on the Social Media Paradigm: Viz-Wiz and Beyond

Jeffrey P. Bigham [6, 7] purports to take social media a whole step beyond basic as he hooks into the internet and gets volunteers to act as guides to the VIPs who are sending near real-time image query to the service for identifying objects and answer specific questions related to the query images. The ‘wiz’ refers to the Wizard-of-Oz approach that utilizes human input in lieu of true artificial intelligence which is not yet available to the system. A team in the City College Visual Computing Laboratory has developed a prototype test system for streaming smartphone video of a blind user to crowd volunteers in order to provide step-by-step navigation guidance [61, 43].

4.5 Discussion

It would seem that social media holds some promise in aiding VIPs with mobility issues, but they are yet to be truly realized. Relying on crowd-sources works so far as crowds are willing to participate; and – in the case of Open Street Map, etc. – a concerted effort needs to made to ensure true coverage, and can we truly depend on crowds to always fully update the data? Much has to do with the motivation behind any crowd based service, and it remains to be seen if these solutions will bear fruit in the long run. Indeed Brady’s survey [9] suggests that the VIP community is reluctant to utilize social media, and thus some re-education may be necessary before pushing forward social media solutions to VIP mobility.
5 Data

When Scoble and Israel discuss data, they talk of the ubiquity of data; how being everywhere it endangers our privacy, and thus they consider ‘The Age of Context’ to be a threat rather than a promise. As VIPs go, data no more relates directly to them than it does to non-VIPs.

Let us then meditate on what data means to us, data as we discussed in Section §4 is electronically represented bits that have some meaning, however, beyond considering 10 as an integer or ‘Sam’ as an alphanumeric value, without some sort of context to give meaning to these values, they are meaningless. However, if given context, they themselves then become context to other data, and so forth.

5.1 Big Data

Big data refers to the vast quantities of data that are out there, are increasing everyday and are multiple modalities, structured or unstructured, video, images, text, etc. They are supposed to be there for use – not necessarily by everyone, but for those willing to dig or to pay - for sure. Computing power has improved exponentially over the past several decades and storage has become progressively cheaper and easier to access. With ‘Cloud’ technologies, data has become more readily accessible from almost anywhere, and to anyone with access.

We are not overly concerned with the details here, save to say that we must consider Big Data as a tool we might utilize to help our VIPs with their mobility issues. This is particularly true when a lot of multimedia data, such as geo-tagged videos of scenes and activities, images of scenes and faces, tweets and chats with rich information on locations of interests, are posted everyday on the web every corner of the world, thus providing a great potential to use them for assisting mobility of the blind and visually impaired.
5.2 ‘The Miracle of Little Data’

This ‘miracle’ is coined by Scoble and Israel [75] where they assume any data can be found, no matter how large the dataset. It is a bit of an overstatement; however, [70] shows us how we can utilize graph database technology. Through data retrieval techniques, including data mining techniques, textual data – at least – has been ‘searchable’ for quite a while. In is not in the purview of this paper to discuss this in depth.

As for image data, that which most concerns us, there has been progress in searching relevant images - any quick search of Google Images will attest to that. However, making sense of untagged image data is a little harder.

In a 1998 paper[34], and again in a recent paper [55] both teams talk about the need for better data in aids for VIPs. Both papers suggest sonic output as a means of outputting usable data to VIPs, but both papers emphasize the need for better data exploitation solutions for VIPs. Indeed, this may hint at why the white cane and guide dog are still the only real travel aids utilized by VIPs.

5.3 Audio and Braille Enhanced Devices and Maps

Digital maps (such as Google Map and various mobile GPS navigation systems) are widely used for normal sighted people for travel planning and step-by-step directions. However there are great challenges for visually impaired to use such systems if not enhanced. ClickAndGo Wayfinding [42] provide services to visually impaired people with audio-annotated maps with important landmarks for navigation. Brock [10] has incorporated both audio tags as well as braille labels onto her tactile maps, this is a solid example of how extra data can be added to a device such as a tactile map. It certainly is a step in the right direction, but will it be adopted? This still remains to be seen.

---

5.4 Enhanced GPS locators

OpenStreetMaps is a project developed in Britain [28] which collects data from users and builds up a database of street data; it is used by VIPs in the form of the iOS application: BlindSquare. In a later paper [27] the authors evaluated the success of crowd sourced data for OpenStreetMaps.

5.5 Image Data

Our main interest is in how we can harness processed image data in our quest to aid VIPs with mobility. Several technologies, while not specific to aiding VIPs, hold promise in this endeavor.

5.5.1 Geo-Tagging

A form of identifying images is geo-tagging, a system by which the camera – e.g. on an Apple™ iPhone – collects GPS data as an image is captured, and thus the image has this geographic information embedded into the image for future extraction. In [74], the team describes how geo-tagging is accomplished through detecting and matching repeated patterns in an urban environment, and [51] provides a survey on how geo-tagging is applied to multimedia applications. These have potentials to be used for assisting VIPs: geo-tagging is obviously useful information, and might be incorporated into any system that wishes to use images as a means of localizing a subject.

5.5.2 Vision Algorithms for Location

An example of how big data has been utilized is in [91], where the authors utilize the vast dataset of Google street images, and by using SIFT features, they attempt to identify where a new image belongs based on a comparison with the labelled images.

In another concept, the ‘Bag of Words’ (BOW) idea has been proposed [20] (Fig. 10) and may be useful to us. We can break an image into smaller blocks of image data, each
Figure 10: An example of the Bag of Words concept (Source: [20])

several pixels wide and high. The basic idea is that given small enough image sections – called **textons** – search algorithms, not unlike text-based searches, can be carried out on images. The textons would be found by matching against the image then through a histogram of resulting matches, a signature of sorts would be defined for an image. This still needs some serious development for its use in VIP applications.

### 5.5.3 3D Data Used for Orientation

In a book chapter on the use of 3D point clouds, [39], they look at the limitations of the data collected in 3D point clouds.

At time of writing, I approached one of the authors – Prof. Ioannis Stamos – with a question, “Given a 3D model and an 2D image taken in the vicinity of the 3D model, could we figure out how the image plane falls across the 3D model?”, to which he responded that it was an issue of calibrating the camera pose, which led to the idea of being able to orient a VIP in a given environment. His particular solution to this problem was based on his work with deriving lines in both the 2D image and the 3D point cloud and matching the through clustering along vanishing point orientations. [50, 79, 80]

### 5.6 Discussion

While there is no definitive winners in applying a vast amount of computer vision algorithms in VIP tasks, the possibilities described in this section do hold an optimistic
outlook to getting there. The visual algorithms that traditionally have been geared to robotic applications will be able to provide location of the users, identities and locations of objects, and activities in the environments, all vital for mobility of VIPs.

6 Integration of the Five Forces

At this stage we want to have a look at how all the five forces come together. Scoble and Israel [75] called the accumulated effect of their five forces a ‘technological storm’ - a harbinger of our loss of privacy. However, we take a very different view on this, seeing the potential imbuement in the integration of these five forces as it applies to aiding VIPs with mobility issues.

6.1 Beyond Tactile Maps

Anke Brock has looked at the problem of navigation and orientation through the use of tactile maps. Noting the limitations of the traditional maps, she has explored ways in which to enhance the usage in a way that provides better context to the VIP traveller. In a Wizard of Oz study [11] the use of haptic response delivered to the user by way of vibrating wristbands was explored. In her later paper [10] the use of audio response to touch input on a touch-sensitive electronic tactile map, providing audio labels was explored. In this work, maps for location aware is integrated with hatpic and auditory senses for feedback.

6.2 An Indoor Navigation System

In a PhD thesis paper, [5] describes the system used to aid VIPs with indoor location. They use a three-layered system, each providing some context for the next layer, thus enabling accurate localization and navigation support.

In the first level, GPS is utilized to identify the building in question - meta-localization.
The second level utilizes a barometer to identify the floor level of the subject and then applies Wi-Fi triangulation to get a rough location on the floor. The final level of the system utilizes a digital camera (in their diagram, this is a smartphone worn on the chest from a necklace), combined with motion sensors in the building which then give a high accuracy localization. This is a complex system that integrates multiple sensors (GPS receivers, barometers, Wi-Fi localizer, digital camera and motion sensors) and it might be more readily applied if less infrastructure could be utilized.

6.3 The BlindNavi Project

The BlindNavi project[13] has some basic ideas for integrating mobile and location awareness with social media. If the BlindSquare (See Section §4.3) application utilizes crowdsourced street mapping with a Foursquare engine, this project (designed for a Taiwanese environment) has students who are VIPs use previous data collected by themselves or other students to aid them with navigating certain pathways. This idea was based on interviews the researchers did with a group of VIPs and a mobility instructor. An interesting point made is that VIPs are loath to go out on unknown paths without company. This product desires to provide that company.

6.4 The Wayfinder Project

Wayfinder[43] is a collaboration between the RLSE[44] Youth Forum and a company called ustwo[45]. In their words:

Using existing iBeacon (BLE) and smartphone technology, and drawing on the experience of simulated journeys with limited sight on the network, ustwo developed a location-tracking system that provided audio guidance via bone-

---

43 Wayfinder project: [http://www.rlsb.org.uk/campaigns/wayfinder](http://www.rlsb.org.uk/campaigns/wayfinder) the page contains an interesting promotional video discussing the background process toward arriving at the project
44 Royal London Society for Blind People
45 global digital product design studio
conduction headphones. Refining and testing the product with the RLSB Youth Forum throughout the exercise led to real advances in understanding the challenges faced by young vision-impaired people, and the growing possibilities of an affordable tech solution.

As a summary, this work integrates BLE technology, smartphone sensors and bone-conduction headphone suitable for young generations.

### 6.5 Sensor Enhanced Mobile Devices

In [64] authors have described their implementation of a “portable pedestrian guidance system”, which integrates two kinds of distance sensors (i.e., infra-red and ultra-sound) similar to the ideas discussed in Section 1.4. Here they create a PCB [46] which is attached to a smartphone. They then integrate infra-red receivers with ultra-sonic sensors and the phone’s GPS to interact with LEDs augmented into the environment.

### 6.6 A Interactive Approach

In [88] the authors state:

> Talking Points 3 is a location- and orientation-aware smartphone-based system that provides information to users about nearby points of interest in the environment.

The 3 in Talking Point 3 tells us that this is part of an on-going series of projects. However, if a #4 version has been created, we haven’t been able to find it. While this is largely a Wizard of Oz study, it is interesting in it’s approach - that being that the system can dynamically adapt to the user’s needs. Their conclusional remarks are of interest:

> Our observations from the TP3 user study indicate that providing users with specific types of information about the environment along with tools for ac-

---

46 Printed Circuit Board
cessing it is helpful for supporting general spatial awareness among individuals with visual impairments. This increased spatial awareness is, in turn, helpful for supporting wayfinding in a broad sense for travelers with visual impairments.

We note that Talking Points 3 includes four of the five forces: mobile devices, sensors, location awareness and data. Social media could be added in such a system for providing more relevant and current information for VIPs.

6.7 Discussion

We find that the integration of the five forces has been explored with reference to aiding VIPs with mobility issues. The Yang article [88] brings us closest to seeing where the research has brought us in this regard, and it is definitely a positive start.

This, however, is a limited approach to a genuine interactive environment. We are considering a more robust solution – that is, an intelligent environment that has good data on itself, that can process input and provide intelligible results to the inputting system. This would then require a system that could interact with the environment.

7 Conclusions and Discussions

So, thus far we have tried to see how the “technological storm” described by Scoble and Israel [75] has affected VIPs with respect to aiding or hampering their ability to “get around”. In general the consensus seems to agree that it is helping, but somehow it is still limited. In [59] review how this technology has been applied to the general disabled population. They conclude:

The technology progress is constantly transforming the meaning of disability to the state of just a particular human condition that needs to be addressed.
This conceptual turn can eventually transform the people with disabilities from a minority group to equal standing society members.

7.1 The Impact of the Five Forces on VIPs

Here we would like to summarize the possible impact of each of the five forces:

- **Mobile Devices**
  Mobile devices are a key aspect in supporting the mobility of VIPs. Without mobile devices, digital aid for mobility would be far less than it can be in this digital world; in particular, the ability of mobile devices to communicate externally is the driving facet of this technology. Moreover, the relative ease with which applications can be developed and employed makes the mobile device a good basis for development.

- **Sensors**
  Without sensors most of the other four forces will be rendered useless. These sensors include on-board sensors, sensors in the environment, and feedback approaches (haptic, voice, etc.).

- **Location Based Services**
  Location information is key for navigation and orientation. Location Based Services are made possible by all the other four forces: sensors for obtaining and displaying location data; social media for communicating this information; mobile devices for front-end delivery systems; and, data being the actual digital information entity.

- **Social Media**
  The power of social media has started to attract more and more attention from researchers and developers for servicing people in need. This should also be viewed in the more general sense of communication (see below) and considered for its use by VIPs.
• Data

Data, in particular Big Data over the Internet with video, images and geo-tagged information will hold significant promise in assisting VIPs with mobility. The key is real-time data analytics.

7.2 Looking Forward

We might say that the storm has not been fully harnessed; beyond the use of the smartphone – in particular the iPhone – there do not seem to be major breakthroughs with respect to aiding VIPs with the mobility issue.

We might consider another “force” overlooked by Scoble and Israel, that being communication. It perhaps seems trivial in their book, or could be included in social media, but it may lead to possibilities not as of yet researched, for example, in augmenting cities, traffic networks and facilities with sensors that talk with each other and communicate with personal mobile devices.

[05] have looked at how communication principles may be used to aid VIPs. As with any good relationship, communication is what alleviates problems and we are considering how we might integrate communications into systems that can better handle the exchange of pertinent information in a timely fashion to aid VIPs while on their way from A to B.
References


[73] Torsten Sattler, Bastian Leibe, and Leif Kobbelt. Improving image-based localization


[80] Ioannis Stamos, Lingyun Liu, Chao Chen, George Wolberg, Gene Yu, and Siavash Zokai. Integrating automated range registration with multiview geometry for the


[90] Mengmeng Ye, Ping Li, and Qing Li. An intelligent mobile news reader for the visually impaired people. 2014.
