Pattern Recognition in Service of People with Disabilities

L. Coetzee and E. Barnard

Information Society Technologies Centre (ISTC),
CSIR, P.O. Box 395, Pretoria, 0001
{louis.coetzee}/{etienne.barnard}@csir.co.za

Abstract

South Africa has a large community of people living with a variety of disabilities. Very often they can be productive members of society, but are excluded due to the high cost of assistive technologies. These assistive technologies are based on a variety of pattern-recognition techniques and algorithms. This paper analyses a number of disabilities, their assistive devices and the associated technologies in order to highlight the role pattern recognition plays in enabling accessibility and improving human computer interaction for people living with disabilities. In addition we point out some areas where pattern-recognition research can beneficially be adapted to address the needs of people with disabilities, and argue for the use of open-source technologies to improve accessibility for a larger part of the disabled community in South Africa.

Keywords: pattern recognition; disability; accessibility; information; human computer interface; open source

1. Introduction

According to Statistics South Africa, at least 5.9% of the people living in South Africa live with one or more disabilities[1]. People with disabilities often cannot contribute to the economy, even though many have the desire and ability to; similarly, they are often prevented from participation in other spheres of society to the extent of their abilities and desire[2]. The South African government has recognised this fact and has introduced legislation in an attempt to increase the fraction of people living with disabilities that are economically active, and to ensure that increased support is made available to improve the quality of life experienced by people with disabilities[3].

Depending on the specific disability, various assistive devices and technologies are required to empower a person living with a disability to become a productive member of society. These devices vary tremendously with respect to factors such as technological sophistication and user friendliness. However, a few common trends characterise the majority of such devices:

- They tend to be imported and thus expensive, placing them out of reach of most disabled people.
- These devices and applications are proprietary and closed.
- These devices and applications are not localised to allow for the cultural and linguistic variety that characterises South African society.
- Finally, almost all of these devices utilise a variety of pattern-recognition techniques and algorithms.

The current paper is an overview of the National Accessibility Portal (NAP), an initiative between the CSIR, the Office of the Status of Disabled Persons (OSDP), the Independent Living Centre (ILC), the SA National Council for the Blind (SANCB), the Deaf Federation of South Africa (DEAFSA) and the National Council for Persons with Physical Disabilities in SA (NCPPDSA). NAP's primary aim is to provide a national networking and communication system based on Internet technologies for people with disabilities; and to improve accessibility to information about or for people with disabilities in a cost effective way.

We analyse a number of disabilities and their associated assistive technologies (with the focus on technologies enhancing accessibility to information using personal computers) in order to identify the utilised pattern-recognition algorithms. It is envisioned that this understanding will ultimately direct us to develop localised South African alternatives and allow us to contribute to the few open-source alternatives.

In the next section, we present technologies associated with low vision and blindness. In Section 3 we analyse technologies associated with deafness. Section 4 contains a discussion regarding physical disabilities, and is followed by a discussion of the required and actual characteristics of pattern-recognition applications aimed at people with disabilities, and an open-source perspective on technical applications (Section 5). A conclusion is presented in Section 6.

2. Low Vision and Blindness

A number of different categories classifying visual impairments exists. They include:

- *Low Vision, severe* – where visual tasks are performed at a reduced level.

- *Low Vision, profound* – gross visual tasks are performed with difficulty.

- *Near Blind – vision* is classified as being unreliable, and

- *Blind* – the person is totally without sight.

A Web browser installed on a standard personal computer has become one of the best tools to access information via the Internet for people who are not visually impaired. For a person suffering from low vision a screen magnifier is an important tool used to magnify regions of the user's desktop. Using a screen magnifier the user has access to the standard applications such as web browsers and email clients, thus providing the person with unlimited access to information. People suffering from colour blindness can benefit from using high contrast themes.

For a person with severe visual limitations (e.g. blindness) the situation is more complex. To enable access to the personal computer, specialised hardware (such as braille keyboards and displays) or software (such as screen readers) is required. Screen readers normally plug into the operating system and receive events and other information from the desktop. The screen reader interprets the received information and provides audible prompts of what is happening on the desktop. The audible prompts are generated using text-to-speech engines.

Commercial screen readers for Windows-based machines are quite common. They interface with Microsoft's SAPI and provide good text-to-speech prompts for languages such as English . These readers interact well with applications, allowing for a manageable desktop. Unfortunately, they are expensive (very often more than R10 000 for a single user license), and do not allow for the use of indigenous languages.

3. Deafness

The most common forms of deafness can be categorised into 3 broad categories:

- *Conductive hearing loss* – caused by damage to the outer or middle ear. Sufferers may benefit from the use of hearing aids.

- *Sensorineural hearing loss* – caused by damage in the hair cells of the inner ear or nerves. Sufferers do not benefit from the use of hearing aids.

- *Combination* of conductive and sensorineural hearing loss.

Sign language is commonly used as communication method by deaf people. Due to the nature of the disability deaf people have extreme difficulty communicating with other people, or absorbing information contained in multimedia (e.g. TV). Literate deaf people have few problems obtaining information or interacting with computers. However, a large percentage of deaf people are also illiterate, thus severely limiting their ability to interact and share information.

An exciting prospect is the application of virtual reality avatars, used in conjunction with human language technologies, to create sign language from interpreted text. Another exciting prospect is the use of speech recognition to automatically generate subscripts on multimedia (video) footage.

4. Physical disabilities

A vast number of physical disabilities exists. They include:

- *Paraplegia* – complete paralysis of the lower half of the body including both legs, usually caused by damage to the spinal cord.
- *Quadriplegia* – complete paralysis of the body from the neck down.
- *Cerebral palsy* – a disorder usually caused by brain damage occurring at or before birth and marked by muscular impairment. Often accompanied by poor coordination, it sometimes involves speech and learning difficulties.
- *Muscular dystrophy* – a group of progressive muscle disorders caused by a defect in one or more genes that control muscle function and characterised by gradual irreversible wasting of skeletal muscle.

Access to computers using the standard input devices requires a certain amount of mobility, the very element a physical disability impacts the most. Paraplegics typically have full mobility in the upper halves of their bodies, thus empowering them to use standard pointing and input devices, provided that the environment is ergonomically appropriate.

Quadriplegia has a far greater impact on mobility, preventing usage of normal character and pointing devices. In this domain, the use of head-mounted pointing devices (requiring sophisticated target tracking algorithms and image processing capabilities), in conjunction with on-screen keyboards (in combination with predictive text) and a variety of switches (mouse-click simulation devices), improves accessibility significantly. If the user also suffers from a speech disability, the above scenario is extended with the addition of a text-to-speech output device – thus allowing the person to actively communicate. Alternate *command-and-control* mechanisms use speech recognition as input device, allowing interaction with applications installed on the computer.

Sufferers of cerebral palsy and other disorders have extreme mobility limitations. Standard input devices are impractical. The use of large keys on keyboards in conjunction with applications providing synthesised voice output has proven to be beneficial as learning aids, and for information access.

5. Technological tools: characteristics, challenges and the open-source approach

The tools described in the previous section along with similar tools aimed at assisting people with disabilities, make extensive use of pattern-recognition algorithms to compensate for physical or sensory disabilities. Examples of such algorithms include:

- *Visual recognition and tracking* algorithms for gaze-tracking systems, and for use in gesture recognition[4, 5].
- *Speech recognition* to provide deaf users with transcriptions of spoken material, and to respond to commands of people with disabilities[6].
- *Language-processing and speech-synthesis* algorithms to generate spoken output for blind people – for example, in screen readers[7].
- Algorithms for *context interpretation*, for use in virtual-reality avatars and in command-and-control interfaces for various applications.

Together these are exciting and challenging applications of pattern recognition; we now discuss some of the characteristics of this domain from a pattern-recognition perspective and highlight some of the limitations of current approaches.

Some of the most salient aspects of these applications of pattern recognition are:

- Users of such systems are highly cooperative repeat users – thus, the pattern-recognition algorithm is not required to perform with perfect accuracy, but it does need to be highly robust and predictable. Users of an eye-tracking system, for example, will be tolerant of failures in tracking performance, as long as the failures do not generate false key presses.

  This also implies that user-adaptive systems are of great importance in this domain. Since a user of known identity will repeatedly interact with a given system, it is possible to refine the performance of the system for that particular user, both during an initial training phase and during on-going usage; this should produce significantly more accurate behaviour.

- Since these systems invariably function with a human in the loop, it is possible to combine human and machine intelligence in ways that leverage the strengths of both. Specifically, humans are very good at extracting and processing semantically relevant information, whereas pattern recognition performs well at the syntactic (structural) and lower levels. This implies, for example, that an optimal speech recogniser for command-and-control applications can operate with a relatively small number of keywords, and rely on the user of the system to combine those keywords to operate the system in an acceptable fashion. (Therefore, sophisticated attempts at natural-language processing are not likely to be of great value in such applications.)

- People from all language groups suffer from disabilities, and to assist them in their own languages is of crucial importance in many applications (e.g. for monolingual users, or when the cognitive load of operating in something other than the user's first language is not tolerable). Consider, for example, a text-to-speech device that is used to assist a speech-impaired person in communicating with her family: even if family members are conversant with several languages, support for their home language is required for acceptance of the device.

  It is therefore imperative that all these systems be designed within a multi-lingual framework from the outset, and that this framework be populated with as many languages as possible.

Off-the-shelf pattern-recognition algorithms are generally not optimal for these conditions, which implies that there is much scope for algorithm development or refinement[8, 9]. Current speech-recognition algorithms, for example, are strongly biased towards "normative" or "standard" speech. Highly adaptive algorithms, which are able to handle idiosyncratic pronunciations produced by users with various speech defects, would be of great practical value. Similarly, sophisticated eye-tracking algorithms should adapt themselves to user characteristics and behaviour in a transparent fashion.

In addition, general design paradigms are required in order to assist developers in creating systems with an appropriate assignment of responsibilities to the user and system, respectively. To this end, abstract models of task domains, user characteristics and system characteristics must be developed.

Although much research and development in this area remain to be done, a substantial range of assistive devices have already been developed. Current commercial products in this domain are often of great use, but are usually prohibitively expensive for the majority of potential users in a developing country. Fortunately, a number of open-source alternatives exist. At the forefront of current development is the *Gnopernicus* suite of applications[10]. It consists of a screen reader, a screen magnifier and braille input and output interfaces. Gnopernicus interacts with the Gnome desktop through the AT-SPI (assistive technology service provider interface)[11]. The screen reader utilises the gnome-speech API to synthesise voice. Gnome-speech abstracts a number of different TTS engines, which includes Java FreeTTS[12] and Festival[13].

GOK (the gnome on-screen keyboard) is used as an alternative input mechanism[14]. GOK provides complete control over the gnome desktop, again by interacting with the AT-SPI. Using only a pointing device, the end user is thus able to generate general-purpose keyboard input. GOK utilises predictive text for faster text generation. Figure 1 depicts a screen grab of GOK with predictive text providing choices for *pattern* or *patterns* based on the input *patt*.

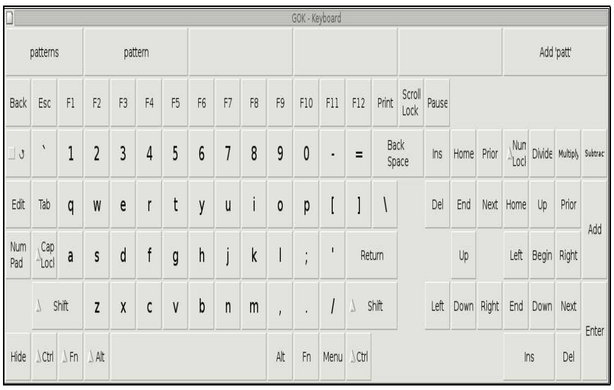

Figure 1: GOK – Gnome on-screen keyboard

For people with physical disabilities usage of normal pointing devices is problematic. Utilities to control the NaturalPoint trackIR and SmartNAV (as presented in Figure 2) are currently under development[15]. These utilities use advanced target tracking algorithms to track a specific reflective element in continuous frames. The tracked object is then translated into on-screen mouse movement.

Figure 2: TrackIR

Perlbox[16] uses a combination of Festival[13] and Sphinx[17] to provide a front end providing command-and-control capability to the Linux desktop.

The above mentioned technologies provide open API's allowing developers to customise and localise based on the needs of the disabled person. In addition up-front costs are mostly limited to hardware (computer and other devices) with software and associated device drivers based on open source applications.

6. Conclusion

In this paper, we presented a view into the world of disabilities, the assistive devices and technologies normally associated with improving accessibility, with the aim of highlighting the underlying pattern-recognition algorithms. A vast number of disabilities exist, each with their own specialised needs. We focused on three disabilities: visual, hearing and physical. It is clear that human language technologies (more specifically speech recognition and speech synthesis) as well as a variety of image processing algorithms play an important part; however, the specific characteristics of users with disabilities imply that there is much scope to improve the operation of these algorithms by careful consideration of the strengths and limitations of the target users.

Acknowledgements

We would like to thank Marelie Davel as well as the NAP team for their research, inputs and suggestions.

7. References

[1] Statistics South Africa, Census 2001. Primary tables South Africa. Census '96 and 2001 compared, 2004. Report No. 03/02/04 (2001). `http://www.statssa.gov.za/census01/html/c2001primtables.asp`.

[2] Disabled People South Africa – Pocket Guide On Disability Equity. An Empowerment Tool. Published by the DPSA Parliamentary Office on behalf of DPSA. P. O. Box 15, Cape Town 8000.

[3] Integrated National Disability Strategy White Paper. Office of the Deputy President–South Africa, November 1997. `http://www.polity.org.za/html/govdocs/white_papers/disability1.html`.

[4] E. R. Davies. *Machine Vision: Theory, Algorithms, Practicalities*. Academic Press, San Diego, California, USA, 2nd edition, 1997.

[5] D. H. Ballard and C. M. Brown. *Computer Vision*. Prentice Hall, Englewood Cliffs, New Jersey, 1982.

[6] S.Young. Large Vocabulary Continuous Speech Recognition. *IEEE Signal Processing Magazine*, 13(5):45–57, 1996.

[7] T.Dutoit. *An Introduction to Text-to-Speech Synthesis*. Kluwer Academic Publishers, Dordrecht, 1997.

[8] R. O. Duda and P. E. Hart. *Pattern classification and scene analysis*. Wiley, New York, 1973.

[9] C.M.Bishop. *Neural networks for pattern recognition.* Oxford University Press, Oxford, 1995.

[10] Gnopernicus Assistive Technology. http://www.baum.ro/gnopernicus.html.

[11] The GNOME Accessibility Project. http://developer.gnome.org/projects/gap/.

[12] FreeTTS. http://freetts.sourceforge.net/.

[13] The Festival Speech Synthesis System. http://www.cstr.ed.ac.uk/projects/festival/.

[14] GOK – Gnome On Screen Keyboard. http://www.gok.ca/.

[15] trackIR for Linux. http://trackir.superlucidity.net/.

[16] Perlbox.org – Linux Speech Control and Voice Recognition. http://perlbox.sourceforge.net/.

[17] The CMU Sphinx Group Open Source Speech Recognition Engine. http://cmusphinx.sourceforge.net/html/cmusphinx.php.