How the New Approaches on Cloud Computer Vision can Contribute to Growth of Assistive Technologies to Visually Impaired in the Following Years?

Henrique da Mota Silveira[1]*, Luiz César Martini[1]

[1] *University of Campinas (UNICAMP), School of Electrical and Computer Engineering (FEEC), BRAZIL*

***Corresponding Author:** henriquedamota@gmail.com

ABSTRACT

This paper brings to light an important discussion on how to create better assistive technologies for visually impaired people with the new advances in the computer vision field. Until very recently, assistive technology solutions required large and expensive hardware, as well as complex software. However, this scenario seems to be changing with web services on cloud that turns available computer vision features, offering image and video content analysis. The challenge is on how the results of these analyses will be processed and which will be the action or interaction displayed afterwards, showing the importance of appropriate user interface to visually impaired.

Keywords: assistive technology, computer vision, visually impaired, accessibility

INTRODUCTION

According to Terven (2014), applications based on computer vision can be a support to mobility, orientation, object recognition, access to printed material, and support to promote social interaction to the visually impaired. This study is a discussion of how the new computer vision approaches hosted on cloud can support accessibility as assistive tools to visually impaired. It started as a partial requirement of "Smart Cities and Cognitive Cities", a discipline from the university's postgraduate program and, nowadays, it is the basis for the development of a doctoral thesis.

MATERIALS AND METHODS

Visual impairment refers to an irreversible state of diminished visual capacity of a person, originated by congenital (pathogenic) or environmental factors (pathologies, lesions, tumors etc.). The deficiency remains even taking clinical (therapies) or surgical procedures, even with the use of conventional optical items (glasses, contact lenses) (Costa, 2004). To Vivarta (2003), the concept of inclusion has a close relationship with another concept, known as accessibility. Expanding the usual concept that people have about what accessibility is, an idea refers to changes on urban and architectonic elements planning. To qualify a society as accessible it is important to check its adequacy according to the six basic requirements below, emphasizing that technological accessibility is not another type of accessibility - the technological aspect must permeate all those described, with the exception of attitudinal accessibility (Vivarta, 2003).

● Architectural accessibility: there are no physical barriers in houses, buildings, urban spaces and individual or collective means of transportation;

● Communicational accessibility: there are no barriers in interpersonal (face to face, sign language), written (including laptop using) and virtual communication (digital accessibility);

● Methodological accessibility: there are no barriers in studying methods and techniques (school), work (professional), community action (social, cultural, artistic) and child education (family);

● Instrumental accessibility: there are no barriers in the study, work and leisure or recreation (community, tourism or sports) instruments and tools;

● Programmatic accessibility: there are no invisible barriers embedded in public policies (laws, decrees) and norms (institutional, business);

● Attitudinal accessibility: there is no prejudice, and discriminations.

Given the need to apply the mentioned accessibility types, according to Bersch (2008), Assistive Technology (AT) is "still a new term, used to identify the whole arsenal of resources and services that contribute by providing or expanding functional abilities of disabled people, thus consequently promoting an independent life and inclusion".

Cook and Hussey (1995) define AT as "a wide range of equipment, services, strategies and practices conceived and applied to decrease the functional problems faced by individuals with disabilities".

Some types of ATs for the visually impaired can be based on software or hardware. Usually, software solutions provide accessibility options on operating systems or programs such as screen readers. The hardware solutions present products as navigation aids and environments with multimodal hardware (Rodrigues, 2006).

APPROACHES OF CLOUD COMPUTER VISION PLATFORMS

Szeliski (2010) says the computer vision can be seen as the research area in which the computer has a powerful understanding of images and videos, through mathematical techniques for recovering the three-dimensional shape and appearance of objects.

Microsoft Cognitive Services (2017) discuss the recent advances in the machine learning, artificial intelligence and computer vision areas are allowing computer scientists to create smarter applications that can identify sounds, words, images and even facial expressions.

The tools are designed so that developers can insert the computer vision services in their applications without necessarily implementing them, themes like machine learning, artificial intelligence and computer vision (Microsoft Cognitive Services, 2017).

The following solutions (Microsoft Cognitive Services and IBM Watson) are paid platforms based on cloud computing services that can be accessed through REST APIs on the Internet. They deserve to be studied as computer vision platforms that can offer resources for future assistive technologies for visually impaired people.

Microsoft Cognitive Services

Initially as Project Oxford and now named Microsoft Cognitive Services, the initiative provides tools that help developers achieve capabilities such as recognition of faces and emotions, tracking, proposing characteristics, among others functions (Microsoft Cognitive Services, 2017). The **Figure 1** show examples of the feedback system (from the web site of Microsoft Cognitive Services https://www.microsoft.com/cognitive-services), being all the interactions results acquired in JSON format too.

IBM Watson

According to IBM (2017), the IBM Watson platform was built to do advanced natural language processing, information retrieval, knowledge representation, automated reasoning and machine learning technologies. In a simplified way, it can be considered a question and answers computer system. More than a hundred different techniques are used to analyze natural language, identify sources, find and generate hypotheses, find and mark proofs, merge and classify hypotheses (IBM Watson, 2017).

As Microsoft, IBM also has an API for image and video recognition called Visual Recognition API, which allows users to understand the content of an image or video frame by answering the question "What's in this picture?" (IBM Watson, 2017).

CONCLUSIONS

There are several tools and technologies that can support the visually impaired obtain greater perceptions using computational intelligence and vision as the basis for this new horizon. According to Terven (2014), the computer vision area can help as a substitute for human vision, thus becoming an area that can offer powerful tools to the
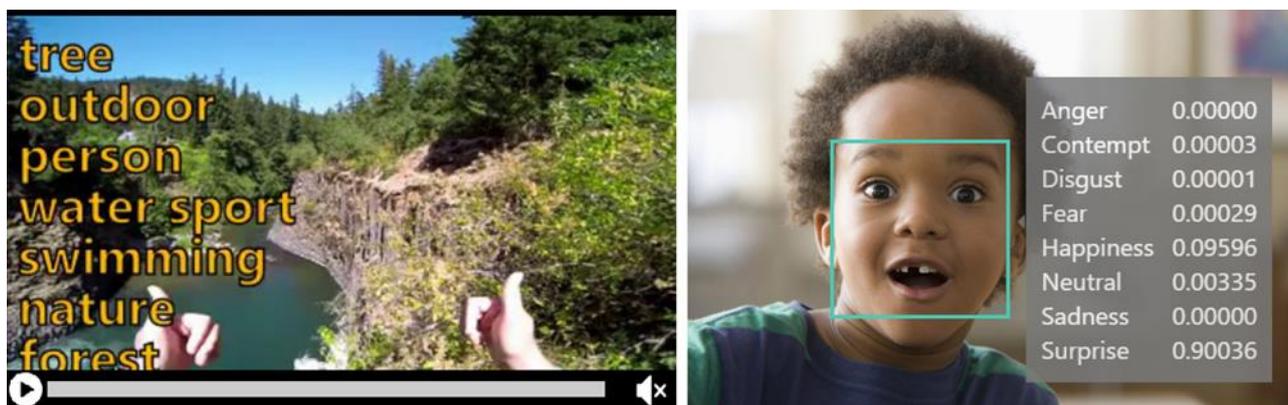
Figure 1. In the left there is a test playing a video, using the Computer Vision API (Microsoft Cognitive Services). Small fragments of the video are sent to the server and the API returns its interpretation on the media. In this example the video shows a person's view of nature environment and the give the feedback with what the machine understands from the analysis of the content printed: tree, outdoors, person, water sport, swimming, nature, forest. In the right there is a test with an image of a face, using the Emotion API (Microsoft Cognitive Services). The Emotion API retrieves the facial expressions of the image and returns a set of emotions, and each of them with a corresponding value that represents the weight of the emotion on the face displayed. The emotions detected are: anger, contempt, disgust, fear, happiness, sadness, surprise and neutrality.

development of technologies for the visually impaired. The current trend is to create devices and applications that exploit a wide range of technologies. In this sense, the computer-based assistive systems are in an intermediate stage (Terven, 2014).

To grow beyond this stage, it's necessary to use advanced algorithms that can reliably interpret visual information, refining it enough to understand the content of an image, a scene, or the attitude of the listener. A combination of multiple technologies such as computer vision, GPS, wireless Internet and voice recognition in a wearable platform similar to Google Glass can deliver a single, versatile, hands-free assistive device that can learn from the user and provide multiple functions (Terven, 2014).

The mentioned computer vision platforms are hosted on Internet in a cloud infrastructure, which opens new paths for those who want to create assistive technologies and are interested in how to establish communication with the visually impaired. The challenges that will need to be clarified will be on issues around how dialogues and feedbacks will be and concerns about new user interface templates for the visually impaired people obtain information about mobility; guidance; people; scenes; object recognition; printed material; and even deal with social interaction.

REFERENCES

Bersch, R. (2008). Introduction to Assistive Technology. *Specialized Center for Child Development. Porto Alegre (RS - Brazil).* Available at: http://www.assistiva.com.br/Introducao_Tecnologia_Assistiva.pdf (Accessed 30 January 2017).

Cook, A.M. and Hussey, S.M. (1995). *Assistive Technologies: Principles and Practices.* St. Louis, Missouri. Mosby - Year Book, Inc.

Costa, L.G. (2004). *Technological appropriation and teaching: Information and communication technologies and the teaching of physics for people with visual impairment.* Porto Alegre (RS - Brazil). Ph.D Thesis. Federal University of Rio Grande do Sul - UFRGS.

IBM Watson (2017). Available at: http://www.ibm.com/watson (Accessed 30 January 2017).

Microsoft Cognitive Services (2017). Available at: https://www.microsoft.com/cognitive-services (Accessed 30 January 2017).

Rodrigues, C.E.M. (2006). *A Haptic Aid for Navigation for the Visually Impaired.* Bachelor's Degree in Computer Science. Federal University of Pernambuco (Brazil).

Szeliski, R. (2010). *Computer Vision: Algorithms and Applications.* Springer-Verlag New York Inc.

Terven, J.R., Salas, J. and Raducanu, B. (2014). New Opportunities for Computer Vision–Based Assistive Technology Systems for the Visually Impaired. *Computer*, 47(4), pp. 52-58.

Vivarta, V. (2003). Media and disability. Brasilia (Brazil), News Agency for the Rights of the Child, The Banco do Brasil Foundation, pp. 24.