ScienceDirect

Neurocomputing

Volume 606, 14 November 2024, 128284

Survey Paper

Assistive systems for visually impaired people: A survey on current requirements and advancements

Preeti Kathiria, Sapan H. Mankad, Jitali Patel, Mayank Kapadia, Neel Lakdawala

Show more

https://doi.org/10.1016/j.neucom.2024.128284

Get rights and content

Abstract

In this survey, we provide a comprehensive study on the assistive technological devices which help visually impaired persons in their day-to-day lives. With various forms of disabilities such as visual, auditory, mobility, or cognitive impairment affecting a significant number of people worldwide, the aim of this study is to enable the blind persons to achieve independence. The primary objective of this work is to provide a concise description of assistive technologies for the visually impaired, along with an exploration of approaches for object detection, text detection, and text-to-speech synthesis. We present detailed discussion on these approaches

using both traditional as well as deep learning based techniques. Finally, we conclude with some open opportunities where further research can be carried out.

Introduction

Visual impairment [1] is a term experts use to describe any kind of vision loss, whether it is someone who cannot see at all or someone who has partial vision loss. Some of the challenges blind people face in daily routine are access to information, overly helpful individuals, societal stigma, leisure, often living in isolation and navigating around places.[1] Fig. 1 shows a geographical map of some organizations around the world that help and offer assistance to visually impaired people.

Computer vision technology [2] has the potential to improve accessibility for visually impaired people. When navigating through real and virtual spaces, they face a number of challenges with many public spaces still demonstrating a disregard for accessibility.

Several studies have been conducted over the last few decades on efforts to assist visually impaired people by means of methodologies for object detection, text detection, and text-to-speech conversion. However, our survey is different from others in a way that we mainly focus on how visually impaired people are assisted through computer-aided technologies. Assistive systems for the visually impaired encompass a broad range of technologies, including mobility aids, access to information tools, educational support, employment solutions, and social inclusion enhancements. These systems aim to address diverse needs such as navigation, communication, health monitoring, and skill development, ultimately fostering independence and participation in various aspects of life. For that, we have explored Emergency Technical Aids (ETA) for visually impaired people, and established the connection between the three approaches, namely, object detection, text detection, and text-to-speech synthesis. A typical scenario of using wearable glasses to assist visually impaired people is depicted in Fig. 2.

Elmannai et al. [3] discussed various equipment developed to assist visually impaired people. They provided a detailed description of the equipment, the arrangement of the hardware, and their working in a nutshell, along with their benefits and drawbacks. However, their survey does not discuss techniques for object detection and text detection as their scope was limited only to sensors.

Sumathi et al. [4] presented different text detection approaches for identifying text from images, signboards, pdf documents; and other techniques for extracting text from scene text images. However, their survey does not include how visually impaired people are

assisted through text detection, as they only prioritize text extraction. Han Lin et al. [5] demonstrated a variety of traditional and deep learning based text detection techniques. They categorized traditional methods as sliding window and connected component. On top of that, they provided various deep learning models and briefed about various loss functions, datasets, and accuracy for each technique.

Zhu et al. [6] presented various video object detection techniques. They classified various techniques into two broad categories. Each classification technique is further subclassified based on the use of temporal information and the combination of features extracted from video snippets. Each classification technique is compared using mean average precision (mAP). Their entire focus is on object detection, without consideration of text mining and various ETAs.

Although these works have contributed significantly, none of the above-mentioned surveys studied text-to-speech synthesis which can be more seamless and effortless for visually challenged people.

Our survey refers numerous research works contributed from various regions across the world. These contributions are visualized through a filled map as shown in Fig. 3, where each region is represented by the number of papers submitted from that particular location. A noteworthy observation regarding in this map is that certain regions may not display numerical values due to their limited geographical area on the map, rendering is impractical to accommodate text within those areas. However, the relative values can be inferred based on the shading intensity of the respective regions.
Many governments and WHO have also laid down certain guidelines and standards for use of web-based, desktop-based and wearable assistive technologies for physically disabled people.

Most people who use assistive technology use more than one product, making integrated services important. The WHO and UNICEF Global report on assistive technology offers recommendations specifically intended to steer governments and other stakeholders toward achieving universal access to assistive technology and set up ten recommendations underline the ongoing efforts required to enhance access to assistive technology for all in need.[2] The Global Report on Assistive Technology in 2022, published by WHO can be accessed from this link.[3]

In India, the Ministry of Family and Health Care [4] provides a comprehensive guidelines for accessibility when dealing with persons with low vision and blindness. This is basically for medical staff who is interacting or communicating with the disabled person. The

national list of essential assistive products (NLEAP) from Indian Council of Medical Research is available at this link[5] which focuses on policies for promotion and use of various assistive technologies for the needy people.
Some attempts are available in the literature which consider the collaborative study of object detection, text detection, and text-to-speech synthesis as components for visually assisted systems. Rapid research advancement in contemporary years, in the field of computer aided systems for visually impaired people inspired us to conduct in-depth survey by exploring, selecting and summarizing concerned studies. Following is the motivation for this survey.

1. The devices being utilized presently to give blind individuals eyesight are addressed in this work. In addition, we have also included details regarding the underlying components of such devices, their construction, and the benefits and drawbacks. This information can aid researchers in future by encouraging novel solutions for assisting visually impaired people.

2. Three main aspects of enabling computer vision for visually challenged people: object detection, text detection, and text-to-speech synthesis—are primarily studied to help the blind. Further, we have listed their benefits and drawbacks to assist researchers in selecting the best strategies for a certain domain.

Fig. 4 provides an overview of the structure and organization of major sections discussed in this work.

This paper exhibits an extensive survey of computer-assisted technologies for visually impaired people.

- Application domains in which visually impaired people can be assisted are highlighted.
- A comprehensive description of various prototypes discovered till now for the assistance of visually impaired people is presented in this study.
- In-depth explanations of several text and object detection methods are provided along with the corresponding metrics.
- Several text-to-speech conversion approaches are discussed.

- To encourage researchers working in this area, current research opportunities and challenges are also described.

The structure of this research study is as follows: A Systematic survey methodology is presented in Section 2. The application domains for assistive systems for visually impaired people are discussed in Section 3. In Section 4, the equipment used to assist people with vision impairments are described. Section 5 describes the methodology for object detection. The details of various text detection methodologies are provided in Section 6. Various approaches to text-to-speech synthesis are discussed in Section 7, followed by the research opportunities and Conclusion in Section 8 and 9, respectively.

Section snippets

Survey methodology

This section demonstrates the methodology used to carry out this survey. …

Application domains

Any device or system that aids a person with a disability falls under the category of assistive technology. This section describes numerous application sectors where visually impaired people are given assistance (Fig. 7).

…

Equipment used for helping visually impaired people

An assistive system for visually impaired people is a collection of technologies and tools meticulously crafted to augment the independence, accessibility, and overall quality of life for individuals confronting visual impairments or blindness. These systems harness a diverse range of technologies to furnish auditory, tactile, or haptic feedback, thereby empowering individuals to navigate their environments, retrieve information, and execute daily tasks that would pose challenges without visual …

Methodology used for object detection

Object detection is used to find occurrences of objects in pictures or videos. To generate useful results, object detection algorithms frequently incorporate machine learning or deep learning. Humans can quickly identify and locate objects of interest when viewing images or videos. With a computer, object detection aims to simulate this intelligence.

The chart depicted in Fig. 11 illustrates the distribution of papers contributed to the field of object detection from 1999 to the present day.

The …

Methodology used for text detection

Text detection has many applications, but their main goals are to determine whether text is present in a given image and, if so, to locate, detect, and recognize it [106]. Text localization seeks to identify potential text locations in images. Text extraction focuses on localization and binarization, and text detection uses localization and verification techniques to determine whether or not there is text [107]. The chart depicted in Fig. 14 illustrates the distribution of papers contributed to …

Text-to-speech synthesis (TTS) approaches

Modern research and applications in speech communication rely heavily on digital speech processing. Speech is mostly used for communication, which refers to message transfer between humans and machines. Using a speech synthesizer, text is transformed into voice by a text-to-speech system (TTS). A voice synthesizer is a type of computer system used for this purpose, and it can be implemented in both software and hardware forms [154].

Fig. 17 illustrates the classification of text-to-speech …

Research opportunities

Following the review of numerous publications in each of the three fields—text detection, object detection, and text-to-speech synthesis, we discuss some open issues and opportunities wherein there is scope for improvement in existing systems. …

Conclusion

In conclusion, this survey has provided a comprehensive overview of assistive technologies, exploring their diverse applications and impact on enhancing the quality of life for individuals with disabilities. Through the synthesis of existing research and advancements in the field, we have highlighted the wide array of assistive technologies available across various domains.

We presented a comprehensive survey of three major technologies required for assisting visually impaired people (namely, …

CRediT authorship contribution statement

Preeti Kathiria: Conceptualization, Formal analysis, Methodology, Supervision, Writing – review & editing. **Sapan H. Mankad:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Jitali Patel:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – review & editing. **Mayank Kapadia:** Formal analysis, Methodology, Validation, Writing – original draft. **Neel Lakdawala:** Formal analysis, Methodology, Writing – original draft. …

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. …

Preeti Kathiria is working as an Assistant Professor at the Computer Science and Engineering Department at Nirma University. She has over 20 years of professional experience, including academia and industry. She has published several Scopus/SCIE-indexed research papers in international journals, conferences and book chapters. She is a member of ISTE. She has also provided her services in various capacities at National and International conferences. She has completed a research project, …

…

…

Recommended articles

References (190)

SainarayananGopala *et al.*
Fuzzy image processing scheme for autonomous navigation of human blind
Appl. Soft Comput. (2007)

TongKang *et al.*
Rethinking PASCAL-VOC and MS-COCO dataset for small object detection
J. Vis. Commun. Image Represent. (2023)

ZouZhengxia *et al.*
Object detection in 20 years: A survey

(2019)

ArakeriMegha P. *et al.*
Assistive technology for the visually impaired using computer vision

ElmannaiWafa *et al.*
Sensor-based assistive devices for visually-impaired people: Current status, challenges, and future directions
Sensors (2017)

SumathiC.P. *et al.*
A survey on various approaches of text extraction in images
Int. J. Comput. Sci. Eng. Survey (2012)

LinHan *et al.*
Review of scene text detection and recognition
Arch. Comput. Methods Eng. (2020)

ZhuHaidi *et al.*
A review of video object detection: Datasets, metrics and methods
Appl. Sci. (2020)

NovikovaTatiana *et al.*
Large-lexicon attribute-consistent text recognition in natural images

IfukubeTohru *et al.*
A blind mobility aid modeled after echolocation of bats
IEEE Trans. Biomed. Eng. (1991)

View more references

Cited by (0)

Preeti Kathiria is working as an Assistant Professor at the Computer Science and Engineering Department at Nirma University. She has over 20 years of professional experience, including academia and industry. She has published several Scopus/SCIE-indexed research papers in international journals, conferences and book chapters. She is a member of ISTE. She has also provided her services in various capacities at National and International conferences. She has completed a research project, “Development of Advanced Algorithms for Land Use Land Cover Classification using Deep Learning Techniques”, funded by the Indian Space Research Organization and some projects funded by Nirma University. Currently, she is working on a research project funded by Nirma University in Machine/ Deep learning. She works in the areas of Machine Learning, Deep Learning, Data Mining and analytics. She has conducted ISTE-approved STTP in the field of Web Services using PHP. She also has delivered Expert Lectures at Training Program for SAC-ISRO on Graph- based Data Analysis with Spark and Neo4j and at several Short Term Training programs.

Sapan H. Mankad received the B.E. degree in Information Technology from Saurashtra University, Rajkot, India in 2004 and M.Tech. degree in Computer Science and Engineering from Nirma University, Ahmedabad, India in 2011. He obtained the Ph.D. degree in the field of Speech Processing using Machine Learning from Nirma University in 2021. He is currently working as an Assistant Professor in Computer Science and Engineering Department at Institute of Technology, Nirma University, Ahmedabad, India. His research interests include Audio and Speech Processing, Voice Biometrics, Machine Learning and Music Information Retrieval. He is a member of IEEE, IEEE – Signal Processing Society, Asia Pacific Signal and Information Processing Association (APSIPA) and Indian Society for Technical Education (ISTE).

Dr Jitali Patel is working as an Assistant Professor in Computer Science and Engineering Department. She has more than 10 years of teaching experience. She received her BE degree from CU Shah College of Engineering and Technology, Wadhwan and ME degree from Dharmsinh Desai University, Nadiad. Currently, she is pursuing her Ph.D. from Ganpat University. Her area of specialization includes Machine Learning and Information Retrieval. She has been contributing to research in the said domain with good quality publications in international conferences and journals.

Mayank C. Kapadia holds a B. Tech. in Computer Science and Engineering from the Institute of Technology, Nirma University, India and is currently pursuing a Master of Science in Data Analytics from San Jose State University, USA. His research focuses on data science-related fields. Mayank is proficient in multiple programming languages, including Python, Java, C, SQL, HTML, NoSQL, HSQL, Pig, and R, and is skilled with a variety of tools such as OpenCV, Scilab, Eclipse, NetBeans, LaTeX, Cisco Packet Tracer, Anaconda Navigator, Visual Studio Code, Wireshark, Dev C + + , Excel, MS Office, Orcit, Jira, OCP, Trifacta, Toad, Tableau, and Power BI. He has experience as a teaching assistant and completed a one-year stint at Bank of America, where he developed strong analytical and communication skills.

Neel M. Lakdawala holds a Bachelor's degree in Computer Science and Engineering from the Institute of Technology, Nirma University, India. He is currently pursuing a Master of Science in Computer Science at the University of Texas at Arlington (UTA). He has work experience as a software developer, with a focus on front-end development. His research interests include machine learning, computer vision, and deep learning.

View full text