Wearable Devices for Digital Interaction to Assist Blind and Low-vision People

Wenhao Xue[1], Yi Ren[1], Yi Tang[1], Ziqi Gong[1], Tianfang Zhang[2], Zuobing Chen[2], Xiaonan Dong[3,4], Xuezhi Ma[5,6 †], Ziyu Wang[6 †], Heng Xu[7 †], Jiaqing Zhao[8 †], and Yuan Ma[1 †]

[1]Department of Mechanical Engineering, The Hong Kong Polytechnic University, Hong Kong, China.
[2]Department of Rehabilitation Medicine, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China.
[3]Life and Health Institute, China Resources Research Institute of Science and Technology Company Ltd., Hong Kong S.A.R., China.
[4]School of Biomedical Sciences, LKS Faculty of Medicine, the University of Hong Kong, Hong Kong S.A.R., China.
[5]A*STAR Quantum Innovation Center (Q. InC), Agency for Science, Technology and Research, Singapore.
[6]Institute of Materials Research and Engineering, Agency for Science, Technology and Research, Singapore.
[7]CVTE Guangzhou Shiyuan Electronics Company Ltd., Guangzhou, China.
[8]Institute of Nuclear and New Energy Technology, Tsinghua University, Beijing, China.

[†]e-mail:
ma_xuezhi@imre.a-star.edu.sg, wang_ziyu@imre.a-star.edu.sg, hengxu@u.northwestern.edu, jqzhao@mail.tsinghua.edu.cn, y.ma@polyu.edu.hk

Abstract | Digital information has permeated all aspects of life, and diverse forms of information exert a profound influence on social interactions and cognitive perceptions. In contrast to the flourishing of digital interaction devices for sighted users, the information interaction needs of blind and low-vision (BLV) users have not been adequately addressed. Current assistive devices often lead to frustration among BLV users due to limited information delivery efficiency, reliability, and increased cognitive load. With the expected rise in BLV prevalence due to global population aging, there is an urgent need for assistive devices that can deliver information non-visually and effectively tackle the challenges this community currently faces. Here, we present three potential directions to address this issue: multisensory learning and integration, gestural interaction control, and the synchronization of tactile feedback with large-scale visual language models (LVLMs). We also explore future trends in assistive devices, focusing on metrics for text delivery efficiency and the enhancement of image content delivery, thereby enriching the lives of BLV users in the digital age.

Introduction

Vision impairment is a major global issue, impacting 2.2 billion people (27.5% of the world's population. Of these individuals, 70% have low vision and 2% are blind[1]. With its close link to age, the prevalence of blind and low-vision (BLV) is expected to increase rapidly as population aging rates accelerate globally[2]. For the rapidly increasing BLV population, accessing the already ubiquitous digital information for education, social interaction, and entertainment can be challenging as nearly all current digital interactions rely on vision[3,4]. This reliance results in unintended bias and unfair treatment towards BLV groups. Therefore, there is an urgent need for assistive technologies that enable digital interaction for BLV individuals. These devices could potentially enhance their daily quality of life, improve their employment opportunities and prospects, and promote greater social inclusivity and equality.

For BLV individuals, interacting with information poses challenges. With impaired vision, BLV individuals need to utilize non-visual information delivery (NVID) approaches, including auditory, tactile, and multimodal interactions to facilitate content comprehension. Although current devices offer plausible substitutes for BLV users, the representation of visual information by these devices is often incomplete or even inaccurate[3,5,6]. Another critical limitation is that these devices typically convert visual elements into text descriptions, which are often not timely and may fail to fully capture the richness of visual information[5,7,8], preventing BLV users from selectively comprehending key information. Such inadequacies in information representation arise from an insufficient understanding of users' personalized interaction needs and the complexities of multimodal information presentation. Therefore, the research and development of intelligent assistive devices that provide efficient and rich information delivery is particularly urgent and has significant application value.

In this Perspective, we will review and summarize the developments and challenges of assistive devices in terms of delivery efficiency and cognitive burden, as shown in **Fig. 1**, and propose design ideas for future wearable assistive devices that may break through the knowledge gap, including optimizing multisensory learning, synchronizing haptic large-scale visual language models (LVLMs), and combining gesture interaction control. Our aim is to provide theoretical and technical insights for the development of assistive wearable devices in NVID for BLV people and seeks to raise societal awareness regarding the needs of the BLV community in digital information.

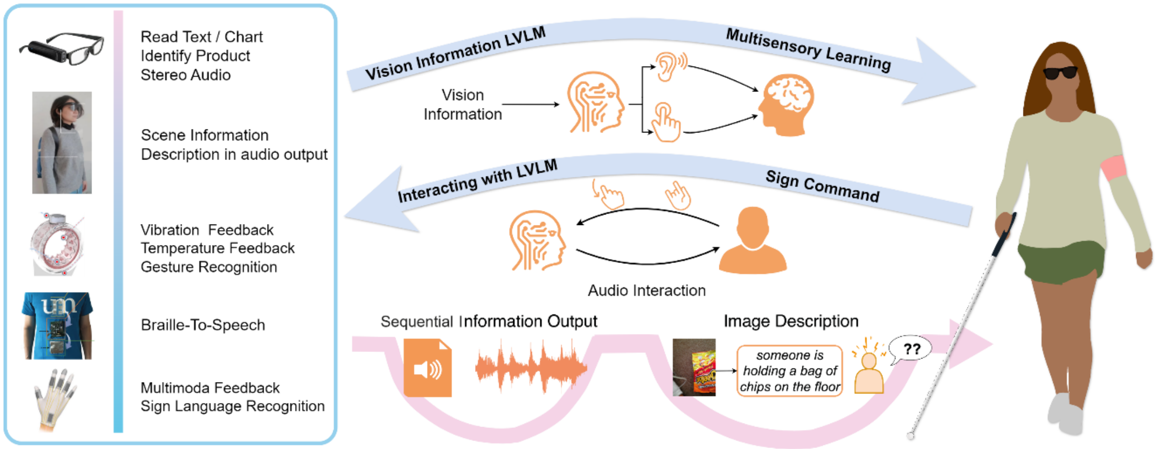

Figure 1 | **Overview of challenges and perspectives for current wearable devices in achieving digital information interaction for BLV people**[9–13]**.** Current wearable devices provide multimodal feedback such as audio descriptions of scenes and braille-to-speech translations, but their limitations in sequential information output, reliability issues, and low-quality descriptions pose significant challenges for BLV users. Utilizing synchronized haptic LVLM, multisensory learning, and user gesture interaction may help alleviate these difficulties.

Recent advancements in assistive wearable devices research for BLV people

Current assistive hardware

The rapid advancement of wearable devices aligns with the growing demand for a convenient and vibrant life, bringing new application opportunities across various fields[14,15]. Wearable devices, integrating sensors and wireless communication technologies into glasses, headgear, smartwatches, fabrics, and e-skins, enable real-time and continuous collection of users' physiological information[16]. These devices can monitor various parameters, including body temperature[17], respiration[18], posture monitoring[19], attention tracking[20], and safety assurance[21]. This multifunctionality has spurred applications in virtual and augmented reality (VR & AR)[9,22], education[15], communication[23], and assistive technologies[24].

Among these advancements of wearable devices, a significant emphasis has been placed on assisting BLV people to improve their navigation and mobility. These researches use various sensors to perceive the environment and relay information to users[25,26]. A typical example of these devices is the LightGuide cap[27], which employs a set of sensors and light cues for guiding BLV people with a certain level of light sense. These types of aids bring improvements to the mobility of BLV users. Compared to the significant efforts to navigation and mobility, however, only a relatively small number of research focused on the issues of the accessibility of information for BLV people. These devices that help information accessibility come in the form of auditory, surface tactile, audio-tactile, or wearable tactile devices, as illustrated in **Fig. 2**. The typical function of auditory devices are audio description and assistive reading, which help BLV users recognize text and convert it to audio output. Wearable tactile interaction represents an innovative form of assistive technology that greatly enhances the portability of devices and provide greater operational freedom. For example, the wearable Braille platform, as shown in **Fig.**

2d[28], utilizes the sequential inflation of air capsules to mimic the sensation of sweeping a finger across a Braille letter. Additionally, some currently non-wearable technologies such as braille displays[29–31] and tactile graphics[32–34], show potential to become wearable in the future. These tools are crucial for helping BLV people access text and images and could be further developed to enhance their usability and convenience. Besides, audio-tactile devices, primarily designed for educational and creative environments, are also effective in enhancing user engagement and understanding. A11yBoard[35], for instance, enhances the accessibility of creating digital artworks, while T-Pad[36] and PantoGuide[37] assist BLV students to understand and interpret graphics. While these assistive techniques and productions can provide real-time assistive service for BLV users, they suffer from deficiencies like the inherent inefficient information delivery rate.

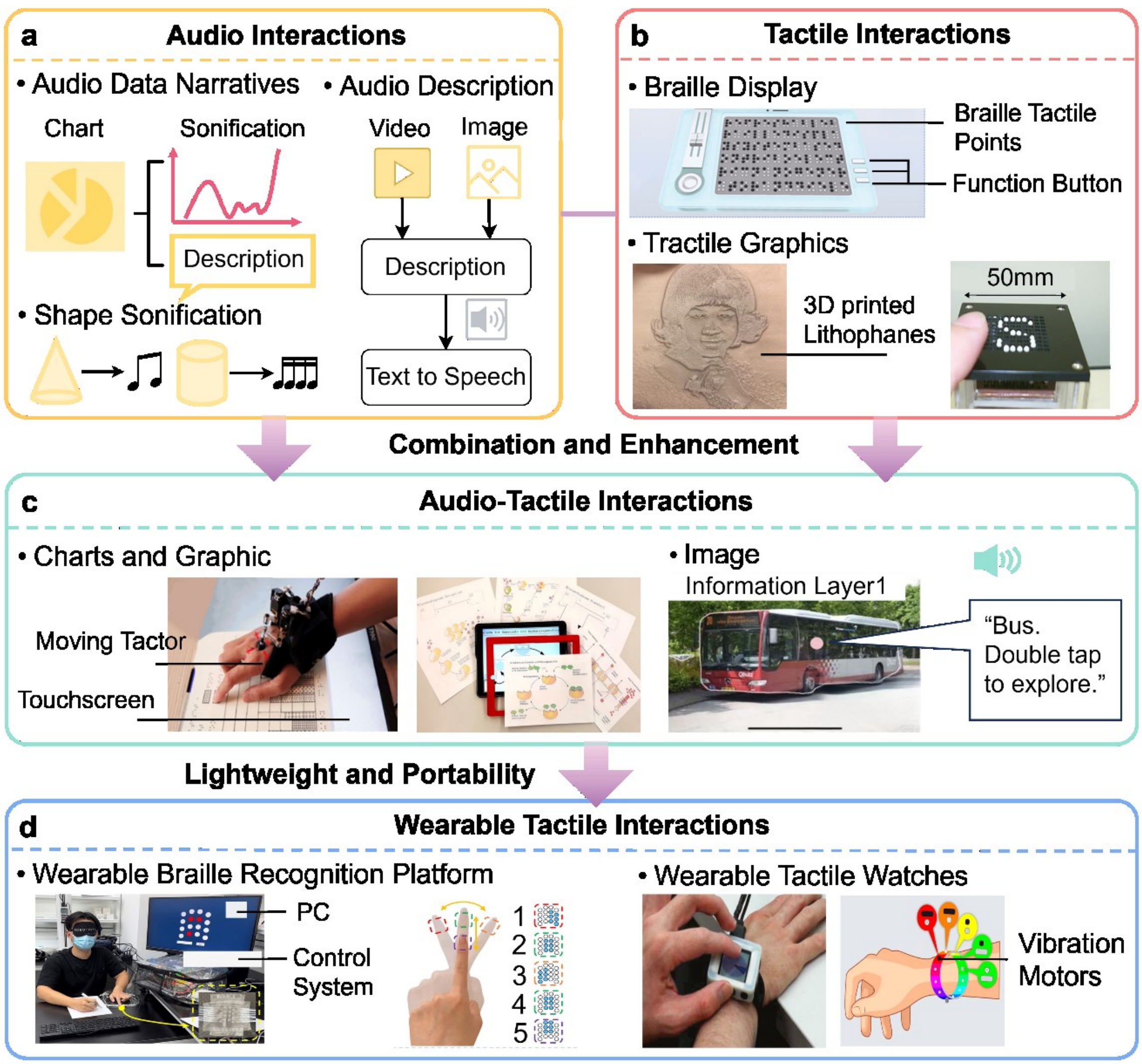

Figure 2 | **Overview of current methods in achieving information interaction for BLV people. a,** Scheme for audio interactions, including audio data narratives for charts[38], audio description for videos and images[39] and object shape perception with sonification[40]. **b,** Scheme for tactile interactions, including refreshable braille displays[30,41], lithophanes[42], and tactile displays[43]. **c,** Scheme for audio-tactile interactions, including the accessibility of charts and graphic in educational environments[36,37,44] and image exploration systems in social webpages[45]. **d,** Scheme for wearable tactile interactions, including the wearable braille recognition platform[28] and wearable tactile watches[46,47].

Software design for BLV accessibility

Accessible visualization software for BLV people has been developed to display charts, images, and artworks. Current research can convey charts and images through sonification and alternative texts (Alt Texts), which are generated descriptive texts to alternatively describe visual items. These visualization schemes are compromised solutions since Alt Texts does not support autonomous user interaction and exploration, and sonification imposes an additional cognitive burden on users. Some chart visualization schemes propose structured navigation and conversation interaction to compensate for these shortcomings[7,48–50]. For image information, researchers have explored various algorithms to enhance the visual experiences for BLV users. Notably, ImageExplorer[45] system has been developed for touch-based exploration, and the efficacy of utilizing music and haptic elements for the accessibility of artwork has also been demonstrated. Furthermore, with the development of smartphones and artificial intelligence (AI), various types of visual assistance apps, such as Seeing AI[51] and Be My Eyes[52], have become available for BLV users[53]. These apps rely on capturing images to perform AI-driven image recognition, providing real-time descriptions of the environment and objects. However, these accessible tools also suffer from challenges that will be elucidated in the next section.

Challenges in non-visual information delivery

Low delivery efficiency caused by sequential information output

Current assistive wearable devices mainly utilize speech and braille to convey information[11,54,55], which present information as sequential letter streams, posing challenges for BLV users who need to efficiently locate specific content or skim through large volumes of data. Such sequential navigation demands extensive time for key information extraction and requires BLV users to retain and mentally manipulate more information than is typically feasible for humans[56].

When engaging with content that embodies spatial, temporal, and structural elements, such as images, videos, charts, and maps, BLV users are confined by the analytical capabilities of assistive technologies because they hinder users' ability to independently perform in-depth analysis or explore information at various levels of detail[5,8].

In addition, BLV users also require time to learn and adapt to the information delivery methods of different assistive devices. This learning burden is further complicated by the lack of comprehensive accessibility tutorials for many assistive devices, thus impeding independent utilization[57]. NVID is symbiotic with assistive devices, and the effectiveness of information delivery is dependent on the capabilities of the particular assistive device. Although current NVID methods are plausible solutions for improving information accessibility for BLV users, it is unreasonable to expect these users to learn different systems to meet varying information needs.

The cognitive burden caused by low-quality Alt texts

BLV users often depend on Alt Texts to decipher images and complex non-textual information[58]. However, previous studies have shown that Alt Texts often lack quality, miss crucial details, and struggle to accurately describe complex images or tables[6,59,60]. Such inadequacies impose considerable cognitive demands on users.

Current Alt Texts are designed to be over-simplistic and lack substantial meaning. The mainstream Alt Texts, such as image captioning and scene description, are typically generated by computer vision techniques[61–63]. Such algorithmic methods generally adopt a one-size-fits-all description approach, resulting in descriptions with insufficient details. This may hinder contextual understanding of images and exacerbate cognitive burden.

For poor-quality images, the robustness of image description algorithms becomes another issue. Images taken by BLV people without visual feedback often suffer from common issues, including improper lighting, focusing, blur, and occlusion of areas of interest[64–66]. These issues may result in BLV users receiving inaccurate and incomplete descriptions, potentially leading to dangerous situations, or compelling them to expend considerable effort to verify and understand the information.

Alt Texts also struggle to reflect the structure of the spatial layout and the interaction among the image elements, making it challenging for users to form vivid mental visualizations[67]. Images convey information in an inherently non-sequential way, and even if the quality of Alt Texts improves, the loss of information when sequentially representing this content is inevitable. This is particularly true in expressing the spatial structure of the image, the relationships between elements, and the emotional atmosphere conveyed by the images[8,45,60]. Such an information gap not only impedes full comprehension of the narrative and context but also restricts the imaginative capacity of the user. To bridge this gap, BLV users may have to depend on additional sources, such as background knowledge or prior knowledge, to mentally reconstruct and interpret the image, further increasing cognitive complexity.

Designing assistive wearable devices in achieving informative interaction for BLV people

Boosting the delivery efficiency through multisensory stimulation

Multimodal sensory stimulation, which leverages the principle of multisensory integration, can effectively modulate attention within specific modalities[68,69]. Such an approach may significantly enhance the ability of BLV individuals to acquire a larger volume of information while simultaneously reducing the learning efforts. The following table provides an overview comparison of the learning costs, cognitive burdens, and amount of information required with current assistive technologies.

Table 1 | **Comparison of the assistive technologies dedicated to BLV people**

Method	**Learning cost***	**Cognitive burden****	**Information volume*****	**Data form**	**Interaction form (A=Audio, T=Tactile,)**
Audio narratives[38]	Low	High	High	Charts	A
Image sonification[70]	Moderate	Moderate	Moderate	Natural Images	A
HILM[71]	Low	Moderate	Moderate	Natural Images	A
Chart Reader[48]	Moderate	Moderate	Moderate	Charts	A
Voxlens[72]	Moderate	Low	High	Charts	A
Tactile line[73]	NR	Moderate	Moderate	Natural Images	T
3D-print map[74]	Moderate	Moderate	Moderate	Tactile Graphic	T
Tactile heatmaps[75]	Low	Moderate	High	Charts	T
Braille Key[36]	High	Low	Moderate	Tactile Graphic	T
Digital Key[36]	Moderate	Moderate	Low	Tactile Graphic	T
ATR[76]	Low	Low	High	Charts	A+T
A11yBoard[35]	Low	Low	NR	Digital Artboard	A+T
ImageExplorer[45]	High	High	High	Natural Images	A+T
Angel's Girl[77]	Low	Low	High	Digital Artboard	A+T
AccessComics[78]	Low	Low	NR	Comic Images	A+T
TPad System[36]	Low	Low	High	Tactile Graphic	A+T
InChorus[79]	Low	Moderate	High	Charts	A+T

*According to the quantitative analysis or qualitative descriptions related to learning cost in the corresponding research.

**According to the quantitative analysis or qualitative descriptions related to cognitive burden or mental load in the corresponding research.
***According to the quantitative or qualitative discussions related to the information volume delivered to participants in the corresponding research. We listed the specific evaluation methods used in these studies to assess learning costs, cognitive burden, and information volume, along with the discussions of our ratings in **supplementary table 1**. We aim to compare the characteristics of different methods, with no intention of providing any objective comments on the studies.

Table 1 shows that assistive technologies designed for multimodal interaction typically result in a lower learning cost and relatively high information volume for users. Multimodal stimulation activates multiple senses concurrently and enhances information processing more effectively than single-sense stimuli. Stein et al.[80] elucidate the neuroscientific principles of this enhancement, showing that multisensory processing activates more complex neural activities than single-sensory input. This collaborative neural activity bolsters perception and attention, furnishing BLV users with a more comprehensive analytical framework for decision-making—a vital support considering their reliance on non-visual sensory input[80]. The amalgamation of tactile and auditory stimuli has also been shown to improve BLV individuals' spatial awareness and comprehension of data representations[81–83].

According to the above findings, multisensory learning could effectively improve the efficiency of information processing. However, simply combining sensory inputs may not fully account for their interactions or harness their potential synergistic effects. Fox et al.[84] demonstrated that modulating auditory input frequencies can improve tactile perception accuracy through an auditory-tactile frequency modulation mechanism, suggesting that sensory interactions can be finely optimized. By tuning special frequencies for multisensory integration, one can enhance perception accuracy and consistency, which is crucial for the design of assistive devices to ensure efficient information processing and comprehension[84].

Moreover, the balance between complementary information and sensory occupancy should be carefully considered. Vercillo et al.[85] showed that attention to auditory stimuli significantly improves auditory accuracy and indirectly affects haptic accuracy when performing spatial localization tasks, resulting in an optimized perceptual output during the integration of auditory and haptic senses. This evidences the premise that sensory complementarity and the intentional direction of modality-specific attention can fortify the veracity of integrated sensory information. In the context of assistive device design, the delivery of information should balance between primary and secondary modalities: auditory output can serve as the principal channel for holistic information description, while haptic feedback can localize and underscore critical data, thereby facilitating information transmission and comprehension. Furthermore, the assistive device should be able to dynamically adjust the weight of auditory and haptic inputs according to the BLV user's attention to further optimize the balance of the multisensory integration process.

Relieving the cognitive burden through large-vision-language-model

The advent of AI, particularly through the deployment of large vision models (LVMs) and large language models (LLMs), presents a number of potential solutions to alleviate the cognitive burden challenges inherent in the information interactions of the BLV people[86,87]. These systems facilitate a range of functions such as assisted reading[49,88], target scene recognition[51,52] appreciation[89,90], and creating arts[77,91]. With the prevalence of large models, commercial assistive software and devices integrated with AI have demonstrated exciting prospects, such as Be My Eyes[52], Seeing AI[52], and OrCam eyes Pro[92]. Although these AI vision assistive methods have been recognized by many BLV people to be helpful in daily life, they are still not satisfied with the performance when reading nonplanar, unformatted information[53,93]. It often captures the main elements of an image without sufficient details. These systems frequently fail to provide feedback according to the prompt for the region of interest, and even require users to re-upload images, particularly when the initial images are of low quality.

Spatial awareness can also impose an extra cognitive burden, which is a critical concern for BLV people. Prior researches[8,45,60,94] have indicated that spatial awareness of autonomous exploration allows users a more profound sense of spatial distribution and element size, fostering a sense of agency that enhances their ability to vividly imagine the imagery. Haptic feedback has been demonstrated as an effective method for enhancing spatial awareness[95–97]. Therefore, it is essential to design a large visual-language model framework for assistive wearable devices that supports synchronized haptic output and is robust to low-quality images. This integration not only delivers audio output of image descriptions but also empowers users to actively investigate the spatial arrangement of elements via haptic feedback, helping them comprehend the size and spatial relationships within image scenes.

Enhancing the interaction efficiency and reliability through gesture-based interaction technologies

Gesture-based interactions offer a unique and significant advantage in the realm of information exchange[98], especially for BLV individuals. Most of the currently available assistive devices and smartphone apps require the use of both hands, which is impractical during many daily operations, particularly when a white cane is required. Previous research also demonstrated that BLV users prefer not to have both hands fully occupied while traveling[99,100]. A key benefit of gesture interactive controls is that they can alleviate the constraints of having both hands occupied. For example, a simple swipe or pinch gesture with a single hand can be used to accelerate or pause audio content, providing an efficient interaction form and keeping one hand free. Compared to audio or haptic feedback, gestures can also mitigate privacy concerns and minimize attracting undesired attention in public spaces [99,101,102]. For example, for BLV individuals, operating those currently available assistive devices often relies on sounds or vibrations, which may cause discomfort or self-consciousness, especially in noisy environments or noise-sensitive environments such as crowded transportation hubs and libraries. Consequently, gesture-based interactive controls within assistive devices are essential, allowing BLV users the capacity to operate the device with one-handed gestures, thereby facilitating efficient access to crucial information.

Gestures provide a more convenient and simpler mode of human-machine interaction. By integrating assistive devices with gesture recognition technologies, BLV people can interact with devices effortlessly, enhancing interaction efficiency. For example, when engaging with audio output, BLV users could employ gestures to adjust the speed and the position of reading or to command repetitions. This integration requires the utilization of sensitive, adaptable, and reliable sensors, as well as high-speed synchronized communication, to deliver an immersive and seamless feedback experience.

To enhance the user experience quality for BLV individuals, integrating haptic feedback into gesture interactions is helpful, as it provides more degrees of freedom (DoFs) for information exchange between users and devices, which can significantly improve interaction efficiency and reliability. This real-time feedback can prompt users whether the gesture was successfully recognized by the device[103,104]. By ensuring consistent and accurate interpretation of user inputs, haptic feedback increases the device's dependability across diverse situations. Additionally, haptic feedback guidance facilitates the learning of new gestures. By providing tangible cues associated with specific actions and prediction prompts, users can understand and memorize the desired action more quickly. This provides BLV users with a more intuitive and user-friendly assistive experience.

Assistive wearable devices in delivering text and image

We used the word per minute (WPM) metric as the evaluation standard (reading English as an example to compare the reading efficiency through different assistive technologies, as presented in **Fig. 3**. As a reference, sighted individuals' average reading speed is around 251 WPM, with an upper limit of 500 WPM due to the physical limitations of the human eye's information processing capacity[105]. Braille users typically read paper-based materials at around 120 WPM, with more skilled individuals reaching up to 170 WPM by employing refined hand movement techniques and precise localization[106,107]. Most braille displays, however, are constrained by the limited text lines with 12-80 braille units, which further restrict reading speeds to 30-50 WPM[108]. Meanwhile, researchers have found that BLV individuals can develop a more sensitive sense of hearing, enabling them to adapt faster audio output speeds, reaching up to 278 WPM[109,110] (As a reference, people typically speak at an average speed of around 150 words per minute[111]). By utilizing Text-to-Speech (TTS) technology, they can achieve higher reading efficiency while maintaining comprehension. However, to reach or even exceed the visual upper limit of 500 WPM, current assistive technologies need to move beyond a single mode of information delivery and incorporate capabilities for text compression and reasoning.

LLMs utilize their deep understanding of language and context to identify key details and remove redundancies, showing the potential to enhance reading efficiency[112,113]. This capability to distill information into concise summaries improves readability, making information consumption faster and more accessible. Consequently, assistive wearable devices integrated with LLMs are expected to further optimize WPM metrics and improve the efficiency and accuracy of textual information delivery. These devices can also customize information based on user preferences and context. Ultimately, they may reach or even exceed the visual upper limit of information extraction.

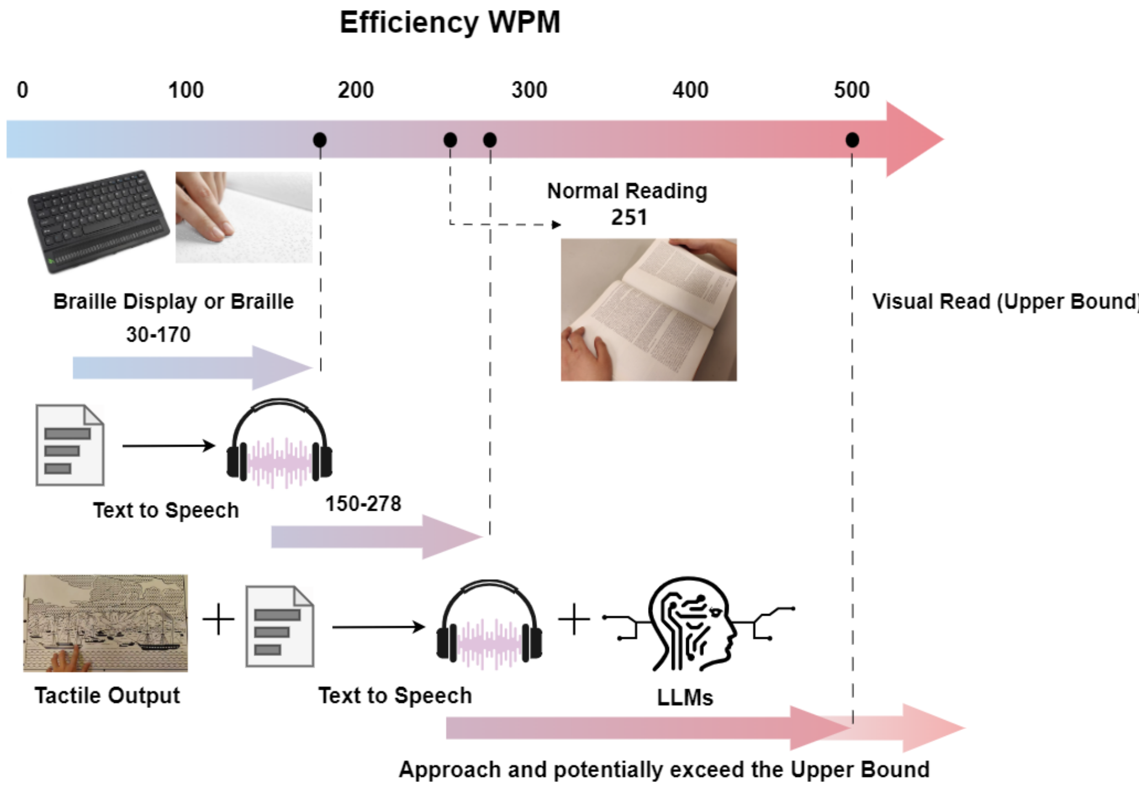

Figure 3 | **Comparison of the reading efficiency (WPM) for assistive technologies.** Braille and braille displays have significantly lower reading efficiency compared to sighted individuals, and TTS allows BLV people to read at speeds comparable to or exceeding those of sighted people. With the combination of the tactile, TTS, and LLMs, the WPM may approach and even exceed the upper bound of optimal reading speed.

Another important task for assistive devices is the delivery of image information content. Images are at the heart of information in digital communications, and BLV users need to rely strongly on image descriptions to access them and stay up-to-date. Coverage of image descriptions on mainstream websites was estimated to be approximately 72%[114]. As research advances in optimizing image descriptions, studies have shown that BLV users prioritize elements such as emotion, location, color, activity, and objects[8]. Additionally, they emphasize the significance of including the number of people, facial expressions, and gestures in social media descriptions[6]. Although the above studies provide guidelines on how image descriptions should be presented, they overlook the information preferences of different BLV users. Research found that one-size-fits-all image descriptions do not meet the accessibility needs of images, and that contextual information is essential for users to understand images[115,116]. To address these shortcomings, one approach used a multimodal model that integrates textual and visual information from social media to generate more accurate and contextually relevant descriptions[117]. However, this approach is limited to images with accompanying social media posts and struggles with contextless images and text recognition errors.

Considering the above findings, future assistive wearable devices should integrate context-awareness capabilities to recognize key visual elements in images. This involves enhancing the collection and utilization of contextual information and supporting real-time analysis of image content and contextual information from different sources. The visual elements of foreground and background content that assistive devices need to highlight on different types of images are summarized in **Fig. 4**. In scenes without characters, the foreground comprises the designated target and its relative spatial positioning, while the background encompasses locational context and accompanying textual details. Conversely, in event-type scenes characterized by the presence of characters, the foreground is dominated by the characters and their interactions. In such images, users typically prioritize the characters, necessitating a detailed description that encompasses aspects such as gender, celebrity status, and facial expressions. Background elements in these scenes often include architectural characteristics and any relevant text information. This could enable BLV users to access not only the factual narrative of the image, but also to understand the spatial layout, the interaction of the elements in the scene, and the emotion or atmosphere conveyed, thus making the visual information conveyed by the text accessible and meaningful. The cognition difference between congenital BLV and acquired BLV people should also be considered. For congenital BLV individuals, descriptions should emphasize non-visual details and highlight the functionality of the object. Conversely, descriptions for those with acquired BLV can be more effective by leveraging their visual memories and familiar visuals.

To further help BLV individuals to connect to society, social media and crowdsourcing platforms can be connected to future assistive wearable devices via APIs to automatically gather relevant information from multiple sources. This would include metadata such as the temporal and spatial details of image postings, as well as user comments. These metadata can be used to infer image themes and more comprehensive contextual information through LVLMs. By collecting and analyzing user preferences and past interactions, as an extra benefit, LVLMs can provide more relevant information and personalized content. This can achieve customized services for users.

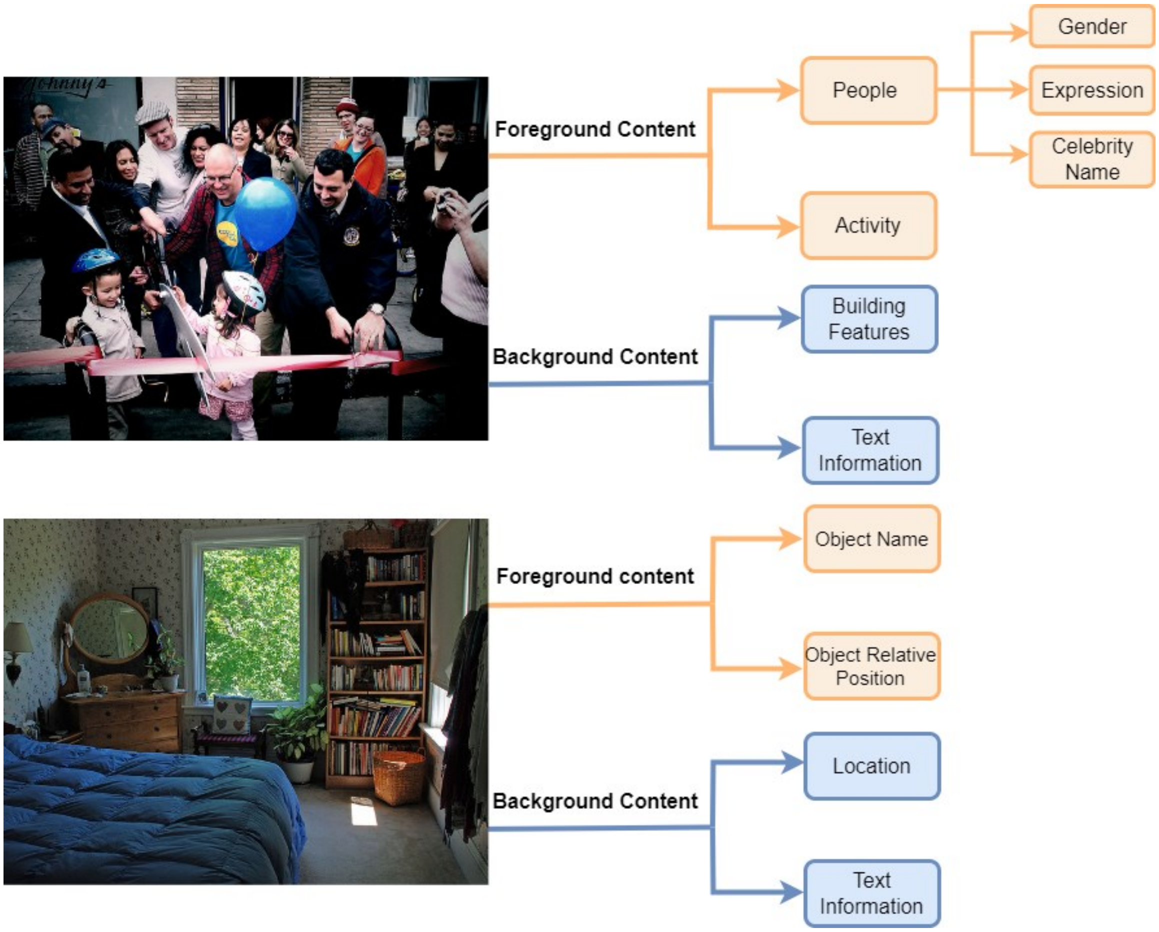

Figure 4 | **The Image information content presented from the assistive wearable devices.** In images depicting events (with characters) or scenes (without characters), assistive wearable devices should emphasize distinct visual elements of both foreground and background contents.

Conclusions and outlook

In this perspective, we discussed current assistive devices and accessible visualization tools for NVID and summarized the challenges faced by BLV people. Although the advancement in multimodal interaction aids and data visualization demonstrate expectations for addressing these challenges, their performance has not yet fully met expectations in terms of efficiency, cognitive load, and reliability. We hope our perspective can evoke and foster broader interest in design concepts for future wearable assistive devices that may break through the bottlenecks. Additionally, we aim to discuss efficiency metrics for assistive devices in delivering textual information while emphasizing key areas in conveying image-based information.

We believe that assistive devices designed to overcome current limitations will significantly enhance the digital information experience for BLV users, facilitating their various pursuits in education, entertainment, and employment. In the near future, their enhanced productivity may in return help alleviate the productivity deficit caused by global population aging and provide resilient solutions during economic slowdowns. We also anticipate that these devices will have a broader impact on future elder care. For example, these devices can support elderly individuals with deteriorating vision by providing haptic feedback or audio prompts, thereby improving their ability to access daily information and enhancing their autonomy and quality of life.

We forecast that continued innovations in hardware and algorithms can make wearable devices more seamless and informative. Achieving this integration requires collaboration between distinct domains and BLV communities. As the population ages and technology advances, the market size and demand for assistive wearable devices are expected to grow. Major technology giants like Microsoft, Apple, and Meta have already invested substantially in enhancing the accessibility of their wearable products, driving the development of smarter and more personalized assistive devices.

References

1. World Health Organization. Vision impairment and blindness. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (2023).
2. Bourne, R. *et al.* Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *Lancet Glob. Health.* **9**, e130–e143 (2021).
3. Kim, N. W., Joyner, S. C., Riegelhuth, A. & Kim, Y. Accessible Visualization: Design Space, Opportunities, and Challenges. *Computer Graphics Forum.* **40**, 173–188 (2021).
4. Choi, J., Jung, S., Park, D. G., Choo, J. & Elmqvist, N. Visualizing for the Non-Visual: Enabling the Visually Impaired to Use Visualization. *Computer Graphics Forum.* **38**, 249–260 (2019).
5. Schaadhardt, A., Hiniker, A. & Wobbrock, J. O. Understanding Blind Screen-Reader Users' Experiences of Digital Artboards. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems 1-19*, https://doi.org/10.1145/3411764.3445242 (2021).
6. Salisbury, E., Kamar, E. & Morris, M. Toward Scalable Social Alt Text: Conversational Crowdsourcing as a Tool for Refining Vision-to-Language Technology for the Blind. *HCOMP.* **5**, 147–156 (2017).
7. Zong, J. *et al.* Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum.* **41**, 15–27 (2022).
8. Morris, M. R., Johnson, J., Bennett, C. L. & Cutrell, E. Rich Representations of Visual Content for Screen Reader Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1–11, https://doi.org/10.1145/3173574.3173633 (2018).
9. Sun, Z., Zhu, M., Shan, X. & Lee, C. Augmented tactile-perception and haptic-feedback rings as human-machine interfaces aiming for immersive interactions. *Nat. Commun.* **13**, 5224 (2022).
10. Wen, F., Zhang, Z., He, T. & Lee, C. AI enabled sign language recognition and VR space bidirectional communication using triboelectric smart glove. *Nat. Commun.* **12**, 5378 (2021).
11. Li, Z. *et al.* Electrostatic Smart Textiles for Braille-To-Speech Translation. *Adv. Mater.* **36**, 2313518 (2024).
12. Liu, H. *et al.* HIDA: Towards Holistic Indoor Understanding for the Visually Impaired via Semantic Instance Segmentation with a Wearable Solid-State LiDAR Sensor. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)* 1780–1790, https://doi.org/10.1109/ICCVW54120.2021.00204 (2021).
13. Ou, W. *et al.* Indoor Navigation Assistance for Visually Impaired People via Dynamic SLAM and Panoptic Segmentation with an RGB-D Sensor. Preprint at http://arxiv.org/abs/2204.01154 (2022).
14. Yin, R., Wang, D., Zhao, S., Lou, Z. & Shen, G. Wearable Sensors-Enabled Human–Machine Interaction Systems: From Design to Application. *Adv. Funct. Mater.* **31**, 2008936 (2021).
15. Shen, S. *et al.* Human Machine Interface with Wearable Electronics Using Biodegradable Triboelectric Films for Calligraphy Practice and Correction. *Nano-Micro Lett.* **14**, 225 (2022).
16. Lin, M., Hu, H., Zhou, S. & Xu, S. Soft wearable devices for deep-tissue sensing. *Nat. Rev. Mater.* **7**, 850–869 (2022).
17. Han, S. *et al.* Battery-free, wireless sensors for full-body pressure and temperature mapping. *Sci. Transl. Med.* **10**, 4950 (2018).
18. Yang, Y. *et al.* Artificial intelligence-enabled detection and assessment of Parkinson's disease using nocturnal breathing signals. *Nat. Med.* **28**, 2207–2215 (2022).

19. Sardini, E., Serpelloni, M. & Pasqui, V. Wireless Wearable T-Shirt for Posture Monitoring During Rehabilitation Exercises. *IEEE Trans. Instrum. Meas.* **64**, 439–448 (2015).
20. Malasinghe, L. P., Ramzan, N. & Dahal, K. Remote patient monitoring: a comprehensive study. *J. Ambient Intell. Human Comput.* **10**, 57–76 (2019).
21. Zu, L. *et al.* A Self-Powered Early Warning Glove with Integrated Elastic-Arched Triboelectric Nanogenerator and Flexible Printed Circuit for Real-Time Safety Protection. *Adv. Mater. Technol.* **7**, 2100787 (2022).
22. Zhu, M., Sun, Z. & Lee, C. Soft Modular Glove with Multimodal Sensing and Augmented Haptic Feedback Enabled by Materials' Multifunctionalities. *ACS Nano.* **16**, 14097–14110 (2022).
23. Zhou, Q., Ji, B., Hu, F., Luo, J. & Zhou, B. Magnetized Micropillar-Enabled Wearable Sensors for Touchless and Intelligent Information Communication. *Nano-Micro Lett.* **13**, 197 (2021).
24. Slade, P., Tambe, A. & Kochenderfer, M. J. Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision. *Sci. Robot.* **6**, eabg6594 (2021).
25. Mai, C. *et al.* Laser Sensing and Vision Sensing Smart Blind Cane: A Review. *Sens.* **23**, 869 (2023).
26. Tapu, R., Mocanu, B. & Zaharia, T. Wearable assistive devices for visually impaired: A state of the art survey. *Pattern Recognition Lett.* **137**, 37–52 (2020).
27. Yang, C. *et al.* LightGuide: Directing Visually Impaired People along a Path Using Light Cues. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **5**, 1–27 (2021).
28. Wang, Y. *et al.* Multiscale haptic interfaces for metaverse. *Device* **2**, 100326 (2024).
29. Yang, W. *et al.* A Survey on Tactile Displays For Visually Impaired People. *IEEE Trans. Haptics.* **14**, 712–721 (2021).
30. Qu, X. *et al.* Refreshable Braille Display System Based on Triboelectric Nanogenerator and Dielectric Elastomer. *Adv. Funct. Mater.* **31**, 2006612 (2021).
31. Kim, J. *et al.* Braille Display for Portable Device Using Flip-Latch Structured Electromagnetic Actuator. *IEEE Trans. Haptics.* **13**, 59–65 (2020).
32. Ray, A., Ghosh, S. & Neogi, B. An Overview on Tactile Display, Haptic Investigation towards Beneficial for Blind Person. *Int. J. Eng. Sci. Technol. Research* **6**, (2015).
33. Mukhiddinov, M. & Kim, S.-Y. A Systematic Literature Review on the Automatic Creation of Tactile Graphics for the Blind and Visually Impaired. *Processes* **9**, 1726 (2021).
34. Holloway, L. *et al.* Animations at Your Fingertips: Using a Refreshable Tactile Display to Convey Motion Graphics for People who are Blind or have Low Vision. in *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* 1–16, https://doi.org/10.1145/3517428.3544797 (2022).
35. Zhang, Z. (Jerry) & Wobbrock, J. O. A11yBoard: Making Digital Artboards Accessible to Blind and Low-Vision Users. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–17, https://doi.org/10.1145/3544548.3580655.
36. Melfi, G., Müller, K., Schwarz, T., Jaworek, G. & Stiefelhagen, R. Understanding what you feel: A Mobile Audio-Tactile System for Graphics Used at Schools with Students with Visual Impairment. in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* 1–12, https://doi.org/10.1145/3313831.3376508 (2020).
37. Chase, E. D. Z. *et al.* PantoGuide: A Haptic and Audio Guidance System To Support Tactile Graphics Exploration. in *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* 1–4, https://doi.org/10.1145/3373625.3418023 (2020).
38. Siu, A., S-H Kim, G., O'Modhrain, S. & Follmer, S. Supporting Accessible Data Visualization Through Audio Data Narratives. in *CHI Conference on Human Factors in Computing Systems* 1–19, https://doi.org/10.1145/3491102.3517678 (2022).
39. Bodi, A. *et al.* Automated Video Description for Blind and Low Vision Users. in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* 1–7, https://doi.org/10.1145/3411763.3451810 (2021).
40. Huang, J., Siu, A., Hanocka, R. & Gingold, Y. ShapeSonic: Sonifying Fingertip Interactions for Non-Visual Virtual Shape Perception. in *SIGGRAPH Asia 2023 Conference Papers* 1–9, https://doi.org/10.1145/3610548.3618246 (2023).
41. Başçiftçi, F. & Eldem, A. An interactive and multi-functional refreshable Braille device for the visually impaired. *Displays* **41**, 33–41 (2016).
42. Koone, J. C. *et al.* Data for all: Tactile graphics that light up with picture-perfect resolution. *Sci. Adv.* **8**, eabq2640 (2022).
43. Matsunaga, T., Totsu, K., Esashi, M. & Haga, Y. Tactile display using shape memory alloy micro-coil actuator and magnetic latch mechanism. *Displays* **34**, 89–94 (2013).
44. Engel, C., Konrad, N. & Weber, G. TouchPen: Rich Interaction Technique for Audio-Tactile Charts by Means of Digital Pens. in *Computers Helping People with Special Needs (ICCHP)* 446–455, https://doi.org/10.1007/978-3-030-58796-3_52 (2020).

45. Lee, J., Herskovitz, J., Peng, Y.-H. & Guo, A. ImageExplorer: Multi-Layered Touch Exploration to Encourage Skepticism Towards Imperfect AI-Generated Image Captions. in *CHI Conference on Human Factors in Computing Systems* 1–15, https://doi.org/10.1145/3491102.3501966 (2022).
46. Twyman, M., Mullenbach, J., Shultz, C., Colgate, J. E. & Piper, A. M. Designing Wearable Haptic Information Displays for People with Vision Impairments. in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI)* 341–344, https://doi.org/10.1145/2677199.2680578 (2015).
47. Carcedo, M. G. *et al.* HaptiColor: Interpolating Color Information as Haptic Feedback to Assist the Colorblind. in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* 3572–3583, https://doi.org/10.1145/2858036.2858220 (2016).
48. Thompson, J. R., Martinez, J. J., Sarikaya, A., Cutrell, E. & Lee, B. Chart Reader: Accessible Visualization Experiences Designed with Screen Reader Users. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–18, https://doi.org/10.1145/3544548.3581186 (2023).
49. Gorniak, J., Kim, Y., Gwon, S., Wei, D. & Kim, N. W. VizAbility: Multimodal Accessible Data Visualization with Keyboard Navigation and Conversational Interaction. Preprint at http://arxiv.org/abs/2310.09611 (2023).
50. Moured, O. *et al.* Chart4Blind: An Intelligent Interface for Chart Accessibility Conversion. Preprint at https://doi.org/10.1145/3640543.3645175 (2024).
51. Seeing AI - Talking Camera for the Blind. https://www.seeingai.com/ (2024).
52. Be My Eyes - See the world together. https://www.bemyeyes.com/ (2024).
53. Pundlik, S., Shivshanker, P. & Luo, G. Impact of Apps as Assistive Devices for Visually Impaired Persons. *Annu. Rev. Vis. Sci.* **9**, 111–130 (2023).
54. Dai, X. *et al.* A phonic Braille recognition system based on a self-powered sensor with self-healing ability, temperature resistance, and stretchability. *Mater. Horiz.* **9**, 2603–2612 (2022).
55. Liu, W. *et al.* Enhancing blind-dumb assistance through a self-powered tactile sensor-based Braille typing system. *Nano Energy* **116**, 108795 (2023).
56. Thiele, J. E., Pratte, M. S. & Rouder, J. N. On perfect working-memory performance with large numbers of items. *Psychon. Bull. Rev.* **18**, 958–963 (2011).
57. Li, J., Yan, Z., Shah, A., Lazar, J. & Peng, H. Toucha11y: Making Inaccessible Public Touchscreens Accessible. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–13, https://doi.org/10.1145/3544548.3581254 (2023).
58. Understanding Success Criterion 1.1.1 | Understanding WCAG 2.0. https://www.w3.org/TR/UNDERSTANDING-WCAG20/text-equiv-all.html (2023).
59. MacLeod, H., Bennett, C. L., Morris, M. R. & Cutrell, E. Understanding Blind People's Experiences with Computer-Generated Captions of Social Media Images. in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* 5988–5999, https://doi.org/10.1145/3025453.3025814 (2017).
60. Nair, V., Zhu, H. 'Hazel' & Smith, B. A. ImageAssist: Tools for Enhancing Touchscreen-Based Image Exploration Systems for Blind and Low Vision Users. in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* 1–17, https://doi.org/10.1145/3544548.3581302 (2023).
61. He, X. & Deng, L. Deep Learning for Image-to-Text Generation: A Technical Overview. *IEEE Signal Process. Mag.* **34**, 109–116 (2017).
62. Liu, S., Zhu, Z., Ye, N., Guadarrama, S. & Murphy, K. Improved Image Captioning via Policy Gradient optimization of SPIDEr. in *2017 IEEE International Conference on Computer Vision (ICCV)* 873–881, https://doi.org/10.1109/ICCV.2017.100 (2017).
63. Stefanini, M. *et al.* From Show to Tell: A Survey on Deep Learning-Based Image Captioning. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 539–559 (2023).
64. Gurari, D. *et al.* VizWiz Grand Challenge: Answering Visual Questions from Blind People. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 3608–3617, https://doi.org/10.1109/CVPR.2018.00380 (2018).
65. Tseng, Y.-Y., Bell, A. & Gurari, D. VizWiz-FewShot: Locating Objects in Images Taken by People With Visual Impairments. Preprint at http://arxiv.org/abs/2207.11810 (2022).
66. Mandal, M., Ghadiyaram, D., Gurari, D. & Bovik, A. C. Helping Visually Impaired People Take Better Quality Pictures. *IEEE Trans. Image Process.* **32**, 3873–3884 (2023).
67. Murphy, E., Kuber, R., McAllister, G., Strain, P. & Yu, W. An empirical investigation into the difficulties experienced by visually impaired Internet users. *Univers. Access Inf. Soc.* **7**, 79–91 (2008).
68. Ohshiro, T., Angelaki, D. E. & DeAngelis, G. C. A Normalization Model of Multisensory Integration. *Nat. Neurosci.* **14**, 775–782 (2011).
69. Talsma, D., Senkowski, D., Soto-Faraco, S. & Woldorff, M. G. The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* **14**, 400–410 (2010).

70. Toffa, O. K. & Mignotte, M. A Hierarchical Visual Feature-Based Approach For Image Sonification. *IEEE Trans. Multimed.* **23**, 706–715 (2021).
71. Yuksel, B. F. *et al.* Human-in-the-Loop Machine Learning to Increase Video Accessibility for Visually Impaired and Blind Users. in *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS)* 47–60, https://doi.org/10.1145/3357236.3395433 (2020).
72. Sharif, A., Wang, O. H., Muongchan, A. T., Reinecke, K. & Wobbrock, J. O. VoxLens: Making Online Data Visualizations Accessible with an Interactive JavaScript Plug-In. in *CHI Conference on Human Factors in Computing Systems* 1–19, https://doi.org/10.1145/3491102.3517431 (2022).
73. Panotopoulou, A., Zhang, X., Qiu, T., Yang, X.-D. & Whiting, E. Tactile line drawings for improved shape understanding in blind and visually impaired users. *ACM Trans. Graph.* **39**, (2020).
74. Holloway, L., Marriott, K. & Butler, M. Accessible Maps for the Blind: Comparing 3D Printed Models with Tactile Graphics. in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1–13, https://doi.org/10.1145/3173574.3173772 (2018).
75. Engel, C., Müller, E. F. & Weber, G. Tactile Heatmaps: A Novel Visualisation Technique for Data Analysis with Tactile Charts. in *Proceedings of the 14th PErvasive Technologies Related to Assistive Environments Conference (PETRA)* 16–25, https://doi.org/10.1145/3453892.3458045 (2021).
76. Melfi, G., Schölch, L., Schwarz, T., Müller, K. & Stiefelhagen, R. Audio-Tactile Reader (ATR): Interaction Concepts for Students with Blindness to Explore Digital STEM Documents on a 2D Haptic Device. in *2022 IEEE Haptics Symposium (HAPTICS)* 1–6, https://doi.org/10.1109/HAPTICS52432.2022.9765568 (2022).
77. Liu, H. *et al.* Angel's Girl for Blind Painters: An Efficient Painting Navigation System Validated by Multimodal Evaluation Approach. *IEEE Trans. Multimed.* **25**, 2415–2429 (2023).
78. Lee, Y., Joh, H., Yoo, S. & Oh, U. AccessComics: an accessible digital comic book reader for people with visual impairments. in *Proceedings of the 18th International Web for All Conference (W4A)* 1–11, https://doi.org/10.1145/3430263.3452425 (2021).
79. Srinivasan, A., Lee, B., Henry Riche, N., Drucker, S. M. & Hinckley, K. InChorus: Designing Consistent Multimodal Interactions for Data Visualization on Tablet Devices. in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* 1–13, https://doi.org/10.1145/3313831.3376782 (2020).
80. Stein, B. E. & Stanford, T. R. Multisensory integration: current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**, 255–266 (2008).
81. McDonald, J. J., Teder-Sälejärvi, W. A. & Ward, L. M. Multisensory Integration and Crossmodal Attention Effects in the Human Brain. *Science* **292**, 1791–1791 (2001).
82. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–417 (2008).
83. Noppeney, U. Perceptual Inference, Learning, and Attention in a Multisensory World. *Annu. Rev. Neurosci.* **44**, 449–473 (2021).
84. Foxe, J. J. Multisensory Integration: Frequency Tuning of Audio-Tactile Integration. *Curr. Biol.* **19**, R373–R375 (2009).
85. Vercillo, T. & Gori, M. Attention to sound improves auditory reliability in audio-tactile spatial optimal integration. *Front. Integr. Neurosci.* **9** (2015).
86. Wang, J., Wang, S. & Zhang, Y. Artificial intelligence for visually impaired. *Displays* **77**, 102391 (2023).
87. Fei, N. *et al.* Towards artificial general intelligence via a multimodal foundation model. *Nat. Commun.* **13**, 3094 (2022).
88. Mukhiddinov, M. & Cho, J. Smart Glass System Using Deep Learning for the Blind and Visually Impaired. *Electronics* **10**, 2756 (2021).
89. Ahmetovic, D., Kwon, N., Oh, U., Bernareggi, C. & Mascetti, S. Touch Screen Exploration of Visual Artwork for Blind People. in *Proceedings of the Web Conference 2021 (WWW)* 2781–2791, https://doi.org/10.1145/3442381.3449871 (2021).
90. Holloway, L., Marriott, K., Butler, M. & Borning, A. Making Sense of Art: Access for Gallery Visitors with Vision Impairments. in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* 1–12, https://doi.org/10.1145/3290605.3300250 (2019).
91. Bornschein, J., Bornschein, D. & Weber, G. Blind Pictionary: Drawing Application for Blind Users. in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* 1–4, https://doi.org/10.1145/3170427.3186487 (2018).
92. OrCam MyEye 3 Pro - Revolutionize Your Vision with Cutting-Edge AI Technology. OrCam Techologies https://www.orcam.com/en-us/orcam-myeye-3-pro (2024).
93. Granquist, C. et al. Evaluation and Comparison of Artificial Intelligence Vision Aids: Orcam MyEye 1 and Seeing AI. *J. Vis. Impair. Blind.* **115**, 277–285 (2021).
94. Goncu, C. & Marriott, K. in *Human-Computer Interaction – INTERACT 2011* (eds. Campos, P. et al.) 30–48 (Springer, 2011).

95. de Jesus Oliveira, V. A., Brayda, L., Nedel, L. & Maciel, A. Designing a Vibrotactile Head-Mounted Display for Spatial Awareness in 3D Spaces. *IEEE Trans. Vis. Comput. Graph* **23**, 1409–1417 (2017).
96. Peiris, R. L., Peng, W., Chen, Z. & Minamizawa, K. Exploration of cuing methods for localization of spatial cues using thermal haptic feedback on the forehead. in *2017 IEEE World Haptics Conference (WHC)* 400–405, https://doi.org/10.1109/WHC.2017.7989935 (2017).
97. Tang, Y. *et al.* Advancing haptic interfaces for immersive experiences in the metaverse. *Device* **2**, 100365 (2024).
98. Tashakori, A. *et al.* Capturing complex hand movements and object interactions using machine learning-powered stretchable smart textile gloves. *Nat. Mach. Intell.* **6**, 106–118 (2024).
99. Liu, G. *et al.* Keep the Phone in Your Pocket: Enabling Smartphone Operation with an IMU Ring for Visually Impaired People. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **4**, 1–23 (2020).
100. Feng, C. Designing Wearable Mobile Device Controllers for Blind People: A Co-Design Approach. in *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)* 341–342, https://doi.org /10.1145/2982142.2982144 (2016).
101. Wang, R., Yu, C., Yang, X.-D., He, W. & Shi, Y. EarTouch: Facilitating Smartphone Use for Visually Impaired People in Mobile and Public Scenarios. in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* 1–13, https://doi.org /10.1145/3290605.3300254 (2019).
102. Ye, H., Malu, M., Oh, U. & Findlater, L. Current and future mobile and wearable device use by people with visual impairments. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 3123–3132, https://doi.org /10.1145/2556288.2557085 (2014).
103. Young, G. *et al.* Designing Mid-Air Haptic Gesture Controlled User Interfaces for Cars. *Proc. ACM Hum.-Comput. Interact.* **4**, 81:1-81:23 (2020).
104. Tchantchane, R., Zhou, H., Zhang, S. & Alici, G. A Review of Hand Gesture Recognition Systems Based on Noninvasive Wearable Sensors. *Adv. Intell. Syst.* **5**, 2300207 (2023).
105. Reading Speed Statistics – WordsRated. https://wordsrated.com/reading-speed-statistics/ (2023).
106. Wright, T., Wormsley, D. P. & Kamei-Hannan, C. Hand Movements and Braille Reading Efficiency: Data from the Alphabetic Braille and Contracted Braille Study. *J. Vis. Impair. Blind.* **103**, 649–661 (2009).
107. Legge, G. E., Madison, C. M. & Mansfield, J. S. Measuring Braille reading speed with the MNREAD test. *Vis. Impair. Res.* **1**, 131–145 (1999).
108. McCarthy, T., Holbrook, C., Kamei-Hannan, C. & D'Andrea, F. M. Speed and Accuracy Measures of School-Age Readers With Visual Impairments Using a Refreshable Braille Display. *J Spec. Educ. Technol.* **38**, 423–433 (2023).
109. Guerreiro, J. & Gonçalves, D. Faster Text-to-Speeches: Enhancing Blind People's Information Scanning with Faster Concurrent Speech. in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS)* 3–11, https://doi.org /10.1145/2700648.2809840 (2015).
110. Choi, D., Kwak, D., Cho, M. & Lee, S. 'Nobody Speaks that Fast!' An Empirical Study of Speech Rate in Conversational Agents for People with Vision Impairments. in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* 1–13, https://doi.org /10.1145/3313831.3376569 (2020).
111. Tauroza, S. & Allison, D. Speech Rates in British English. *Appl. Linguist.* **11**, 90–105 (1990).
112. Tuckute, G. *et al.* Driving and suppressing the human language network using large language models. *Nat. Hum. Behav.* **8**, 544–561 (2024).
113. Mahowald, K. *et al.* Dissociating language and thought in large language models. *Trends Cogn. Sci.* **28**, 517–540 (2024).
114. Guinness, D., Cutrell, E. & Morris, M. R. Caption Crawler: Enabling Reusable Alternative Text Descriptions using Reverse Image Search. in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1–11, https://doi.org /10.1145/3173574.3174092 (2018).
115. Stangl, A., Verma, N., Fleischmann, K. R., Morris, M. R. & Gurari, D. Going Beyond One-Size-Fits-All Image Descriptions to Satisfy the Information Wants of People Who are Blind or Have Low Vision. in *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)* 1–15, https://doi.org /10.1145/3441852.3471233 (2018).
116. Stangl, A., Morris, M. R. & Gurari, D. 'Person, Shoes, Tree. Is the Person Naked?' What People with Vision Impairments Want in Image Descriptions. in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* 1–13, https://doi.org /10.1145/3313831.3376404 (2020).
117. Srivatsan, N., Samaniego, S., Florez, O. & Berg-Kirkpatrick, T. Alt-Text with Context: Improving Accessibility for Images on Twitter. Preprint at https://doi.org/10.48550/arXiv.2305.14779 (2024).

Highlighted references

9. Sun, Z., Zhu, M., Shan, X. & Lee, C. Augmented tactile-perception and haptic-feedback rings as human-machine interfaces aiming for immersive interactions. *Nat. Commun.* **13**, 5224 (2022).

This paper demonstrates tactile perceptions and haptic-feedback rings with multimodal sensing and feedback capabilities, providing an interactive metaverse platform with cross-space perception capability.

11. Li, Z. *et al.* Electrostatic Smart Textiles for Braille-To-Speech Translation. *Adv. Mater.* **36**, 2313518 (2024).

This paper reports a fabric-based wearable Braille-to-speech translation system, advancing assistive technology to improve the lives of disabled individuals.

16. Lin, M., Hu, H., Zhou, S. & Xu, S. Soft wearable devices for deep-tissue sensing. *Nat. Rev. Mater.* **7**, 850–869 (2022).

This review surveys the engineering of soft wearable devices that sense signals in deep tissues, discussing remaining challenges and research directions.

24. Slade, P., Tambe, A. & Kochenderfer, M. J. Multimodal sensing and intuitive steering assistance improve navigation and mobility for people with impaired vision. *Sci. Robot.* **6**, eabg6594 (2021).

This paper reports an augmented cane, equipped with advanced sensors and intuitive feedback, significantly improves mobility and navigation for people with impaired vision, enhancing their independence and quality of life.

53. Pundlik, S., Shivshanker, P. & Luo, G. Impact of Apps as Assistive Devices for Visually Impaired Persons. *Annu. Rev. Vis. Sci.* **9**, 111–130 (2023).

This review discusses the availability and capability of current vision assistance apps, providing significant insights to the developers of assistive technologies.

83. Noppeney, U. Perceptual Inference, Learning, and Attention in a Multisensory World. *Annu. Rev. Neurosci.* **44**, 449–473 (2021).

This review elucidates how the brain integrates and adapts multisensory signals, particularly under causal uncertainty and in complex environments, offering significant insights into sensory processing and attention mechanisms.

87. Fei, N. *et al.* Towards artificial general intelligence via a multimodal foundation model. *Nat. Commun.* **13**, 3094 (2022).

This paper presents a foundation model pre-trained with large multimodal data, adaptable for various downstream cognitive tasks.

98. Tashakori, A. *et al.* Capturing complex hand movements and object interactions using machine learning-powered stretchable smart textile gloves. *Nat. Mach. Intell.* **6**, 106–118 (2024).

This paper reports a dynamic and accurate tracking of hand movements using washable and wireless textile gloves, unlocking new avenues in human-computer interaction and assistive technologies.

Acknowledgments

Y.M. acknowledges financial support from the Hong Kong Research Grants Council (ECS 25228722). X.M. acknowledges the funding from MTC Young Individual Research Grants (YIRG M23M7c0129).

Author contributions

W.X and Y.R. researched data for the article. W.X., Y.R., Y.T., Z.G., X.M., J.Z., and Y.M. contributed to the writing and editing of this manuscript. Y.T., X.M., Z.W., H.X., T.Z., C.Z., X.D., and Y.M. contributed to the discussion, reviewing the manuscript before submission. All the authors contributed substantial contributions to the discussion of content.

Competing interests statement

The authors declare that they do not have any competing interests.

Key points

- BLV individuals rely on non-visual information delivery (NVID) when interacting with digital information, leading to low information efficiency, reliability issues, and extra cognitive burden.
- Limitations of NVID are related to the sequential information output and low-quality alternative texts.
- There are three design ideas for future wearable assistive devices that may break through current bottlenecks: multisensory learning, gesture interaction, and synchronization of tactile and LVLMs.
- Future wearable devices will enhance text and image information delivery for BLV users by integrating multimodal interaction, contextual reasoning, and personalized content.

ToC blurb

Blind and low-vision individuals face significant challenges in digital information interaction. This perspective highlights the bottlenecks in wearable devices for information accessibility, design ideals for future assistive devices, and outlook for improving text and image information delivery.