A Novel Wearable and Wireless Device to Investigate Perception in Interactive Scenarios

L. Schiatti[1], G. Cappagli[1,2], C. Martolini[1], A. Maviglia[3], S. Signorini[2], M. Gori[1], and M. Crepaldi [3]

***Abstract*— The aim of the present work is to introduce a novel wearable device suitable to be used to investigate perception in interactive tasks, on individuals with and without sensory disabilities. The system is composed by small units embedded with sensors and actuators that allows emitting different kind of stimuli (light, haptic, sound) and to record the user response, thanks to a capacitive sensor. We validated the system by implementing an interception task in three different sensory modalities: visual, tactile and auditory. Six subjects with normal sight were asked to tap either a static or a moving stimulus generated by 6 units placed on their forearm. Results suggest that the system can effectively provide new insights in characterizing how perception principles vary when perceptual judgement occurs through different senses. This confirms the device potential in contributing to the design of rehabilitation protocols rooted on neuroscientific findings, for people with sensory impairments.**

I. INTRODUCTION

Impairments in the typical functioning of a sensory modality, such as problems related to vision, can affect a child's overall development. This includes spatial and locomotor competences, interpretation and integration of input from other senses, cognitive and social skills [1]. Given the importance of vision in almost all instrumental and social activities of daily living, there is a big effort towards the development of technologies for quantitative assessment of visual impairments and to enhance rehabilitation procedures. Indeed, although it is generally assumed that a visual deficit may be compensated by increased auditory and tactile skills, it was demonstrated that vision impairments could negatively affect also the development of other sensory modalities [2]. A visual rehabilitation therapy based on appropriate stimulation often improves a child's functional vision, i.e. functional activities that are normally vision-dependent, thus facilitating the development in all areas [3].

In this context, we present a novel technological system that can be used both for assessment of developmental stages of perceptive mechanisms (e.g. unimodal and multimodal perception), and to design rehabilitation protocols based on sensory stimulation, targeting in particular children who present visual impairments. In this work, we focused on the perceptive mechanism of interception. Interception is a fundamental ability in daily activities that requires consolidated spatiotemporal skills, since it is based on the ability to predict and integrate individual's movement and the movement of the object to be intercepted, in order to produce a timed motor response [4]. While most of the research related to interception has mainly focused on the role of vision in guiding motor actions to reach an object [5], recent studies demonstrated that it is possible to localize and intercept a moving object even relying on other senses such as touch and hearing. For instance, it has been shown that interception of a tactile target can be modeled in a similar way to interception of visual targets, resulting in an overall interception accuracy dependent on haptic judgments on position and velocity of the moving stimulus [6]. Similarly, it has been demonstrated that sound contains prospective information for the purposeful control of goal-directed behavior, more specifically individuals can vary the velocity profile of their intercepting movements based on the dynamics of the moving sound object [7], [8].

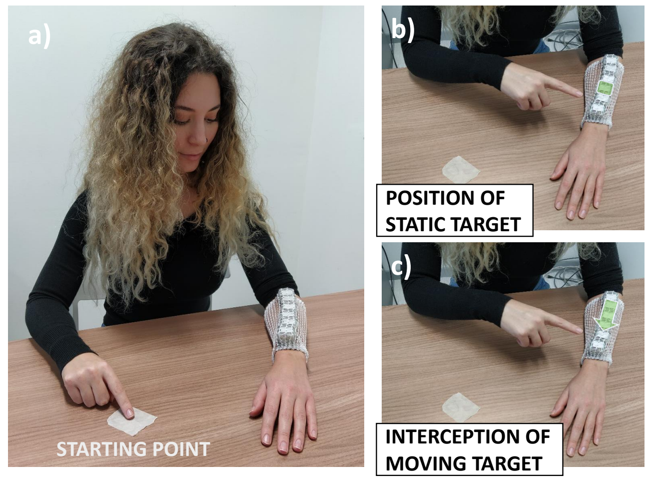


Fig. 1: Experimental setup for the TechARM system validation (a); Position (b) and Interception (c) tasks protocols.

Despite the relevance of understanding such perceptual mechanisms, extensive data assessing a direct comparison of interception accuracy across senses (vision, touch, hearing) on the body are still missing to date. To this purpose, we introduce a novel technological device, called TechARM, suitable to be used to investigate perception during interactive tasks. We validated the system during a manual interception task performed on a group of sighted adult subjects. Thanks to sensors and actuators embedded on the device, it was possible both to deliver to the subjects visual, auditory and tactile stimuli, either static or moving across the subject's arm, and to directly record users' responses, without the need of complex equipment, e.g. a motion tracking system.

[1]Unit for Visually Impaired People (U-VIP), Istituto Italiano di Tecnologia, Genoa, Italy (Email: lucia.schiatti@iit.it).
[2]Center of Child Neuro-Ophthalmology, IRCCS Mondino Foundation, Pavia, Italy.
[3]Electronic Design Laboratory (EDL), Istituto Italiano di Tecnologia, Genoa, Italy.

978-1-7281-1990-8/20/$31.00 ©2020 IEEE

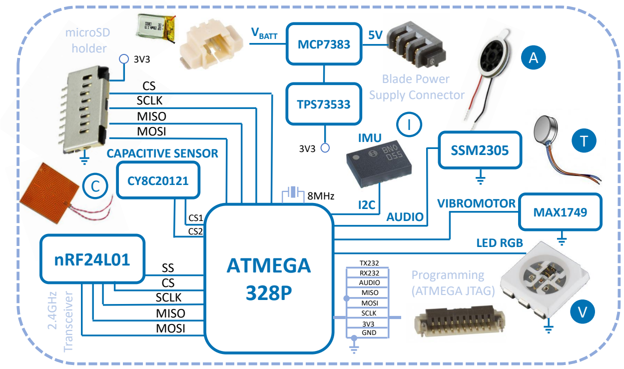


Fig. 2: TechARM architecture, including actuators and sensors technical specifications.

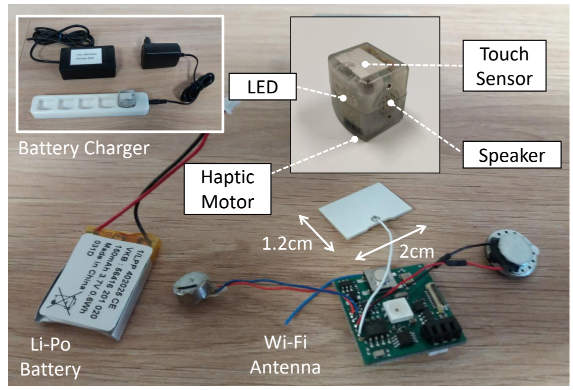


Fig. 3: TechARM unit final prototype. All components are enclosed within a biocompatible semi-transparent plastic (top–right). The unit is accessible from the bottom by a customized battery charger (top–left).

II. METHODS AND MATERIALS

A. TechARM system:

Technological ARM (TechARM) is a portable and battery powered device comprising sensors and actuators to enable visual, haptic and auditory interaction, including touch input from the user.

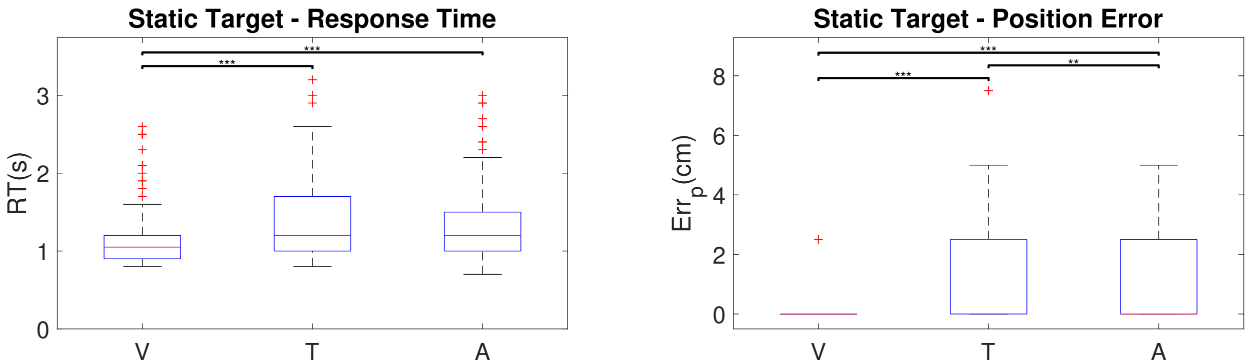
a) Hardware: Fig. 2 shows the block scheme of the wearable device. The platform, implemented at firmware-level at bare-metal to minimize latency, is based on an ATMEGA328P Micro-Controller Unit (MCU) that interfaces a two-channel CY8C20121 touch sensor (C), a BNO055 I2C Inertial Module Unit (IMU) useful to acquire inertial data (I), a MAX1749 haptic moto driver (T), controlled using a single GPIO, a nRF24L01 2.4 GHz ISM transceiver through a dedicated Synchronous Peripheral Interface (SPI) bus, a WS2812 Red Green Blue (RGB) LED (V), a microSD card holder and a SSM2305RMZ speaker driver (A). Both the touch sensor and the IMU are input peripherals while the others are output peripherals. TechARM includes an internal Li-Po 160 mAh accumulator. Battery charge is handled using an on-board MCP73831T that accepts a 5 V input from a dedicated connector. The external voltage is converted to the system 3.3 V supply (powering all the sub-modules) with a TPS73533 voltage regulator. Up to five TechARM can be recharged ($\sim$2 hr) using a full-custom charger that integrates an array of blade type power connectors that provide a 5 V stable supply. Fig. 2 shows the final TechARM prototype including the battery charger and the enclosure obtained using a photo-polimerizing resin.

b) Software: The firmware that runs on the MCU is conceived to enable immediate message-based event-driven operation with maximum availability: the system can be polled anytime to read-out the input peripheral data (touch detection and IMU data), and at the same time it can provide output events to the vibromotor (haptic), speaker (audio) and LED (visual) peripherals. The firmware implements commands using such atomic coarseness to enable a high-level programming using a Matlab^{TM}or Python environment, through a serial port connection with a 2.4 GHz transceiver. Complex experiments can be designed based on a predetermined sequence of commands on single or multiple units. Moreover, the possibility of recording touch events is novel compared to previously released systems [9]. The firmware also enables the play of audio files in the microSD card of the module, stored in a WAV format. Device addressing can be achieved in parallel up to 255 units, and for each unit the internal input/output peripheral can be addressed. All the commands have a constant length of 8 bytes, and based on the command type TechARM can reply with 2, 4 or 32 bytes. Latency (measured in the range 2-5 ms), does not depend on networking because the RF transceiver work on a 2.4 GHz ISM band with proprietary encoding and modulation hence, not subjected to, e.g., TCP/IP network loads.

c) Interactive tasks: In order to use TechARM for implementing an interception protocol, the units were programmed such that during each trial it was possible to arbitrarily activate audio, visual, and haptic peripherals of each unit with a preset delay (50ms resolution), and to record touch responses on each unit for a time up to six seconds, with a resolution of 100ms. These reads were than sent to the PC through serial communication after each trial.

B. Experimental Protocol:

We implemented two experiments, as shown in Fig.1: Position and Interception. The protocol for both experiments followed the guidelines present in the related literature [6]. Six sighted adults (average age 27 ± 3.28) took part in both experiments, after giving their written consent with the experimental procedure, approved by the local Ethics Committee (ASL3 Genovese). Participants were sitting with their hands laying on a table, with six TechARM units placed in a row on their left/ right arm (right-handed/ left-handed). They were asked to perform a tapping task, with the goal of matching the position of a static stimulus on their arm (Position experiment), or intercepting it while it was moving along the arm (Interception experiment). The stimuli, either static or moving, was delivered by the TechARM units in three distinct conditions: visual (V), tactile (T), and auditory (A), in the form of a white LED lighting, a vibration, and a white noise sound, respectively. Each trial started with the subjects placing their dominant-hand index finger at a



(a) (b)

Fig. 4: Position experiment results (distributions among subjects) in terms of response time RT (a) and Position Error Err_p (b), comparing V, T and A conditions. Each box indicates the distribution median (red line), the 25th and 75th percentiles, and outliers ('+'). Stars highlight significant differences between conditions according to paired permutation t-tests ($*p \leq 0.05$, $**p \leq 1e-2$, $***p \leq 1e-3$).

starting point indicated on the table (see Fig. 1a). After a beep, the stimulation started and the subject was instructed to match the position of the stimulus as fast as possible, by tapping the top surface (embedded with the capacitive sensor) of the active TechARM unit.

a) Position: This task was implemented to take into account individual baseline differences in position detection accuracy for stimuli of different sensory modalities. This allowed us to take in account such differences when interpreting the interception task results. Six TechARM units covered a distance of 15 cm (2.5 cm each unit) on the person's forearm. During the trial, one out of the six units delivered a $200ms$ stimulation, in one of the three sensory modalities, and subjects had to tap the target as soon as they felt it. For each position on the arm, there were three repetitions in each sensory modality, resulting in $6x3x3 = 54$ trials. Performance was evaluated in terms of Response Time (RT), i.e. the time between the start of trial and the first touch detected, and Position Error (Err_p), i.e. the distance between the selected unit and the active unit.

b) Interception: During the interception experiment, the TechARM units were programmed to deliver a certain stimulation (either V, T or A), with a progressive delay from one end of the row to the opposite one, thus simulating a moving stimulus. Participants were instructed to intercept the stimulus by tapping on the active unit, as fast as possible after the trial started. Three stimuli durations were implemented, namely $100ms$, $200ms$, and $300ms$, leading to an apparent speed of approximately $25cm/s$ (high), $12.5cm/s$ (medium) and $8.3cm/s$ (low) respectively. Two starting positions were possible, at each end of the units row. Thus the stimulus could travel in the elbow-to-wrist direction or the opposite one, resulting in a total of $3x3x2 = 18$ conditions. For each condition there were 5 repetitions, leading to a total of 90 trials per participant. Performance was evaluated in terms of Response Time (RT) and Time Error (Err_t), i.e. the time interval between the tap and the time in which the selected unit was active ($Err_t > 0$ if the unit was tapped after the end of its activation time, $Err_t < 0$ if tapped before its activation time, $Err_t = 0$ if tapping matched time of unit's activation).

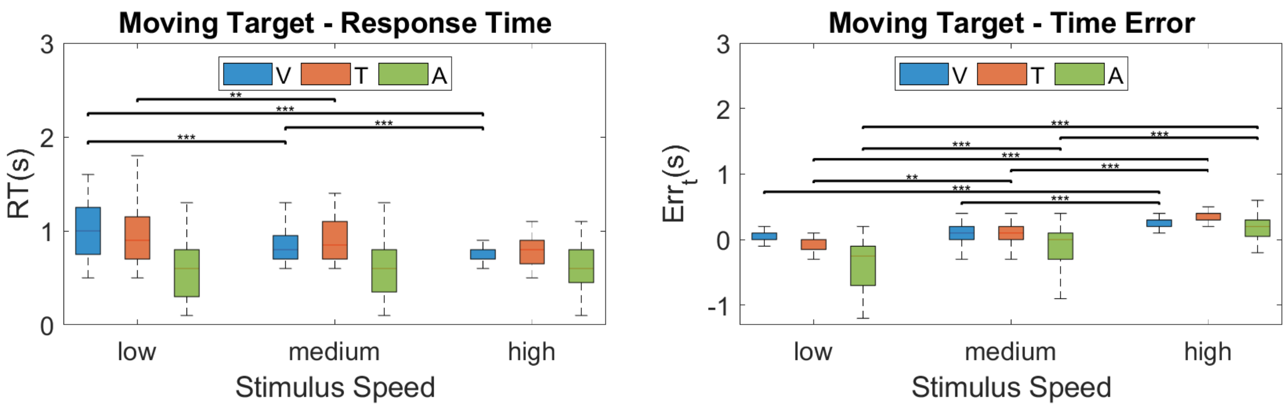
III. RESULTS AND DISCUSSION

A. Position experiment:

For what concerns static targets, we found that the ability to localize stimuli on the arm strongly depends on the sensory information provided (see Fig. 4b). More specifically, after assessing that data don't follow a normal distribution (Shapiro-Wilk test), we applied paired permutation t-tests with Bonferroni correction for multiple comparisons. Results suggest that participants tend to localize less accurately T compared to V stimuli ($p = 1.92e-19$), A compared to V stimuli ($p = 5.48e-14$) and T compared to A stimuli ($p = 0.01$), therefore participants seem to be overall more impaired in tactile localization of static stimuli on the arm. On the contrary, as shown on Fig. 4a, no significant difference exists between audition and touch when considering reaction times ($p = 0.26$), while V significantly differs from T ($p = 4.00e-6$) and A ($p = 4.27e-4$). Overall, the visual modality leads to faster and more precise responses. This can be explained by taking into consideration the spatial and temporal resolution of the visual system, being dependent on the resolution of the fovea. On the other hand, individuals similarly react to A and T sensory stimulation in terms of velocity but they better identify the location of the stimulation when using the A sensory modality. Such observation could be explained both by the type of stimulus, i.e. a vibration that is transmitted through the whole TechARM unit bottom surface, and by the resolution of skin receptors on the arm.

B. Interception experiment:

For what concerns dynamic targets, a permuted two-way ANOVA analysis revealed for both RT and Err_t a significant effect of the two within-subjects factors, i.e. stimulus type ($p < 2.20e-16$) and speed ($p < 2.20e-16$), as well as of their interaction ($p = 1.00e-3$). For all sensory modalities except the auditory one RT decreases with increasing stimulus velocity (see Fig. 5a). This confirms previous literature findings on tactile and visual interception [6], and supports the idea that our device is suitable for the purpose of investigating perception. Results from Err_t, shown in Fig. 5b, indicate that also interception accuracy



(a) (b)

Fig. 5: Interception experiment results (distributions among subjects) in terms of response time RT (a) and Time Error Err_t (b), for low, medium and high speed of moving target, comparing V, T and A conditions. Outliers are not shown. Stars highlight significant differences between conditions according to paired permutation t-tests ($*p \leq 0.05$, $**p \leq 1e-2$, $***p \leq 1e-3$).

varies depending on the stimulus velocity: while participants correctly intercept the moving stimulus across all sensory conditions for medium velocity (200 ms), they manifest a motor delay resulting in a positive temporal error for high velocity (100 ms) and a motor anticipation resulting in a negative temporal error for low velocity stimuli (300 ms) in case of acoustic stimuli. This can be interpreted in the context of the existing literature on the topic, showing that individuals intercept auditory target more quickly when fast compared to slow stimuli are presented [7].

Another interesting result comes from the observation that, while all sensory modalities show the same interception accuracy for a medium stimulus velocity, V modality behaves similarly to T and better than A for slow targets, and similarly to A and better than T for fast targets. This suggest that a multisensory audio-tactile feedback could improve the interception capabilities of children with visual disabilities, by providing more robust information than with unimodal feedback. Several studies investigated how interception abilities change when sensory information is temporarily and/or partially unavailable and demonstrated that in such cases a-priori knowledge is fundamental to target interception [10]. Results from our study suggest that auditory and tactile information could indeed be effective in rehabilitating interception capabilities, in presence of a vision deficit.

IV. CONCLUSIONS

We presented a novel wearable device suitable to be used in investigating perceptive capabilities during interactive tasks, such as interception of a moving target. The experimental validation of the system confirmed its effectiveness in reproducing literature results on similar tests. Furthermore, we implemented a novel paradigm, comparing interception accuracy across visual, auditory and tactile sensory modalities. With the present study, we demonstrated that auditory and tactile information, other than vision, can be used to intercept a target on the arm and that the velocity of dynamic stimuli strongly affects interception accuracy. These findings shed light onto the role played by predictive abilities in motor actions planning and therefore they highlight the importance of developing and testing interception abilities in children with sensory disabilities such as blindness. To confirm such results, in future work we plan to test the presented protocol and system on a large sample of both sighted and visually impaired subjects.

ACKNOWLEDGMENT

This research was supported and funded by the Istituto Italiano di Tecnologia, in partnership with the Center of Child Neuro-Ophthalmology, IRCCS Mondino Foundation. The TechARM system was realized by the Unit for Visually Impaired People unit and the Electronic Design Laboratory.

REFERENCES

[1] L. Hyvärinen, "Considerations in evaluation and treatment of the child with low vision," *American Journal of Occupational Therapy*, vol. 49, no. 9, pp. 891–897, 1995.

[2] M. Gori, G. Sandini, C. Martinoli, and D. C. Burr, "Impairment of auditory spatial localization in congenitally blind human subjects," *Brain*, vol. 137, no. 1, pp. 288–293, 2014.

[3] A. Colenbrander, "Assessment of functional vision and its rehabilitation," *Acta ophthalmologica*, vol. 88, no. 2, pp. 163–173, 2010.

[4] E. Brenner and J. B. Smeets, "Continuously updating one's predictions underlies successful interception," *Journal of neurophysiology*, vol. 120, no. 6, pp. 3257–3274, 2018.

[5] N. L. Port, D. Lee, P. Dassonville, and A. P. Georgopoulos, "Manual interception of moving targets i. performance and movement initiation," *Experimental Brain Research*, vol. 116, no. 3, pp. 406–420, 1997.

[6] J. Nelson, G. Baud-Bovy, J. Smeets, and E. Brenner, "Accuracy of intercepting moving tactile targets," *Perception*, vol. 48, no. 8, pp. 685–701, 2019.

[7] N. Komeilipoor, M. W. Rodger, P. Cesari, and C. M. Craig, "Movement and perceptual strategies to intercept virtual sound sources," *Frontiers in neuroscience*, vol. 9, p. 149, 2015.

[8] D. M. Shaffer, I. Dolgov, E. Mcmanama, C. Swank, A. B. Maynor, K. Kelly, and J. G. Neuhoff, "Blind (fold) ed by science: A constant target-heading angle is used in visual and nonvisual pursuit," *Psychonomic bulletin & review*, vol. 20, no. 5, pp. 923–934, 2013.

[9] M. Gori, A. Bollini, A. Maviglia, M. B. Amadeo, A. Tonelli, M. Crepaldi, and C. Campus, "Msi caterpillar: An effective multisensory system to evaluate spatial body representation," in *2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE, 2019, pp. 1–6.

[10] G. Bosco, S. Delle Monache, and F. Lacquaniti, "Catching what we can't see: manual interception of occluded fly-ball trajectories," *PLoS One*, vol. 7, no. 11, 2012.