Visuo-tactile AR for Enhanced Safety Awareness in Human-Robot Interaction

Matti Krüger Martin Weigel Michael Gienger

Honda Research Institute Europe
Offenbach/Main, Germany
{matti.krueger, martin.weigel, michael.gienger}@honda-ri.de

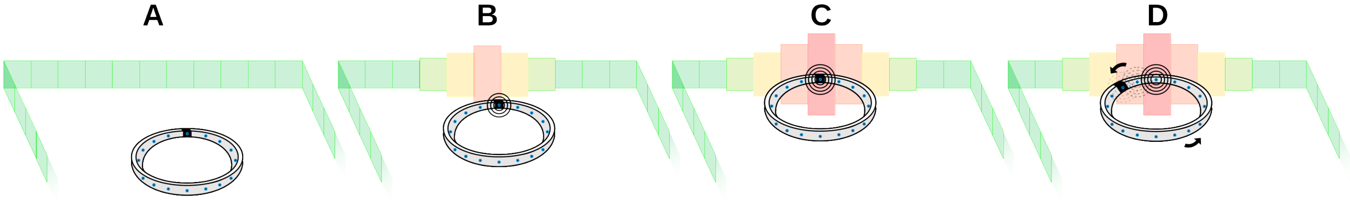


Figure 1: Basic functionality of the visuo-tactile safety augmentation: (A) When located inside an area considered to be safe, the tactile interface produces no output and the visualization shows the boundary-elements of the virtual fence in green. (B) When approaching a border, the border elements located most closely to the user start growing and changing color such that elements grow and turn from green to yellow to red with decreasing distance. The tactile interface starts vibrating at the tactor that is located most closely to the border at a low intensity. (C) When reaching the border, the color change, element size and vibration intensity are at maximum. (D) Rotating the belt shifts activity from the previous actuator to the actuator now located closest to the border.

ABSTRACT

In this workshop submission, we describe our approach for developing a multimodal AR-system that combines visual and tactile cues in order to enhance the safety-awareness of humans in human-robot interaction tasks. Motivated by a competition for attentional resources between the need for safety-maintenance and achievement of a primary task, we employ multimodal cues that inform a user about unsafe proximities to dangerous areas. The system augments a scene with both visual output provided via AR-glasses and tactile stimuli produced by vibration motors embedded into a belt. The tactile belt allows the user to focus visual attention on a primary task while keeping him or her safety-aware. The visual representation that is additionally rendered into the scene provides visual grounding. This feedback is beneficial to a user as well as to external observers in training and supervision scenarios. We tested the system with informed and naive users to iterate over the design and to gain first insights into the utility of our multimodal approach.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Interface design prototyping**; *Interaction techniques.*

KEYWORDS

Virtual fence, assistive technologies, augmented senses, multimodal feedback, augmented reality, head-mounted display, tactile belt, wearable devices

1 INTRODUCTION

Work involving human-robot interaction usually requires the human's visual attention for a successful execution of a task. Especially when such interactions involve certain dynamics such as in a collaborative manipulation of objects (e.g. [21]) or take place in an environment with multiple independent actors such as a production facility, it becomes critical to also monitor the environment for potential safety hazards. Despite the illusion of a momentarily available complete visual scene, human visual information acquisition is spatially constrained by anatomy and thus partially of sequential nature [47]. Any additional load on the visual modality created by the need to monitor the environment for hazards therefore results in a trade-off between information acquisition for safety maintenance and information acquisition for task completion.

We present *AwareWear*, an approach to enhance the safety awareness in human-robot interaction tasks using visuo-tactile feedback. The core idea of this approach lies in creating congruent visual and tactile stimuli that contain information about how safe the current location of a person is and, assuming the location surpasses a danger threshold, information about which direction would increase the danger and thus indirectly also information about the directions considered to be safe for movement. We argue that an informed embedding of direction and safety information into the introduced visual and tactile stimuli can provide intuitive safety augmentation that is compatible with simultaneous human-robot interaction tasks.

In the following we focus on three contributions:

- Background information on the topics of multimodal perception, properties of visual and tactile stimuli and prior art on tactile and visual augmentation of safety awareness.

VAM-HRI 2020, The Second International Workshop on Virtual, Augmented and Mixed Reality for Human-Robot Interaction
2020.

- A description of our user-centered approach to prototype experiences of potential system variants. These prototypes are guided by the previously introduced findings and our design goals of creating a system which (a) reduces the burden on the visual system for monitoring tasks, (b) offers quick and intuitive understanding, and (c) can make safety-hazards apparent to external observers.
- Implementation details on *AwareWear*, a safety awareness augmentation system for human-machine interaction scenarios.

2 TACTILE AND VISUAL SAFETY PERCEPTION

While vision may be the most commonly used sensory modality for interaction tasks and environment monitoring, it is not necessarily the only one capable of doing so. In fact, one's perception generally appears to be the result of information that is integrated from multiple senses [3, 14, 36]. Stimuli that are perceivable via multiple sensory modalities have even been found to cause faster reactions than unimodal signals [2, 10, 26, 42, 56]. Further, not only perception speed but also bandwidth [25, 59] benefits from multisensory integration which suggests an ability to reduce the risk of sensory overload by enabling people to divide perceptual tasks among different senses. Visual and tactile response properties of neurons identified in macaque premotor areas, parietal areas and putamen [4, 6, 15, 23, 24, 49] indicate a close coupling of information from different modalities for events that occur within peripersonal space. These neurons have even been assumed to encode a mental *safety margin* around the body and support the coordination of defensive behavior [22]. This makes multimodal input a promising candidate solution for circumventing the tradeoff between interaction- and safety-maintenance tasks and potentially building upon existing neural representations.

Here we introduce an approach which augments a scene with safety-relevant information provided via both visual and tactile stimuli in order to take advantage of peoples' multimodal sensory capabilities.

Tactile stimuli thereby allow us to exploit a variety of beneficial properties:

- They are easy to localize and an effective method to display location and direction information (see e.g. [13, 28, 29, 32, 33, 37]) due to the ubiquity of tactile receptors across the skin.
- They are automatically attention capturing [50] and thus require no need for a conscious scan of new input.
- They can be used to display information without putting additional load on the visual modality [25, 51, 59].
- They are less affected by background noise than for instance the auditory channel.
- They allow for information to be provided to a specific target person without directly affecting other people in the environment.
- As warning signals they allow for faster reaction times than auditory [13] and even visual [46, 52] warnings (see [37]).
- Aviation studies have further revealed that tactile cues do not interfere with performance of concurrent visual tasks [53] (see [34]).

Visual perception in turn has a variety of complementary properties such as an almost instantaneous recognition of complex patterns [55], saliency- and top-down- based attention guidance [43, 45, 57, 62], *pop-out* effects [35], gestalt perception [44, 61] and a high spatial and temporal resolution [27, 31].

Previous approaches for a tactile augmentation of the perception of safety-relevant information have encoded spatial information about objects of interest in tactile stimuli (e.g. [5, 39, 58]). In some cases a dimension of the stimulus has been used to encode the spatial [1, 5, 7, 16, 48] or the temporal [32, 33] distance to an object of interest. The so called "haptic radar" [7] for instance was motivated by the idea of a spatially extended touch sensitivity and introduced whisker-like properties through a headband equipped with vibromotors which scale their amplitude and frequency according to input by adjoining proximity sensors.

For the creation of visual stimuli which have the purpose to support safety maintenance, a large selection of examples exists out of which many have been integrated into everyday life: Laser or camera-based floor projection techniques are employed in industrial settings. The *Chaperone Bounds* [8] system is a concept for virtual reality (VR) environments to display a grid that indicates the boundary of a workspace once the VR device gets close. In the following we will present our approach to extend such boundary representations for applications in augmented reality (AR) environments and to encode additional features.

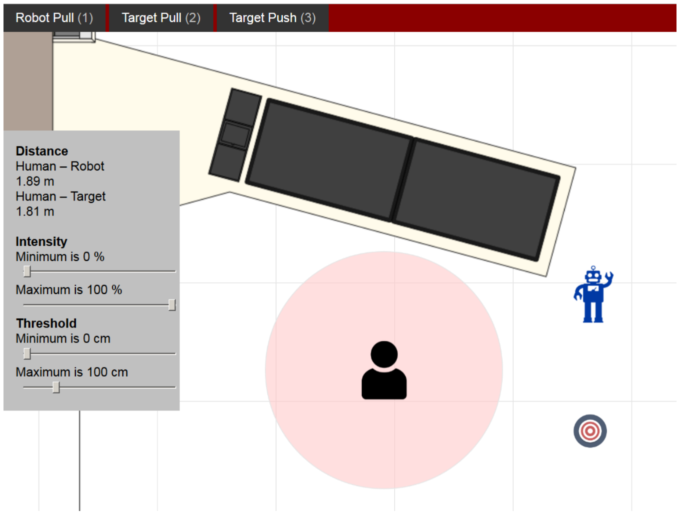
We believe that the utilization of this approach can improve safety awareness and exemplify an advancement in human-computer integration [40] that is characterized by a seamless confluence of sensory capabilities of artificial and biological systems.

3 EXPERIENCE PROTOTYPING

In a first step, we build a Wizard of Oz (WOZ, see [9]) system to explore how informed and naive users respond to different tactile stimuli that encode location information. As an interface we chose to use a *tactile belt* from the feelspace GmbH (see section 4.1 for more details), which is easy to wear, enables 360° stimulation around the wearer's core, and can be triggered wirelessly.

To control the belt we created a graphical user interface (GUI) that interfaces with an application for belt control (Figure 2). The GUI allows a user to change the 2D positions of multiple entities in an environment presented from a birds-eye-view. Depending on a set of adjustable stimulus parameters, our application for belt control then uses the relative arrangement of these entities as the basis for stimulus generation. More specifically, distances between entities thereby determine stimulus intensities and angles between objects determine stimulus locations.

The GUI-controlled application allowed us to investigate various potential system features that rely on location information without having to fully integrate the system with a functional sensor setup. Instead, sensory capabilities of a human controller could be employed for location updates in a Wizard of Oz like manner. When testing ideas with a human user wearing the belt, the human controller matches the locations of the user and possibly other objects (i.e. attractor or robot) in the GUI environment with the actual location of the user or respective entity. Alternatively the virtual



Distance
Human – Robot
1.89 m
Human – Target
1.81 m

Intensity
Minimum is 0 %
Maximum is 100 %

Threshold
Minimum is 0 cm
Maximum is 100 cm

Figure 2: Prototyping interface for WOZ control.

entities can be dragged around according to any object-decoupled protocol implementing for instance a repelling invisible fence.

From the users' feedback, we concluded that vibrations which represent repelling characteristics, by being provided from the direction the user should avoid, were favoured over a signaling from the direction towards a target location. Similarly, stimulus strength, which thereby decreased when moving away from the stimulus direction, was preferred over an increasing stimulus strength. Overall the meaning of tactile stimuli was quickly understood and the pretended functionality considered to be useful. For these reasons we decided to expand the system and enable faster and more accurate feedback by substituting the human controller with realtime location and orientation measurements from artificial sensors.

4 VISUO-TACTILE SAFETY AUGMENTATION

This section describes the implementation of our *AwareWear* system.

The purpose of this system is to intuitively communicate information about the safety of a user's location and further provide guidance to support a user in reaching a safe state in case the safety should fall below a set threshold. Here we conceptualize these properties in the form of a virtual fence which marks the boundaries of a region classified as safe. The definition of these boundaries thereby depends on the given context. For instance, in a joint task with a robot (see figure 8c), the area in which a human can safely move without entering the operation space of the robot, reaching regions in which the robot would be unable to safely support the interaction, or crashing into other objects in the environment, is restricted. These restrictions would define the respective safety boundaries that provide the basis for the function of our system. The system can use tactile and visual stimuli to inform a user and multiple observers about a current location safety and danger direction. In the following sections the interfaces used for stimulus generation as well as the properties of the respective stimuli are illustrated in detail.

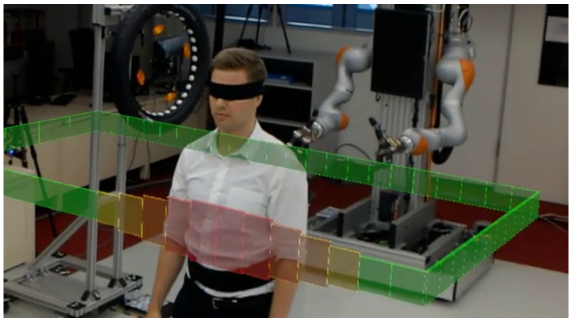


Figure 3: HoloLens output showing the *HoloFence* from the perspective of an external observer. Individual elements are colored and sized according to a person's proximity to the zone boundaries.

4.1 Tactile Interface

The vibrotactile interface consists of a belt with 16 equally spaced vibromotors (feelspace GmbH, further referred to as *tactors*, see [41]) spanning the length of the belt such that the locations of individual tactors can be aligned with directions relative to the wearer's body. This allows for an approximate matching between direction encoding and stimulus position which should facilitate an intuitive understanding of the directional component in signals. To further improve the spatial resolution of the tactile interface, we implemented the option to express an angle as the combination of the relative intensities of two neighboring tactors. Such a co-activation of neighboring actuators can produce a tactile illusion known as *funneling effect* [30] which is experienced as only one stimulus located in between the two and shifted towards one or the other according to the respective relative intensity. The belt uses eccentric rotation mass motors with a maximum amplitude of 2.2 g and a frequency spectrum of 50–240 Hz (0.45 – 3.3 V) triggered with a 50 ms latency. Frequency and amplitude scale almost linearly with voltage.

The belt furthermore contains a digital compass which updates at 10 Hz. We use this compass[1] to determine a wearer's current orientation in the room and thus adapt the mapping of actuators to zone boundaries (see Figure 1D and section 4.3.2).

4.2 Visual Interface

Visual information about the task is given to the user or external observers by means of augmented reality glasses (Microsoft HoloLens). Before starting the task, the glasses are calibrated to a common lab reference using an optical marker, so that the Holograms can be displayed at the spatially consistent locations. The human user, relative to whom safety-measures are calculated, is additionally equipped with a set of infrared markers that are attached to the waist. This allows to precisely track the user with an optical marker tracking system (we use a VICON system with a 12-camera setup).

The human's pose estimate is acquired and transformed into the the lab reference to be consistent with the holographic representation in about 60 Hz. Occlusions are dealt with using a Kalman filter.

---

[1] Orientation information can also be provided by our means of people tracking described in sections 4. However, the independence from these methods allowed us to better cope with temporal loss or increased uncertainty of location information through e.g. visual occlusion.

The stabilized pose estimate is provided to the Stimuli computation module through a ROS interface. A virtual fence AR hologram (*HoloFence*) is computed based on the incoming stimulus information. The fence consists of a set of individual boxes that are aranged according to their supplied environment coordinates. The stimulus information is determining both color and shape of the individual elements, so that a proximity to the safety zone can be overlaid effectively (see Figure 3). Our system is realized as a multi-user system that allows to employ several HoloLens devices sharing the same spatially consistent content. This allows to provide the AR visualization to additional persons, for instance to demonstrate or explain the system to people not involved in the task or to aid a supervisor in monitoring safety in an environment.

4.3 Stimuli

Visual and tactile stimuli are congruently triggered according to location and orientation information in relation to the virtual fence. The rationale of this relationship is to provide no intrusive feedback as long as there is no imminent safety violation and scale saliency-modulating parameters of the stimuli according to the proximity to violating safety- or, more generally, location-conditions.

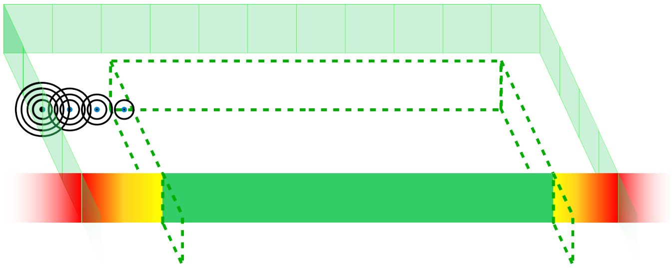


Figure 4: Stimulus Scaling.

4.3.1 Salience. Figure 4 illustrates the relationship between location and stimulus scaling. The area surrounded by a dashed wall (inner fence) represents the area within which no tactile stimuli are generated and no visible deformation or color change of the virtual fence takes place. Note that this area is by default hidden in the visualization. The area between the virtual fence and the dashed wall represents the area in which stimuli scale from a low to a high saliency. Thereby the respective stimulus parameter increases as a function of the spatial proximity to the (outer) virtual fence:

$$\text{Magnitude} = max\left(1 - \frac{d(\text{User}, \text{Fence}_{\text{outer}})}{d(\text{Fence}_{\text{outer}}, \text{Fence}_{\text{inner}})}, 0\right) \qquad (1)$$

where $d(x, y)$ is the euclidean distance $||y - x||$.

Figure 1 illustrates the location-dependent scaling of saliencies[2] by the two visual stimulus parameters of color and size according to Equation 1. A step from green to yellow and subsequent gradient to red is applied as a traffic light metaphor that further increases color contrast to non-violated zone elements and thus local saliency as the proximity to the wall increases. Additionally element size encodes the same property.

Figures 1 and 4 also depict the relationship between wall-proximity and tactile stimuli through the concentric rings surrounding one of the tactors embedded in the belt. The increase in the radius of these rings as the wall is approached (from A to C in Figure 1) represents an increase in tactile stimulus magnitude and presumably saliency. Our tactile interface only allows for a joint scaling of frequency and amplitude through voltage. We interpret the joint change in those components as a change in stimulus magnitude or intensity. The perceived stimulus magnitude has been found to scale logarithmically with physical stimulus magnitude for various senses [12]. Expressed in a power law relation, exponent values can differ between senses and stimulus sites [54]. Reference (Stephens's [54]) exponents for vibrations in the sub 240 Hz range on the body have been found to range from 0.75 to 0.97 [38] which approximates linear scaling. Because initial tests with exponents 0.75, 0.83 and 1.0 revealed the two smaller exponents to feel rather irregular in scaling we decided to apply an exponent of 1.0. Note that the same interface and intensity scaling was used by Krüger et al. [32] to encode a temporal measure.

Outside the virtual fence the saliency stays at a maximum (see e.g. Figure 5). Through this encoding, the stimuli can be understood to act as an error signal with respect to a user's location. This error increases when leaving a safe area and maintains a constant high level outside such an area. An alternative to using a constant saliency outside the safety fence is to add an additional proximity-contingent scaling with reversed directionality. The advantage of this option is that it facilitates entering of a safe area by reducing stimulus saliency as the area is approached and thus providing feedback about whether a person's action was correct, i.e. saliency-reducing (see Figure 6). A separation of the redundant error-encoding features (i.e. visible segment size and color) between inside and outside cases may support a user in quickly recognizing the current state correctly.

4.3.2 Direction. Another property aimed towards facilitating location error corrections is a directional component in the stimuli. Both the visual and the tactile interface allow for localized stimulus generation by (a) altering a visual feature at a particular location in the visual case and (b) activating one tactor or a subset of tactors from a set of tactors distributed across the belt. Building further on the location error metaphor we thereby chose to provide stimuli that appear to originate from the direction of danger/error. In the visual case this means to apply the previously introduced saliency scaling to the virtual fence elements that are at the risk of being violated (user inside the zone near the fence) or have already been violated (user outside the zone). However, leaving the safe zone leads to a sudden reversal of visual stimulus directionality from the perspective of the user (see Figure 5a) which carries the risk of complicating re-entry because the previous appearance metaphor no longer holds outside the fence. To avoid such a directionality reversal we propose to visualize this form of location error by letting it visibly disturb the structural integrity of the virtual fence at the closest location in one of the forms illustrated by Figures 5b

---

[2] In Figure 4 the inner fence is a downscaled version of the outer fence and also shares its center and orientation, resulting in equal stimulus scaling for all directions. These conditions are however not fixed and thus by changing shape, position or orientation of the inner fence relative to the outer fence, stimulus scaling ranges can be made to vastly differ across directions when necessitated by corresponding asymmetric safety constraints.

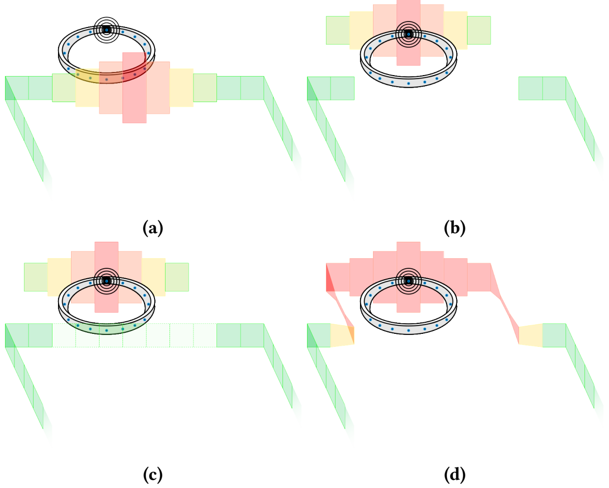

(a) (b) (c) (d)

Figure 5: Variants for a visual representation of being located outside a safe zone: 5a: Virtual fence appearance matches that of being located exactly at the zone boundary. Mismatch between tactile and visual stimuli. 5b: Violated fence portion sticks to the user and opens up an "entrance" to the safe zone. No multimodal mismatch. 5c: A transparent replacement for the broken fence is added to indicate the entry and need for closure. 5d: Connecting segments that bend with increasing distance to the safe zone are added to signal affiliation and pull from elastic elements. The uniform coloring depicts a more logical color progression as all parts of the person and belt are located outside the safe zone.

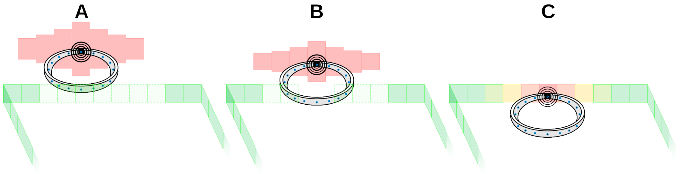


Figure 6: Feature combination for an extended encoding range and higher situation awareness.

to 5d. Thereby the part of the fence closest to the user is removed from its original location and attached to the user at the side facing away from the remaining fence. Additional supporting features that can facilitate the understanding of a need for entering the area are e.g. an introduction of stretching connection elements with the elasticity producing additional "force" (Figure 5d) or a display of transparent placeholder elements (Figure 5c).

In the case of the tactile stimuli the orientation information of the belt is used in order to determine the current orientation of the user relative to the virtual fence and thereby, in combination with location information, the (set of) tactors that should be activated. Tactor activation is updated whenever the location or the orientation of the user changes (see Figure 1D). Inside the virtual fence, these tactors would be those located the closest to the (outer) fence. Outside, the tactors located the furtherst away would be chosen in order to preserve the error-origin-direction or pushing metaphor.

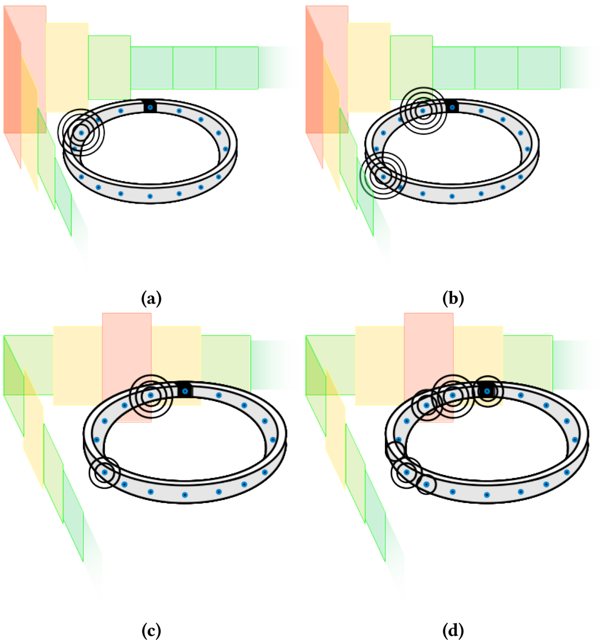

(a) (b) (c) (d)

Figure 7: Encoding options for multiple location errors. 7a: Average location and direction. 7b: Hybrid - Visual average, tactile max. of local cluster.7c: True visual correspondence, tactile simplification. 7d: True visual and tactile correspondence.

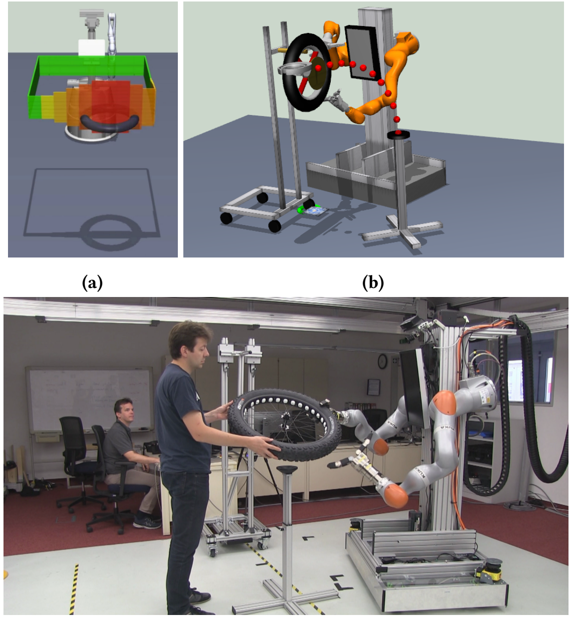
4.3.3 Encoding multiple location errors. A variety of options for selecting the most suitable set of zone elements and tactors exist and the optimal choice depends on the respective application scenario and zone geometry. Figure 8 illustrates a selection of different options for the encoding of tactile and visual stimuli. In the first example (Figure 7a), a constraint is applied which ensures that only one neighborhood of the virtual fence can be subject to variation at a time and that only one direction may be indicated by tactile stimuli at a time. In this example, fulfillment of the constraint is realized by outputting signals derived from measures of central tendency. While the actual distance scenario corresponds to that displayed in Figure 7c, summarizing the vectors for the error directions results in a new direction in between, which corresponds to a corner. While this approach yields incomplete information it can be advantageous for zones with more complex shapes. Another motivation for this constraint is to reduce stimulus complexity to a minimum and avoid possible confusion created by the simultaneous signaling of multiple errors. Figure 7b shows a variant which follows the same averaging approach for visual feedback but for tactile stimuli utilizes the original underlying data. Such a hybrid approach could help disambiguation in corner cases while maintaining advantageous properties of both encodings but also carries the risk of disrupting contingencies in multimodal input. Figures 7c and 7d are two examples for a commitment to the actual data for both modalities. They differ in that for the system of Figure 7c only the closest element of a straight wall is communicated whereas for Figure 7d each tactor with a sub-threshold distance to a wall-element is activated. The limited variant can be particularly advantageous for box-like fence configurations where no additional information can be gained from the activity of neighboring actuators and may better fit human sensory limitations for the simultaneous processing of multiple tactile stimuli (see [18–20]).

4.3.4 Complementarity and Redundancy. The visual and tactile stimuli described here provide partially complementary and partially redundant information. Complementary features are included to exploit modality specific properties. Specifically we target visual 3D perception capabilities by rendering zone boundaries at their corresponding absolute physical location (from the perspective of the observer). As vision depends on an active sampling through eye movements and focus adjustment, permanent rendering of these boundaries has only little potential for distraction. Tactile perception around the core of the body on the other hand can be considered to be more ego-centric, passive, and attention capturing which is why it may be of very little practical use for permanent rendering of 3D information but very suitable for event-based (low resolution 2D) alerting. As tactile stimuli are provided relative to a user's location and orientation, additionally an intuitive mapping to one's state and actions should be facilitated. Because this enables active environment exploration through movement, these stimuli may accordingly also be referred to as *haptic stimuli* [60]. The redundancy has the purpose to provide visual grounding in order to facilitate tactile stimulus understanding and take advantage of multisensory facilitation effects, thus potentially increasing e.g. information processing speed and bandwidth. Further it allows for meaningful independent employment in cases where no tactile interface is utilized. Visual rendering of zone boundaries and local violations at the respective physical location or a representation of it additionally allows external observers to more easily monitor and understand potential safety-hazards.

Note that the benefit of the visual scene augmentation varies between the users, relative to whom safety is measured (i.e. who should be within the area surrounded by the fence) and external users with an observer-role: Both the limited field of display of the AR glasses and the limited field of view of the human eyes make it unlikely for a user inside the virtual fence to see the whole fence at a time. With increasing distance to the fence, this ability becomes more feasible. An external user thus has the advantage of being able to see a much larger portion or even the entirety of a virtual fence at a glance whereas sequential scanning of directions would be needed from the inside. A user inside the zone can however still use the visualization for perceptual grounding in cases of uncertainty in the interpretation of tactile stimuli or for estimating safe motion ranges without requiring exploration through movement.

5 INTERACTION SCENARIOS

We have employed the presented system in two different environments and prototypical scenarios. The emphasis of the first scenario is to illustrate and practice the safety concept with a mobile robot. In the second scenario the concept is applied in a physical human-robot collaboration task. The two environments differ with respect to the sensors that are used for obtaining location information. In the first environment a setup of multiple synchronized *Microsoft Kinect* sensors is used to assign unique identifiers (id) based on appearance and simultaneously track multiple people based on these ids (see [11, 17]). By utilizing a range of coordinates for calibration, this also allows us to associate a person wearing the belt with a tracked id and thus location information. In the second environment a marker tracking solution (see section 4.2) was employed instead. This method allows for higher spatial and temporal accuracy at the expense having to attach visible markers to the user instead of using face features as a marker. Besides differences in tracking speed and accuracy, our system works identically with both setups.


(a) (b)

(c)

Figure 8: Interaction scenarios. 8a: Practice scenario. 8b and 8c: Physical human-robot cooperation scenario. 8b: Illustration of the task. 8c: Image from the setup.

5.1 Practice scenario

Figure 8a shows a visualization of the the system in the practice scenario. Here the boundaries of the safe area are dynamically defined relative to the location of a mobile robot. The user is asked to enter the safe location, explore the boundaries and update his or her location correctly in case the robot moves.

5.2 Physical human-robot cooperation

In this scenario, the human has to perform a physical interaction task cooperatively with a robot. Figure 8b shows an illustration of the scenario. A motorcycle wheel hangs vertically at a wheel support. Human and robot have to place it horizontally on a pole that is located about 1.5 m away from the support. The task requires re-grasping for both robot and human as well as moving sideways for some distance. The robot is equipped with an omni-directional mobile base that allows it to travel this distance. This mobility introduces some safety issues, since robot and human could get into close proximity. Further, the coordination of changing grasp holds requires the human to pay attention to the robot's actions so that the attention is focussed on the task.

6 DISCUSSION

Here we described our approach for developing a multimodal AR-system to support safety-awareness in human-robot interaction tasks. This development was guided by multiple goals such as a reduction of visual load for monitoring tasks, quick and intuitive understanding, and situation transparency to external observers. Our primary approach for a reduction of visual load consists of introducing tactile stimuli that convey information about how safe the current location of a person is through stimulus magnitude and further encode directions towards dangers in the stimulus site. The continuous coupling between stimulus magnitude and location danger thereby allows users to quickly explore contingencies and understand the directionality provided by the stimulus site. In addition to the tactile stimuli, we introduce an adaptive visual representation of safety boundaries and momentary location safety that provides partially complementary and partially redundant information. Combined, this lets the system incorporate complementary advantages from both modalities such as urgency-dependent attention capturing, self-determined information sampling, intuitive referencing, and external transparency.

We experienced the system over multiple weeks and informally tested it with informed and naive users. Even so, a formal user study will be necessary to evaluate the effectiveness of the system and especially our assumptions about the benefit of the design choices illustrated here. Such an evaluation may further inform the development and feasibility of possible alterations.

For instance, although first feedback suggests that the continuous and linearly-scaled tactile and visual stimulation is well suited for effective location-safety communication, other encodings might entail advantages for specific applications.

Other possible alterations concern zone- and stimulus complexity: We constrained the first implementation to rectangular shapes with universal distance-magnitude relationships to simplify visuo-tactile 2D coordinate correspondence for a subset of stimulus encodings (see Figure 8). Nevertheless, the introduced concept is compatible with any convex zone shape and extensions to support arbitrary shapes are easily conceivable. Success with such increased complexity may also be a function of exposure. Therefore one future task will be to investigate how much training is required to achieve the transition from a conceptual understanding to an incorporated and effortless augmented safety awareness.

ACKNOWLEDGMENTS

We would like to thank Simon Manschitz, Manuel Mühlig, Tamas Bates, Stefan Fuchs, Heiko Wersing, Christian Goerick and Andreas Richter for their valuable feedback and help during the development of the system.

REFERENCES

[1] Matthias Berning, Florian Braun, Till Riedel, and Michael Beigl. 2015. ProximityHat: a head-worn system for subtle sensory augmentation with tactile stimulation. In *2015 ACM International Symposium on Wearable Computers*. ACM, 31–38.

[2] Ira H. Bernstein, Mark H. Clark, and Barry A. Edelstein. 1969. Effects of an auditory signal on visual reaction time. *Journal of experimental psychology* 80, 3p1 (1969), 567.

[3] Matthew Botvinick, Jonathan Cohen, et al. 1998. Rubber hands ' feel' touch that eyes see. *Nature* 391, 6669 (1998), 756–756.

[4] Claudio Brozzoli, Tamar R. Makin, Lucilla Cardinali, Nicholas P. Holmes, and Alessandro Farne. 2011. Peripersonal space: a multisensory interface for body-object interactions. In *The neural bases of multisensory processes*, M. M. Murray and M. T. Wallace (Eds.). Taylor & Francis, London, 449–466.

[5] Sylvain Cardin, Daniel Thalmann, and Frédéric Vexo. 2007. A wearable system for mobility improvement of visually impaired people. *The Visual Computer* 23, 2 (2007), 109–118. https://doi.org/10.1007/s00371-006-0032-4

[6] Lucilla Cardinali, Claudio Brozzoli, and Alessandro Farnè. 2009. Peripersonal Space and Body Schema: Two Labels for the Same Concept? *Brain Topography* 21, 3 (2009), 252–260. https://doi.org/10.1007/s10548-009-0092-7

[7] Alvaro Cassinelli, Carson Reynolds, and Masatoshi Ishikawa. 2006. Augmenting spatial awareness with Haptic Radar. In *2006 10th IEEE International Symposium on Wearable Computers*. 61–64. https://doi.org/10.1109/ISWC.2006.286344

[8] Valve Corporation. [n.d.]. *Chaperone*. https://xinreality.com/wiki/Chaperone Accessed: 2020-01-31.

[9] N. Dahlbäck, A. Jönsson, and L. Ahrenberg. 1993. Wizard of Oz Studies — Why and How. *Know.-Based Syst.* 6, 4 (Dec. 1993), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

[10] Adele Diederich and Hans Colonius. 2004. Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics* 66, 8 (2004), 1388–1404. https://doi.org/10.3758/BF03195006

[11] Nils Einecke, Stefan Fuchs, Bram Bolder, and Manuel Mühlig. 2019. Living Lab: A 24/7 Human-Machine-Interaction Space in an Office Environment. In *UbiComp/ISWC '19 Adjunct*. ACM.

[12] Gustav Theodor Fechner. 1860. *Elemente der Psychophysik*. Vol. 2. Breitkopf u. Härtel.

[13] Gregory M. Fitch, Raymond J. Kiefer, Jonathan M. Hankey, and Brian M. Kleiner. 2007. Toward Developing an Approach for Alerting Drivers to the Direction of a Crash Threat. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49, 4 (2007), 710–720. https://doi.org/10.1518/001872007X215782

[14] Eric Fixmer and Sarah Hawkins. 1998. The influence of quality of information on the McGurk effect. In *AVSP'98 International Conference on Auditory-Visual Speech Processing*.

[15] L. Fogassi, V. Raos, G. Franchi, V. Gallese, G. Luppino, and M. Matelli. 1999. Visual responses in the dorsal premotor area F2 of the macaque monkey. *Experimental Brain Research* 128, 1 (01 Sep 1999), 194–199. https://doi.org/10.1007/s002210050835

[16] Marquart Franz, Andreas Zeidler, Marcos dos Santos Rocha, and Cornel Klein. 2008. Vibro-Tactile Space-Awareness. *on ubiquitous Computing* (2008), 117.

[17] Stefan Fuchs, Nils Einecke, and Fabian Eisele. 2019. SmartLobby: using a 24/7 remote head-eye-tracking for content personalization. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 53–56.

[18] Alberto Gallace, Hong Z Tan, and Charles Spence. 2006. Numerosity judgments for tactile stimuli distributed over the body surface. *Perception* 35, 2 (2006), 247–266.

[19] Alberto Gallace, Hong Z Tan, and Charles Spence. 2008. Can tactile stimuli be subitised? An unresolved controversy within the literature on numerosity judgments. *Perception* 37, 5 (2008), 782–800.

[20] Alberto Gallace, Sophia Zeeden, Brigitte Röder, and Charles Spence. 2010. Lost in the move? Secondary task performance impairs tactile change detection on the body. *Consciousness and Cognition* 19, 1 (2010), 215 – 229. https://doi.org/10.1016/j.concog.2009.07.003

[21] Michael Gienger, Dirk Ruiken, Tamas Bates, Mohamed Regaieg, Michael Meißner, Jens Kober, Philipp Seiwald, and Arne-Christoph Hildebrandt. 2018. Human-Robot Cooperative Object Manipulation with Contact Changes. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 1354–1360.

[22] Michael S.A. Graziano and Dylan F. Cooke. 2006. Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia* 44, 6 (2006), 845 – 859. https://doi.org/10.1016/j.neuropsychologia.2005.09.009

[23] Michael SA Graziano, Charles G Gross, Charlotte SR Taylor, and Tirin Moore. 2004. A system of multimodal areas in the primate brain. *Crossmodal space and crossmodal attention* (2004), 51–67.

[24] Michael S. A. Graziano and Charles G. Gross. 1993. A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Experimental Brain Research* 97, 1 (01 Dec 1993), 96–109.

https://doi.org/10.1007/BF00228820

[25] Peter A. Hancock, Joseph E. Mercado, James Merlo, and Jan B.F. Van Erp. 2013. Improving target detection in visual search through the augmenting multi-sensory cues. *Ergonomics* 56, 5 (2013), 729–738. https://doi.org/10.1080/00140139.2013.771219 PMID: 23510197.

[26] Maurice Hershenson. 1962. Reaction time as a measure of intersensory facilitation. *Journal of experimental psychology* 63, 3 (1962), 289.

[27] Joy Hirsch and Christine A Curcio. 1989. The spatial resolution capacity of human foveal retina. *Vision research* 29, 9 (1989), 1095–1101.

[28] Cristy Ho, Nick Reed, and Charles Spence. 2006. Assessing the effectiveness of "intuitive" vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Analysis & Prevention* 38, 5 (2006), 988 – 996. https://doi.org/10.1016/j.aap.2006.04.002

[29] Cristy Ho, Hong Z. Tan, and Charles Spence. 2005. Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour* 8, 6 (2005), 397 – 412. https://doi.org/10.1016/j.trf.2005.05.002

[30] Hamideh Kerdegari, Yeongmi Kim, Tom Stafford, and Tony J Prescott. 2014. Centralizing bias and the vibrotactile funneling illusion on the forehead. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 55–62.

[31] Holle Kirchner and Simon J Thorpe. 2006. Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision research* 46, 11 (2006), 1762–1776.

[32] Matti Krüger, Heiko Wersing, and Christiane B. Wiebel-Herboth. 2018. Approach for Enhancing the Perception and Prediction of Traffic Dynamics with a Tactile Interface. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, ACM Press, New York, NY, USA, 164–169. https://doi.org/10.1145/3239092.3265961

[33] Matti Krüger, Christiane B. Wiebel-Herboth, and Heiko Wersing. 2020. The Lateral Line: Augmenting Spatiotemporal Perception with a Tactile Interface. In *AHs '20: Augmented Humans International Conference*. ACM, ACM Press, New York, New York, USA. https://doi.org/10.1145/3384657.3384775

[34] John D Lee, Joshua D Hoffman, and Elizabeth Hayes. 2004. Collision warning design to mitigate driver distraction. In *SIGCHI Conference on Human factors in Computing Systems*. ACM, 65–72.

[35] Vera Maljkovic and Ken Nakayama. 1994. Priming of pop-out: I. Role of features. *Memory & cognition* 22, 6 (1994), 657–672.

[36] Harry McGurk and John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264 (1976), 746–748.

[37] Fanxing Meng and Charles Spence. 2015. Tactile warning signals for in-vehicle systems. *Accident Analysis & Prevention* 75 (2015), 333 – 346. https://doi.org/10.1016/j.aap.2014.12.013

[38] Sebastian Merchel, M Ercan Altinsoy, and Anna Schwendicke. 2015. Tactile intensity perception compared to auditory loudness perception. In *2015 IEEE World Haptics Conference (WHC)*. IEEE, 356–361.

[39] John Morrell and Kamil Wasilewski. 2010. Design and evaluation of a vibrotactile seat to improve spatial awareness while driving. In *2010 IEEE Haptics Symposium*. 281–288. https://doi.org/10.1109/HAPTIC.2010.5444642

[40] Florian 'Floyd' Mueller, Pedro Lopes, Paul Strohmeier, Wendy Ju, Caitlyn Seim, Martin Weigel, Suranga Nanayakkara, Marianna Obrist, Zhuying Li, Joseph Delfa, Jun Nishida, Elizabeth M. Gerber, Dag Svanaes, Jonathan Grudin, Stefan Greuter, Kai Kunze, Thomas Erickson, Steven Greenspan, Masahiko Inami, Joe Marshall, Harald Reiterer, Katrin Wolf, Jochen Meyer, Thecla Schiphorst, Dakuo Wang, and Pattie Maes. 2020. Next Steps in Human-Computer Integration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM. https://doi.org/10.1145/3313831.3376242

[41] Saskia K. Nagel, Christine Carl, Tobias Kringe, Robert Märtin, and Peter König. 2005. Beyond sensory substitution—learning the sixth sense. *Journal of neural engineering* 2, 4 (2005), R13.

[42] Raymond S Nickerson. 1973. Intersensory facilitation of reaction time: energy summation or preparation enhancement? *Psychological review* 80, 6 (1973), 489.

[43] Aude Oliva, Antonio Torralba, Monica S Castelhano, and John M Henderson. 2003. Top-down control of visual attention in object detection. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, Vol. 1. IEEE, I–253.

[44] Stephen E Palmer. 1999. *Vision science: Photons to phenomenology*. MIT press.

[45] Michael I Posner. 1980. Orienting of attention. *Quarterly journal of experimental psychology* 32, 1 (1980), 3–25.

[46] M. S. Prewett, L. R. Elliott, A. G. Walvoord, and M. D. Coovert. 2012. A Meta-Analysis of Vibrotactile and Visual Information Displays for Improving Task Performance. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, 1 (Jan 2012), 123–132. https://doi.org/10.1109/TSMCC.2010.2103057

[47] Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (1998), 372.

[48] Andreas Riener and Alois Ferscha. 2008. Raising awareness about space via vibro-tactile notifications. In *European Conference on Smart Sensing and Context*. Springer, 235–245.

[49] Giacomo Rizzolatti, Luciano Fadiga, Leonardo Fogassi, and Vittorio Gallese. 1997. The Space Around Us. *Science* 277, 5323 (1997), 190–191. https://doi.org/10.1126/science.277.5323.190

[50] Nadine B. Sarter. 2000. The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: the case of cockpit automation. *The International Journal of Aviation Psychology* 10, 3 (2000), 231–245.

[51] Nadine B. Sarter. 2007. Multiple-resource theory as a basis for multimodal interface design: Success stories, qualifications, and research needs. *Attention: From theory to practice* (2007), 187–195.

[52] J.J. Scott and Robert Gray. 2008. A Comparison of Tactile, Visual, and Auditory Warnings for Rear-End Collision Prevention in Simulated Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50, 2 (2008), 264–275. https://doi.org/10.1518/001872008X250674

[53] Aaron E. Sklar and Nadine B. Sarter. 1999. Good Vibrations: Tactile Feedback in Support of Attention Allocation and Human-Automation Coordination in Event-Driven Domains. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 41, 4 (1999), 543–552. https://doi.org/10.1518/001872099779656716

[54] Stanley Smith Stevens. 2017. *Psychophysics: Introduction to its perceptual, neural and social prospects*. Routledge.

[55] Simon Thorpe, Denis Fize, and Catherine Marlot. 1996. Speed of processing in the human visual system. *nature* 381, 6582 (1996), 520.

[56] John Welhoff Todd. 1912. *Reaction to multiple stimuli*. Number 25. Science Press.

[57] Anne M Treisman and Garry Gelade. 1980. A feature-integration theory of attention. *Cognitive psychology* 12, 1 (1980), 97–136.

[58] Jan B. F. Van Erp, Hendrik A. H. C. Van Veen, Chris Jansen, and Trevor Dobbins. 2005. Waypoint Navigation with a Vibrotactile Waist Belt. *ACM Trans. Appl. Percept.* 2, 2 (April 2005), 106–117. https://doi.org/10.1145/1060581.1060585

[59] Basil Wahn, Jessika Schwandt, Matti Krüger, Daina Crafa, Vanessa Nunnendorf, and Peter König. 2016. Multisensory teamwork: using a tactile or an auditory display to exchange gaze information improves performance in joint visual search. *Ergonomics* 59, 6 (2016), 781–795. https://doi.org/10.1080/00140139.2015.1099742

[60] Ernst Heinrich Weber. 1851. *Die Lehre vom Tastsinne und Gemeingefühle auf Versuche gegründet*. Friedrich Vieweg und Sohn.

[61] Max Wertheimer. 1923. Untersuchungen zur Lehre von der Gestalt. *Psychological Research* 4, 1 (1923), 301–350.

[62] Alfred L Yarbus. 1967. Eye movements during perception of complex objects. In *Eye movements and vision*. Springer, 171–211.