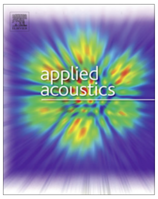
Applied Acoustics 129 (2018) 72–85

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust



Teleoperated path following and trajectory tracking of unmanned vehicles using spatial auditory guidance system

Antonio Vasilijevic *, Kristian Jambrosic, Zoran Vukic

University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia

ARTICLE INFO

Article history:
Received 23 January 2017
Received in revised form 8 June 2017
Accepted 2 July 2017

Keywords:
Human–machine interface
Spatial auditory dispaly
Guidance
Teleoperation
Unmanned vehicles
Supernormal transformations

ABSTRACT

The majority of Human–Machine Interfaces (HMIs) designed for the teleoperation of unmanned vehicles only present information visually. Frequent overloading of the operator's visual channel may result in accidents. To reduce the occurrence of accidents and improve overall operating performance, this paper proposes the extension of the HMI to the hearing modality. The proposed HMI takes the form of a spatial auditory display for complex navigation tasks such as path following (route guidance) or trajectory tracking. The interface incorporates two novel components. The first is the application of a modified "lookahead distance" guidance strategy, as a guidance law for generating a reference. The reference is presented to the operator as a spatial auditory image of the virtual target to be followed. Before being presented to the operator, the reference information is preconditioned through "supernormal" transformations to increase azimuth and distance resolution in regions of interest and thus to improve efficiency and overcome the comparative disadvantages of the hearing channel. A series of experiments show that the proposed guidance strategy provides a comprehensible and effective auditory reference. The enhanced supernormal auditory interface improves performance by reducing the tracking error of both path following and trajectory tracking compared to a non-enhanced interface without the supernormal transformation. Experiment participants report that the novel auditory interface is intuitive and easy to use, providing very good situational awareness. While the auditory interface investigated in this paper is shown to be effective for the teleoperation of unmanned vehicles, it may have other applications such as navigation for the visually impaired.

1. Introduction

Unmanned vehicles (UVs) are robotic vehicles that do not have a human operator aboard. Many of them are teleoperated, meaning that they are remotely operated, controlled, or guided by a person via some control or user interface. The control interface is connected wirelessly or through a cable to the UV. UVs can be found in many scenarios—on the ground, underwater, on the sea surface, and in the air. In this paper, we focus on human-operated marine UVs, although the research results are also applicable to other human-operated UVs. The marine UVs referred to in this paper are remotely operated underwater vehicles (ROVs) and unmanned surface vehicles (USVs).

Over the past decade, ROVs have been increasingly used by the oil & gas industry, the military and scientific communities, and, nowadays, for recreational activities. The market research company Douglas-Westwood quantified this increase in their World ROV Market Forecast 2013—2017, stating that there is currently "...total ROV operations expenditure of $9.7 billion, an increase of nearly 80% over the previous five-year period". Such widespread use of UVs is a consequence of their flexibility in performing different tasks. To accomplish such tasks, UVs are equipped with a variety of payloads, making the UV control room full of screens presenting everything from navigation data and multiple video streaming to data acquired by various sensors and sonars. This information is almost exclusively presented in visual form. The human operator, often performing multiple tasks simultaneously, handles an enormous quantity of information dispersed across different screens [33]. As a result, the operator's visual channel may become overloaded, which may prevent important information related to the particular task from being perceived. These unique operator-related issues often result in failed missions or even accidents, and represent a significant problem in UV operations. However, little research effort has been focused on overcoming the problem of an overloaded visual channel [25,9].

Similar issues can be found in many other situations when the visual channel is required for concurrent tasks, e.g. when a pilot

---

* Corresponding author.
E-mail address: antonio.vasilijevic@fer.hr (A. Vasilijevic).

http://dx.doi.org/10.1016/j.apacoust.2017.07.001

must navigate an aircraft while tracking targets, or when the visual channel is unavailable, as in the case of visually impaired, firefighters in a smoke-filled building or military personnel operating in darkness [61]. Humans prefer to receive spatial information visually because the spatial acuity of the visual channel is much better than that of the auditory channel [51]. However, in extreme situations when the visual channel is overloaded, the introduction of a display using an unused human sense may result in both the unloading of the visual channel and some additional advantages specific to that human sense [34,41]. Analysing alternative display modalities, the auditory channel is an obvious choice. Humans naturally use the auditory modality for the development and maintenance of situational awareness in their environment. They are able to determine the location of a sound source anywhere in the 360° space around them, monitor events at multiple locations simultaneously, and switch the focus of attention between virtual sound sources at will [3]. By exploiting these human abilities, it is reasonable to expect the operator's situational awareness to be improved using spatial 3-D sound.

To tackle the problem of an overloaded visual channel, and following the growing interest in research on spatial auditory displays [66] and multi-modal [21] or even multi-user interfaces [16], this paper presents a novel auditory navigation interface that can be employed as a standalone interface or as part of a more complex multi-modal human–machine interface (HMI). The aim of the interface is to support UV operations by improving the operator's perception of the operational environment and reducing the "eye-off-the-road" time. In Section 2, the motivation behind this research is revealed through a discussion of previous work and the present state of knowledge. The basic formulation of the guidance problem and its solution from the control perspective is given in Section 3. Section 4 introduces a novel supernormal auditory interface with enhanced angular and distance perception for high-level guidance. A detailed description of experiments and their corresponding results are presented and discussed in Section 5. Finally, a short conclusion summarises the work presented in this paper.

2. Motivation and related work

A teleoperation system is a human-in-the-loop system that consists of an operator, i.e. controller, a robot, and the HMI presenting the data required for teleoperation. Typically, only a visual interface is used as an HMI, with the exception of several scientific studies. Fig. 1 illustrates a graphical user interface (GUI) of the type commonly used for ROV teleoperation, presenting the vehicle and the actual and desired paths.

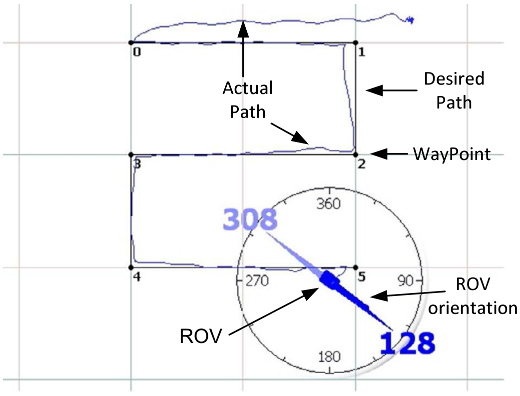


Fig. 1. Visual display presenting the mission layout: desired path (straight path segments connecting waypoints 0–5), ROV position, ROV orientation (in a compass form showing heading of 128 degrees), and actual, i.e. achieved, ROV path (curved line following desired path).

Hearing and vision are somewhat different, meaning that the path and trajectory cannot be presented as an auditory stimulus in the form seen on the visual interface. This raises the first question—how to convey audio data to the operator for efficient teleoperation of a UV? Is verbal or non-verbal communication better? If non-verbal, how should navigation data be incorporated into the sound?

The first audio displays were used by sonar operators to perform target detection and classification [55]. An audio interface can present verbal or non-verbal information. Verbal information is generated using recorded or synthesised speech, assembled so as to provide specific instructions or information to the operator such as "go left" or "10 metres to the target". Non-verbal audio is achieved by sonification, i.e. the transformation of data relations into perceived relations in an acoustic signal for the purpose of facilitating communication or interpretation [28]. Some non-verbal examples are Earcons, Auditory Icons, or spatial auditory cues. A taxonomy and definitions of sonification and auditory displays are given in [24].

Most existing auditory aids for the blind present directions and information via synthesised speech. They combine virtual acoustics with GPS and GIS for navigation [31] or lidar/sonar for obstacle avoidance [10,47]. Although verbal-based navigation systems have been used in some applications, they may place an unnecessary additional load on the operator. In [32], audio interfaces that aid the visually impaired to cross the road, supporting both sonification and speech, are compared. Most subjects preferred sonification over speech. Therefore, there is a need to understand how and when to apply non-verbal systems.

When we use hearing for orientation or guidance in our natural environment, we first detect and spatially locate the source of the sound or target, and then decide how to proceed. Examples of our innate responses, determined by real-life situations, are going toward a target, following a target, or running away from a target. The selected behaviour generates our internal guidance mechanism, composed of speed and steering laws. Combining them in different ways, different motion objectives can be achieved. For example, regarding the behaviour of going toward the target, our steering law would be to turn towards the target, i.e. set the azimuth angle relative to the target to zero, and move forward at a constant speed. When following the target, we would generate the same steering law, but adjust the speed to maintain a fixed, arbitrary distance from the leading target. When creating an auditory system for human-in-the-loop guidance, these facts should be considered. Consequently, the interface, i.e. auditory display, presenting the spatial virtual target (VT) to an operator who was previously advised to follow the target does not require any further instruction, training, or adaptation.

The Instrument Landing System (ILS) is an internationally accepted system for navigation of aircrafts upon the final approach for landing, which has been in use for nearly 7 decades. The ILS usually consists of a Localiser (it provides lateral guidance to the runway centreline), Glide Path (it provides the aircraft with a glide angle), and Markers (they mark a momentary distance from the aircraft to the runway's threshold). The signals of the localiser and glide slope consist of two intersected radiation patterns, modulated at 90 and 150 Hz. These signals are arranged left-right with respect to the runway centreline for localiser and on top of each other for glide slope. The onboard indicator of the ILS system provides vertical as well as horizontal guiding. The indicator track bar is deflected in the direction of the stronger signal. In the point where both signals 90 Hz and 150 Hz have the same intensity, the track bar is in the middle i.e. the plane is located exactly in the approach axis. The marker signals are made up by a Morse code