# TriControl – multimodal ATC interaction

The TriControl multimodal controller working position demonstrates a novel concept for natural human-machine interaction in Air Traffic Control (ATC) by integrating speech recognition, eye tracking and multi-touch sensing.

A multimodal Human-Machine Interface (HMI) is the part of a computer system capable of interpreting information from various sensors and communications channels. Multimodal HMIs emphasise the use of richer and more natural ways of interaction, such as speech, gestures and gaze. The primary goal of TriControl is to ensure that a human operator can input data more quickly and intuitively. This capability is a prerequisite for advancing human-machine systems to the point where computers and humans can truly act as a team.

## Advantages of multimodal interaction in ATC

An approach air traffic controller is responsible for merging several streams of traffic into a single sequence for the runway. Highly automated decision support tools like Arrival or Departure Managers have been developed to support human operators in this challenging task. These systems need to adapt to the air traffic controller's intentions to provide useful support. Therefore, these systems require knowledge about – and input from – their human operators. Multimodal HMIs represent a new class of user-machine interface, which is expected to offer faster, easier and more intuitive methods for data entry.

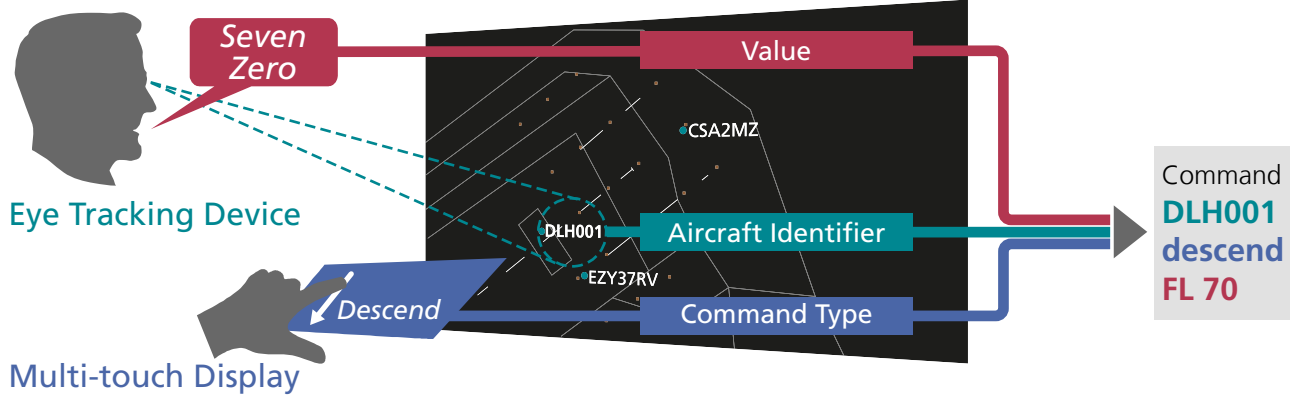These systems have the potential to enhance human-computer interaction in a number of ways by:

- supporting improved efficiency through faster task completion, especially when working with graphical information
- enhancing error avoidance
- accommodating differences between individual operators by permitting the flexible use of input modes and providing alternative methods for user interaction
- adapting to a wider range of users, tasks and situations.

Implementation of the multimodal human-machine interaction concept is still in its early stages in ATC. DLR gained international attention within the air traffic management community with its leading research and successful implementations in the field of eye-tracking [1], multi-touch [2], and speech recognition [3].

Motivated by the above mentioned advantages of multimodal interaction and promising research results on multimodal systems, DLR developed TriControl, a prototype of a multimodal controller working position for approach controllers, focusing on the integration of the three most promising interaction technologies: automatic speech recognition, sensing of multi-touch gestures, and eye gaze detection.

**Automatic Speech Recognition**

*Seven Zero*

Value

CSA2MZ

**Eye Tracking Device**

DLH001

Aircraft Identifier

EZY37RV

**Multi-touch Display**

*Descend*

Command Type

Command
**DLH001
descend
FL 70**

Schematic view of the communication modes and information obtained through them at the TriControl working position

## Interaction concept in TriControl

One of the main tasks of an approach air traffic controller is monitoring the radar situation display. Within TriControl it is assumed that the aircraft label the air traffic controller is looking at is the focus of attention. Eye gaze measurement is used to continuously calculate the position of the air traffic controller's eye gaze and correlate it with aircraft label positions on the display. In TriControl, eye gaze measurement is used to select aircraft labels.

Even though data link technology (CPDLC) is now in use, the main information exchanges between air traffic controllers and pilots are carried out using voice communication. Speech recognition algorithms are capable of detecting commands from the ATC vocabulary with acceptable reliability. TriControl specifically uses algorithms to detect spoken numbers.

In combination with gestures on a multi-touch display, the air traffic controller can give commands to the selected aircraft. Through a set of single- and dual-touch gestures, the air traffic controller selects the type of command – whether it concerns the altitude, speed or course of the aircraft. The direction of the gestures indicates whether the aircraft should, for example, speed up or slow down.

As the three interaction modes can be used simultaneously, the input of the air traffic controller's intention into the ATC system is faster. As a consequence, the controller is able to work more efficiently.

## Next steps

Other combinations of interaction modes can be tested and evaluated against each other. As ATC must satisfy stringent safety standards, the reliability and accuracy of these input modes needs to be very high. In order to accommodate individual differences and preferences, in a next step the user chooses the modes that he or she wants to use to interact with the system. Errors made by an air traffic controller can be avoided by using redundant data streams for input – for example cross-checking a spoken call sign against the visual attention of the air traffic controller.

The infrastructure available at the DLR Institute of Flight Guidance allows the setting-up of air traffic management workplaces with these novel interaction modes and the running of realistic high-quality simulations. Thus, initial results and empirical evidence about the usefulness of multimodal interaction for air traffic control can be gained in an early development phase.

## References

[1] Moehlenbrink & Papenfuss (2014): 'Eye-data metrics to characterize tower controllers' visual attention in a multiple remote tower exercise', ICRAT Istanbul.

[2] Uebbing-Rumke, Gürlük; Jauer, Hagemann & Udovic (2014): 'Usability evaluation of multi-touch displays for TMA controller working positions', 4th SESAR Innovation Days, Madrid.

[3] Gürlük, Helmke, Wies, Ehr, Kleinert, Mühlhausen, Muth & Ohneiser (2015): 'Assistant based speech recognition – another pair of eyes for the Arrival Manager', 34th DASC, Prague.