Speech-Augmented Cone-of-Vision for Exploratory Data Analysis

Riccardo Bovo, Daniele Giunchi, Ludwig Sidenmark, Joshua Newn, Hans Gellersen, Enrico Costanza, and Thomas Heinis

Imperial College London, University College London, Lancaster University

Published at CHI23

Definition

Mutual awareness of visual attention: The ability to identify collaborators' visual attention. Important for successful collaboration.

Visual attention cues: visual elements indicating where collaborators' visual attention is on the workspace. Ex: Eye-tracking cursor.

- Mono-directional: Only the collaborators see the cues.
- Bi-directional: Everyone, including the user, see the cues.

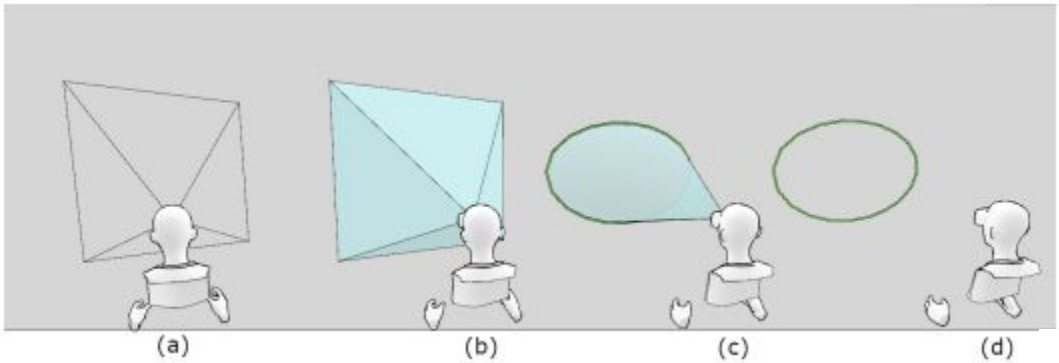
Mutual alignment: Orienting themselves along the general direction of collaborators.

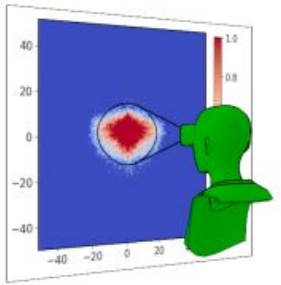
Contributions

(1) A new visual attention cue: Speech-augmented Cone-of-Vision
(2) A study comparing this cue to 2 baseline:
    (a) Classic Cone-of-Vision
    (b) Eye tracking cursor
(3) A dataset from the study

Speech-augmented Cone-of-Vision

First, what is Cone-of-Vision?

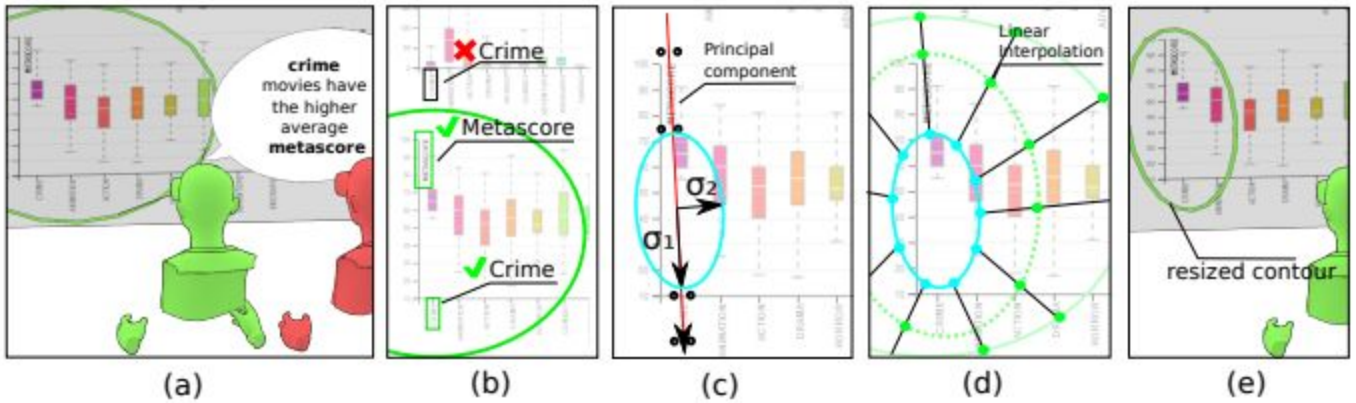




(e) 70% PBC to CoV

Source: Cone of Vision as a Behavioural Cue for VR Collaboration. Bovo et al. 2022

Speech-augmented Cone-of-Vision



The study

Condition: CoV + Speech (Speech-augmented Cone-of-Vision), CoV (classic Cone-of-Vision), Eye-gaze cursor

Task: Finding insights in a multi-visualisation environment (Exploratory data analysis). Two participants in two separate room => Remote collaboration

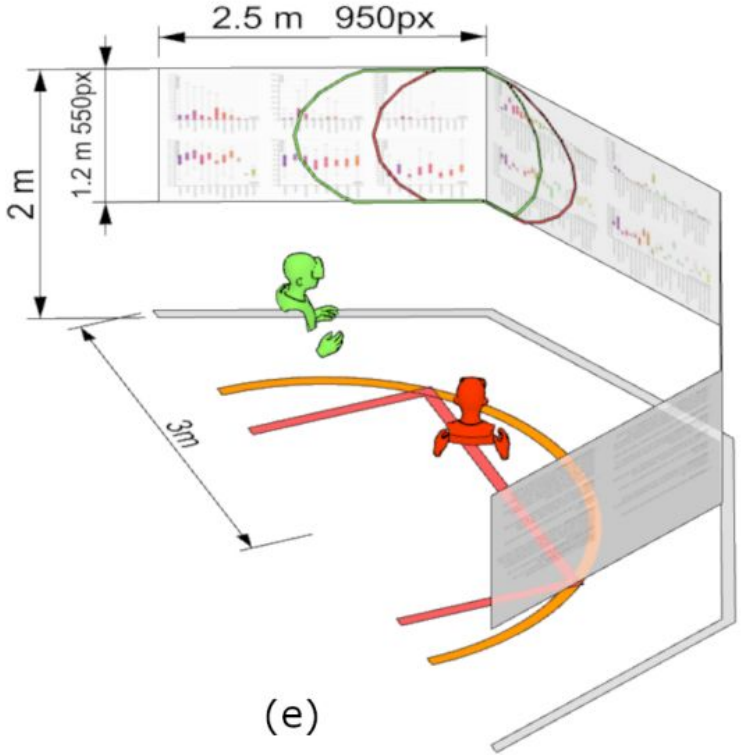
10 pairs of participants

Measures: Number of insights, Visual attention, non-standard questionnaires

Post-Analysis on speech

Note: All participants wore an eye-tracker.

A quick view of the environment



(e)

Overall results

No difference in insights

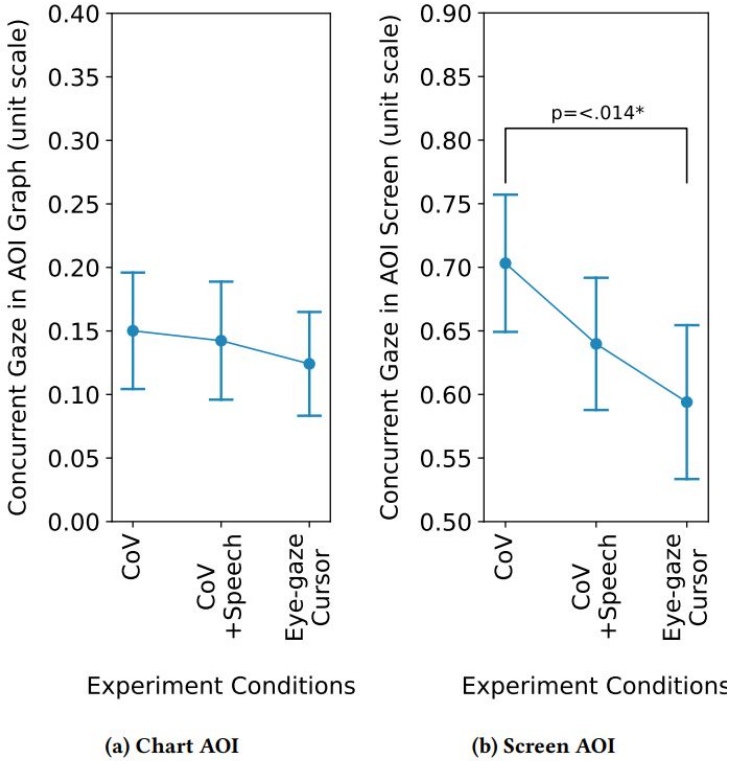
Small difference in VA

Some differences in questionnaire

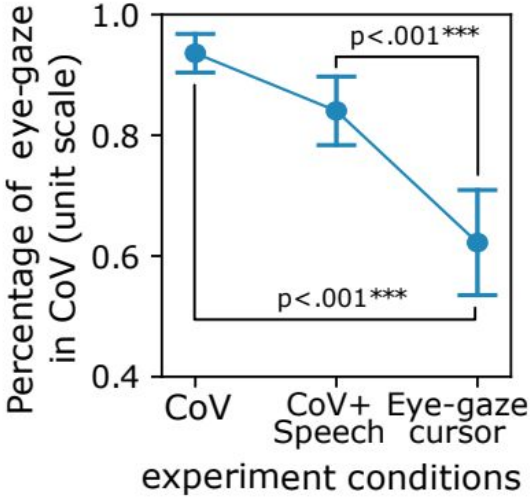
Interesting comments on the use of visual attention cues in general

Big Issue: Real-time speech recognition too laggy to be usable (=> post analysis on speech)

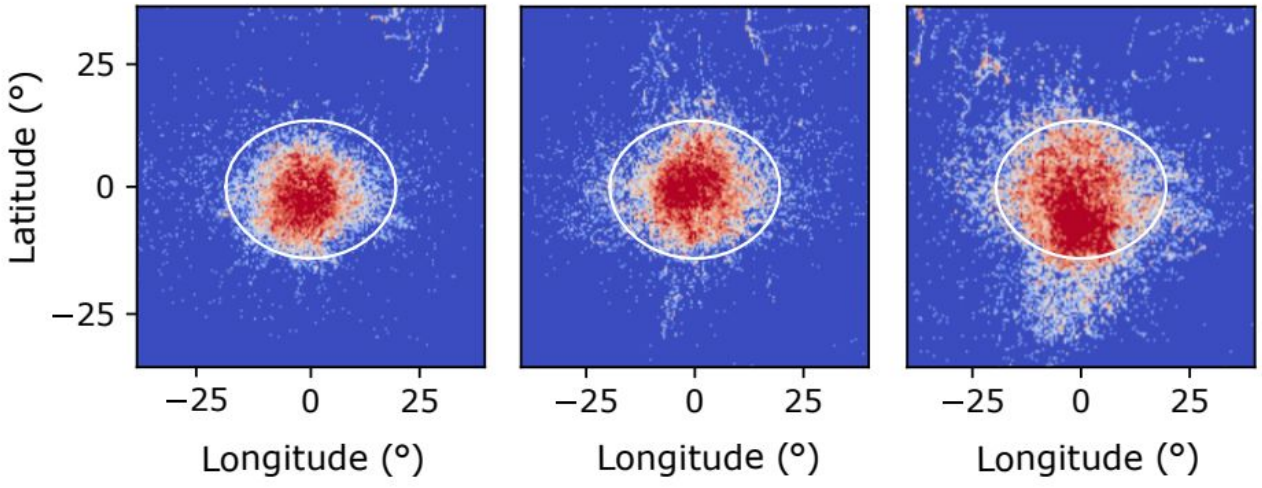
Visual Attention (concurrent)



Visual Attention (individual)



(a) Eye-gaze inside CoV



(b) Gaze in CoV

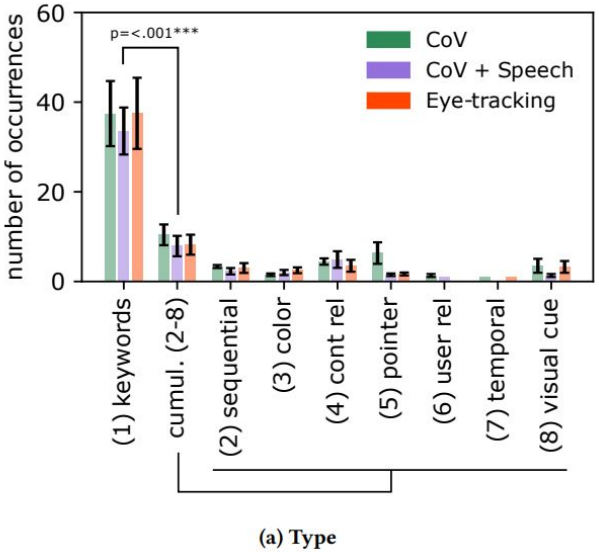
(c) Gaze in CoV+Speech

(d) Gaze in eye-cursor

Speech post analysis

Transcription and coding of the oral communication (with video recording)

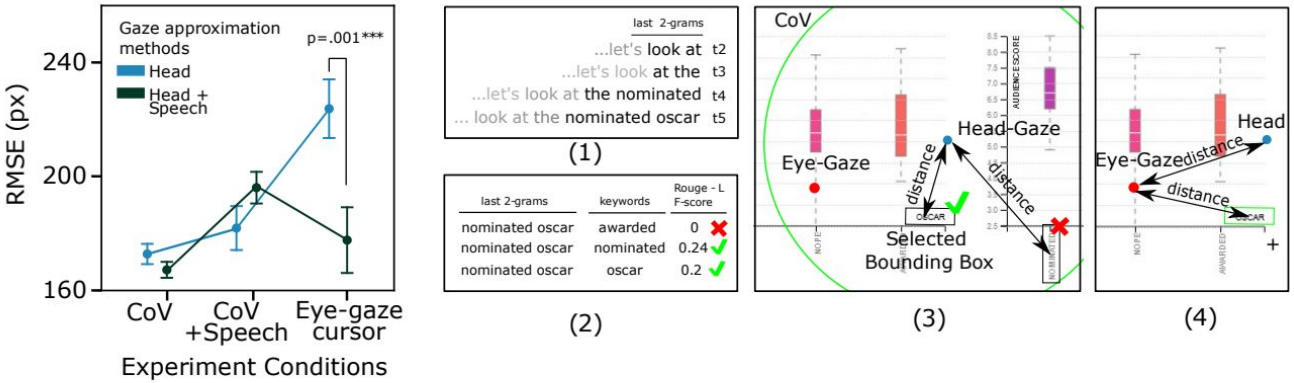
Identification of words used to collaborate (taxonomy adapted from previous one)



(a) Type

Speech post analysis

Computation a posteriori of the CoV+Speech and of the distance between this zone and the real position of the gaze of the participant. Comparison with the classic CoV.



(a) RMSE gaze approximation methods

(b) Head+Speech method

Interview with participants

Participants preferred eye gaze cursor when talking about something specific but CoV good for mutual alignment.

Bi-directional CoV => more focused, collaboration more coupled and less wandering.

Lots of use of verbal communication, to refine accuracy of cues or as fallback when they did not work.

CoV+Speech: Encouraging comments from participants, but laggy. They want to be able to control it directly (“go up”, “go to the corner of the graph”, etc.).

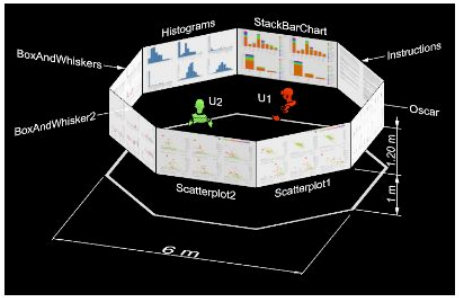
The dataset

README.md

Collaborative Speech Gaze and Head Behaviour Datatset

This dataset contains the verbal, head, and eye behaviour of ten pairs of participants performing a collaborative explorative data analysis task in VR; for more information about this, read the section Data. Furthermore, this dataset also contains the segmented and semantically annotated visual context of the explorative data analysis; for more information about this, read the section Visualizations. This dataset has been released with the paper Speech-Augmented Cone-of-Vision for Exploratory Data Analysis at CHI 2023.

```
@article{Bovo:2023,
    author  = "Riccardo Bovo, Daniele Giunchi, Ludwig Sidenmark, Hans Gellersen, Enrico Costanza, Thomas Hein
    year    = 2023,
    journal = "TBD",
    doi = "10.1145/3544548.3581283"  }
```



Link: https://github.com/Collaborative-Immersive-Visual-Toolkit/Speech-Gaze-Head-Datatset

Take-away message

(1) CoV in general act as a container of the visual attention
(2) Overall good feedback for bi-directional visual attention cues
(3) Keywords are mostly used as collaborative communication for EDA
(4) To augment CoV with speech could be a good alternative