

SPAZIALIZZAZIONE BINAURALE CON HEAD TRACKING PER L'ACCESSIBILITÀ DI INTERFACCE GRAFICHE

Giuseppe Bergamino, Michael Fioretti, Leonardo Gabrielli, Stefano Squartini

Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche
via Breccie Bianche 12, 60131, Ancona

Uno screen reader è un'applicazione che intercetta le informazioni testuali presenti su un desktop e le traduce in output vocale, consentendo a persone non vedenti o ipovedenti di interagire con interfacce grafiche altrimenti inaccessibili. Poiché gli elementi dell'interfaccia vengono annunciati in sequenza, l'utente deve memorizzare lunghe liste di controlli e ricostruirne mentalmente la disposizione; ciò accresce il carico cognitivo [1] e compromette l'usabilità e accessibilità di software complessi, specie negli ambienti di produzione musicale digitale [2]. Studi precedenti hanno affrontato questo problema introducendo stimoli acustici non verbali –icone sonore ed *earcons* [3]– che fungono da punti di riferimento, oppure impiegando screen readers sperimentali. Tali soluzioni non hanno fornito evidenze sulla riduzione del carico cognitivo e non si integrano con i lettori di schermo nativi dei sistemi operativi. Il presente lavoro propone di spazializzare in tempo reale l'output di VoiceOver, il lettore di schermo integrato in macOS, collocando il feedback verbale associato a ogni elemento grafico in un punto preciso di un panorama binaurale sferico, esplorabile tramite movimento attivo della testa

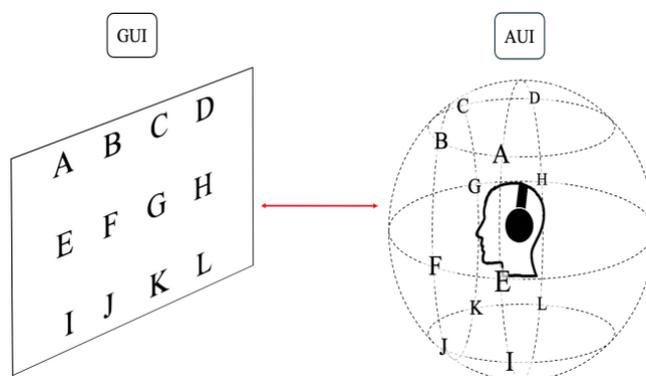


Fig. 1 – Rappresentazione dello schema di spazializzazione dell'interfaccia

Il sistema, realizzato nel framework JUCE, intercetta l'output vocale di VoiceOver e lo instrada in Reaper, dove una catena di plug-ins (IEM Suite) lo codifica in Ambisonics di settimo ordine per poi decodificarlo in binaurale per un ascolto in cuffia. Un sensore IMU MPU-9250 fissato alle cuffie rileva i movimenti della testa dell'utente; la scena ambisonica viene fatta ruotare in senso opposto, così che i messaggi di VoiceOver restino ancorati a posizioni fisse e facilmente memorizzabili. Il mapping acustico riflette la geometria dell'interfaccia: i controlli sono distribuiti fra $[-120^\circ; +120^\circ]$ in azimuth e $[-30^\circ; +60^\circ]$ in elevazione.

Allo studio hanno partecipato 32 volontari vedenti, bendati, suddivisi in due gruppi: il gruppo di controllo ha utilizzato VoiceOver nella sua versione standard, il gruppo sperimentale ha impiegato la variante spazializzata proposta. I partecipanti hanno seguito un protocollo suddiviso in 4 fasi: i) tutorial sull'utilizzo del sistema, ii) esplorazione dell'app di prova, iii) ricostruzione del layout su una griglia tangibile, iv) compilazione di un questionario su usabilità (SUS), carico di lavoro percepito (NASA-TLX) e con 2 domande aperte. Durante l'esperimento sono stati registrati numero di click, tempi di esplorazione e ricostruzione, ed errori nella ricostruzione calcolati come distanza di Manhattan.

Tabella 1: Media e deviazione standard dei parametri registrati durante il test

| | Numero Click | Tempo Esplorazione | Tempo Ricostruzione | Errori Ricostruzione |
|---------------------|--------------|--------------------|---------------------|----------------------|
| Gruppo controllo | 51 ± 15 | 241 ± 86 s | 210 ± 123 s | 26.6 ± 22.8 |
| Gruppo sperimentale | 59 ± 16 | 291 ± 108 s | 193 ± 69 s | 17.5 ± 12.7 |

Non sono emerse differenze significative in tempi e numero di click; tuttavia, l'errore medio di ricostruzione si è ridotto da 26,6 a 17,5 per il gruppo con l'audio spazializzato, con ricostruzioni più accurate e meno elementi mancanti. Sebbene i questionari indichino usabilità e carico di lavoro percepito comparabili, le 2 risposte aperte hanno rivelato strategie mnemoniche più strutturate nel gruppo sperimentale.

L'integrazione di uno screen reader con head tracking, catena Ambisonics di ordine elevato e decodifica binaurale ha migliorato l'accuratezza con cui i volontari dell'esperimento hanno ricostruito la disposizione dei controlli dell'interfaccia, senza impattare negativamente su carico cognitivo e usabilità. Le prossime fasi di ricerca coinvolgeranno persone non vedenti e ipovedenti e introdurranno migliorie rispetto alle criticità emerse durante questo studio pilota, per consolidare l'efficacia e usabilità del sistema.

Bibliografia

- [1] A. D. N. Edwards, The design of auditory interfaces for visually disabled users, SIGCHI Conference on Human Factors in Computing Systems, Washington D.C., USA, 1988
- [2] Payne W. C. et al., "How Blind and Visually Impaired Producers Leverage Music Technology," ASSETS '20, 2020
- [3] S. Brewster, The design of sonically-enhanced widgets, Interacting with Computers, Vol. 11, 1998.