

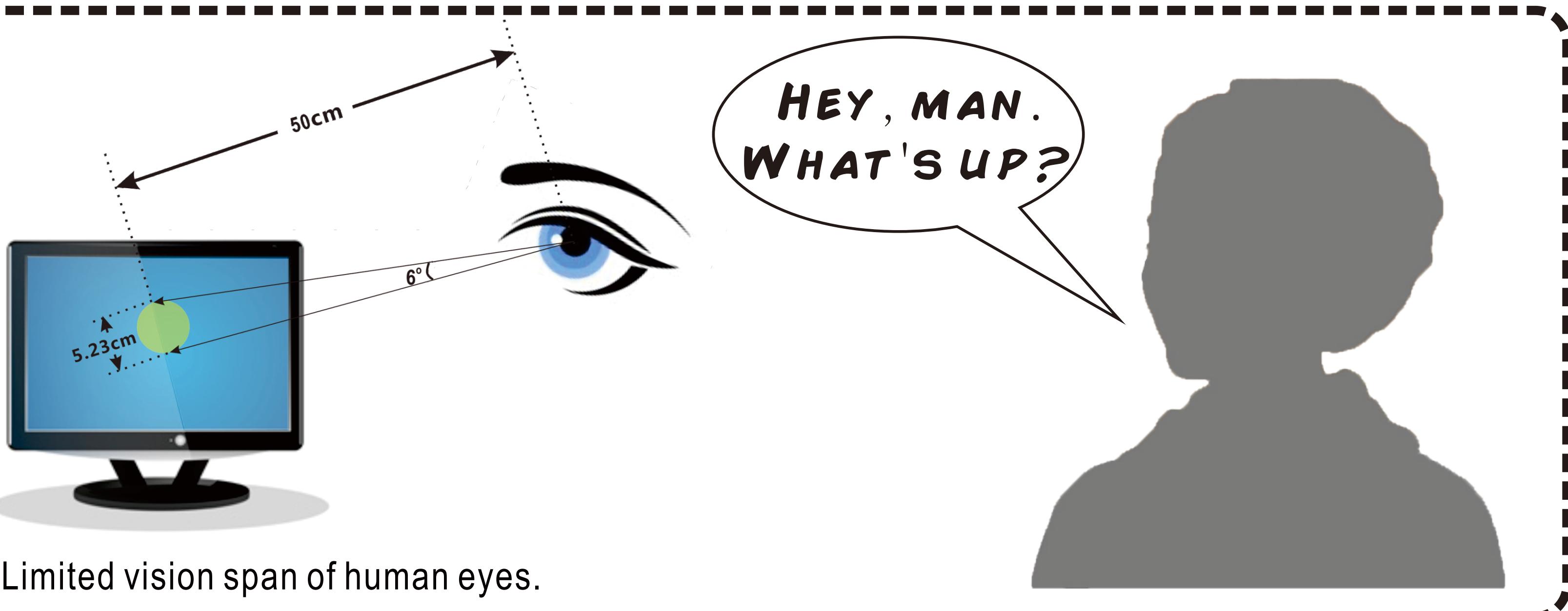


Enhancing Video Viewing Experience via Speaker-following Subtitles

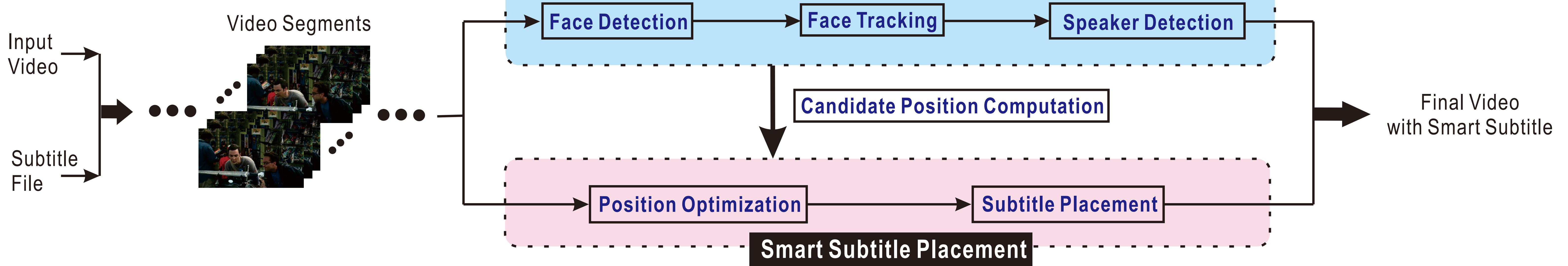
Yongtao Hu

Motivation

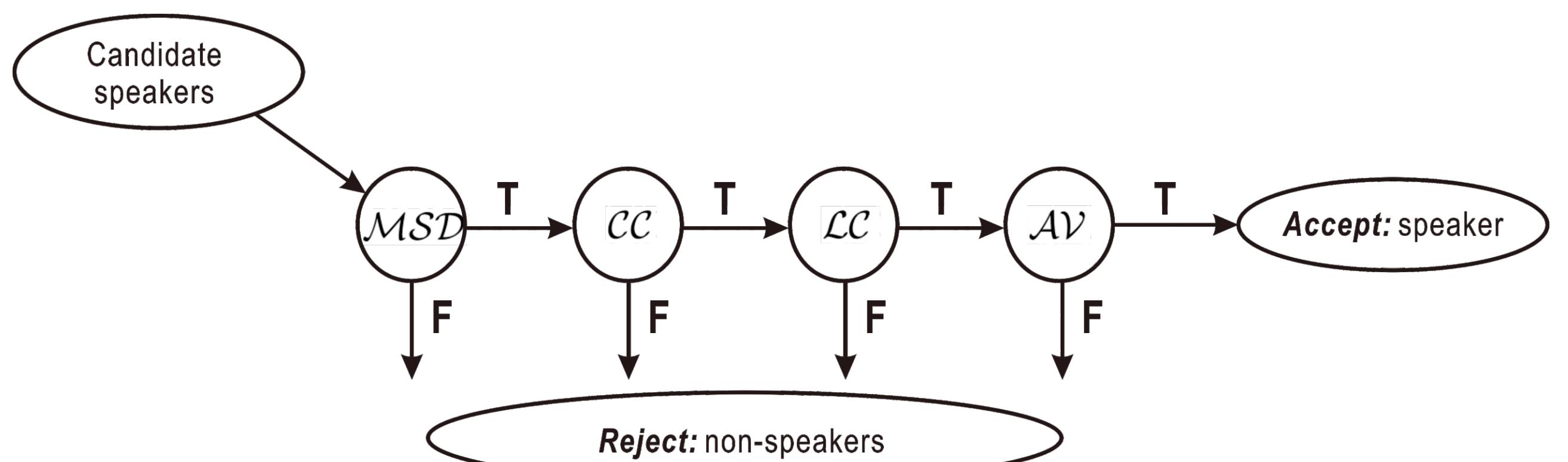
- Subtitles are essential
 - Subtitles usually appear at a fixed position
 - ▶ Miss the flow of visual contents
 - ▶ Especially when speaking fast
 - Limited vision span of human eyes



Framework



Speaker Detection



Smart Subtitle Placement

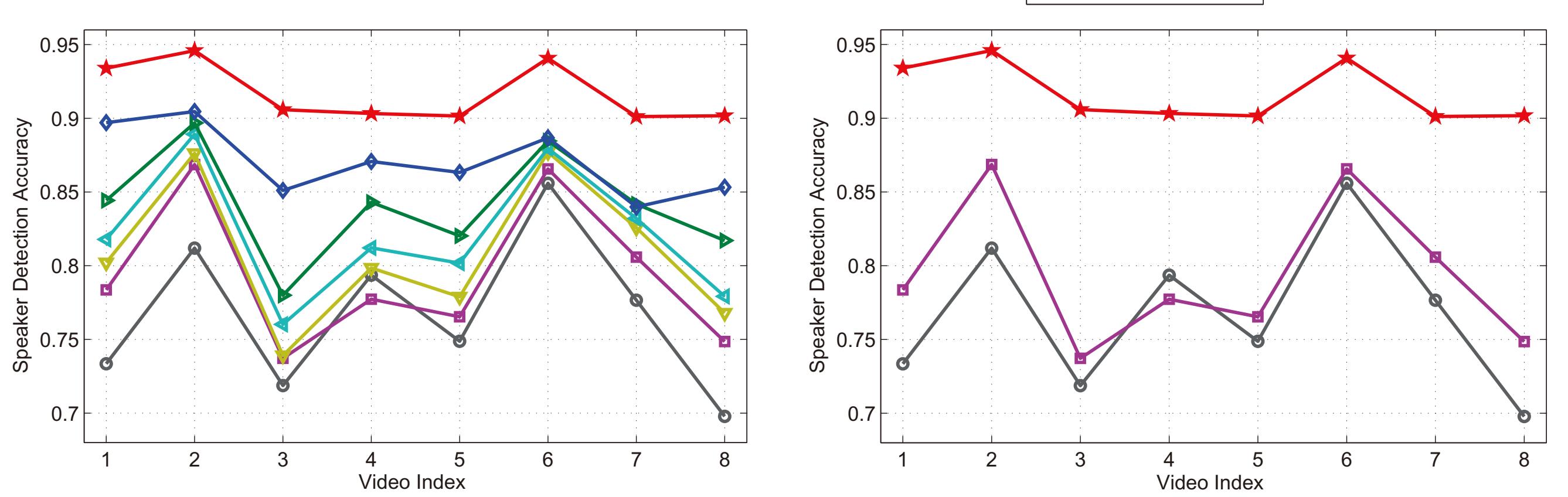
$$\min E = w_1 \cdot E_{local} + w_2 \cdot E_{global} + w_3 \cdot E_{layout}.$$

close to the speaker and far away from non-speakers

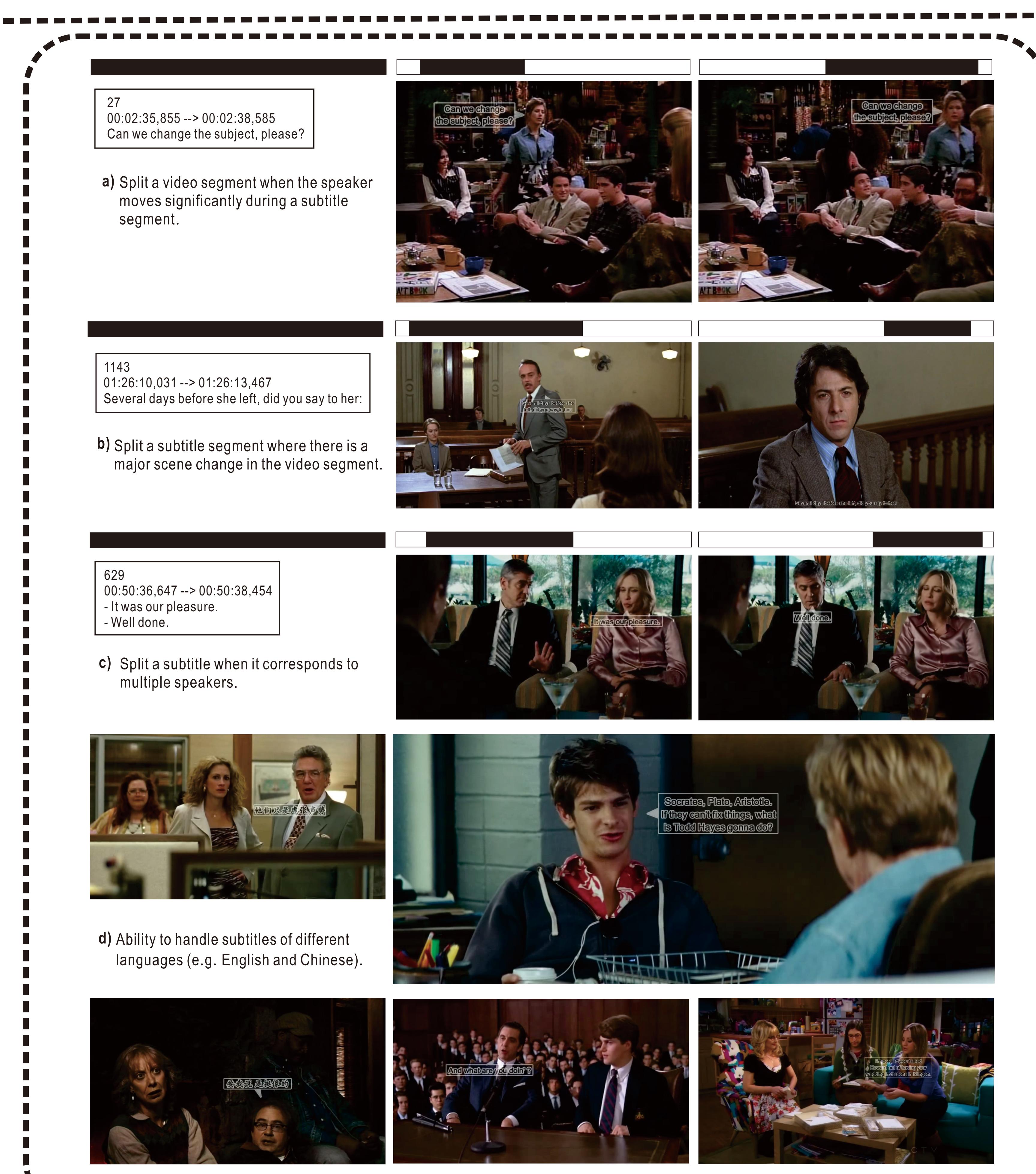
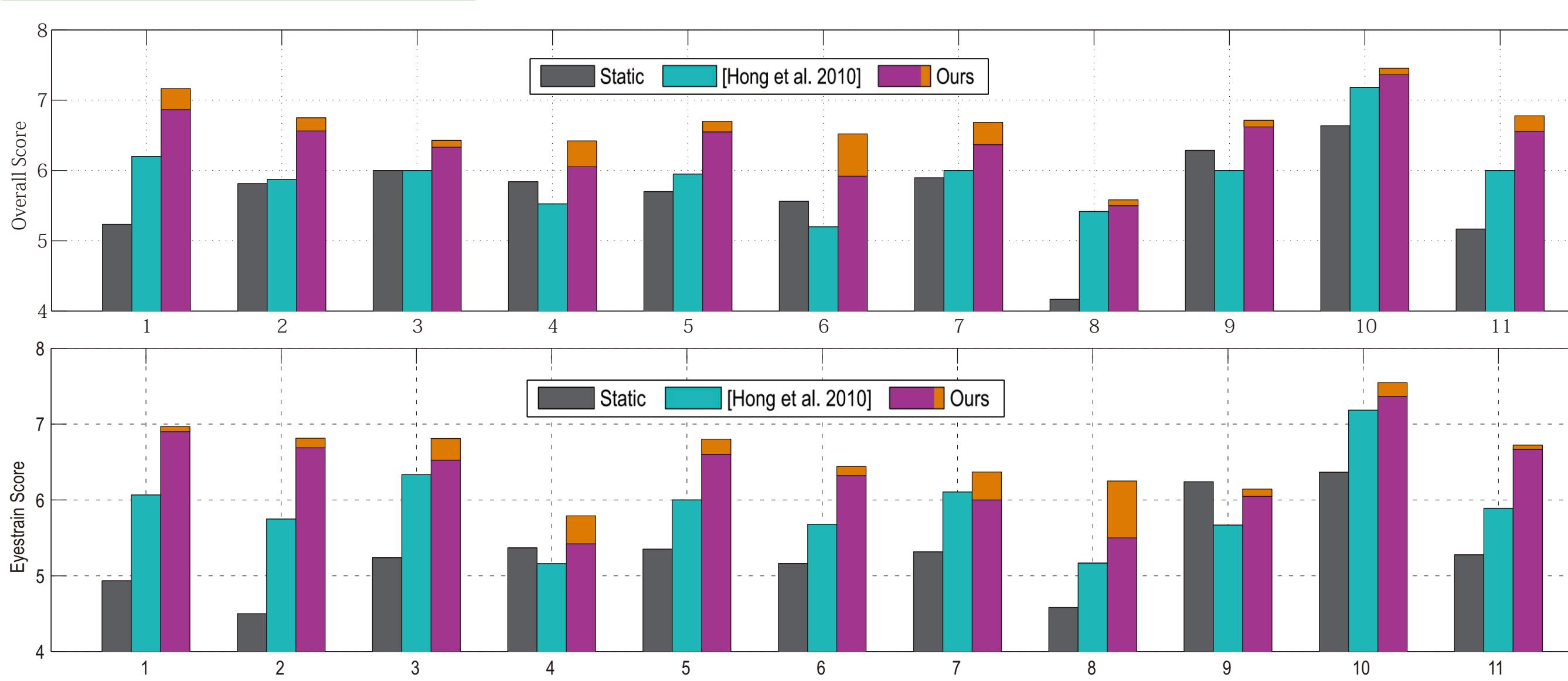
time consistence

away from boundary

Comparison



User Study



Reference

- [1] HU, Y.-T., KAUTZ, J., YU, Y.-Z., AND WANG, W.-P. 2014. 608 Speaker-following video subtitles. ACM Transactions on Multimedia Computing, Communications, and Applications (ACM TOMM), to appear.
 - [2] MONACI, G. 2011. Towards real-time audiovisual speaker localization.
 - [3] HONG, R., WANG, M., XU, M., YAN, S., AND CHUA, T.-S. 2010. Dynamic captioning: video accessibility enhancement for hearing impairment. In Proceedings of the international conference on Multimedia. ACM 421–430