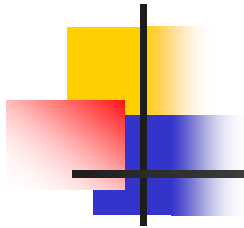




# SPEAKER INDEXING

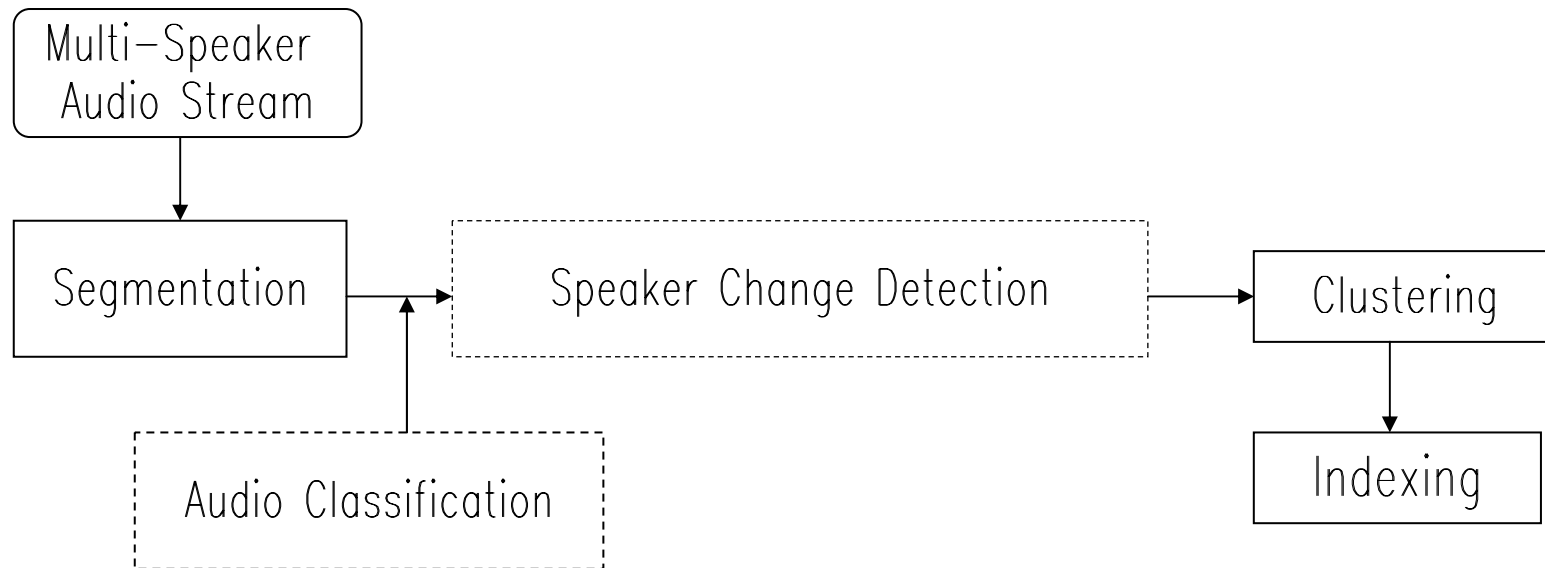
---

Review of ICSLP2002 Papers



# General Diagram

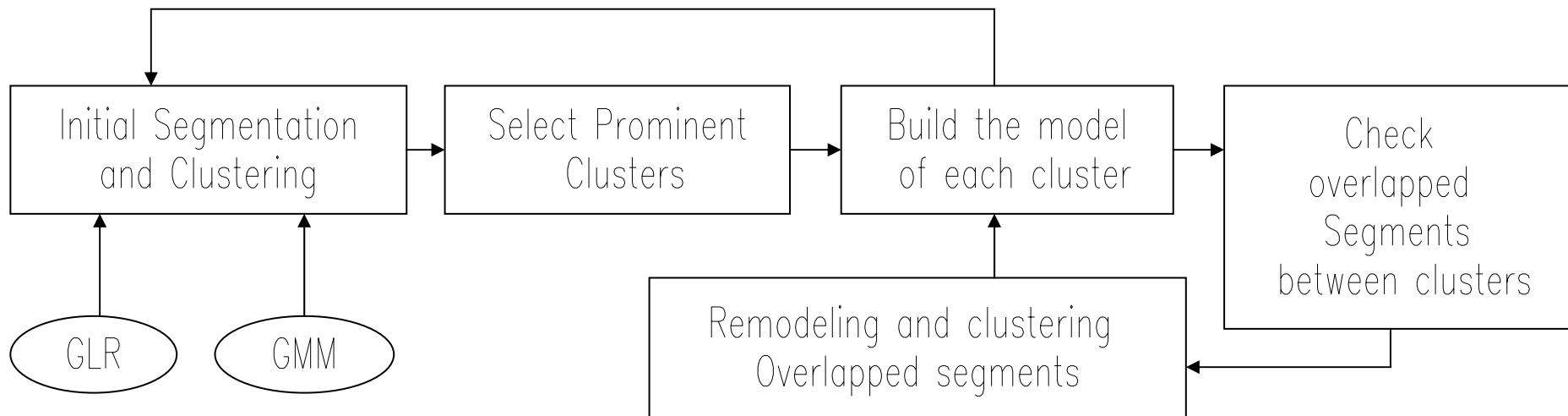
---



# Unsupervised Speaker Segmentation of Telephone Conversations (1)

- Main Point

: Iterative Segmentation and Clustering.



# Unsupervised Speaker Segmentation of Telephone Conversations (2)

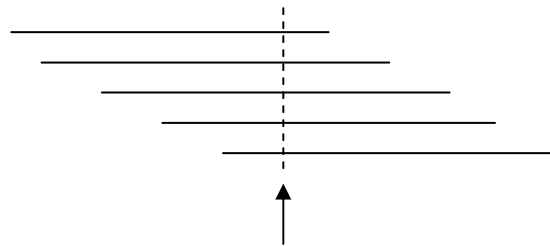
- Problem

1. Fixed length window : 1.6 seconds

→ hard to resolve speaker changes within a segment.

2. Window shift : 0.2 seconds

→ several windows contain the same speaker change. (hard to cluster)



# Unsupervised Speaker Segmentation of Telephone Conversations (3)

3.



This segment can be recognized as ‘speaker A’, but the model of ‘speaker A’ will be corrupted.



# An Effective Unsupervised Scheme for Multiple Speaker Change Detection (1)

---

- Main Point

1. Adaptive window length

- window length is grown until a speaker change is detected or length is over a maximum value.

- lower computational complexity.

2. Refinement step

- computed at a higher resolution rate.



## An Effective Unsupervised Scheme for Multiple Speaker Change Detection (2)

---

- Problem

1. Adaptive window length algorithm is very complicated.
2. Too many arguments should be determined through a training.
3. Is refinement step enough for overlapped segments? (.. see slide 5 )



# Conclusion

---

1. Try to delete silence and background noise before or within a segmentation step.
2. In segmentation, refinement step can be useful for overlapped segments.
3. Simplified adaptive window length algorithm can be useful.