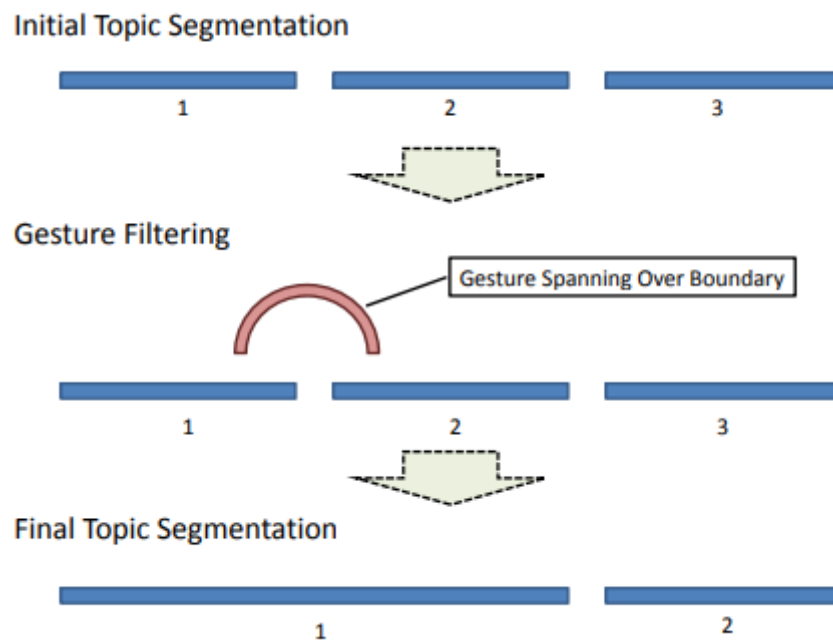


目前从 [ACL Anthology](#) 和 [谷歌学术](#) 没有找到关于 **多模态分词** 的顶会论文。如下两篇是关于多模态文本分割的论文，时间也比较久远：

## 1. [Multimodal dialogue segmentation with gesture post-processing](#)

- 接收时间：2014
- 任务：本文利用手势进行多模态对话分割。具体过程分为两阶段：首先使用已有的文本分割算法 [TextTiling](#) 和 [U00](#) 将对话进行主题分割，然后基于单个手势只能关联一个对话主题的假设，将多个分割的主题进行合并。如下图所示：



## 2. [Gestural Cohesion for Topic Segmentation](#)

- 接收年份：2008
- 任务：本文利用手势的连贯性对演讲进行主题细分。具体的过程首先是用 **Bayesian generative model** 对演讲稿的每条句子进行主题细分，然后结合视频中演讲者手势的连贯性和相似性将主题合并，当手势变化较大时，代表主题发生了改变。