audio ———

tempo estimation ———→

↑

score with
lyrics →

```
                           ┌─────────────────────────────┐
      ──────────────────▶  │  subsequence-DTWsearch      │  ──
                           └─────────────────────────────┘
                                          ▲
                                          │

   ▸ tempo  ──▶  query  ──────────────────▶
                            ▲
                            │
   ┌──────────────────────┐
   │  query construction  │
   └──────────────────────┘
```

n-best
candidate segments

Hierarchical HHM
decoding

ranked
result segments