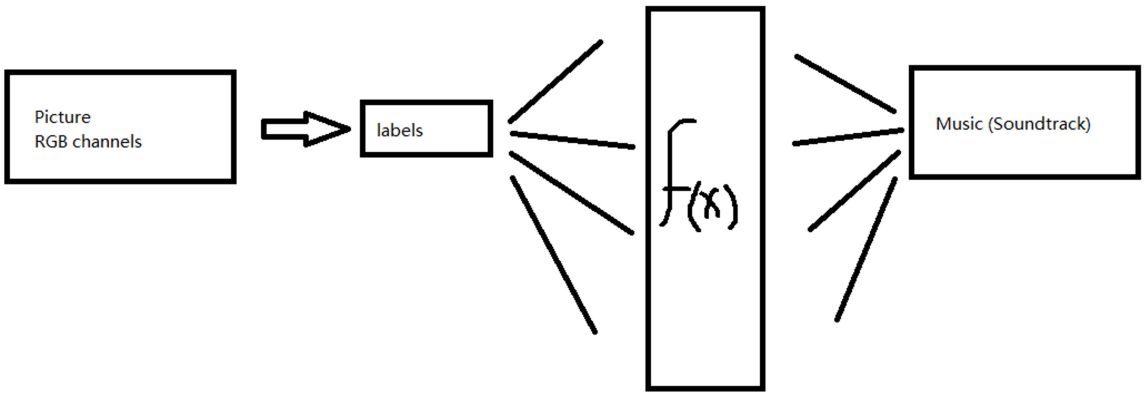
Introduction

2019年7月8日　0:13

Video -> Music（selection / generation）

Picture -> Music（虽然没有什么需求）



timeline带来的信息（frame之间的关系）：

shot transition

事件

Selection: 视频使用哪几个要素来选择合适的音乐?

1. 场景转换 -> 鼓点
2. 标签：风格，物品，事件 -> 节拍，调性...

Fx: L2R

PICASSO – Automated Soundtrack Suggestion for Multi-Modal Data

Generation: 视频使用哪几个要素来生成音乐?

1. 场景转换 -> 鼓点
2. 标签：风格，物品，事件 -> 节拍，调性...

Videos -> music genertaion / selection?

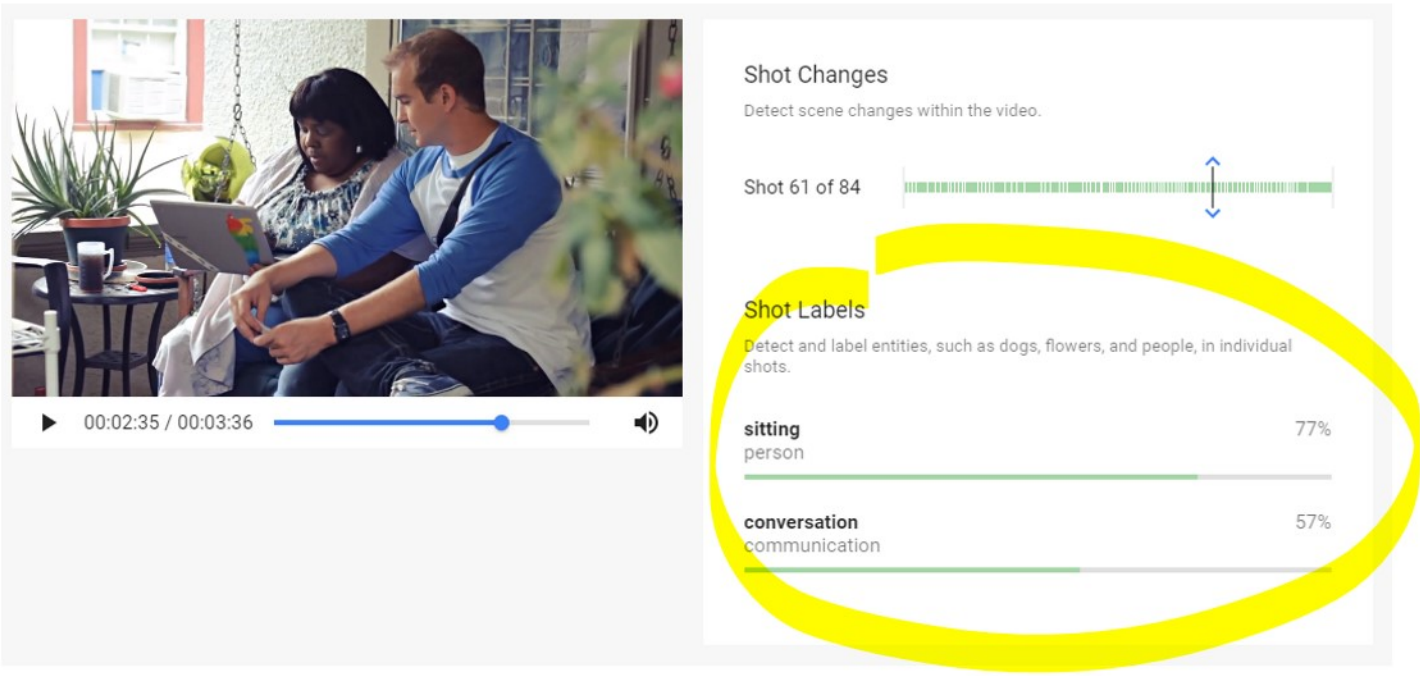
videos -> labels or shot changes -> music Generation / selection

Shot: a series of interrelated consecutive pictures taken contiguously by a single camera and representing a continuous action in time and space

Shot boundary detection

2019年7月5日　17:17

Shot changes detection (shot transition detection / shot boundary detection)

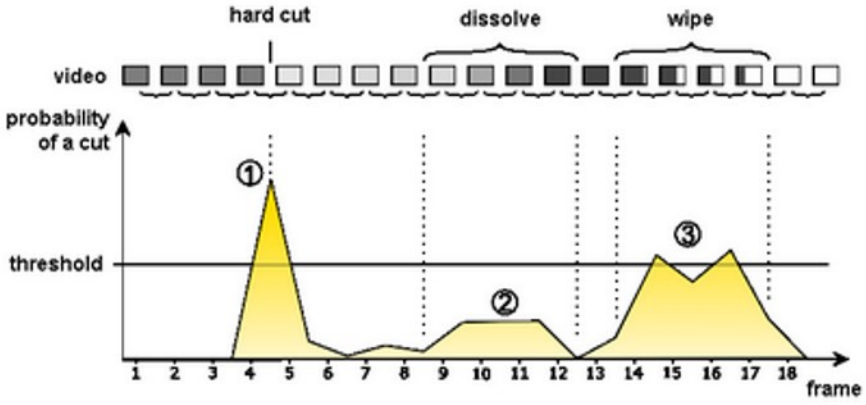


hard cut (abrupt transitions) & soft cut (gradual transitions)

Traditional Shot Transition Detection Method: two-phase-principle, Scoring & Decision

Scoring: Each pair of consecutive frames of a digital video is given a certain score that represents the similarity/dissimilarity between these two fra mes.
Decision: All scores calculated previously are evaluated and a cut is detected if the score is considered high.

Threshold: fixed / adatptive



Shot-Boundary Detection: Unraveled and Resolved?
Comparison of video shot boundary detection techniques
Performance Characterization of Video-Shot-Change Detection Methods

evaluation

<CNN-Based Shot Boundary Detection and Video Annotation>

Tags can describe the contents of the frame well. Tags can be used for shot detection, cause frames in the same shot tend to share similar tags.

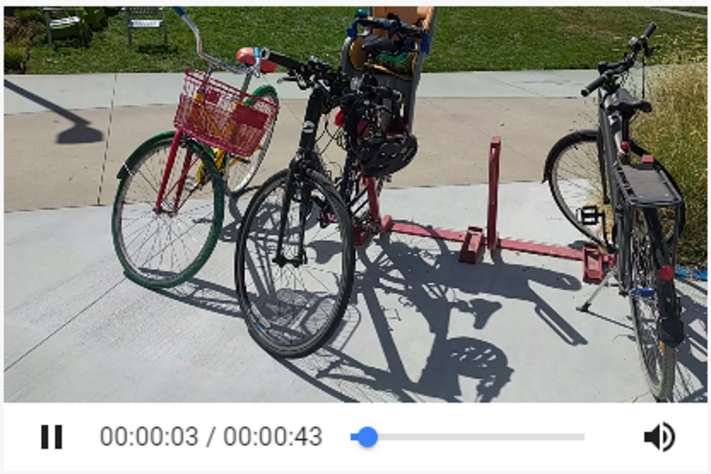
Using a CNN model similar to the ImageNet to get labels of each frame.

Most frames are non-boundary frames, so we pick candidate segmentation selection first. This step can increase both accuracy and speed of the further steps. This step is based on the fact the consecutive frames within one shot always have high correlations.

Video annotation

2019年7月8日　0:13

Video labels



Video Labels

Detect and label entities, such as dogs, flowers, and people, throughout the entire video.

dinosaur	80%
vehicle	79%
tree plant	40%
land vehicle vehicle	36%
bicycle vehicle	36%