Music generation

2019年7月22日 22:47

Baseline:

1. emotion -> tonality -> chord -> rhyme
2. Shot motion intensity for every shot -> tempo & meter

TSC
Dataset: music video

Method:

1. TSC -> tempo
2. Encoder + decoder

Image captioning

2019年7月21日 10:04

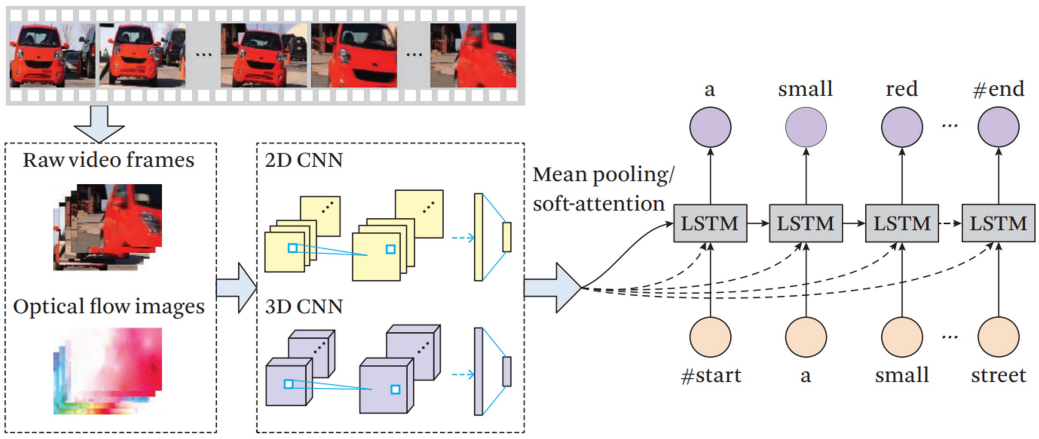
Image captioning

1. Based on searching: using CNN to compute a distance between caption and image
    a. A model to evaluate track of music generated and video?
2. Based on language template: getting objects in picture and using SVO template creating a sentence.
3. Based on sequence learing: CNN + RNN (encoder + decoder).
    a. Image captioning with X
    b. Music generator with LS structure and self-attention mechanisms

Video captioning:

1. Template-based language models
2. Sequence learning models: directly learn a translatable mapping between video content and sentence.

1. Prime -> LSTM



Timeseries classification

2019年7月21日 23:00

1 neasrest neighbor (K = 1) and Dynamic Timewarping is very difficult to beat.

LSTM

Event:

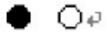
[event_code, start frame, duration, speed]

[transition, start frame, duration frames,]
[camera_shift, start frame, duration frames, speed]
[camera_zoom, start frame, duration frames, speed]
[camera_spin, start frame, duration frames, speed]
[motion, start frame, duration frames, intensity]

Type 1:
1. Volume: 2
2. Speed: 2
3. Tonality: 23

Type 2:
3. Tonality: 2

Tempo -> beat

拍号		含义	强弱规律
$\frac{2}{4}$		以四分音符为一拍每小节二拍	
$\frac{3}{4}$		以四分音符为一拍每小节三拍	
$\frac{4}{4}$		以四分音符为一拍每小节四拍	
$\frac{3}{8}$		以八分音符为一拍每小节三拍	
$\frac{6}{8}$		以八分音符为一拍每小节六拍	
$\frac{5}{4}$	$\frac{3}{4}+\frac{2}{4}$	以四分音符为一拍每小节五拍	
	$\frac{2}{4}+\frac{3}{4}$		