計算機は音楽のオーディオとビデオの相関を学習できるか？
Can we learn the correlation between music audio and video?
深層注意モデルによる音声と映像のクロスモーダル音楽検索に関する新たな研究
New research on cross-Modal Music Retrieval Between Audio and Video by Deep Attention Model

Donghuo Zeng(SOKENDAI), Yi Yu, Keizo Oyama

背景 Background

深層学習は、異なるデータモダリティ間の統合表現を学習するには、優れた性能を示してきました。しかしながら、オーディオやビデオなどのデータモダリティに重要である時間的構造は、クロスモーダル相関学習に関するほとんどの研究には考慮されていなかった。

Deep learning has successfully showed excellent performances in learning joint representations between different data modalities. Unfortunately, little research focuses on cross-modal correlation learning where temporal structures of different data modalities such as audio and video should be taken into account.

目標 Target

音楽とビデオのクロスモーダル検索を目指し、我々は、音楽のオーディオ信号とビデオ信号の時間的構造の特性を利用して、その深層シーケンス相関モデルを学習します。

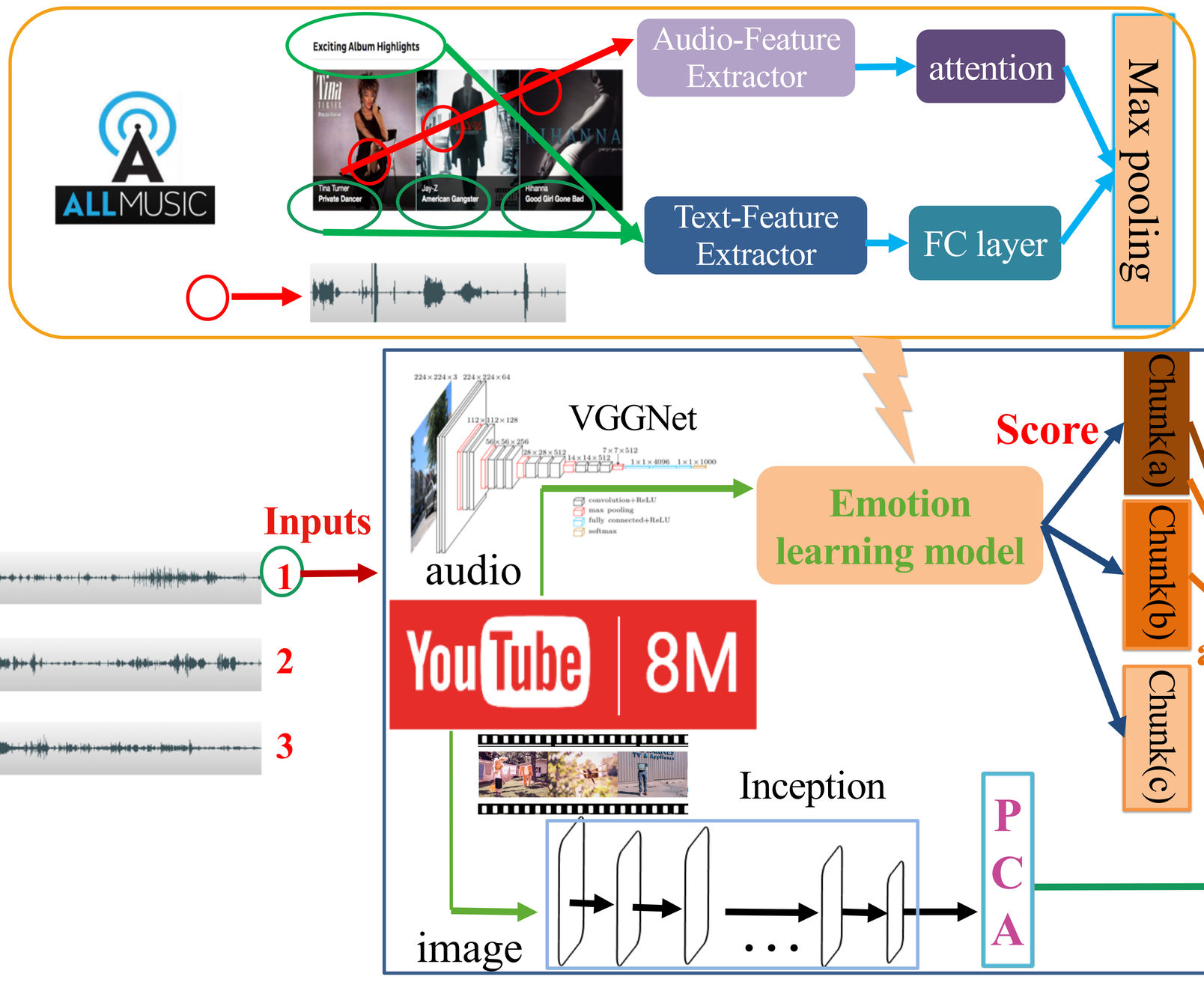
Aiming at the cross-modal retrieval between music audio and video, we try to exploit the temporal structure of audio and video signal, and learn a deep sequential correlation model between them.

内容 Contents

このポスターでは、音楽に関するマルチモーダルデータセットとして知られているYouTube-8Mからオーディオとビデオのペアをサブセットとして抽出する方法、時系列信号をそれぞれの注意特性を持つチャンクに分け、さらにチャンクごとにその時系列的性質を抽出する深層注意モデルの学習、および、2つのモダリティの相関性を評価する目的関数としてマルチリニア部分空間学習モデルを使用する方法について説明します。予備実験中得られた、Recall @ K (1, 5, 10, 20) などの結果により、我々が提案した深層注意モデルが音楽オーディオとビデオとの間のクロスモーダル検索には有効であることを確認しました。

In this poster, we will demonstrate how we extract a subset with the pairs of audio and video from YouTube-8M as music multimodal content-based dataset, how we learn deep attention model to infer the position information of all chunk sequence and the temporal sequence property inside chunk, and how we use multi-linear subspace learning as objective function to map the two modalities. Some preliminary results such as Recall@K (1, 5, 10, 20) show our proposed deep attention model can be applied to cross-modal retrieval between music audio and video.

研究アーキテクチャ Research Architecture



a: position=3; **a+b**: position=1
Recall@k = #position / #rank items

a, a+b

3, 1

2, 2

連絡先：YU Yi (ユイ) ／ 国立情報学研究所 コンテンツ科学研究系
TEL：03-4212-2574 FAX：03-4212-2035 Email：yiyu@nii.ac.jp