Problem Statement Problem Statement

Dataset Dataset

Approach

Neural Style Transi on images Cost Function

Cost Function

How Mel

Spectogram wor

Results

Any Questions

NEURAL STYLE TRANSFER ON AUDIO SIGNALS

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY VADODARA

Course co-ordinator: Dr Jignesh Bhatt



Contents

1 Problem Statement

- Problem Statement
- Motivation
- Dataset
 - Dataset
- 3 Approach
 - Neural Style Transfer on images
 - Cost Function
 - Mel Spectogram
 - How Mel Spectogram works?
 - Style Transfer
 - Results
 - References
 - Any Questions?

Problem Statement

Problem Statement Problem Statement Motivation

Datase Dataset

Approach
Neural Style Transfer
on images
Cost Function
Mel Spectogram
How Mel
Spectogram works?
Style Transfer

How Mel Spectogram work Style Transfer Results References Any Questions? Neural Style transfer is very popular for generating novel arts. In this project we try to extend the application of Neural Style transfer to audio signals as well. We intend to create new music genre by applying style of an existing genre on another.

Motivation

Problem Statement Problem Statemer Motivation

Dataset
Approac

Approach
Neural Style Transfel
on images
Cost Function
Mel Spectogram
How Mel
Spectogram works?
Style Transfer
Results
References
Any Questions?

"Style Transfer" of images are very popular and an active research topic. Its everyone's favourite tool to make photos look like surreal paintings. Convolutional Neural Networks (CNNs) are adaptable to a great variety of tasks. We want to extend and modify the algorithm for audio signals and use the power of CNNs to generate new audio from a style audio. This can be a tune or a beat, or even simply someone speaking the lyrics of a song.

Dataset

Problem Statement Problem Statemen Motivation

Datase Dataset

Neural Style Transfer on images Cost Function Mel Spectogram How Mel Spectogram works? Style Transfer Results

- GTZAN Dataset Music Genre Classification
 - o 30 seconds .wav files of Jazz and Classical genres
- Other audio clips taken from different sites to try out combinations
 - Music from games, ex Mario, etc
 - Popular movies theme songs, ex Interstellar, Harry Potter, etc

Neural Style Transfer on images

Problem Statement Problem Statemer Motivation

Dataset

Approach

Neural Style Transfer
on images

Cost Function

Mel Spectogram

How Mel Spectogram works? Style Transfer Results

Results
References
Any Questions?

Content Image

- The image on which style transfer is applied.
- Style Image
 - o The style we want to transfer to content image.
- Generated Image
 - The final blend of content and style image.
- Loss Function
 - The Loss Function is made up of Content Loss Function and Style Loss function.

Neural Style Transfer on images

Problem Statement

Statement

Date

A -----

Neural Style Transfer

on images

cost i unction

How Mel

Spectogram work

Results

Any Opertion













Neural Style Transfer on images

Problem
Statement
Problem Statemer
Motivation

Dataset

Approach

Neural Style Transfer on images

Cost Function

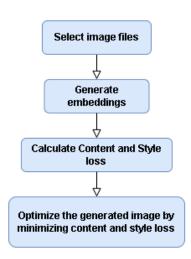
Mel Spectogram

How Mel

Style Transfer

Results

References



Content Loss Function

Cost Function

Using hidden layer 'l' to compute content cost.

- Use pre-trained ConvNet. (E.g., VGG network)
- Let $a^{[l](c)}$ and $a^{[l](G)}$ be the activation of layer I on the images
- If $a^{[l](C)}$ and $a^{[l](G)}$ are similar, both images have similar content

$$\mathcal{L}_{content}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2$$

p The original image

 \vec{x} The generated image

l Layer

 F_{ii}^l Activation of the i^{th} filter at position j in the feature representation of \vec{x} in l

 $P_{ii}^{\vec{l}}$ Activation of the i^{th} filter at position j in the feature representation of \vec{p} in l

Style Loss Function

Problem Statement Problem Statemen Motivation

Dataset Dataset

Approach
Neural Style Transfon images
Cost Function
Mel Spectogram
How Mel

Mel Spectogram
How Mel
Spectogram work
Style Transfer
Results
References
Any Questions?

- Style is defined as correlation between activations across channels
- Let $\mathbf{a}_{i,j,k}^{[l]} =$ activation at (i, j, k) and layer 'l'. $G^{[l]}$ is $\mathbf{n}_{\mathrm{c}}^{[l]} \times \mathbf{n}_{\mathrm{c}}^{[l]}$
- Calculated as the distance between the gram matrices of the generated image and the style reference image.
- Difficult to compare intermediate features of both images hence Gram matrix used.
- Style loss defined as -

$$\begin{split} L_{style} &= \sum_{l} w^{l} L_{style}^{l} \text{ where,} \\ L_{style}^{l} &= \frac{1}{M^{l}} \sum_{ij} (G_{ij}^{l}(s) - G_{ij}^{l}g)^{2} \text{ where,} \\ G_{ij}^{l}(I) &= \sum_{k} A_{ik}^{l}(I) A_{jk}^{l}(I). \end{split}$$

Total Loss Function

Problem Statement Problem Statement

Dataset

Approach

Neural Style Transfon images

Cost Function

How Mel Spectogram works

Style Transfer

Poforoncoc

Any Questions

$$\mathcal{L}_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{content}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{style}(\vec{a}, \vec{x})$$

 $\pmb{\alpha}$ and $\pmb{\beta}$ are the weighting factors for content and style reconstruction

Mel Spectogram

Problem Statement Problem Statement Mativation

Dataset Dataset

Approach Neural Style Transfer

Neural Style Transfe on images Cost Function

Mel Spectogram

Spectogram works? Style Transfer Results

Any Questions

- What is Spectogram?
- Why Mel Spectogram?

$$m = 2595log(1 + \frac{f}{500})$$

■ How Mel Spectogram works?

How Mel Spectogram works?

Problem
Statement
Problem Statement
Motivation

Datase Dataset

Approach
Neural Style Transfe
on images
Cost Function
Mel Spectogram

How Mel Spectogram works?

Results

- Apply STFTs on audio signals.
- Convert Amplitude to Decibels(DBs).
- Convert frequencies to Mel Scale.

Style Transfer

Problem
Statement
Problem Statement

Dataset

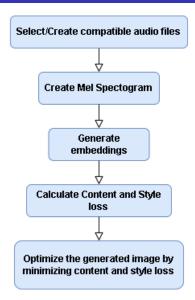
Annroach

Neural Style Transf on images Cost Function Mel Spectogram

How Mel

Style Transfer

Results



Results

Results





Style



Generated Audio

References

Problem
Statement
Problem Statement
Motivation

Datase Dataset

Approach

Neural Style Transfer on images Cost Function Mel Spectogram

How Mel Spectogram work

Style Transfer Results

References
Any Questions?

- Neural Style Transfer on Audio Signals
- Neural Style Transfer for Audio Spectograms
- TensorFlow

Any Questions?

Problem

Duckley Casas

Motivation

Dataset

A

Neural Style Transfe

. . . .

How Mel

Spectogram work

Results

References