ART & ML Final Project
4/30/2019

Painterly: An Art Agent That Paint With Music

Teammates: Qingyi Dong, Zhuona Ma, Vera Schulz

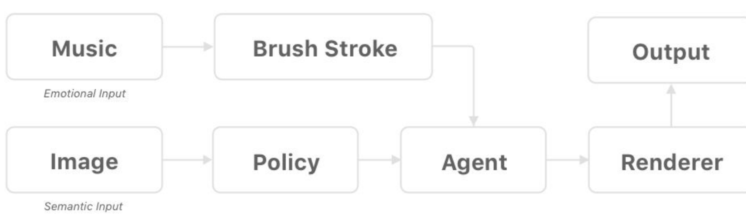
Github: https://github.com/CandyDong/Painterly

Concept

Painterly is an art agent that is able to learn to paint with human like strokes while actively "listening" to music. It intakes the emotional measurements from music pieces to render its brush, and then, paints semantics contents with abstraction.

Although, there already exists a lot of work using machine learning to produce paintings, we focused on two aspects that could improve the performance of machine learning algorithms on producing authentic human art. First, we wanted the machine to mimic the process that a human artist goes through when deciding what art to generate and then painting his or her vision. Since we had a specific focus, first, instead of examining images from pixels, we wanted to use brush strokes to create paintings. Second, since the human process of painting is mostly highly defined by the emotions they have before, during and after they paint, we also wanted the brush strokes to carry emotion information similar to that of humans .

Technique

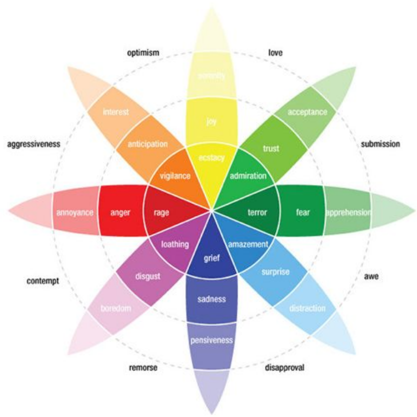


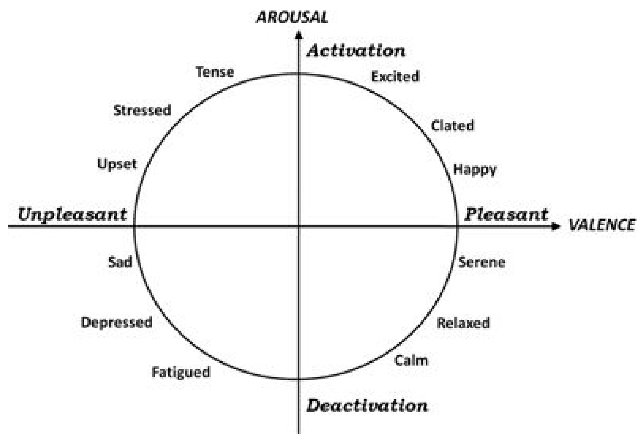
Model and Pipeline

Process

I. Music

In order to interpret the musical elements as emotional indicators, we analyzed music based on the Arousal-Valence Emotion Model, i.e. Circumplex Model. We used a subset of a DEAM dataset that contains 2,000 music pieces and then mapped them based on arousal, valence, tempo, pitch, etc to digital brushstrokes. We then, applied deep reinforcement learning with a neural renderer to create a progressive digital painting.



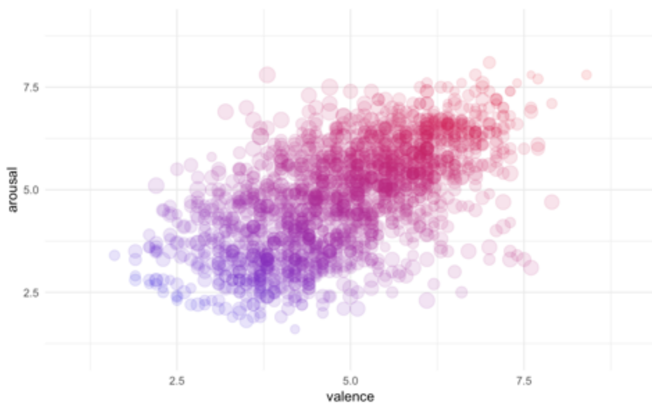


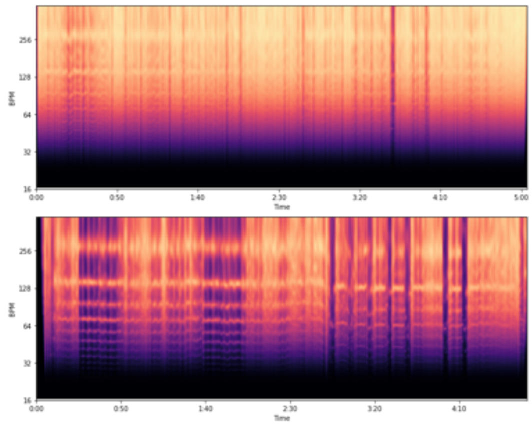
One way to interpret the music is to input it in terms of emotions, to embedded it and evaluate it based off of the Circumplex Model, which maps emotions into two dimensions: valence and arousal. A state of arousal represents the vertical axis and valence represents the horizontal axis, where the center of the circle represents a neutral valence and a medium level of arousal. In this model, emotional states can be represented at *any* level of valence and arousal, or at a neutral level of one or both of these factors. We used a subset of DEAM dataset, which contain a corpus of valence/arousal labeled music pieces by psychologists.

Since our goal is to extract compositional features that are representative of human emotions, we extracted a list of 18 features and their mean, standard deviation, and variance. Some key features are listed below:

Contrast	**Rolloff**	**Tempo**	**MFCC**	**Chroma**	**RMSE**	**Frame**	**Harm**

After we computed all the features for a subset of 4,000 music pieces, we decided to select the most representative features by looking at, both, the extreme valence/arousal values and their corresponding features to evaluate the relevant representativeness of each feature. This enabled us to map the emotional attributes to brushstroke attributes like distribution, shape, etc.





II. Painting

Neural Renderer:

Instead of using images files of different brush strokes, we used a neural network to generate different brush strokes using parameters specified in the *Brush Stroke Design* section below. We used *OpenCV* to generate random training samples. The neural renderer network is then, trained with supervised learning to map a sequence of stroke parameters to a rendered stroke on canvas.

Painting Agent:

We decomposed the target image to stroke representations before feeding it into our model which is based on a reinforcement learning network.

We used a Markov decision process to model the task, with state space S, action space A, transition function P, and reward function R.

The state space S consists of three parameters: canvas, target image, and the step number. Given a target image and canvas Ct, the goal of the painting agent is to choose a sequence of parameters that describes a brush stroke, and then let the trained neural renderer produce canvas Ct+1 with the new brush stroke rendered on the original canvas.

The action space A contains a set of sequences of parameters that define brush strokes that would be painted onto the canvas at the current step.

The reward function R is defined as follows,
r(st, at) = Lt − Lt+1,

Where st is the current state, at is the action chosen, Lt is the loss between the target image and the previous canvas Ct, and Lt+1 is the loss between the target image the current canvas Ct+1. It's designed to be able to model the difference between the target image and the current canvas.

Dataset:

We trained the painting agent on the officially center-cropped celebA dataset which contains 200000 images of celebrity faces. We resized all of the training image to 128x128 before feeding them into the neural network.

We then assigned selected artworks to certain music pieces, to simulate the connection between art, music and then emotion. We wanted to show how intense the bond can be between art and music, and show that the creation of art while listening to different types of music can highly impact the feeling and imagery of the final painted product.

Brush Stroke Design:

The brush stroke design defines the emotion contents of a target painting. Our brush stroke structure is represented by the Quadratic Bezier Curve (QBC) as follows:

B = (x0, y0, x1, y1, x2, y2, z0, w0, z1, w1, R, G, B, S),

Where (x0, y0), (x2, y2) are two endpoints of one continuous stroke, and (x1, y1) is the third point that defines a Quadratic Bezier Curve; z0, z1 each represents the radius of the brush at two endpoints; w0, w1 each represents the transparency of the brush at two endpoints; R, G, B are the RGB color values of the brush stroke; S represents the basic shape of a brush stroke, for example, circle, rectangle, triangle, eclipse, etc.

We mapped the valence and arousal parameters of a music piece to different stroke transparency based on radius, and shape.

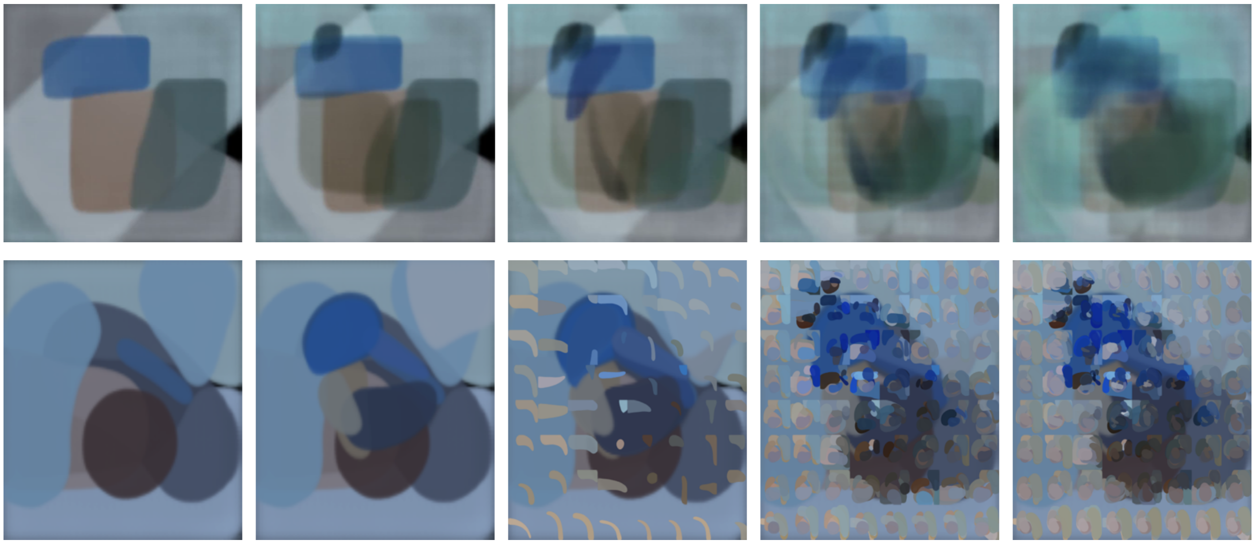
Training

We trained the neural renderer for 50,000 steps with a batch size of 64. We trained the painting agent for 200,000 steps with a batch size of 96.

Prediction

The results shown in the *Results* section is produced by the painting agent taking a maximum of 40 steps. The first 20 steps are made on the original canvas size (128x128), and the rest of the steps are made in parallel on 16 smaller canvas to produce results with finer details.

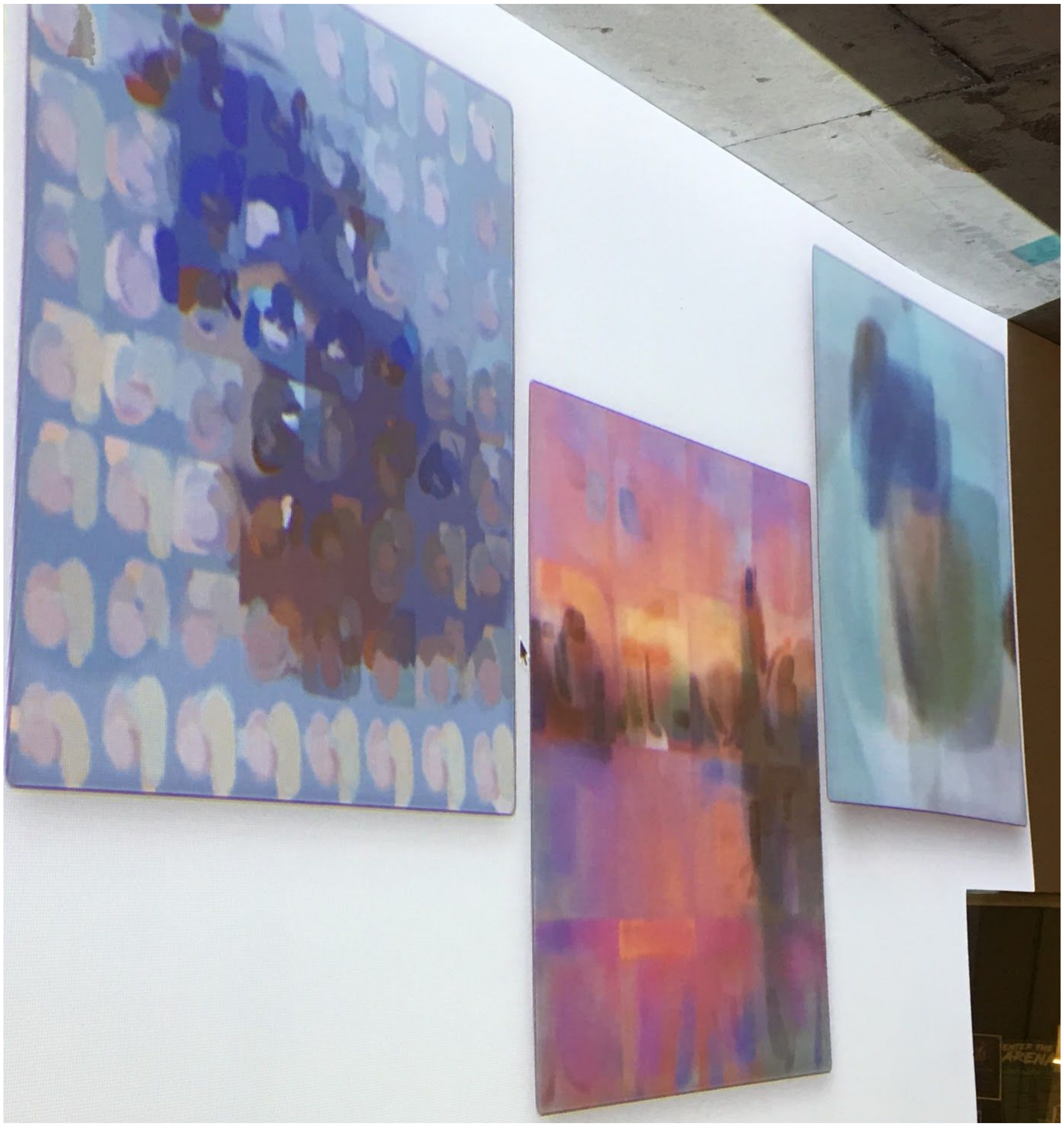
Results



Reflection

Currently, we are creating a digital painting based on specific brushstrokes and emotion that is generated based on what type of music is played: happy or sad. We would like this research to eventually lead to giving machines the opportunity to derive their own emotions from music and then output an individual painting based off of the different brush stroke styles we fed into the system. Ideally, the machine should be able to listen to a specific piece of music which is associated with a certain emotion and then generate its own painting and style, based off of the emotional input.

The final step would be letting the machine generate its own emotions and then paint how it felt, but in order to do this we would have to figure out how to define the levels of machine independence.Eventually, we will use the step-wise V/A labels to create brushstrokes that varies during the music, which could be a good simulation to human artists who “paint while listening to music.”



Individual Contributions

Zhuona Ma
Music Dataset Cleaning, Feature Extraction, Emotional Measurements Evaluation

Qingyi Dong
Brush stroke generation, RNN model research/training, Output video generation.

Vera Schulz
Artistic inquiry and concept, Preliminary and primary dataset research and identification of abstract brushstroke, Music and artwork datasets, Technical documentation

References

[1] White, Tom, University of Wellington School of Design, Title: Perception Engines

[2] Elgammal, Ahmed; Liu, Bingchen, Art & AI Laboratory - Rutgers University, Title: CAN: Creative Adversarial Networks Generating “Art” by Learning About Styles and Deviating from Style Norms

[3] Sundararajan, Louise, Journal of Creative Behavior, Title: Mind, Machine, and Creativity: An Artist’s Perspective

[4] Campos-Bueno, J. J., Spanish Journal of Psychology (2015) Title: Emotional Dimensions of Music and Painting and their Interaction

[5] Wieczorkowska1, Alicja; Synak, Piotr; Title: Extracting Emotions from Music Data., Polish-Japanese Institute of Information Technology.

[6] Mohammad, Saif M; Kiritchenko, Svetlana; Title: WikiArt Emotions: An Annotated Datatset of Emotions, National Research Council Canada