

Computationally Generating Images for Music Albums

Miguel Cruz, Paul Hardman, João Miguel Cunha

CISUC, Department of Informatics Engineering,

University of Coimbra

migmcruz@gmail.com, {phardman,jmacunha}@dei.uc.pt

Abstract

Nowadays, music albums are seen as a set of unrelated pieces rather than a whole. We propose an approach for a system which produces images for songs of music albums, with the goal of establishing graphic cohesion. It uses semantic analysis of lyrics and semiotic properties to visually represent the meaning and emotions of songs.

Introduction

In the last 20 years there were a lot of changes in the music industry, due to the tremendous technological growth. The new paradigm, dominated by singles and playlists, changed the concept of music album from something which was normally seen as a whole into a simple set of individual and unrelated pieces. Moreover, the relation between the music album and its related concerts has been reduced – more importance is given to festivals with several music artists rather than to individual concerts (Tschmuck 2012).

The main goal of this project is to use computational means to integrate all the materials and events related to a music album, making it work both individually and as a whole. In order to achieve this, we propose a computational approach that produces images from the song lyrics of an album and transforms them according to sound variables of the songs. These images can then be used to (i) create the graphic materials of the album, (ii) produce videos for the songs and (iii) visual effects for the corresponding concerts. This way, all of the materials and events will be integrated in one single graphic system, fully related to the music album.

Our Approach

The relation between text, image and sound serves as a basis to the proposed computational system, which is divided into six layers (L1-L6). This modular architecture allows an iterative development towards our final goal: having images that illustrate the songs' lyrics and are distorted according to the sound of the voice and instruments, using semiotic properties to visually represent the emotions of the songs.

L1 Lyrics Analysis The first layer is responsible for the analysis of the lyrics, which retrieves words that can be searched in an image database. In order to do this, the system uses the Processing RiTa library to divide the lyrics in nouns, adjectives and verbs.

L2 Lyrics Preparation In this layer, the user is asked to determine the time in the song in which each retrieved word appears, in order to know when the images should be created or distorted. This achieves synchronization between song and image.

L3 Image Retrieval The nouns retrieved by L2 are used in a search process conducted in an image database (e.g. Unsplash), gathering images that illustrate them. Our goal is to follow a semi-automatic approach, in which the user has the possibility of selecting which gathered images should be used.

L4 Image Preparation This layer uses the images provided by the previous layer and applies a filter to them, making them better suited for visual transformations. Currently, we are using OpenCV Canny Edge Detector to transform the images into a set of ellipses. We plan to explore other possibilities in the future: e.g. line representation or even clustering-techniques.

L5 Illustration Production The graphic elements, produced by the previous layer from the gathered images, are used in combination with semantic analysis to produce illustrations. This process uses the verbs and adjectives retrieved using RiTa and uses them to apply visual transformations, based on semantic-semiotic mappings (e.g. words related to movement affect element positioning).

L6 Background Variation The background of the illustration will be affected by sound variables of the instrumental. One example is changing the colour according to the musical scales or chords that are being played at each moment: colder colours if the scales/chords are minor and warmer colours if they are major. In order to do this, we will orientate the study to the piano to make a MIDI analysis in Pure Data. The program could have another component consisting in an estimated analysis of the full instrumental.

Future Work Currently, the user is manually searching the images, as the image retrieval is not yet implemented. In the future we intend to explore with other filters for the Image Preparation layer.

References

Tschmuck, P. 2012. Creativity and innovation in the music industry. In *Creativity and Innovation in the Music Industry*. Springer. 225–251.