Sound Feature-Driven Painterly Rendering

Kyeong A Kwon and Hae Won Byun*

Dept. of Computer science, Sungshin Women's University
{kyeonga, hyewon}@sungshin.ac.kr

Abstract

In this paper, we propose a new painterly rendering technique for visualizing sound, based on correlation between sight and auditory. We analyze the sound feature to create an image of an intense atmosphere or a relaxing atmosphere which is dependent upon the input sound. We analyze the tempo and the frequency which are the basic characteristics of sound features. Sound features are used to determine strokes shape and color which are the crucial elements to complete the pictorial rendering. In addition, during the rendering process, we use a multiple layer system to overlap the layers. From this study, we can create a non-photorealistic image which is suitable for the mood of the music.

Keywords: *Non-Photorealistic Rendering, Painterly Rendering, Sound Feature*

1. Introduction

In recent years, there have been plenty of researches tackling how to connect sight and sound. For example, a musical fountain where the water moves by music and a media player which displays a wavelength based on the music. To visualize the sound, a musical fountain is focused on making the shape of the fountain by the rhythm and the pitch and a media player is focused on changing the shape of the wavelength.

In this paper, we use a painterly rendering technique to visualize the sound. To express the synesthesia that harmonizes with visual and auditory senses, we use sound features to determine the brush strokes. The tempo and the frequency which are the basic features used to categorize the music genre are used as sound features. Then each sound feature is used to select the shapes and the colors of the brush strokes. So we obtain the result which is suitable for the atmosphere of the input music.

2. Related Work

The ultimate objective of painterly rendering is to create non-photorealistic images by placing brush strokes according to some goals. Painterly rendering was developed in the early '90s by Paul Haeberli [1]. Haeberli proposed a painterly rendering by brush strokes from user input. Litwinowicz [2] created the brush strokes by using direction and texture of the strokes. Later, the painterly rendering using the curved β -spline strokes were proposed by Hertzmann [3]. He also presented a multi-layer system and used the different size of strokes in each layers. Hays [4] further arranged the brush strokes in motion layers. Hays used an interpolation using Radial Basic Function (RBFs) to make the direction of the strokes. Decarlo and Santella [5] provided an eye-tracking-based user interface to allow for a user-guided specification of regional importance which enables adaptive control of the level of abstraction. In recently,

* Corresponding Author

Shugrina, *et al.*, [6] present an interactive painterly rendering whose appearance adapts in real time to reflect the perceived emotional state of the viewer. They analyze facial expressions of user through the detection of facial action units derived from the FACS scheme.

In this paper, we propose a new painterly rendering technique for visualizing sound and create the brush strokes using sound features from input music. Our goal is to create the suitable results for the mood of inputted music. Consequently we can express the synesthesia that is harmonized with visual and auditory.

3. System Overview

Figure 1 shows the configuration of our system. The proposed system can be roughly divided into the sound feature phase and the rendering process phase. In the sound feature phase, we calculated the music tempo by using BPM (Beat Per Minute) and analyze the frequency of the input music. The tempo and the frequency are used to determine strokes shape and color in the rendering process phase. In the rendering process, first of all we apply a Gaussian blur to reduce the noise in the input image and then determine the shape and the color of the strokes. After that, the strokes are rotated following the edge orientation by using ETF(Edge Tangent Flow) [7] and placed into a grid-based. Finally, using a multiple-layer system, we can obtain the final results.

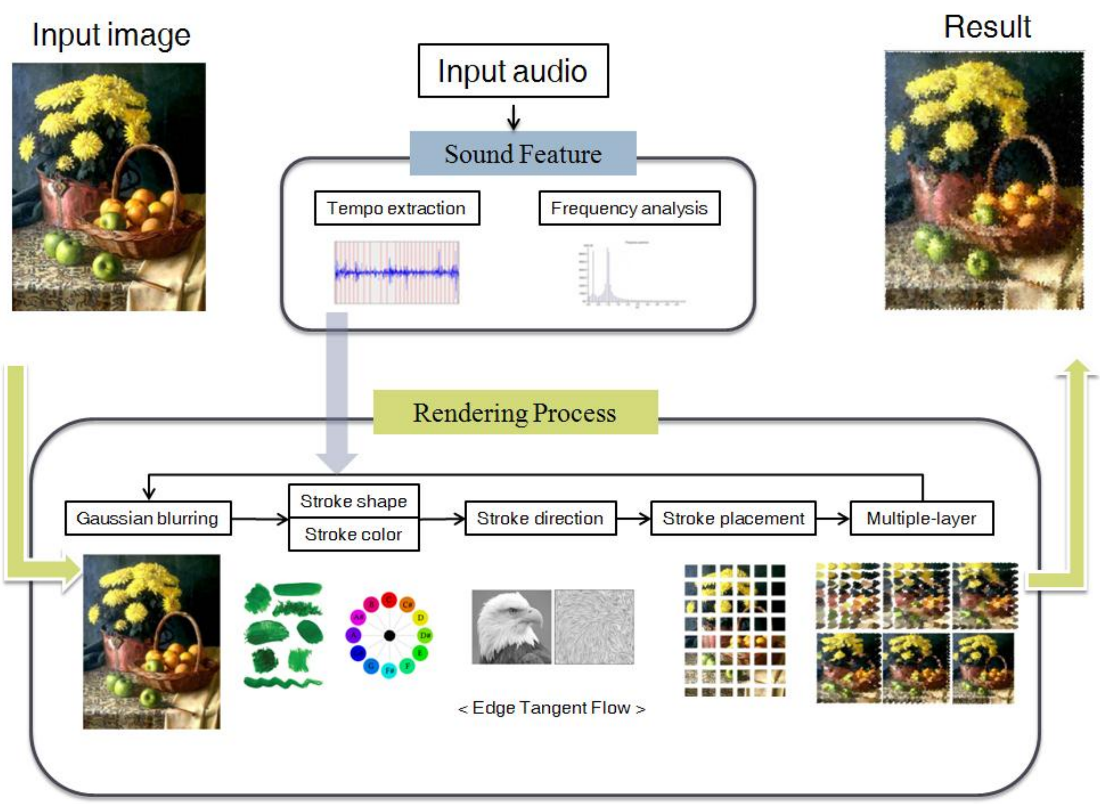

Figure 1. Overview

4. Sound Feature Extraction

4.1. Tempo Extraction

To calculate the tempo from the input sound, we use Beats per minute (BPM). BPM is a unit typically used as a measure of tempo in music and heart rate. In musical terminology, tempo is the one of the most basic feature of the music configuration and it usually means fast of the music. However, a lexical meaning of the tempo is an atmosphere of the music. In general, the classic music has a slow tempo and the rock music has a fast tempo. So tempo is a crucial element to categorize a music genre. And in this paper, tempo is used to determine the shape of the stroke.

4.2. Frequency Analysis

A fast Fourier transform (FFT) is an efficient algorithm to compute the discrete Fourier transform (DFT) and its inverse. To obtain the frequency from the sound, we used FFT. Then determine the colors of the strokes by the frequency. The frequency of the sound decides the pitch. High frequency makes a high pitch and low frequency makes a low pitch.

5. Painterly Rendering

5.1. Stroke Shape

The shape and the color of the strokes are used as elements in painterly rendering. The shapes of the strokes are chosen by matching the extracted tempo from the source music. The atmosphere generated with a slow tempo matches the sound of its kind, giving it a simple and smooth stoke. Vice versa with music that has a fast and upbeat tempo the atmosphere rendered will be seemingly suitable for that particular type. We adopted some brush examples from the brush dictionary [8] (see Figure 2).

Figure 2. Brush Strokes

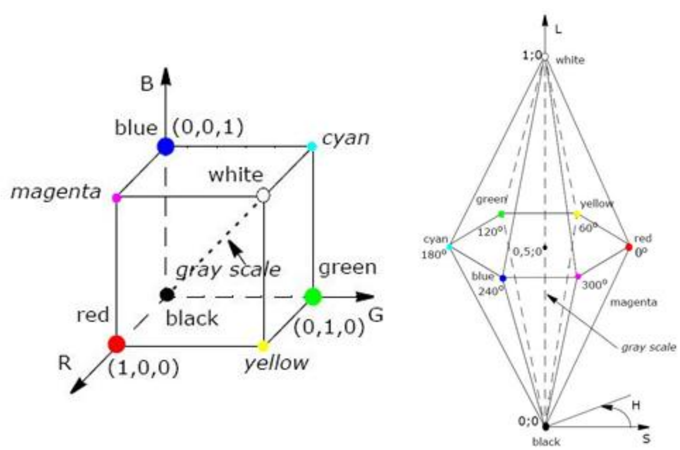

Figure 3. RGB Color Model and HSI Color Model

5.2. Stroke Color

In this study, we use a HSI color space rather than a RGB color space; because in RGB color space there will be correlations between the different channels' regarding the color displacement, values (see Figure 3). The HSI color model represents every color accurately using three components. H stands for the Hue, S for the Saturation and I for the Intensity. H can be chosen by frequency of the sound. We converted an RGB color image to HSI color space by using Equations 1. The original colors of the brush strokes are chosen from blurred source image and can be transferred by frequency of the input sound.

$$\mathrm{H} = \cos^{-1}\left[\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(\mathrm{R}-\mathrm{G})^2+(\mathrm{R}\text{-}\mathrm{B})(\mathrm{G}\text{-}\mathrm{B})}}\right]$$

$$\mathrm{S} = 1 - \frac{3}{(\mathrm{R+G+B})}[\min(\mathrm{R,G,B})] \qquad (1)$$

$$\mathrm{I} = \frac{(\mathrm{R}+\mathrm{G}+\mathrm{B})}{3}$$

$$(\text{If B} > G, H = 360° - H)$$

In this paper, we adopt the sound to color conversion to determine the color of strokes [9]. The sound to color conversion is a conversion method to convert sound into color, based on the likelihood in the physical frequency information between sound and color. Human can recognize the various colors according to the energy distribution of light 390THz (=3.9x10^12Hz)~750THz(=7.5x10^14Hz). Sound manifested the frequency (Hz) and the range of hearing which human can hear is approximately 20Hz~20kHz. Figure 4 shows the frequency mapping between audible and the visible frequency bands.

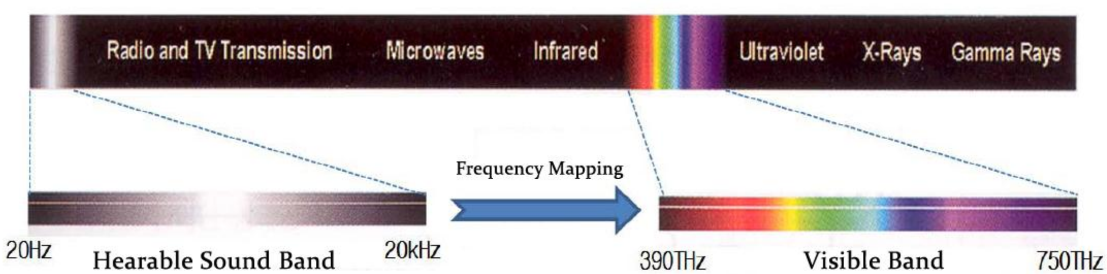

Figure 4. The Frequency Mapping between Audible and Visible Frequency Bands

We extract the sound feature for the frequency mapping between sound and color. The color and the sound physically share the attribution of a wave motion as sympathetic vibration, amplification, interference and offset. And a ratio of Do, Mi, Sol and a ratio of red, green, blue are conformable to each other. Using this principle, the audible frequency range is converted to visible frequency (see figure 5).

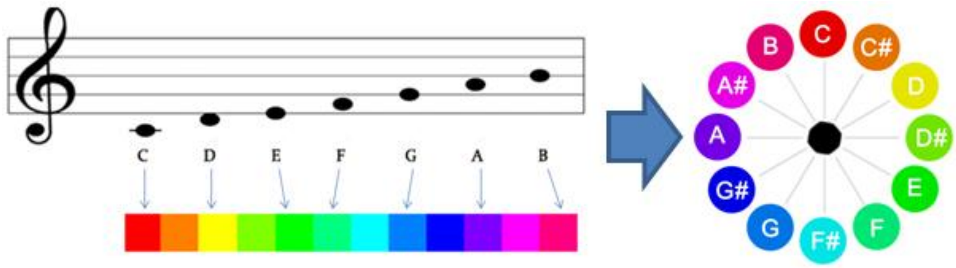

Figure 5. The Process of Converting Musical Scale Elements into Color

5.3. Stroke Placement

We use edge tangent flow (ETF) to obtain the flow field and rotate the strokes which are guided by ETF. During the rendering process, we use a multiple layer system. We start painting with a large brush for the first layer and paint again over the foreground layer with a smaller brush to add detail.

6. Results

Figure 6 shows rendering results obtained from the source image in Figure 6 (a). Figure 6 (b) is the basic rendering result that has no input music. We used the default brush and color was chosen from blurred source image. In Figure 6 (c) the input music is soft music, so we gave an example stroke of simple and smooth. In Figure 6 (d) the music expressed the fast and intense atmosphere. Thus we gave a rough brush and the red-shift colors used in the rendering.

(a) (b) (c) (d)

Figure 6. Results

7. Conclusion and Future Work

We have presented a new automatic system for painterly rendering from images. To choose the shapes and the colors of the strokes, sound features were extracted from the input music. Through the connection between sound features and rendering process, we obtain the results of the various styles.

Currently, we only extract the tempo and the frequency as sound features from input sound, but in the future, the painterly rendering will be considered the meaning of the lyrics of the song to apply to the rendering process. So that painterly rendering system will be able to create a result which is not only suitable for the mood of the song but also the meaning of the lyrics of the song. Our system is based on PC. However, in the future, if we extend the system for mobile phone, it can be more convenient for user.

Acknowledgements

This work was supported by the Sungshin Women's University Research Grant of 2012.

References

[1] P. Haeberli, "Paint by numbers: abstract image representations", In Proceedings of ACM SIG-GRAPH, vol. 4, **(1990)**, pp. 207–214.
[2] P. Litwinowicz, "Processing images and video for an impressionist effect", In Proceedings of ACM SIGGRAPH, **(1997)**, pp. 407–414.
[3] A. Hertzmann, "Painterly rendering with curved brush strokes of multiple sizes", In Proceedings of ACM SIGGRAPH, **(1998)**, pp. 453–460.
[4] J. Hays and I. Essa, "Image and video based painterly animation", In Proceedings of 3rd ACM Sympos. on NPAR, **(2004)**, pp. 113–120.
[5] D. Decarlo and A. Santella, "Abstracted painterly renderings using eye-tracking data", In Proceedings of ACM SIGGRAPH, **(2002)**, pp. 769–776.
[6] M. Shugrina, M. Betke and J. Collomosse, "Empathic painting: interactive stylization through observed emotional state", In Proceedings of the 3rd international symposium on Non-photorealis-tic animation and rendering (NPAR 2006) (Annecy, France, 2006), ACM Press, **(2006)**.
[7] H. Kang, S. Lee and C. K. Chui, "Flow-based image abstraction", IEEE Transactions on Visuali-zation and Computer Graphics, vol. 15, no. 1, **(2009)**, pp. 62-76.
[8] K. Zeng, M. Zhao, C. Xiong and S. -C. Zhu, "From image parsing to painter-ly rendering", ACM Transactions on Graphics (TOG), vol. 29 no. 1, **(2009)** December, pp. 1-11
[9] J. I. Kim and J. S. Jung, "A basic study on the system of convering color image into sound", In Pro-ceedings of KIIS Spring Conference, vol. 20, no. 1, **(2010)**, pp. 251–256.

Authors

Kyeong A Kwon

She received the BS degree in computer science from Sungshin Women's University, Korea, in 2010. She is currently working toward the MS in computer science at the Sungshin Women's University. Her research interests include non-photorealistic rendering, computer graphics, and mobile application.

Hae Won Byun

She received the BS degree in computer science from Yonsei University, Korea, in 1990 and MS and PhD degrees in computer science from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1992 and 2004, respectively. She is currently a professor in the School of Information Technology at Sungshin Women's University, Seoul, Korea. She's primary research interests and facial motion capture and expression retargeting, performance based animation, non-photorealistic rendering, and texture synthesis.