Music Generation from Moving Image

Shinji Hashimoto and Dingeling Chang
Advanced Research Center for Science and Engineering, Waseda University
Okubo 3-4-1, Shinjuku, Tokyo 169 Japan
Email: shinji@cisw.waseda.ac.jp, dcheng@cisw.waseda.ac.jp

ABSTRACT: This paper proposes a new system to generate music from moving image for multimedia applications. The system consists of an image analyzer to extract the features of the scene from the standard TV signal, the media harmonizer to associate the image features with musical parameters and the music generator to generate the harmonized sounds. The system is flexible and programmable to perform various music driven by moving image.

1 Introduction

Graphical methods are often employed in computer music composition. The typical examples are found in UPIC [Marin et al] and MAX [Puckette & Zicarelli]. Graphical image is used to express the temporal evolution of the sound as a sort of score in UPIC system. Graphical chart is used to design sound processing patches in MAX. Still picture is in a convenient medium to record and to describe music in the batch process of composition, while a moving image can be considered as a rich source of sound and music because it has both spatial and temporal attributes like music. Some musical systems to use video signal have been already reported [Morita et al.] [Boulanger][Vincen]. However, most of them use the image processing technology to analyze the performer’s gesture. In the paper, we propose a new system to generate music from moving video image for multimedia applications.

2 System Overview

The system consists of an image analyzer, sound generator and media harmonizer. As shown as Figure 1, the image analyzer is to extract the features of the scene in real time from the standard TV signal.

![Figure 1: System Overview](image-url)
signal. The media harmonizer associates the image features and musical elements and synchronizes the moving image and the generated music. The user can define the musical primitives and the rules of the harmonization. The music generator generates sounds according to the output of the media harmonizer using the FM sound source.

3 Moving Image Analysis

Although a lot of information is contained in moving image, we extract the image features which are useful for music generation and can be detected in real time [Pratt][Chang et al]. Scene understanding or object recognition is not available for real time implementation at present because it requires complicated operations. Two kinds of image features are extracted from every frame at first. One is structural information such as the number of edge points, spatial frequency and average brightness. Another is color information such as averages of RGB components, hue and saturation. These feature parameters are measured in every 8x8 sub-areas of the frame and represent the spatial characteristics of the scene. Then the choreographs of these features are examined to obtain the temporal features of the scene activities such as the scene change, object movements and temporal frequency.

4 Media Harmonizer

To associate the moving image with music, the media harmonizer has some rules of composition and music knowledge data base. The composition rules are selected and changed according to the extracted image features. Although user can define the rule of the harmonization, the spatial features are mainly associated with the musical timbre and harmony while the temporal features are related to the tempo and rhythmic aspect of music at present.

The synchronization between image and music is another important role of the media harmonizer. Therefore the progression of image features are watched to schedule the musical performance that is smooth and suitable to the trend of the scene.

The media harmonizer can accept the extra signals from the external devices as the foot switches and touch sensors to expand the applications.

5 Conclusion

Although the image analyzer does not understand the meaning of the scene, the proposed system can be used to compose the algorithmic music driven by moving image. Moreover it is useful for the automated back ground music generation in multimedia applications. The media harmonizer is flexibly programmable to realize a personal music scene. One can listen to the images in his own sensation and recipe.

References


ICMC PROCEEDINGS 1992

362