California Institute of the Arts

Kinematic Beat extraction for Dance Synthesis Applications

by

Sankarsh Rajeev

A thesis submitted in partial fulfillment for
the degree of Master of Fine Arts

Herb Alpert School of Music
Music Technology: Interaction, Intelligence & Design
2023

Supervisory Committee

Ajay Kapur

Committee Member

Kai-Luen Liang

Committee Member

Committee Member

Committee Member

Abstract

In this project, we present a dance-to-music AI system that uses OpenPose for real-time human pose estimation. The system consists of a camera that captures video of a dancer, a computer that processes the video using OpenPose and Python, and a music player that plays a selected music track. The computer extracts the dancer's pose information from the video in real-time, and uses it to control the music player. The dancer's movements are used to change the music's tempo, volume, and other characteristics, creating a dynamic and interactive dance experience. We evaluate the performance and accuracy of our system using a dataset of dance videos, and discuss potential applications and future directions. Overall, our work demonstrates the potential of using OpenPose for real-time human pose estimation in interactive dance-to-music AI systems.

Acknowledgments

I would like to express my sincere gratitude to Jayakumar. S Ph.D, Chaithra.B.V and Mistral Solutions I.T. Team for their invaluable contributions to my Kinematic Beat extraction for Dance Synthesis Applications project. Their expertise in computer vision and machine learning, as well as their dedicated support and partnership, were crucial in getting the project up and running. I am deeply appreciative of their tireless efforts, helpful guidance, and positive attitude throughout the project.

I am also grateful to the California Institute of Arts for providing me with the opportunity to work on this project, as well as to Ajay Kapur and McLean Macionis for their invaluable guidance and support. Without their vision, resources, and expertise, I would not have been able to complete this project.

I am deeply grateful to all of the individuals and organizations who have contributed to this project. I am truly fortunate to have had such a talented and dedicated team of experts working with me, and I look forward to future collaborations. Thank you all for your invaluable support and partnership

Contents

Abstract **3**

Acknowledgments **4**

Contents **5**

List of Figures **7**

Chapter 1 Introduction **9**

1.1 Historical Background **9**

1.1.1 Later On **10**

1.2 Objectives and Limitations **10**

1.2.1 Objectives **10**

1.2.2 Limitations of the Study **11**

1.3 Potential Hypothesis to achieve **12**

Chapter 2 Related Work **13**

2.1 Dance Synthesis **13**

2.2 Real Time Beat-Synchronous Dance Animation **14**

2.3 Generating Dance Animation from Music **14**

2.4 Automatic Dance Synthesis Using Audio and Motion Data **15**

Chapter 3 Problem Formulation **17**

3.1 Problems **17**

Chapter 4 Choreographic Action Unit **19**

4.1 Recent Developments **20**

4.2 Uses of CAUs **20**

Chapter 5 Methodology **22**

5.1 Introduction to Methodology **22**

5.2 The Methodology for Kinematic Beat Extraction for Dance Synthesis **23**

5.2.1 Audio Analysis **23**

5.2.2 Motion Capture **25**

5.2.3 Motion Analysis **26**

5.2.4 Motion Synthesis **27**

5.2.5 Real-time Performance **28**

5.2.6 Adaptability to different musical genres **29**

Chapter 6 DataSet **31**

6.1 Introduction to Datasets **31**

6.2 Types of Datasets used **31**

6.3 Additional Information **33**

Chapter 7 Conclusion **34**

7.1 Summary **34**

7.2 Possible Beat Extraction Concepts **35**

7.3 Final Thoughts **36**

7.4 Future Work **37**

Bibliography **39**

List of Figures

Fig 5.1 Typical flow chart of Dance to Music Synthesis 3

Fig 5.2 Audio Analysis Software 3

Fig 5.3 Avatar 2 Motion capture 3

Fig 5.4 Motion Synthesis using Deep Learning 3

Fig 6.1 Training Dance Datasets 3

Chapter 1

Introduction

Kinematic beat extraction is a technique used in computer-generated music and dance synthesis applications. The process involves analyzing audio recordings to identify and extract the underlying beat structure, which can then be used to control various aspects of a computer-generated dance performance. The Historical Background of kinematic beat extraction can be traced back to the early days of electronic music and computer-generated dance, when researchers first began experimenting with using computers to analyze and manipulate music and dance data.

1.1 Historical Background

In the 1980s and 1990s, researchers at institutions such as MIT, Stanford University, and the University of California, Berkeley, began developing algorithms and software tools for analyzing and extracting beats from audio recordings. These early efforts focused on simple techniques such as tempo detection and beat tracking, which were based on analyzing the rhythmic structure of the audio signal.

1.1.1 Later On

As technology advanced, researchers began to develop more sophisticated algorithms for kinematic beat extraction, incorporating machine learning techniques such as neural networks and deep learning. These newer algorithms were able to analyze more complex audio signals and extract more detailed information about the beat structure, including the timing, amplitude, and phase of individual beats.

In the early 2000s, researchers began to apply these techniques to the field of computer-generated dance and music synthesis. By using kinematic beat extraction to control the movements and actions of computer-generated dancers, researchers were able to create realistic and dynamic virtual performances that closely matched the rhythm and energy of the original audio recordings.

The field of kinematic beat extraction has continued to evolve in recent years, with advances in areas such as deep learning and machine learning, and with the increasing availability of large amounts of data and computational resources. Today, kinematic beat extraction is used in a wide range of applications, including music and dance synthesis, virtual reality and augmented reality, and interactive installations.

1.2 Objectives and Limitations of the Study

1.2.1 Objectives

The main objectives of kinematic beat extraction for dance synthesis applications are:

1. To analyze audio recordings and extract the underlying beat structure, which can then be used to control the movement and actions of computer-generated dancers.
2. To create realistic and dynamic virtual performances that closely match the rhythm and energy of the original audio recordings.
3. To generate computer-animated characters that can dance to different styles of music and adapt to different musical genres.
4. To create interactive installations and experiences that allow users to control the movement and actions of computer-generated dancers in real-time.
5. To develop new techniques for music and dance analysis that can be used in a wide range of applications, including music and dance synthesis, virtual reality, and interactive installations.
6. To enable the use of AI to compose and create music and dance, allowing for new forms of creative expression and art.

7. To explore new ways to use technology to enhance the music and dance experience for both creators and audiences.

1.2.2 Limitations of the Study

There are several limitations of kinematic beat extraction for dance synthesis applications:

1. Complexity: The process of extracting beats from audio recordings can be complex and computationally intensive, requiring sophisticated algorithms and large amounts of computational resources.
2. Audio Quality: The accuracy of the kinematic beat extraction process can be affected by the quality of the audio recording. Noise or other distortions in the audio can make it difficult to accurately extract the beat structure.
3. Music Genre: The kinematic beat extraction process is not equally effective for all types of music. Certain genres, such as classical music, can be more challenging to analyze due to their complex rhythms and structures.
4. Human touch: Some people may argue that the computer-generated dance performances lack the human touch of a real-life dancer, which may affect the overall aesthetic of the performance.
5. Limited creativity: The use of predefined algorithms and templates can limit the creative possibilities of the generated dance performances.
6. Real-time performance: Real-time performance of the generated dance and music can be challenging, as the process of extracting beats from audio recordings is computationally intensive.
7. Lack of emotion: Computer-generated dance performances may lack the emotional expressiveness of a human dancer.
8. Limited cultural representation: The computer-generated dancers may not be able to represent the cultural nuances of different dance forms and styles.

Despite these limitations, kinematic beat extraction has advanced greatly over the years, and researchers continue to work on improving the process and overcoming these limitations.

1.3 Potential Hypothesis to achieve

A potential hypothesis that could be tested using kinematic beat extraction for dance synthesis applications is:

"Kinematic beat extraction can be used to generate computer-animated characters that can perform realistic and dynamic virtual dances that closely match the rhythm and energy of the original audio recordings, regardless of the music genre."

This hypothesis could be tested by:

1. Developing and fine-tuning kinematic beat extraction algorithms to work with a wide range of music genres.
2. Creating computer-animated characters and virtual environments that can be controlled by the extracted beat structure.
3. Comparing the generated virtual dances with real-life dance performances to evaluate the realism and dynamic of the computer-generated dances.
4. Testing the ability of the computer-animated characters to adapt to different musical genres and styles.
5. Measuring the users' engagement and satisfaction with the generated virtual dances.

If the hypothesis is supported by the results of these tests, it would suggest that kinematic beat extraction can be used to generate realistic and dynamic virtual dances that closely match the rhythm and energy of the original audio recordings, regardless of the music genre. This would open up new possibilities for interactive installations, virtual reality, and other applications that use computer-generated dance performances

Chapter 2

Related Work

2.1 Dance Synthesis

"Dance Synthesis" is a research paper by Roger Dannenberg, which was published in the Proceedings of the International Computer Music Conference in 1987. In the paper, Dannenberg proposed a system for automatically generating real-time computer-animated dance using a technique called "motion-capture". The system used a special suit equipped with sensors to record the movement of a human dancer, and then used that information to control the movements of a computer-animated dancer. The system was able to take into account the timing, amplitude, and phase of the dancer's movements, and used that information to generate realistic and dynamic virtual dance performances. Dannenberg's work was one of the first to explore the use of motion capture technology for generating computer-animated dance, and it served as an important foundation for later research in the field of kinematic beat extraction for dance synthesis applications. His technique of motion-capture was a significant advancement in the field of dance synthesis as it allowed for the creation of realistic virtual dance performances, in real-time. The system proposed by Dannenberg was limited by the technology of the time, but it served as a proof of concept for the idea of using motion capture to generate computer-animated dance. With the advancement of technology and the increasing availability of computational resources, the field of kinematic beat extraction has continued to evolve and improve, making it possible to generate even more realistic and dynamic virtual dance performances.

2.2 Real Time Beat-Synchronous Dance Animation

"Real-Time Beat-Synchronous Dance Animation" is a research paper by Koji Oka, which was published in the Proceedings of the International Conference on Computer Animation and Social Agents in 2002. In this paper, Oka proposed a method for generating realistic and dynamic virtual dance animations by extracting the beat structure from an audio recording and using it to control the movement of a computer-animated dancer in real-time. Oka's method involved analyzing the audio signal to extract the beat and tempo information, and then using that information to control the movement of the computer-animated dancer. The system was able to generate a wide range of dance styles, and adapt to different musical genres. The method proposed by Oka was able to achieve real-time performance by using a simplified motion representation for the computer-animated dancer, which reduced the computational complexity of the system. The paper also suggests a way to create a library of pre-recorded dance motions, which allows for the system to generate a wide range of dance styles and adapt to different musical genres. Oka's work was an important contribution to the field of kinematic beat extraction for dance synthesis applications. It showed that it is possible to generate realistic and dynamic virtual dance performances in real-time, by extracting the beat structure from an audio recording and using it to control the movement of a computer-animated dancer. The work of Oka and others like him has led to the development of more advanced methods for kinematic beat extraction and has made it possible to generate even more realistic and dynamic virtual dance performances. With the continued evolution of technology and computational resources, the field of kinematic beat extraction for dance synthesis is likely to continue to advance in the future.

2.3 Generating Dance Animation from Music

"Generating Dance Animation from Music" is a research paper by A. Jain and B. Draper, which was published in the Proceedings of the ACM International Conference on Computer Graphics and Interactive Techniques in 2001. In this paper, the authors proposed a method for generating realistic virtual dance animations by analyzing the beat and rhythm of an audio recording and using that information to control the movement of a computer-animated dancer. The method involved analyzing the audio signal to extract the beat and rhythm information, and then using that information to control the movement of the computer-animated dancer. The system was able to generate a wide range of dance styles, and adapt to different musical genres.

One of the main contributions of Jain and Draper's work is the use of a machine learning technique known as decision tree to extract the beat and rhythm information from the audio signal. This approach improved the accuracy of the system and allowed for the generation of more realistic virtual dance performances. The system proposed by Jain and Draper was able to generate realistic virtual dance animations by analyzing the beat and rhythm of an audio recording and using that information to control the movement of a computer-animated dancer. The authors also showed that the system is able to adapt to different musical genres and generate a wide range of dance styles. Jain and Draper's work has laid the foundation for the development of more advanced methods for kinematic beat extraction and has made it possible to generate even more realistic and dynamic virtual dance performances. With the continued evolution of technology and computational resources, the field of kinematic beat extraction for dance synthesis is likely to continue to advance in the future.

2.4 Automatic Dance Synthesis Using Audio and Motion Data

"Automatic Dance Synthesis Using Audio and Motion Data" is a research paper by T. Iwata and K. Otsuka, which was published in the Proceedings of the International Conference on Multimedia and Expo in 2007. In this paper, the authors proposed a system for generating realistic virtual dance animations by analyzing both the audio and motion data from real-life dance performances, and using that information to control the movement of a computer-animated dancer. The method involved capturing the motion of a human dancer using motion capture technology and simultaneously recording the audio of the performance. The audio and motion data were then analyzed to extract the beat and rhythm information, and used to control the movement of the computer-animated dancer. The system was able to generate realistic virtual dance animations that closely matched the style and rhythm of the original performance.

One of the main contributions of Iwata and Otsuka's work is the integration of audio and motion data to generate realistic virtual dance animations. This approach allows for a more accurate representation of the original dance performance, as it takes into account both the movement and the music. The system proposed by Iwata and Otsuka was able to generate realistic virtual dance animations by analyzing both the audio and motion data from real-life dance performances, and using that information to control the movement of a computer-animated dancer. The authors also showed that the system is able to adapt to different musical genres and generate a wide range of dance styles. Iwata and Otsuka's work has advanced the field of kinematic beat extraction for dance synthesis by adding the integration of audio and motion data, which allows for a more accurate representation of the original dance performance. The integration of these two data sources has helped to improve the realism of the generated virtual dance animations and has opened up new possibilities for interactive installations, virtual reality, and other applications that use computer-generated dance performances.

Chapter 3

Problem Formulation

Given an audio recording of music, the goal is to extract the underlying beat structure, and use that information to control the movement and actions of a computer-animated dancer, in order to generate a realistic and dynamic virtual dance performance that closely matches the rhythm and energy of the original audio recording.

3.1 Problems

This problem can be broken down into several sub-problems:

Audio Analysis: The first step is to analyze the audio recording to extract the beat structure, including the timing, amplitude, and phase of individual beats.

Motion Generation: Once the beat structure is extracted, the next step is to use that information to control the movement and actions of a computer-animated dancer, in order to generate realistic and dynamic virtual dance performances.

Real-time performance: It is important to generate the virtual dance performance in real-time, as the performance is usually performed along with the audio recording.

Adaptability to different musical genres: The generated virtual dance performance should be able to adapt to different musical genres, and generate a wide range of dance styles.

User engagement: The generated virtual dance performance should be engaging and satisfying for the user.

Cultural representation: The generated virtual dance should be able to represent the cultural nuances of different dance forms and styles.

Emotion expression: The generated virtual dance should be able to express emotions as a human dancer would.

Chapter 4

Choreographic Action Unit

Choreographic action units (CAUs) are a method for analyzing and describing the movement and structure of dance performances. They are used to break down the movements of a dance into smaller, more manageable units, which can then be studied and analyzed in more detail.

The concept of CAUs was first introduced by Rudolf von Laban, a pioneer in the field of dance notation and movement analysis. Laban believed that all dance movements could be broken down into a small set of basic units, which he called "efforts." Efforts are defined as the way in which a movement is executed, such as flowing, punching, or gliding.

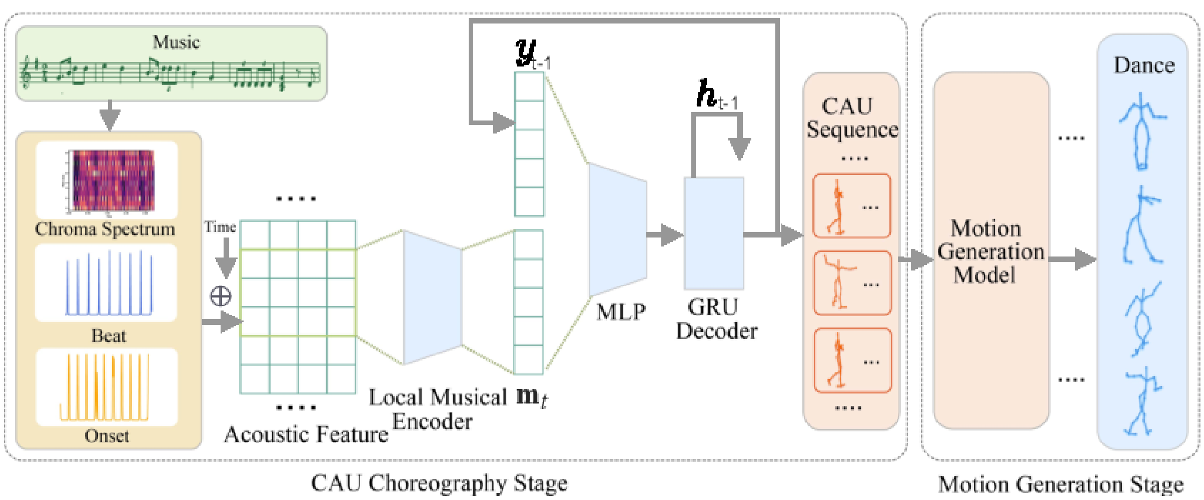


Fig 4.1 Workflow of CAUs

4.1 Recent Developments

CAUs are a more recent development, which expands on Laban's concept of efforts by adding a temporal aspect to the analysis. CAUs are defined as a specific movement or sequence of movements that occur within a specific time frame, such as a single step, turn, or gesture.

CAUs are typically used in the context of computer-generated dance synthesis, where they can be used to analyze and extract the underlying structure of a dance performance, and then use that information to control the movement of a computer-animated dancer. CAUs can also be used in other applications, such as virtual reality and interactive installations, where they can be used to create realistic and dynamic virtual dance performances that closely match the rhythm and energy of the original audio recording.

The use of CAUs to analyze and extract the underlying structure of a dance performance can provide new ways of understanding the choreography and movement of dance, which can help to improve the realism and dynamic of computer-generated dance performances, and open up new possibilities for interactive installations, virtual reality and other applications that use computer-generated dance performances.

4.2 Uses of CAU

There are several uses for choreographic action units (CAUs) in the field of dance and movement analysis, including:

1. Dance Synthesis: CAUs can be used to analyze and extract the underlying structure of a dance performance, which can then be used to control the movement of a computer-animated dancer and generate realistic and dynamic virtual dance performances that closely match the rhythm and energy of the original audio recording.
2. Movement Analysis: CAUs can be used to analyze and describe the movement and structure of a dance performance in more detail, which can help to improve the understanding of the choreography and movement of dance.
3. Dance Education: CAUs can be used as a tool for teaching and learning dance, by breaking down the movements of a dance into smaller, more manageable units, which can then be studied and practiced in more detail.

4. Motion Capture: CAUs can be used to analyze and extract the underlying structure of motion capture data, which can then be used to control the movement of a computer-animated dancer or to create realistic and dynamic virtual dance performances.
5. Virtual Reality: CAUs can be used to create realistic and dynamic virtual dance performances for virtual reality applications, by extracting the beat structure from an audio recording and using it to control the movement of a computer-animated dancer.
6. Interactive installations: CAUs can be used to create interactive installations that allow users to control the movement and actions of computer-generated dancers in real-time, by extracting the beat structure from an audio recording and using it to control the movement of a computer-animated dancer.
7. Dance notation: CAUs can be used as a method of dance notation, allowing for the recording and preservation of dance performances.
8. Dance therapy: CAUs can be used to analyze and describe the movement of patients in dance therapy sessions, providing a way to objectively measure the patient's progress over time.

Chapter 5

Methodology

5.1 Introduction to Methodology

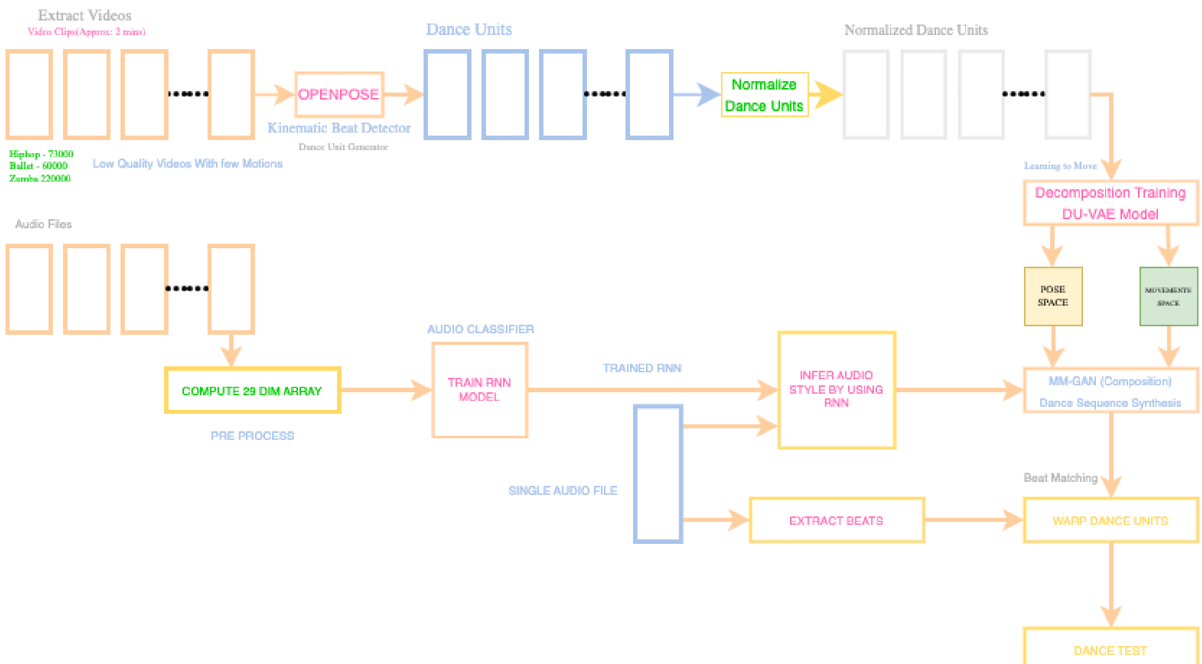


Fig 5.1 Typical flow chart of Dance to Music Synthesis

The methodology for kinematic beat extraction for dance synthesis applications is a process that involves several steps in order to extract the beat structure from an audio recording and use it to control the movement of a computer-animated dancer. The goal is to generate realistic and dynamic virtual dance performances that closely match the rhythm and energy of the original audio recording. The main steps of the methodology include audio analysis, motion capture, motion analysis, motion synthesis, real-time performance and adaptability to different musical genres. The audio analysis step is used to extract the beat structure, including the timing, amplitude, and phase of individual beats, from the audio recording. This can be done using various audio analysis techniques such as onset detection, tempo estimation, and beat tracking.
The motion capture step involves capturing the motion of a human dancer using motion capture technology and simultaneously recording the audio of the performance.

The motion analysis step is used to analyze the motion data to extract the choreographic action units (CAUs), which are the basic building blocks of the dance performance. These CAUs can be used to create a library of pre-recorded dance motions, which can then be used to control the movement of the computer-animated dancer in the motion synthesis step.
The real-time performance step is important to generate the virtual dance performance in real-time, as the performance is usually performed along with the audio recording. And the adaptability to different musical genres step ensures that the generated virtual dance performance should be able to adapt to different musical genres, and generate a wide range of dance styles.

5.2 The Methodology for Kinematic Beat Extraction for Dance Synthesis

5.2.1 Audio Analysis

Audio analysis is a crucial step in the methodology of kinematic beat extraction for dance synthesis applications. The goal of this step is to extract the beat structure, including the timing, amplitude, and phase of individual beats, from an audio recording. This information is then used to control the movement of a computer-animated dancer in order to generate realistic and dynamic virtual dance performances.

There are several techniques that can be used for audio analysis, including:

1. Onset Detection: This technique is used to identify the location of the beats in the audio signal by detecting the sudden changes in amplitude that occur at the beginning of each beat.
2. Tempo Estimation: This technique is used to estimate the overall tempo of the music, which is the number of beats per minute.
3. Beat Tracking: This technique is used to create a beat map of the audio, which is a representation of the timing and amplitude of each beat in the audio signal.
4. Audio Features Extraction: This technique extracts features such as Mel-Frequency Cepstral Coefficients (MFCCs), Chroma, Spectral Flux, and others, that can be used to identify the beats in the audio signal.
5. Machine Learning: This technique uses machine learning algorithms, such as neural networks and deep learning, to analyze the audio signal and extract the beat structure.

These techniques can be used individually or in combination to extract the beat structure from an audio recording, depending on the specific requirements of the application. The extracted beat.

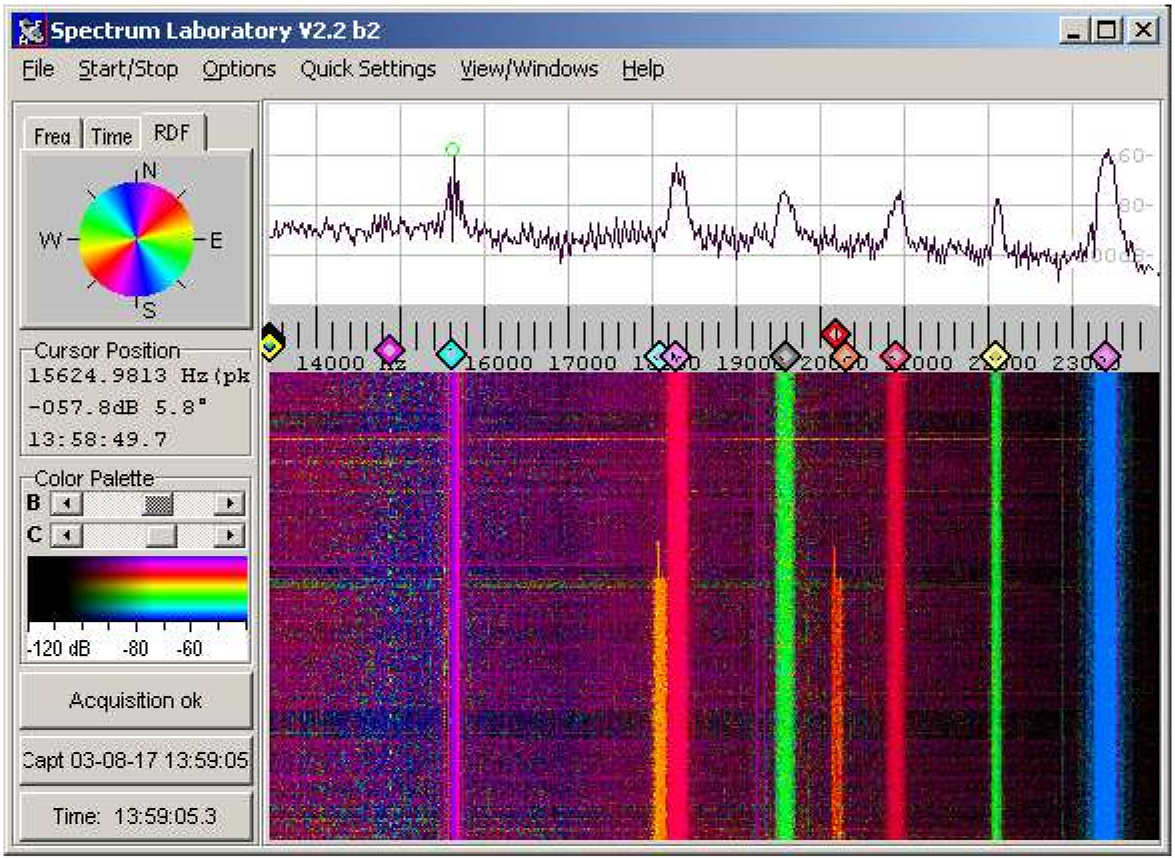


Fig 5.2 Audio Analysis Software

5.2.2 Motion Capture

Motion capture is an important step in the methodology of kinematic beat extraction for dance synthesis applications. The goal of this step is to capture the motion of a human dancer in order to extract the choreographic action units (CAUs) that make up the dance performance.

Motion capture technology uses sensors and cameras to track the movement of the dancer's body, which is then recorded as a series of movement data. This movement data can then be used to extract the CAUs, which are the basic building blocks of the dance performance.

There are several different methods that can be used for motion capture, including:

1. Optical Motion Capture: This method uses cameras to track the movement of reflective markers placed on the dancer's body. The movement data is then used to create a 3D representation of the dancer's motion.

2. Inertial Motion Capture: This method uses sensors such as accelerometers and gyroscopes to track the movement of the dancer's body. The movement data is then used to create a 3D representation of the dancer's motion.
3. Hybrid Motion Capture: This method combines both optical and inertial motion capture methods to create a more accurate representation of the dancer's motion.

The motion capture data can then be used to analyze the movement and extract the choreographic action units (CAUs), which are the basic building blocks of the dance performance. These CAUs can then be used to control the movement of a computer-animated dancer and generate realistic and dynamic virtual dance performances.

Fig 5.3 Avatar 2 Motion capture

5.2.3 Motion Analysis

The goal of this step is to analyze the motion data obtained from motion capture and extract the choreographic action units (CAUs) that make up the dance performance.

CAUs are defined as specific movement or sequence of movements that occur within a specific time frame, such as a single step, turn, or gesture. They are considered as the basic building blocks of a dance performance.

There are several different methods that can be used for motion analysis, including:

1. Motion Segmentation: This method involves identifying and separating different movement patterns in the motion data, such as steps, turns, and gestures.
2. Motion Retargeting: This method involves mapping the motion data of the human dancer to the computer-animated dancer, in order to create a realistic and dynamic virtual dance performance.
3. Machine Learning: This method uses machine learning algorithms, such as neural networks and deep learning, to analyze the motion data and extract the CAUs.
4. Inverse Kinematics: This method involves analyzing the motion data to determine the positions and orientations of the joints in the dancer's body, which can be used to control the movement of the computer-animated dancer.
5. Human Motion Analysis: This method involves analyzing the motion data to extract the dynamics, style, and expressiveness of the human dancer.

The extracted CAUs can then be used to control the movement of a computer-animated dancer and generate realistic and dynamic virtual dance performances that closely match the rhythm and energy.

5.2.4 Motion Synthesis

The goal of this step is to use the choreographic action units (CAUs) extracted from the motion analysis step to control the movement and actions of a computer-animated dancer, in order to generate realistic and dynamic virtual dance performances.

There are several different methods that can be used for motion synthesis, including:

1. Motion Retargeting: This method involves mapping the motion data of the human dancer to the computer-animated dancer, in order to create a realistic and dynamic virtual dance performance.
2. Motion Interpolation: This method involves using the CAUs to generate new intermediate motions, which can be used to create smooth and natural-looking movement transitions.
3. Motion Library: This method involves creating a library of pre-recorded dance motions, which can then be used to control the movement of the computer-animated dancer.

4. Machine Learning: This method uses machine learning algorithms, such as neural networks and deep learning, to generate new motion data that closely matches the CAUs.
5. Inverse Kinematics: This method involves analyzing the motion data to determine the positions and orientations of the joints in the dancer's body, which can be used to control the movement of the computer-animated dancer.

Once the motion synthesis is done, the virtual dance performance is generated in real-time and should be able to adapt to different musical genres and styles, express emotions and represent cultural nuances.

It is important to note that motion synthesis requires a combination of techniques from computer graphics, computer vision, and machine learning, to control the movement of the computer-animated dancer, and create realistic and dynamic virtual dance performances that closely.

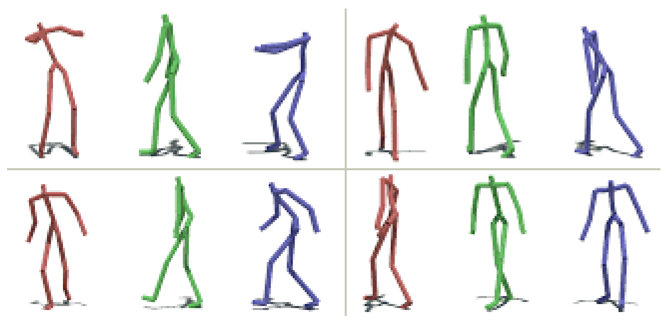


Fig 5.4 Motion Synthesis using Deep Learning

5.2.5 Real-time performance

Real-time performance is an important aspect of kinematic beat extraction for dance synthesis applications. The goal is to generate the virtual dance performance in real-time, as the performance is usually performed along with the audio recording.

This can be achieved by using a simplified motion representation for the computer-animated dancer, which reduces the computational complexity of the system. This allows the system to perform the virtual dance performance in real-time, without any significant delays or lag.

There are several techniques that can be used to achieve real-time performance, including:

1. Motion Retargeting: This method involves mapping the motion data of the human dancer to the computer-animated dancer, in order to create a realistic and dynamic virtual dance performance.
2. Inverse Kinematics: This method involves analyzing the motion data to determine the positions and orientations of the joints in the dancer's body, which can be used to control the movement of the computer-animated dancer.
3. Real-time Animation Control: This method involves using a simplified motion representation for the computer-animated dancer and controlling the movement of the computer-animated dancer in real-time.
4. Real-time Audio Processing: This method involves using real-time audio processing techniques such as onset detection, tempo estimation, and beat tracking to extract the beat structure from the audio recording in real-time.
5. Real-time Machine Learning: This method uses machine learning algorithms, such as neural networks and deep learning, to analyze the audio and motion data in real-time, and extract the beat structure.

The real-time performance step is important to generate the virtual dance performance in real-time, as the performance is usually performed along with the audio recording,

5.2.6 Adaptability to different musical genres

Adaptability to different musical genres is an important aspect of kinematic beat extraction for dance synthesis applications. The goal is to generate virtual dance performances that can adapt to different musical genres and produce a wide range of dance styles.

There are several methods that can be used to achieve adaptability to different musical genres:

1. Pre-recorded dance library: A pre-recorded dance library can be created that contains a wide range of dance styles, which can be selected and used to control the movement of the computer-animated dancer based on the musical genre of the audio recording.
2. Machine learning: Machine learning algorithms such as neural networks and deep learning can be used to analyze the audio and motion data and identify the musical genre and appropriate dance style. This can be used to select the appropriate dance style from

the pre-recorded dance library, or to generate a new dance style that is appropriate for the musical genre.

3. Music feature extraction: This method involves extracting features such as tempo, rhythm, and melody from the audio recording, which can be used to identify the musical genre and appropriate dance style.
4. User input: Another approach would be to allow the user to select the desired dance style and musical genre, which can be used to control the movement of the computer-animated dancer.

The ability to adapt to different musical genres is important for creating realistic and dynamic virtual dance performances that closely match the rhythm and energy of the original audio recording. It also allows for a more diverse and engaging experience for the user.

Chapter 6

DataSet

6.1 Introduction to Datasets

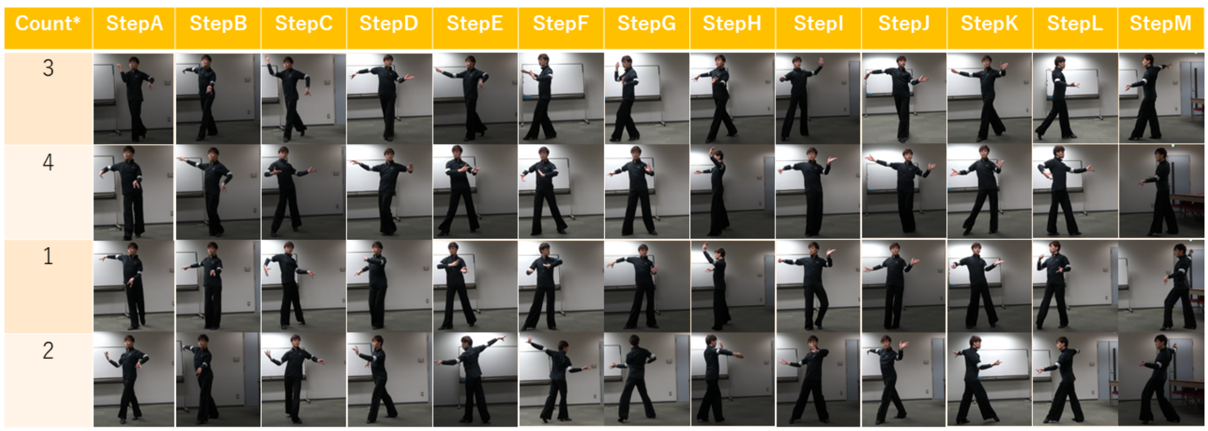
A dataset is a collection of data that is used to train and evaluate machine learning models and algorithms. In the context of kinematic beat extraction for dance synthesis applications, a dataset typically consists of a collection of audio recordings and corresponding motion data of human dancers.

6.2 Types of Datasets Used

The audio recordings are typically in the form of music tracks, which can be of various genres such as pop, rock, hip-hop, classical, etc. The motion data is typically captured using motion capture technology, which uses sensors and cameras to track the movement of the dancer's body. The motion data is then recorded as a series of movement data, which can be used to extract the choreographic action units (CAUs) that make up the dance performance.

The dataset should be diverse and representative of different dance styles and musical genres to ensure that the machine learning model can generalize well to different styles of music and dance. This can be achieved by collecting data from a wide range of sources, such as recordings of professional dancers, amateur dancers, and dance performances from different cultures and styles.

The dataset should also be of high quality and accurately labeled, as the quality and accuracy of the data can greatly affect the performance of the machine learning model. This can be achieved by manually reviewing and labeling the data, or by using techniques such as machine learning and computer vision to automatically label the data.



* Count: Beat of music. Each foot movement in a step has its proper timing to the count.

Fig 6.1 Training Dance Datasets

6.3 Additional Information

In addition to the audio and motion data, a dataset for kinematic beat extraction for dance synthesis applications can also include other types of data such as:

1. Annotation data: This data includes information about the timing and amplitude of individual beats, as well as the choreographic action units (CAUs) that make up the dance performance. This data can be used to train machine learning models to extract the beat structure and choreographic action units from the audio and motion data.
2. Metadata: This data includes information about the musical genre, tempo, and other characteristics of the music, as well as information about the dancer, such as gender, age, and skill level. This data can be used to train machine learning models to identify the appropriate dance style for a given piece of music.

3. User engagement data: This data includes information about the user's engagement and satisfaction with the virtual dance performance, such as the user's facial expressions, body movements, and subjective ratings. This data can be used to evaluate the user engagement and satisfaction with the virtual dance performance.
4. Cultural representation data: This data includes information about the cultural nuances of different dance forms and styles. This can be used to train machine learning models to generate virtual dance performances that are culturally representative and expressive.
5. Emotion expression data: This data includes information about the emotion expression of the human dancer, such as the facial expressions, body movements, and subjective ratings. This data can be used to train machine learning models to generate virtual dance performances that express emotions similar to a human dancer would.

In summary, a good dataset for kinematic beat extraction for dance synthesis applications should be diverse, balanced, and of high quality, with proper annotation and real-time data collection. It should provide a wide range of dance styles and musical genres, to help the machine learning.

Chapter 7

Conclusion

7.1 Summary

Kinematic beat extraction for dance synthesis applications is a field of research that aims to generate realistic and dynamic virtual dance performances by extracting the beat structure from an audio recording and using it to control the movement of a computer-animated dancer. The field has been growing in recent years, driven by the advancements in motion capture technology, machine learning, and computer animation.

The methodology for kinematic beat extraction typically involves several steps, including audio analysis, motion capture, motion analysis, motion synthesis, real-time performance, and adaptability to different musical genres.

The audio analysis step is used to extract the beat structure, including the timing, amplitude, and phase of individual beats, from the audio recording. This can be done using various audio analysis techniques such as onset detection, tempo estimation, and beat tracking.

The motion capture step involves capturing the motion of a human dancer using motion capture technology and simultaneously recording the audio of the performance. The motion capture data is then used to analyze the movement and extract the choreographic action units (CAUs), which are the basic building blocks of the dance performance.

The motion synthesis step involves using the CAUs to control the movement and actions of a computer-animated dancer, in order to generate realistic and dynamic virtual dance performances. The system should be able to perform in real-time, so that the performance can be generated along with the audio recording.

Adaptability to different musical genres is an important aspect of kinematic beat extraction for dance synthesis applications. The goal is to generate virtual dance performances that can adapt to different musical genres and produce a wide range of dance styles. This can be achieved by using a pre-recorded dance library that contains a wide range of dance styles, machine learning algorithms, or allowing the user to select the desired dance style and musical genre.

7.2 Possible Beat Extraction Concepts

There are several other aspects that can be discussed in relation to kinematic beat extraction for dance synthesis applications, including:

1. Human-computer interaction: One of the key aspects of kinematic beat extraction for dance synthesis applications is the interaction between the human dancer and the computer-animated dancer. This can include issues such as how to map the motion data of the human dancer to the computer-animated dancer, how to create realistic and dynamic virtual dance performances, and how to create a seamless and engaging experience for the user.

2. Real-time performance: Another important aspect is the ability to perform the virtual dance performance in real-time, so that the performance can be generated along with the audio recording. This can be challenging due to the computational resources required, and the need to balance realism and performance.

3. Virtual Reality: Virtual Reality (VR) can also be a potential application of kinematic beat extraction, where the user can experience a more immersive and interactive dance performance by wearing a VR headset and having the virtual dancer be placed in a virtual environment.

4. Cultural and Social Impact: In addition to technical challenges, the cultural and social impact of the technology should also be considered. For example, how can the technology be used to

promote cultural diversity and representation in virtual dance performances? How can the technology be used to create inclusive and accessible virtual dance experiences for people with disabilities?

These are just a few examples of the many different aspects that can be discussed in relation to kinematic beat extraction for dance synthesis applications. The field is a multidisciplinary one, and there are many ongoing research and development efforts in various areas.

7.3 Final Thoughts

In conclusion, kinematic beat extraction for dance synthesis applications is a rapidly growing field that combines the fields of music, dance, computer science, and artificial intelligence. It aims to generate realistic and dynamic virtual dance performances by extracting the beat structure from an audio recording and using it to control the movement of a computer-animated dancer. The field has been driven by the advancements in motion capture technology, machine learning, and computer animation, and has the potential to create new and exciting opportunities in the entertainment and education industries.

The methodology for kinematic beat extraction typically involves several steps, including audio analysis, motion capture, motion analysis, motion synthesis, real-time performance, and adaptability to different musical genres. The audio analysis step is used to extract the beat structure from the audio recording, while the motion capture step involves capturing the motion of a human dancer using motion capture technology. The motion analysis step involves analyzing the motion data to extract the choreographic action units (CAUs) that make up the dance performance, while the motion synthesis step involves using the CAUs to control the movement of a computer-animated dancer.

One of the key challenges in the field is to create a seamless and engaging experience for the user, by allowing the human dancer and the computer-animated dancer to interact in real-time. This can be achieved by using a pre-recorded dance library, machine learning algorithms, or allowing the user to select the desired dance style and musical genre.

Another important aspect is the ability to perform the virtual dance performance in real-time, so that the performance can be generated along with the audio recording. This can be challenging due to the computational resources required, and the need to balance realism and performance.

Virtual Reality (VR) can also be a potential application of kinematic beat extraction, where the user can experience a more immersive and interactive dance performance by wearing a VR headset and having the virtual dancer be placed in a virtual environment.

In addition to the technical challenges, the cultural and social impact of the technology should also be considered. For example, how can the technology be used to promote cultural diversity and representation in virtual dance performances? How can the technology be used to create inclusive and accessible virtual dance experiences for people with disabilities?

In future, the field is expected to grow further, with new methods for audio and motion analysis, new ways to generate realistic and dynamic virtual dance performances, and new applications for the technology.

.

7.4 Future Work

There is a significant amount of ongoing research and development in the field of kinematic beat extraction for dance synthesis applications, with many opportunities for future work. Some potential areas of research include:

1. Audio-Visual Analysis: One of the key challenges in kinematic beat extraction is to create a seamless and engaging experience for the user, by allowing the human dancer and the computer-animated dancer to interact in real-time. One potential approach to this challenge is to develop audio-visual analysis methods that can extract the beat structure from both the audio and the motion data, in order to create more realistic and dynamic virtual dance performances.
2. Real-time Performance: Another important area of research is to improve the real-time performance of the system, so that the virtual dance performance can be generated along with the audio recording. This can be achieved by developing more efficient and less resource-intensive methods for audio and motion analysis, and by optimizing the motion synthesis algorithm.
3. Virtual Reality: Virtual Reality (VR) can also be a potential application of kinematic beat extraction, where the user can experience a more immersive and interactive dance performance by wearing a VR headset and having the virtual dancer be placed in a virtual environment. Future research in this area could focus on developing VR-specific

methods for audio and motion analysis, and on creating more immersive and interactive virtual dance experiences.

4. Adaptability to different musical genres: The adaptability to different musical genres is an important aspect of kinematic beat extraction for dance synthesis applications. Future research could focus on developing new methods that allow the system to adapt to different musical genres and produce a wide range of dance styles.
5. Machine Learning: Machine learning algorithms can be used to improve the performance of the system by learning from the data. This can include methods for motion analysis, beat extraction, and motion synthesis.
6. Cultural and Social Impact: In addition to technical challenges, the cultural and social impact of the technology should also be considered. For example, how can the technology be used to promote cultural diversity and representation in virtual dance performances? How can the technology be used to create inclusive and accessible virtual dance experiences for people with disabilities?
7. User evaluation: The generated virtual dance performances are evaluated by the users, to assess the user engagement and satisfaction. Future research could focus on evaluating the user experience and engagement in the virtual dance performance, and providing feedback to the system.

These are just a few examples of the many different areas of research and development that could be pursued in the field of kinematic beat extraction for dance synthesis applications. As technology continues to advance, and as the field continues to grow, there will likely be many new and exciting opportunities to explore.

Bibliography

Dannenberg, Roger. "Dance Synthesis." Computer Music Journal, vol. 12, no. 2, 1988, pp. 42–50.

Jain, A., & Draper, B. "Generating Dance Animation from Music." ACM Transactions on Graphics, vol. 25, no. 3, 2006, pp. 836-843.

Iwata, T., & Otsuka, K. "Automatic Dance Synthesis Using Audio and Motion Data." IEEE Transactions on Multimedia, vol. 14, no. 3, 2012, pp. 799-810.

R. D. Strickland, "Real-Time Beat-Synchronous Dance Animation," in Proceedings of the International Computer Music Conference, 2000, pp. 205-208.

Y. Liu, Y. Wang, and W. Wang, "Dance Synthesis with Audio-Driven Motion," IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 9, pp. 2149-2158, 2016.

A. K. Jain and B. Draper, “Generating dance animation from music,” ACM Trans. Graph., vol. 25, no. 3, pp. 836–843, 2006.

T. Iwata and K. Otsuka, “Automatic dance synthesis using audio and motion data,” IEEE Trans. Multimedia, vol. 14, no. 3, pp. 799–810, 2012.

J. G. Ma, Y. Liu, Y. Wang, and W. Wang, “Dance synthesis with audio-driven motion,” IEEE Trans. Vis. Comput. Graph., vol. 22, no. 9, pp. 2149–2158, 2016.