Music Driven Walk Cycle For Humanoid Character

Aditi Patil (asp270)
Gunjan Singh (gs896)
Karan Pardasani (kp955)

Research Mentor- Jerry Chang

1. Interactive animation of personalized human locomotion

Approach: A personalised walkcycle can be created by controling parameters such as pace, stride length, degrees of freedom of joints. For each locomotion step, they calculate three step constraints from the high-level specification of parameters and attributes: the duration for the leg phases, the leg angles at the end of each step, and the control points which determine the movement of the hip of the stance leg during a step. Then the intermediate values from the previous step to the current step are interpolated.
Benefits: Compared to keyframing, a higher level of control is provided by specifying movements through parameters rather than joint-angles.
Limitations: Everything has to be set manually. Music signal or beats are not used anywhere, as stated in our proposal.

2. Motion Signal Processing

Approach: Motion blending is explored in this paper, along with timewarping algorithm to align the motions. Different joints of humanoid character are taken under consideration to create motion signal, and then by modifying the signal, or mixing two signals, a new motion is created. For example, motion signals of marching and drunk walk cycles were blended to create a new walk cycle. The authors have used spline curves to modify the parameters of the resultant motion signal.
Benefits: Two different motions are blended and animators can control the degrees to which the motions are blended. Modifying the spline curve to a large extent can effectively make the motions unnatural, and reducing the degree results in smooth motions.
Limitations: The dynamics of the new motion might violate the laws of physics. For example, in the newly created walk cycle, the foot might go into the ground if parameters are not chosen correctly.

3. Music-driven motion editing: local motion transformations guided by music analysis

Approach: This paper provides a general framework for synchronising motion curves to music in computer animation. The features extracted from MIDI and analogue audio signals are then visually translated, not by creating entirely new motion curves, but by editing existing motion curves such as keyframe animation and motion capture.
Benefits: Beats are mapped with motion signal, and resultant motion signal is only edited, and not re-created.
Limitations: The modified signal is not converted into an animation. If the motion is not cyclic, the beats to motion mapping is skewed.

4. Music to Motion: Using Music Information to Create Expressive Robot Motion:

Approach: This paper focuses on the developing high-quality motions for a humanoid robot intended for a robot theater. The authors developed a program to exploit melodic and dynamics information from music to create expressive robot motion. The system takes a sound file as its input, and outputs a set of movement instructions for the robot. The system focuses on using the timing and volume of the pitches (notes) that

constitute the prominent melody in a music piece. To simplify the extraction process, the authors used music encoded in MIDI format where the properties of the musical piece are laid out in plain text and thus easier to extract its dynamic information. The extracted information is then used to produce motion data on a robot using a set of production rules.
Benefits: The proposed system was able to capture some of the expressive information in terms of timing, note-on velocity, note duration and pitch, and translate these information into motion properties of: time of execution, range of motion, and acceleration.
Limitations: The extraction and conversion of music information to motion was generated offline, not in real time. There are several MIDI control events that are not being evaluated in the proposed method. They did not use forward or inverse kinematics to use the melodic information on the robot's paths instead of joint angles.

5. Music-Driven Animation Generation of Expressive Musical Gestures:

Approach: The paper proposes a music-driven piano performance generation method using 3D motion capture data and recurrent neural networks. The proposed method uses a long short-term memory (LSTM) network to generate the musical gesture animations.They found Mel-frequency cepstral coefficients (MFCCs) to be an effective feature for music-driven motion Generation. They used Librosa to extract the locations of the beat onsets in the song. The pre-processed data contains the XYZ-locations of 14 joints, resulting in 42 values for every frame and feature extraction using MFCCs. The network is trained on subsequent slices of 100 timesteps or frames, which equals to one second of audio or animation. Evaluation of the results were done using the following metrics: MSE of the validation set, Average Positioning Error (APE), the acceleration and the jerk.
Benefits: From the results they were able to prove that by including the pitch, dynamics and rhythm information additional to MFCCs helps the network to produce more natural looking animations.
Limitations: They did not expand the dataset to contain more different playing styles and instruments. They did not investigate other deep learning approaches such as GANs.

6. Using Music and Motion Analysis to Construct 3D Animations and Visualizations:

Approach: The paper presents a study into music analysis, motion analysis and the integration of music and motion to form natural human motion in a virtual environment. Their system consists of three distinct modules: Music analysis, Motion analysis and Motion synthesis. Their approach to extract beats was to apply a low-pass filter to the audio data. They also used Waveform Similarity based Overlap-Add (WSOLA) algorithm to overcome the problem of connecting the waves naturally without gaps and keeping their pitches. Then they constructed motion graphs to map with the extracted beats information.
Benefits: They found that almost all beats were detected by their beat-detected approach. From the results of experiments, the synthesized motion matched the beats well.
Limitations: The error occurred in the step of finding local maximum because there are more than one local maximum in a period of waveform. They found that the performance in the synthesis was greatly affected by the segment which was generated from the interpolation of the transition of motion. Finding the transition between clips was a hard problem for the authors.

7. Follow the Music: Dance Motion Synthesis Corresponding to Arbitrary Music

Approach: The author first generates a specified professional dance sequence using Style Module. And then it aligns the rhythm and beats from both music and dance sequences to create the final dance video.The author uses an encoder to create a compressed representation of the specific style of continuous dance frames and it gives the encoded representation to LSTM. The LSTM outputs the same style of dance sequences. Then the sequences pass through the Mixture Density Layer and then they are decoded to get the similar dance styles.
Benefits: This model can easily synthesize different styles of dance motion for arbitrary music which no state-of-the-art method can achieve. To do this, the model just needs to train the style module on different

styles of dance motion respectively and use different trained weights. This method also has a great generalization as it has no limitation for the input music since the second step is to warp the pose sequence to match the audio. This means the input music won't affect the quality of the result.
Limitations: The motion poses persistent jitter during moving. The dance poses will still converge to fix poses. Also, did not handle the sudden changes of pose in the videos yet.

8. MAPPING MOTION TO SOUND AND MUSIC IN COMPUTER ANIMATION

Approach: This paper builds a system that focuses on mapping the motion to the sound and generates music that closely relates to the motion. In this paper the main focus was the step in which the author is mapping the motion to sound and music. The system uses keyframing techniques to get the position and orientation data and associated values like joint angles.The sound parameters that the system uses include amplitude, pitch, delay, reverberation, and pitch bend (Doppler shift), are all included. The author uses Timber trees, as hierarchical structures where the nodes represent constant values, variable parameters, mathematical functions or sub trees. To search for the solution in design space, the author uses a genetic approach which is a parametric approach. These parameters are controlled by the user and the sound s generated according to the user requirements.
Benefits: Using this system, the musical soundtracks can be composed and generated in real time.The filtering techniques can be customized to get more diverse environments of motion which are closely related to the music.
Limitations: The system does not have a way to decide which parameters are more appropriate, as a result of which the system has to describe the entire Timbre tree for the sound system.

9. Rhythmic-Motion Synthesis Based on Motion-Beat Analysis

Approach: Given a rhythmic sound signal such as music together with the direction and speed to move, our objective is to generate a new rhythmic motion on the fly. For motion synthesis, the author prepares a library of movement transition graphs each of which represents a collection of rhythmic motions of an identical type. From the library, a movement transition graph is chosen so that the type of motions represented by the graph matches that of the input sound signal. The algorithm traverses the movement transition graph from node to node, guided by the transition probabilities until the sound signal ends, while synthesizing a basic movement at each node. By traversing this graph of movement transition, the author is generating a motion. The author employs the on-line motion blending scheme to generate a basic movement at a node.
Benefits: Using the algorithm described in this paper, the system can generate an on-line and real time motion synthesizing system that can be used in Interactive choreography. In addition to this, The system was also applied to generate rhythmic motions for dance couples in a ball room. They determined the behaviors of the couples and the interaction forces required for the couples to not collide with each other.
Limitations: This scheme cannot handle motion without a priori knowledge on the motion type. Also, the system cannot reflect the continuity between motion phrases, which consist of multiple basic movements.

10. Synthesis of Dancing Character Motion from Beatboxing Sounds

Approach: The algorithm extracts rhythm features from the hand position vector of the dancing character. The algorithm extracts the number of big movements per second and the speed of the hand movement. The beatboxing sounds are separated into three components: 1. Bass drum 2. Snare drum and high hat. These are extracted from the frequency spectrum of sound using FFT. Also, they analyze the number of onset times per second and the average power of the frequency of the drum component sounds in a few seconds Bp. The system uses the rhythm features of beatboxing sounds and sends a query to find an appropriate dancing motion from the motion clip database. The rhythm features extracted from the analysis of beatboxing sounds are matched with the features of motion clips extracted from the onset times and the average volume.
Benefits: The system synthesizes dance motion in real time. We can use the algorithm of this system to extract the beats from the music and use it to generate walk cycle

Limitations: The system described in the paper does not support complex sounds such as disk jockey's scratching sound. The dance motion does not react to the background music.

Research papers

1. BRUDERLIN, A., AND CALVERT, T. Interactive animation of personalized human locomotion. In Graphics Interface '93, Proceedings (May 1993), pp. 17–23.

2. Bruderlin, A. and Williams, L., 1995, September. Motion signal processing. In Proceedings of the 22nd annual conference on Computer graphics and interactive techniques (pp. 97-104).

3. M. Cardle, L. Barthe, S. Brooks and P. Robinson, "Music-driven motion editing: local motion transformations guided by music analysis," Proceedings 20th Eurographics UK Conference, 2002, pp. 38-44, doi: 10.1109/EGUK.2002.1011270.

4. Sunardi, M., Perkowski, M. Music to Motion: Using Music Information to Create Expressive Robot Motion. Int J of Soc Robotics 10, 43–63 (2018). https://doi.org/10.1007/s12369-017-0432-9

5. Alysha Bogaers, Zerrin Yumak, and Anja Volk. 2020. Music-Driven Animation Generation of Expressive Musical Gestures. In Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion). Association for Computing Machinery, New York, NY, USA, 22–26. DOI:https://doi.org/10.1145/3395035.3425244

6. Kuen-Meau Chen, Siu-Tsen Shen & Stephen D. Prior (2008) Using music and motion analysis to construct 3D animations and visualisations, Digital Creativity, 19:2, 91-104, DOI: 10.1080/14626260802037403

7. Mishra, Suneil and James K. Hahn. "MAPPING MOTION TO SOUND AND MUSIC IN COMPUTER ANIMATION AND VE." (2007).

8. Ye, Zijie & Wu, Haozhe & Jia, Jia & Bu, Yaohua & Chen, Wei & Meng, Fanbo & Wang, Yanfeng. (2020). ChoreoNet: Towards Music to Dance Synthesis with Choreographic Action Unit. 744-752. 10.1145/3394171.3414005.

9. Tae-hoon Kim, Sang Il Park, and Sung Yong Shin. 2003. Rhythmic-motion synthesis based on motion-beat analysis. ACM Trans. Graph. 22, 3 (July 2003), 392–401. DOI:https://doi.org/10.1145/882262.882283

10. Kang KK., Kim D. (2007) Synthesis of Dancing Character Motion from Beatboxing Sounds. In: Butz A., Fisher B., Krüger A., Olivier P., Owada S. (eds) Smart Graphics. SG 2007. Lecture Notes in Computer Science, vol 4569. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73214-3_23