A MOBILE APPLICATION FOR CREATING DANCE CHOREOGRAPHY ACCORDING TO MUSICALITY OF INPUTTED AUDIO USING MACHINE LEARNING

Keren Brown[1], David Wei[2]

[1]Woodbridge High School, 2 Meadowbrook, Irvine, CA 92604
[2]Computer Science Department, California State Polytechnic University, Pomona, CA 91768

ABSTRACT

Even the best writers in history were not blessed enough to have a constant surge of inspiration, a never-ending flow of ink on paper, or fingers flying across keyboards. These blocks in creativity are commonly known as writer's block, and many people experience this, since communication is important in every subject. Less commonly known is dancer's block, which is when a choreographer experiences a stop in inspiration while choreographing [1]. These blocks can continue for days or even weeks, before an idea strikes the dancer, setting back the efficiency of choreographing. When due dates come into play, such as choreographing for a production, show, or assignment, a dancer cannot put forth their best choreography, and will be left feeling unsatisfied with the quality of their work. When a choreographer faces a challenge towards creating original choreography, how can one gain inspiration to overcome this block [2]? How can we ensure that the inspiration given is appropriate for the song and style which the choreographer is designing the dance to? This paper develops an application to choreograph visuals in a creative manner, while assessing the musicality of the audio in order to reflect the same emotion in the movement. We applied our application to a jazz piece to be performed as part of a local high school dance show and conducted a qualitative evaluation of the approach. The results show that with the software application, dancers will be able to find inspiration to continue choreographing, pushing them past a barrier of creativity, and allowing them to finish their dance with a quality of choreography that they can be proud to present [3].

KEYWORDS

Dance, Choreography, Machine Learning, Audio Analysis

1. INTRODUCTION

In the art form of dance, movement can be restricted by physical barriers. A human can only bend their back so far before it becomes impossible, we can only move so fast before our muscles reach their limit. These physical obstacles prevent dancers from creating original movement, and dance moves are often re-used. Instead, choreography is made original through the different ways these moves are stringed together, creating interesting variations of previous choreographers' works [4]. Even though the moves in isolation are often not new, a completely different dance can be made through different timing, little adjustments such as a head looking in the opposite

David C. Wyld et al. (Eds): SIPM, ITCA, CMIT, FCST, CoNeCo, SAIM, ICITE, ACSIT, SNLP - 2023
pp. 01-09, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.130901

direction, or changing a pointed toe to a flexed foot. These alterations are utilized to create inventive and original movement, and string the dance moves together to create something visually interesting. What makes new choreography exciting and interesting are not the moves themself, which are commonly re-used, but rather the order in which the moves are presented. In addition to choreographers mainly being restricted to changing the order of moves to create new choreography, they also have to debate which skills they want to include in their piece [5]. It is a common struggle for dancers to decide which combination of moves will be simultaneously aesthetic and newfound, and to constantly access creative ways of thought. Additionally, dance is created to evoke feelings within the audience, and connection with the viewers, and these emotions should reflect that of the music. Choreographing moves that display and narrate the same story as the music poses as yet another challenge. Dancer's facing a block in their creative flow is relevant because dance is an artistic form that can be used to express important concepts, share cultures, evaluate society, and can be used as a stress-relieving activity for the dancers, the choreographers, and the viewers [6].

A common strategy that artists use to overcome this deterrent is to watch videos of choreography that has already been completed, typically within the same category of style. After watching multiple videos, moves or visual effects which leave a strong impression behind will be altered and integrated into their own choreography. An issue which exists in this method is that it is extremely time consuming, since the choreographer must watch multiple videos of dancers. Additionally, inspiration might not be guaranteed if the dancer cannot find a piece of choreography that they feel depicts the same story or evokes the same emotion that they want in their production. Dancers are limited to categorizing the style of choreography that they want with descriptive words such as the style, which still has a broad range. For example, the lyrical style of dance is most commonly either happy and uplifting, or moody and sorrowful, which are complete opposites. A dancer cannot search for results that are tailored to the song which they are choreographing to, since it would not be different enough from existing choreography to be considered a new dance, or their song has not been previously choreographed to. When searching for methods or tools to get past a creative block, recommendations are to take a walk to rest and regain motivation, split the task into smaller groups to make the project less intimidating, and to seek inspiration. There are very few tools to combat this drought of creativity, and these methods require time that choreographers do not always have the privilege of. By approaching choreography with intense thought rather than feeling, the result can look unnatural and forced, connections with other dancers can be ignored, and the rhythm of the dance can look too stagnant or overly variant. Recommended methods for when a dancer is short on time is to identify what might be a distraction in nearby surroundings and to remove the distraction, but this becomes an issue when the distraction is a feeling, such as being overwhelmed or stressed. It is unrealistic to remove every distracting factor in hopes to gain more creativity, and simply ignoring discouraging thoughts will not provide more motivation and energy to overcome a creative block.
In this paper, we follow the same line of research by creating an application that creates choreography to a specific song, allowing a dancer to generate choreography that is tailored to the emotions in the song, and addressing the issue of dancers being limited to categorizing their style of choreography through descriptions, since music is mostly not described unanimously [7]. Additionally, it saves dancers the time that would be used to research different videos of previously choreographed pieces, and it allows them to get immediate inspiration. My tool allows the user to upload a music file, either in mp3 format or wav format, and analyzes the music to generate choreography specifically customized for that specific song [8]. Benefits of my application are that it uses a similar approach to what a human might do, but it is more accurate in the aspect of choreographing in an appropriate style, because it analyzes and matches the users song to similar songs that are already matched with choreography, rather than simply searching by style of dance, which is not always accurate. It also takes inspiration from previously choreographed dances, and strings together these pieces of choreography to create something

original. My method requires less effort from the choreographer, and is both faster at providing inspiration, while also providing choreography that is more accurately represented by the music. In two application scenarios, we demonstrate how the above combination of techniques increases the efficiency of choreographing, both through the decreased amount of time that it takes to be created, and through the quality of the choreography which is created.

First, we show the usefulness of our approach by a case study on a jazz piece for my high school dance team's bi-annual performance. Given different cuts of music, the machine learning application successfully choreographed dance moves which were then modified to add visual layering and transitions between the different cuts of music.

The remaining portions of this paper are organized according to the following format: Section 2 describes the challenges that were faced throughout the process of creating this application, including design and experimentation; Adding to the theme of obstacles, Section 3 constructs solutions that were created in response to the obstacles detailed in Section 2; Moving forward, Section 4 displays experimental details and the evaluation of such details; Section 5 recounts the works which have relation to my topic; To finalize findings through experimentation, Section 6 provides a conclusion and expresses where this project will continue to expand in the future.

2. CHALLENGES

When experimenting with machine learning and artificial intelligence, I predicted a few challenges would arise. Throughout the process there were many questions that needed to be answered, and with limited research completed in the newer artificial intelligence field, extensive research and experimentation was necessary to find satisfactory solutions.Below are elaborations on challenge that were faced, and how they were overcome.

The first challenge was figuring out how to analyze music and compare the songs to find the most similar songs to the inputted audio in the database. In order to choreograph a dance, the moves must match the music, whether it is hitting a sharp and distinct move on a beat or snare in the music, or melting and snaking through movements on a slower and or smooth part of the music. To address the moves matching the music, the librosa library was used to compare individual notes. The similarity levels between notes was used to determine the similarity between music, and by ranking the songs in similarity.

3. SOLUTION

Smart Dance Generator is a machine learning system which utilizes the librosa library to analyze and compare audio, and mediapipe to get landmarks and track movement [9]. Attributes associated with the similarity and differences between audio files include the cosine distance between notes, and also includes the speed of the song through the beats per minute scale. The application is connected to an online database in which pre-choreographed dances are located, with adhering music, in videos. The application takes in an audio clip or a video clip, and extracts just the audio if the format is a video clip such as mp4. The extracted audio is then compared to the extracted audio from the videos in the database, and the three most similar audios from the database are selected [10]. The choreography that is matched with these three audios are then transformed into stick figure images stringed together, creating a stop-motion video. These three stick-figure videos are then cut into smaller sections and concatenated with each other, before being combined with the user's selection of audio. Finally, this result is returned to the user's screen as a stop-motion stick figure video providing choreography that matches the audio provided by the user.

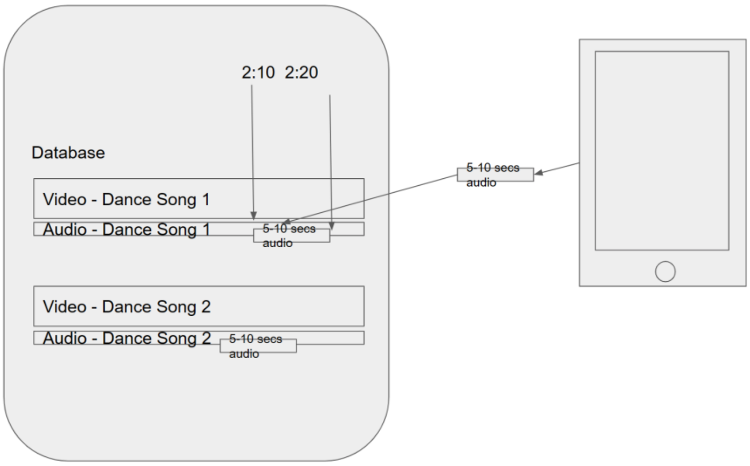

Figure 1. Overview of the solution

The audio file is separated from its suffix file type so that it can be converted to a wav file, and utilizing feature, segment, and display from the librosa library, the cosine distance between notes in the user inputted audio and the notes from the pre-loaded songs in the database are determined [15].

```
def get_matched_song(audio_reference_path, audio_folder, video_range=3):
  cosine_distance = {}
  y_ref, sr = librosa.load(audio_reference_path)
  for song_path in sorted(glob.glob(f"{audio_folder}/*.wav")):  #change mp3 to wav
    y_comp, sr = librosa.load(song_path)
    dm = get_distance_matrix(y_ref, y_comp, sr)
    cosine_distance[song_path] = calculate_average(dm)
    # print(song_path, calculate_average(dm))
  cosine_distance = {k: v for k,v in sorted(cosine_distance.items(), key=lambda item: item[1], reverse=True)}
  return dict(list(cosine_distance.items())[:video_range])
```

Figure 2. Screenshot of code 1

This distance is used to compare songs, by determining songs as more similar the closer the distance between notes are, and ranking songs as less similar as the distance between notes increases. To continue, I created a method that finds different landmarks on an image, in this case landmarks being different points on the human body, for example an elbow. A following method is used to string together the appropriate landmarks with a visual line, since each landmark is assigned to a point on the body, for example the elbow should be connected to the shoulder, but should not be directly connected to the ankle.

```python
def get_landmark(image, landmark):
  landmarks = []
  if landmark :
    for id,lm in enumerate(landmark):
      height,width,c = image.shape
      cx,cy = int(lm.x*width), int(lm.y*height)
      landmarks.append([id,cx,cy,lm.z,lm.visibility])
  return landmarks

def get_position(image, results, id):
  positions = {}
  positions["id"] = id
  positions["pose"] = get_landmark(image,results.pose_landmarks.landmark)
  return positions

def extract_pose_keypoints(video_path):
  id=0
  cap = cv2.VideoCapture(video_path)
  fps = cap.get(cv2.CAP_PROP_FPS)

  #set mediapipe model
  landmark_list = []
  dim = (640,360)
```

Figure 3. Screenshot of code 2

Another method was created to string together different landmark images, therefore creating a video composed of separate images with points along the human body, displaying movement in a stop-motion video of stick figures. This process is used to transform a video of dance choreography to be displayed on a stick figure rather than a human body. This stick figure video is then pieced together with other similar videos to create new choreography.

4. EXPERIMENT

4.1. Experiment 1: Pose Estimate Analysis between MediaPipe and YOLOv7

The COCO dataset, which contains 17 landmark topologies, is used to train Pytorch's YOLOv7 pose, a single-stage multi-person keypoint detector. In addition, segmentation is not directly connected with posture; it supports both CPU and GPU. It differs from MediaPipe, a framework that is only capable of detecting one person. Only CPUs are supported by MediaPipe, and segmentation is built in.

In this experiment, we compare the performance of the FPS on fixed model input size for record inference when just one person is depicted on it using both posture models in the CPU environment. First, since the default image size is 960x960, we updated the YOLOv7 code to forward pass photos that have been scaled to 256x256. Then, for each frame of the captured video, the person's stance was retrieved. According to the findings, MediaPipe outperforms YOLOv7 in CPU inference. While YOLOv7 processes on average 8.1 FPS, the MediaPipe can process the forward pass at a rate of 29.2 FPS, omitting pre- and post-processing time.

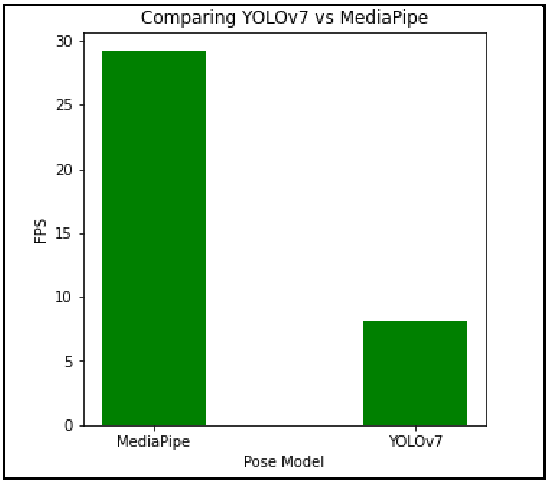

Figure 4. Comparing YOLOv7 and MediaPipe on Fixed Input Size for Record Inference

4.2. Experiment 2: Analysis of the Application Effectiveness

By performing an experiment with 11 participants—a sufficient sample size to allow for any variability—the application is tested for use and convenience. The participants would download the smart dance generator application from the Google Play Store and test out its capabilities, which included uploading some, displaying choreographic and history, for at least five minutes. They used pre-selected song portions in order to adequately assess its effectiveness. The choreography that was created served as the basis for a jazz trio dance that will be performed as part of a high school dance show in the area. Following the testing, participants were given a link to a Google Forms survey where they could provide feedback on the program. The survey replies may be more accurate and consistent if the application is given just after the testing procedure, when the participants' memories of using the application are still fresh. On a scale of 1 to 10, participants were asked to score the application's usability and convenience. Participants could make any further comments in the free-response box at the bottom of the survey, which was optional.

It appears from the table and chart below that participants generally had positive opinions on the convenience and functionality. The ratings for functionality ranged from a maximum of 10 to a minimum of 6, with an average of 7.91; in contrast, the ratings for convenience ranged from a maximum of 10 to a minimum of 5, with an average of 7.27. According to the two average assessments, convenience is generally rated slightly lower than functionality. Given that some participants complained that they were unable to access or read the events page properly, the optional comment appears to provide an explanation for why this is the case. Given that other users of the application do not appear to have run into any problems or errors, it is unclear why this bug is happening. Although the feedback on the interface was nearly entirely good, one participant suggested adding more decoration to make it more visually appealing to users.

The outcomes show that the application is successful in terms of correct feature implementation and feature contributions to the program's main objective, which is the creating choreography. Most participants gave the functionality a score of 6 or above out of 10. This is reasonable given that the program was designed and built with the primary goal of serving as a superior substitute for the experiment, the features underwent extensive testing and revision. The results showed that users liked the application's interface for its convenience and simplicity. We identified the exact components of the interface that did not function and then came up with replacement ideas for improved interfaces.

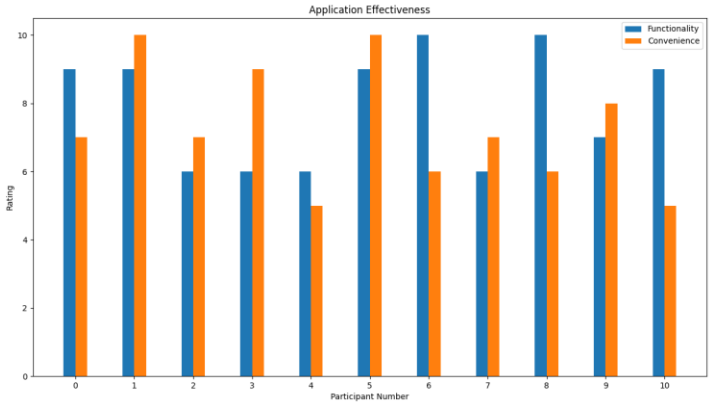

Figure 5. Application Effectiveness

Participant Number	Functionality	Convenience
0	9	7
1	9	10
2	6	7
3	6	9
4	6	5
5	9	10
6	10	6
7	6	7
8	10	6
9	7	8
10	9	5

Figure 6. Smart Dance Generator Application Effectiveness

5. Related Work

Although there is not an extensive amount of research or work completed towards learning how to merge choreography with AI, some machine learning choreography research and experiments are still present.

The first related work completed is “Generative Choreography using Deep Learning”, in which the authors used the system chor-rnn to create choreography specific to a choreographers’ style [11]. The system is trained on capturing data and mimicking the style of a human choreographer to create a basis of choreography that can be used as inspiration. Both of our works present the final choreography in the format of a stick figure, however our system has more accurate proportions and user-friendly movement since we track the movement of a real human dancing, and their system can choreograph for several minutes compared to our application’s thirty second time limit per audio input. Additionally, their system is unique to each choreographer, which makes it more likely for an experienced dancer to use since a style and pre-choreographed pieces must be analyzed, however it appears to repeat the same moves within each video. Our application makes it more accessible by allowing choreographers from all ages to use it without requiring previous choreographing experience, and it has been published on the app store rather than limiting the system for research specifically.

Another finding of work related to our project is the article “Beyond Imitation: Generative and Variational Choreography via Machine Learning”[12]. The priority behind this project was to create completely original choreography, inspired from the words of choreographer Jennings. The authors generated abstract choreography which could be used by a choreographer in a multitude of ways, including the documentation and study of a choreographer’s style, support with instructing movement to dancers by tracking the movement of a dancer as compared to the original choreography, as well as assist with the branching out in terms of style by providing new creative ideas that the choreographer has not yet used. Once again, both this piece of research and the first one referenced in this article are targeted towards more experienced choreographers, who are seeking to expand past previously utilized dance moves and create something completely original. A difference with my application is that it is targeted towards any choreographer or dancer, regardless of how limited their experience may be.

The third related work is “Style Machines”, which analyzed a movement and transformed it to match a specific style [13]. The authors used motion capture to create new choreography, which enabled a person with little to no background experience in dance to be able to choreograph, since the software would adjust the movements to become more defined and stylistic. This can also be helpful with showing a dancer how they can improve, by creating better movements based off of the dancer’s current movements, allowing for them to make a comparison. My work, however, might be more useful when a beginner is learning how to dance since they might not know basic moves to start using as a basis for the software to improve and add style to. Additionally, the output form for my application is a stop motion video, while this work returns motion sequences, therefore my work pieces together the images for the user and makes the fluid movement between poses easier to comprehend and learn.

6. CONCLUSIONS

It is essential that a dancer has choreography before performing or competing so that they feel confident, prepared, and well-rehearsed, which minimizes the amount of mistakes that might occur. In order to ensure that the choreography is well thought out, but balancing this with the limited amount of time that choreographers have to work with dancers, a basic first draft of choreography should be created as quickly as possible. To assist choreographers with pushing through creative blocks when they first begin a new piece this application choreographs dance moves according to music, promising that the choreography is appropriate for the given style.

Despite our achievements, there are still a few limitations which can still be improved upon for optimal use of the application. To name a few, some limitations include the length of the music which choreography can be assigned to, as well as the stop-motion quality of the video which is returned to the user.

For the future we plan to expand the database of videos from which music is compared to and which dance moves are created from, and replace shorter choreographed dance videos with longer versions [14]. The new videos in the database will consist of further choreographed segments. This addresses the limitation of inputted music having a short time limit, since the shortest video in the database will be long, and therefore will not prevent the application from running out of choreographed material when slicing and concatenating videos.

REFERENCES

[1] Gardner, Sally. "The dancer, the choreographer and modern dance scholarship: A critical reading." Dance Research 25.1 (2007): 35-53.
[2] Whalen, Jack, Marilyn Whalen, and Kathryn Henderson. "Improvisational choreography in tele service work." The British journal of sociology 53.2 (2002): 239-258.
[3] Whalen, Jack, Marilyn Whalen, and Kathryn Henderson. "Improvisational choreography in teleservice work." The British journal of sociology 53.2 (2002): 239-258.
[4] Qiu, Zongyan, et al. "Towards the theoretical foundation of choreography." Proceedings of the 16th international conference on World Wide Web. 2007.
[5] Ciolfi Felice, Marianela, Sarah Fdili Alaoui, and Wendy E. Mackay. "How do choreographers craft dance? Designing for a choreographer-technology partnership." Proceedings of the 3rd International Symposium on Movement and Computing. 2016.
[6] Li, Ruilong, et al. "Ai choreographer: Music conditioned 3d dance generation with aist++." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
[7] Wasserman, Anthony I. "Software engineering issues for mobile application development." Proceedings of the FSE/SDP workshop on Future of software engineering research. 2010.
[8] Krause, Amanda E., and Steven Caldwell Brown. "A uses and gratifications approach to considering the music formats that people use most often." Psychology of Music 49.3 (2021): 547-566.
[9] Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349.6245 (2015): 255-260.
[10] Peeters, Geoffroy, et al. "The timbre toolbox: Extracting audio descriptors from musical signals." The Journal of the Acoustical Society of America 130.5 (2011): 2902-2916.
[11] Crnkovic-Friis, Luka, and Louise Crnkovic-Friis. "Generative choreography using deep learning." arXiv preprint arXiv:1605.06921 (2016).
[12] Pettee, Mariel, et al. "Beyond imitation: Generative and variational choreography via machine learning." arXiv preprint arXiv:1907.05297 (2019).
[13] Brand, Matthew, and Aaron Hertzmann. "Style machines." Proceedings of the 27th annual conference on Computer graphics and interactive techniques. 2000.
[14] Han, Heonjong, et al. "TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions." Nucleic acids research 46.D1 (2018): D380-D386.
[15] McFee, Brian, et al. "librosa: Audio and music signal analysis in python." Proceedings of the 14th python in science conference. Vol. 8. 2015.