DanceVision Proposal

Anna Guo Jenna Jacob Katie Wang Natalie Leal Blanco
UT Austin: Computer Vision Course
annayuab@icloud.com

Abstract

Dance is a universal form of expression and entertainment, yet mastering its intricacies can be daunting for beginners. To democratize dance learning and make it more accessible to enthusiasts, we propose a novel approach utilizing computer vision techniques. By breaking down dance videos into vectorized and text-based representations with counts, the complexities of dance routines are simplified, enabling easier comprehension and replication by average fans.

1. Introduction

In recent years, dance has surged in popularity, whether through TikTok trends in younger generations, the growing fanbase for elaborate K-pop dance videos, or the performances in reality competition shows (which themselves cover wide genres from ballet and contemporary to ballroom dances). As a result of this growing interest, many eager dance novices have attempted to self-teach from the abundant dance videos online, available through a variety of platforms. However, such individuals often struggle with breaking down movements due to the speed or intricacy with which they were executed.

We aim to use Computer Vision to partition dance movements into more easily digestible chunks by creating a model that can learn (vectorize) dance choreography from a given dance video, and output it in a more human-readable form, such as by describing multiple steps in a motion and adding counts or timings for context. By developing a tool that provides detailed descriptions of human positions and gestures alongside counts or musical beats, learning dance choreographies could become much more accessible and approachable for a broader audience.

1.1. Existing Attempts

Several existing attempts have been made to tackle the challenge of learning dance from online videos, each with its own approach and limitations. One such approach involves platforms that generate reskinned versions of dance performances, offering visual references for users. These platforms overlay graphics or animations onto the original video, highlighting key movements or positions to provide guidance. While these reskinned versions serve as visual aids, they typically do not offer active instruction or detailed breakdowns of the dance steps. As a result, users may struggle to fully grasp the nuances and intricacies of the movements, limiting the effectiveness of this approach as a learning tool.

Another example is a DancePal Hackathon project, which adopts a more technical approach by comparing an ideal video with the end user's video skeletons. This method involves analyzing the skeletal poses extracted from both videos and identifying discrepancies between them. While this approach has the potential to provide personalized feedback to users based on their own performances, it comes with several challenges. Firstly, it requires an ideal video with specific lighting, background, and perspective conditions to serve as a reference, which may not always be feasible to obtain. Additionally, accurately mapping movements onto predefined definitions can be difficult, especially for complex dance sequences with varying styles and interpretations. As a result, the effectiveness of this approach may be limited by the availability of suitable reference videos and the accuracy of the mapping process.

Unlike reskinned versions of dance performances, our proposed method provides detailed instructional breakdowns of dance movements, guiding learners through each step with clarity and precision. This active instruction facilitates better understanding and retention of choreography. Additionally, our method could be expanded to offer personalized feedback based on the analysis of the learner's own video and that of an original dance video, ideally without the requirement for perfect lighting or camera angle conditions.

1.2. Useful Techniques

Several computer vision techniques offer promising avenues for addressing the challenge of learning dance from online videos by providing deeper insights into the dynamics and nuances of dance movements. One such technique involves exploring 3D models instead of traditional 2D skeletonization methods to create more comprehensive representations of dance movements. Unlike 2D skeletonization, which captures only the spatial coordinates of key body joints in a single plane, 3D modeling techniques can capture the full spatial extent of movements, including depth information. This allows for a more accurate and detailed representation of the dancer's body and movements in three-dimensional space, which is particularly valuable for capturing complex and multi-dimensional dance choreographies.

In addition to 3D modeling, other computer vision techniques such as action recognition, pose estimation, trajectory analysis, and video segmentation and tracking can provide valuable insights into the dynamics of dance movements. Action recognition algorithms can automatically identify and classify different dance gestures and sequences, allowing for automated analysis of choreographic patterns and styles. Pose estimation techniques can accurately estimate the skeletal poses of dancers from video frames, providing a structural representation of their movements that can be used for further analysis and interpretation.

Trajectory analysis techniques can track the trajectories of specific body parts or movements over time, allowing for the analysis of movement patterns, velocities, and accelerations. This can provide valuable insights into the rhythmic and dynamic aspects of dance performances, helping learners to better understand the timing and pacing of movements. Video segmentation and tracking methods can segment dance videos into individual sequences or movements, allowing for more focused analysis and comparison of specific choreographic elements.

By leveraging these advanced computer vision techniques, learners can gain a deeper understanding of the dynamics and nuances of dance movements, facilitating more effective learning experiences. These techniques can help learners to break down complex choreographies into smaller, more manageable components, identify areas for improvement, and refine their technique with greater precision and accuracy. Overall, the integration of computer vision into dance learning has the potential to revolutionize the way dance is taught and learned, making it more accessible, engaging, and effective for dancers of all levels.

1.3. Challenges

Several challenges arise from attempting this project. One is the quality of input data, which encompasses various factors such as lighting, angles, occlusions, frame rates, and blurring. These factors can significantly impact the accuracy of movement analysis, as they can obscure or distort key visual information needed for understanding dance movements. Addressing these challenges may require the development of robust computer vision algorithms capable of handling noisy or low-quality video data, as well as techniques for preprocessing and enhancing the visual information to improve analysis accuracy.

Moreover, handling interactions between multiple dancers in an ideal video presents additional complexities in accurately capturing and representing dance movements. In many dance performances, especially group or partner dances, dancers may interact with each other in complex ways, such as lifts, partner work, or coordinated movements. Capturing these interactions and accurately representing them in the analysis poses challenges for computer vision algorithms, as they must be able to differentiate between individual dancers and track their movements independently while also accounting for their interactions with others. Addressing this challenge may require the development of advanced tracking and segmentation techniques capable of accurately identifying and tracking multiple dancers in a crowded or dynamic environment.

Overall, addressing these challenges requires a multidisciplinary approach that combines expertise in computer vision, dance analysis, and human-computer interaction. By developing robust algorithms, incorporating domain-specific knowledge, and leveraging advanced techniques for handling complex visual data, it is possible to create a tool that effectively supports learning dance from online videos using computer vision. However, continued research and development efforts are needed to overcome these challenges and realize the full potential of computer vision technology in the field of dance education.

1.4. Technical Gap

The primary aim of this project is to reduce the technical gap between online dance learning resources and effective dance instruction. By enabling the correspondence between individuals with different body proportions captured from various angles, this tool seeks to enhance the accessibility and inclusivity of dance education. This interdisciplinary endeavor involves collaboration between experts in dance, biomechanics, and computer vision, with computer vision providing quantitative analysis capabilities to enhance the understanding and teaching of dance movements.

In conclusion, developing a tool for learning dance from online videos using computer vision holds great potential for revolutionizing dance education and making it more accessible to a wider audience. By leveraging advanced computer vision techniques and interdisciplinary collaboration, this project aims to bridge the gap between online dance resources and effective dance instruction, ultimately empowering individuals to learn and appreciate dance in new and innovative ways.

References