Pose Evaluation for Dance Learning Application using Joint Position and Angular Similarity

Jae-Jun Lee, Jong-Hyeok Choi, Tserenpurev Chuluunsaikhan, Aziz Nasridinov
Chungbuk National University, Cheongju, 28644, South Korea
{ leejj, leopard, teo, aziz } @chungbuk.ac.kr

UBICOMP ISWC 2020

Introduction

Background

- Recently, with the worldwide popularity of K-POP, the number of people learning to dance through various video-sharing platforms has drastically increased.
- In the case of learning to dance individually, there is a problem that **it is difficult to obtain objective feedback.**
- Several studies evaluated the dance by **analyzing the similarity of dance gestures** using 3-D joint information.
  - However, **it is difficult to get feedback without evaluation of exact pose** (i.e., position or angle of joints).
  - Besides, considering that these methods **require special equipment,** they have not been widely used for real-life dance learning.

Contributions

- We propose a new method to **compare and evaluate the teacher and the student's dance pose through the AI-based pose estimation** obtained from the 2-D camera of a smartphone.
- In particular, we propose a pose evaluation method that **rectify different body ratios and compare the position and angle of the joints** (see Figure 1).
- Experiment on a smartphone with real-world dataset showed that we could **estimate the dance pose more accurately and provide feedback to the student in real-time.**

Proposed Methods

Joint position similarity evaluation

- To compare the joint positions of teacher *T* and student *S*, we first rectify different body ratios using **P-Affine (Partial-Affine transformation).**
  - P-Affine transforms six body parts of *S* into the same position and ratio as *T*
- The position of each *T* and transformed *S'* joints is then evaluated through OKS to **obtain the final JPSS (Joint Position Similarity Score)**.

Joint angular similarity evaluation

- Further, the joint angular similarity of the dance pose is **evaluated through their angle difference.**
- Here, the difference between the angles obtained from *T* and *S* is used for the evaluation because in *S'* generated through P-Affine the joint angle may change.
- **The final JASS (Joint Angular Similarity Score)** was evaluated using threshold that considers the user level (e.g., beginner - 45°, intermediate - 30°, expert - 15°).

PA similarity evaluation

- Weights are applied to the JPSS and JASS by **reflecting the dance teacher's opinions.**
- Then, **the final PAS (Position and Angle Similarity)** is obtained by summing JPSS and JASS.
- Finally, the average of the PAS obtained for a certain period **is used to provide real-time feedback on the dance moves to the student**.

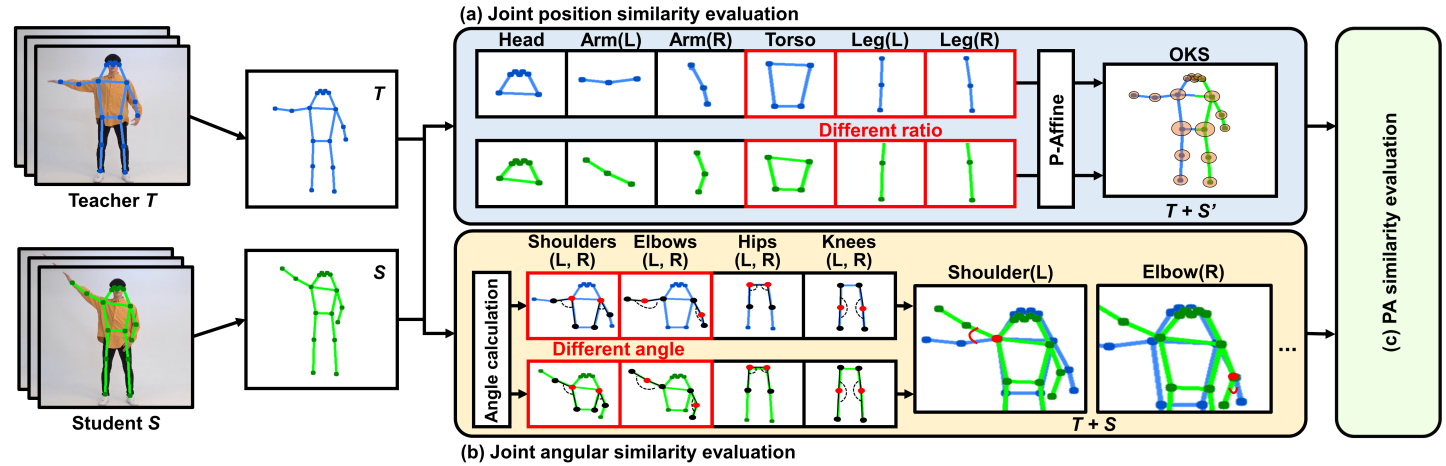

Figure 1 : The process of evaluating dance pose in a dance learning application.

Performance Evaluation

Experimental environment

- We developed **an actual dance learning application** on an iPhone 11 Pro Max (see Figure 2).
- For experiments, we used **PifPaf** and **Tensorflow's PoseNet** to perform the teacher and student pose estimation, respectively.
- For the experiments, **five scenarios (see Table 1) are tested on four methods:** (a) simply compare, (b) normalize and compare body ratios based on torso size, (c) apply only the general affine transform, and (d) the proposed PA similarity method.

Table 1. Experimental scenarios

No	Scenarios
1	Angle and body ratios are the same, but only the joint position is different.
2	Angle is the same, but the ratios of the body and the joint position is different.
3	Body ratios are the same, but the angle and joint position are different.
4	Different people perform different dances.
5	Evaluation based on a real-world dataset.

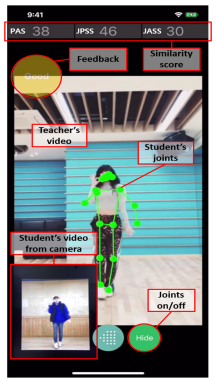

Figure 2 : Dance Learning App.

Result of experiments

- From Table 2, we can observe that the proposed method outperforms the competing methods due to the following reasons:
  - Our PA similarity method was able to consider the body ratios in detail through P-Affine.
  - In the case of angular similarity, the original angle was used, and it thereby showed good performance in various scenarios.
- As for timeliness, PA similarity can be obtained within a short period **(from 0.363 ms to 0.379 ms per frame)**.
- These results show that PA similarity **can be easily used for real-time pose evaluation even on a smartphone and estimate the dance pose more accurately.**

Table 2. The average accuracy scores of dance pose evaluation.

	Scenario(1)			Scenario(2)			Scenario(3)			Scenario(4)			Scenario(5)		
Method	JPSS	JASS	PAS	JPSS	JASS	PAS	JPSS	JASS	PAS	JPSS	JASS	PAS	JPSS	JASS	PAS
(a)	1.12	99.80	50.53	1.17	99.77	50.51	1.12	74.78	37.90	7.89	27.92	17.91	1.82	34.52	18.18
(b)	99.75	99.80	99.78	78.76	99.77	89.30	71.65	74.78	73.18	34.00	27.87	30.96	31.21	34.30	32.77
(c)	97.29	97.84	97.57	85.83	96.40	91.12	71.00	75.53	73.27	32.82	30.07	31.44	33.38	33.53	33.46
(d)	99.75	99.80	99.78	99.73	99.77	99.76	71.61	74.78	73.16	48.36	27.92	38.15	52.09	34.52	43.32