InterPoser: Visualizing Interpolated Movements for Bouldering Training

Keisuke Shiro*
Kazme Egawa*
The University of Tokyo
Tokyo, Japan
shiro@g.ecc.u-tokyo.ac.jp
kazme@g.ecc.u-tokyo.ac.jp

Takashi Miyaki
The University of Tokyo
Tokyo, Japan
miyaki@acm.org

Jun Rekimoto
The University of Tokyo
Sony Computer Science Laboratories, Inc.
Tokyo, Japan
rekimoto@acm.org

ABSTRACT

Bouldering is an urban form of rock climbing that requires precise and complex movement. Similarly to other sports, the simplest way to learn bouldering skill is to mimic professional's motion. However, ordinary beginner boulders cannot learn to coaches, so that they learn by themselves or tutorial videos. Even if they managed, bouldering has a communication difficulty between a trainee and a trainer, that is, climbers cannot mimic the trainer's movement in parallel. Accordingly, we considered a video feedback system would be useful for beginners and suggested InterPoser: a novel visualization system

*The first two authors contributed equally.

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5971-9/19/05.
https://doi.org/10.1145/3290607.3312779

for intermediate motion between a beginner climber and a more experienced. InterPoser receives two videos of different subjects climbing the sample problem and generates an intermediate movement. In addition, this motion is transferred into realistic images of the climber. The proposed system is expected to support beginner to acquire more detailed observation and understanding of the motion.

CCS CONCEPTS

• **Human-centered computing** → **Information visualization**; *Human computer interaction (HCI).*

KEYWORDS

Climbing; sports technologies; motion analysis; expert modeling.

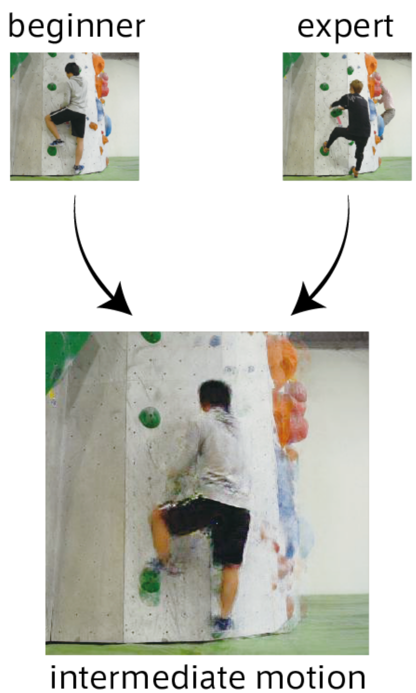

Figure 1: An intermediate motion image generated from videos of a beginner climber and an expert

INTRODUCTION

Learning climbing skill is a demanding task for beginners because climbing is an activity that consists of complicated body movements. To improve climbing skill, beginners have to understand the difference between them and experienced climbers. Especially, bouldering beginners also have this issue. Bouldering is a form of indoor sport climbing. The goal of bouldering is to ascend a pre-defined set of holds, which is called a "problem." Typical bouldering walls have various types of artificial holds. In contrast to other types of climbing, boulders are not equipped with any ropes or carabiners.

Imitating expert's movement is the simplest way of sports training. This is also true in bouldering. However, bouldering needs complex tips and skills to ascend the goal effectively. Thus, professional climbing movements are too complicated for beginners to understand. Moreover, ordinary beginners cannot learn to instructors, because the number of coaches is limited. Most beginner climbers learn themselves or from videos of experienced climbers and this makes high entry barriers.

To overcome this issue, we hypothesized that beginner climbers can observe and understand the professional's movement in detail with interpolated motion between them and experts. Thus we suggest InterPoser: a novel visualizing system for interpolated climbing movement. InterPoser provides interpolated images from two videos of the beginner and the experienced climber climbing the same problem. Our system converts interpolated pose information into realistic images of the beginner using "do as I do" style transfer framework [4] so that beginners can see a synthesized video of themselves climbing in a more effective way. Figure 1 shows example input videos and a resulting intermediate motion image.

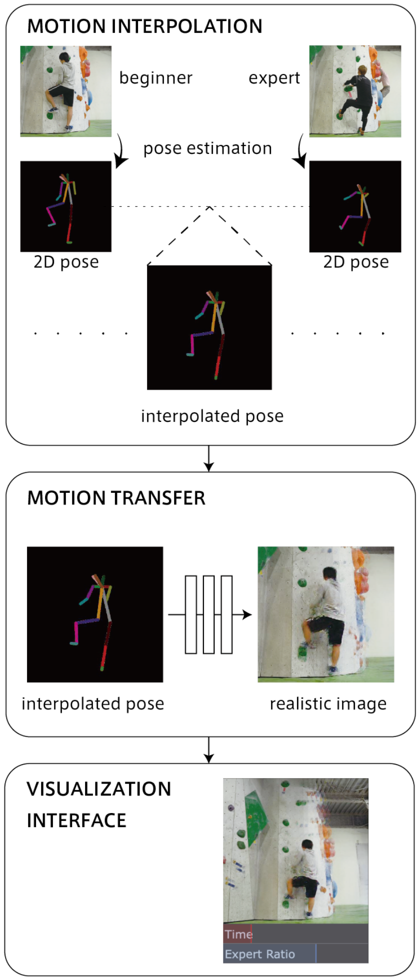

Figure 2: InterPoser system overview: (Top) Pose estimation and interpolation from two video sources (Middle) GANs based motion transfer (Bottom) Visualization interface

RELATED WORK

Human Computer Interaction and Climbing

Mimicking more experienced players is a typical method for sports training. However, climbing trainees cannot utilize this basic training method because instructors cannot perform in parallel on the same wall. Various HCI studies have addressed this communication problem between trainees and instructors in climbing. Currently, these researches have adopted two trending methods: "augmented wall" and "wearable" technologies. Kajastila et al. [7] introduced the *Augmented Climbing Wall* which realizes interactive games on a climbing wall with projection mapping and animation. Wiehr et al. [14] developed the *betaCube*, a self-calibrating camera-projection unit which enables recording and replaying a climbing movement on a wall. Based on this betaCube technology, Kosmalla et al. [9] developed a wall-projection system, which visualizes a third-person view of expert climbers movement, and evaluated various visualization techniques. On the other hand, wearable technologies have targeted climber status tracking. Ladha et al. [10] introduced *ClimbAX* - a climbing performance analysis system. *ClimbAX* assesses climber's skill with a wrist-band-shaped accelerometer and heuristic algorithms. Kosmalla et al.[8] also presented *ClimbSense*, a wrist-worn sensor that automatically recognizes a climbing route. Mencarini et al. [12] explained the em otional difficulties of climbing through elaborate user study and implied design guidelines for wearable technologies that support beginner climbers. In contrast to "augmented wall" and "wearable" methods, Daiber et al. [5] proposed *BourdAR* that utilized AR technology with a smartphone for collaborative training. This approach is convincing because a lot of climbers use a smartphone to record their performance, therefore, *BouldAR* does not need any extra devices for climbers.

Expert Modeling

In other sports, researchers suggested a video feedback training method combined with expert modeling [1, 13], which aims to improve the trainee's performance by showing the correct execution of a specific skill and technique. We consider our interpolation approach will extend the concept of expert modeling in the video training field.

INTERPOSER

InterPoser visualizes an intermediate motion between a beginner climber and an expert. In this section, we describe InterPoser system in detail. Our system consists of three parts: Interpolation, Motion Transfer, and Visualization Interface. Figure 2 describes our system overview.

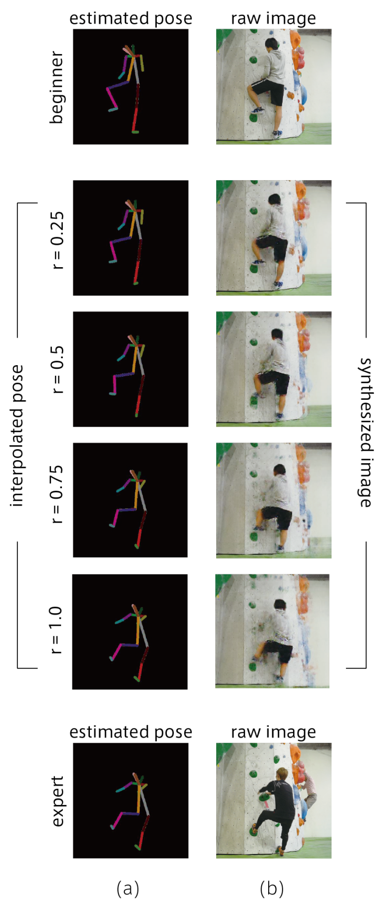

Figure 3: (a) 2D pose images of the beginner, the experienced climber and interpolated
(b) Raw images of the two climbers and synthesized images from interpolated motion

Motion Interpolation

Our system receives two videos of two different subjects climbing the same route. First, we extract 2D pose sequences p with T frames $p_1, ..., p_T$ and p' with T' frames $p'_1, ..., p'_{T'}$ from these videos using OpenPose [2, 3]. OpenPose is a real-time framework for human pose estimation from a single image using CNN-based method. Each p_i and p'_j frame involves N joint points $p_{i1}, ..., p_{iN}$ and $p'_{j1}, ..., p'_{jN}$. To smoothly interpolate them, we synchronize these two motions by selecting the nearest frame for each. We calculate L2 distance between p_i and p'_j

$$d(p_i, p'_j) = \frac{1}{N} \sum_{k=1}^{N} \|p_{ik} - p'_{jk}\|_2 \tag{1}$$

In this research, the number of joints $N = 25$. For each p_i, we select the nearest frame $p'_{t=i}$ in L2 distance. Given division resolution M and division ratio m, the internally divided frame p_{io} between p_i and $p'_{t=i}$ will be

$$p_{io} = rp_i + (1 - r)p'_{t=i} \tag{2}$$

where the internal division ratio is $r : 1 - r$ and $r \in [0, 1]$. We calculate 100 interpolated poses for each p_i. Figure 3(a) illustrates an example of interpolated poses where $r = 0.25, 0.5, 0.75$.

Motion Transfer

To achieve more realistic images, we utilized Everybody Dance Now (EDN) [4] network. EDN is a "do as I do" type of motion transfer method. Given two source videos of A and B dancing, EDN imposes A's motion onto B and generate a fake video of B dancing like A with Generative Adversarial Networks (GANs). Then we applied the EDN method to synthesize a video with the beginner's appearance and with the expert's motion. Figure 3(b) shows examples of conversion results from interpolated poses described in the previous section. In "dance" cases, the target subject performs just standard moves, that is, EDN trains generator network of GANs with the motion irrelevant to the target motion. In contrast, our case has two videos of similar motions, and this makes training for generator network easier.

Visualization Interface

Figure 4 shows the visualization software. The Main view simply visualizes climbing motion. User controls the internal division ratio by slider interface and motion view smoothly changes the image according to the ratio value. This interface was implemented with TouchDesigner.

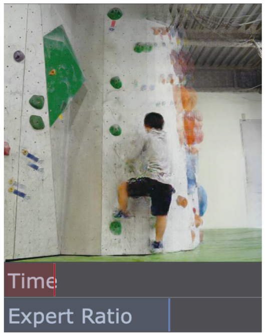

Figure 4: The appearance of InterPoser visualizing software. Users control the internal division ratio with slider interface.

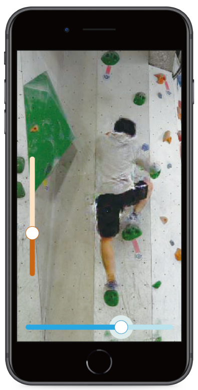

Figure 5: The concept image of the smartphone app version

FUTURE WORK

We hypothesized that visualizing intermediate motion enhances video feedback training and the concept of expert modeling. Therefore we plan to evaluate the effect of this system on actual climbing training. Before evaluation, our system has some limitations to be solved.

Interpolation Method

InterPoser generates a pose with simple linear interpolation and synchronization with L2 distance. This method is a simple and brief way to achieve an intermediate motion from two motions. However, this method cannot reflect joint constraints such as inverse kinematics and position constraints of holds on a bouldering wall. Then we focus on non-linear dimensional reduction method. Style-based inverse kinematics [6] extracts a non-linear latent feature space for human motion which entails kinematic constraints from motion data such as walking or baseball pitching. It also supports style interpolation in the latent space and we consider this method will be suitable interpolation tool for InterPoser.

User Interface

Our visualization system is now based on GUI software on a laptop. Boulders often take a video of themselves climbing for practice. Currently, smartphones are the most popular device for recording and there are a lot of bouldering videos on Instagram. Accordingly, our system should be a smartphone app like Figure 5.

Other Types of Sports

Bouldering is a type of individual sports and its problem has a clear beginning and end. In sports coaching field, bouldering skills are classified into *Closed Skills* [11], which take place in a stable environment like swimming and serving in tennis. We plan to apply our method to analysis on these closed skills of other types of sports beyond bouldering.

CONCLUSION

In this work, we proposed InterPoser: a simple method for synthesizing and visualizing interpolated climbing motion. Our goal is to assist beginner climbers to understand the nuance of expert climber's movement without an instructor. To achieve this, InterPoser synthesizes an intermediate movement between the beginner and the expert. Users can view images which smoothly changes to a skillful movement. In addition, InterPoser renders realistic images with beginner's appearance by motion transfer technology. The proposed InterPoser method can create novel learning sources for beginner climbers, therefore, it is expected to contribute to video feedback training of bouldering.

REFERENCES

[1] E Boyer, R. G. Miltenberger, C. Batsche, and V. Fogel. 2009. Video Modeling by Experts with Video Feedback to Enhance Gymnastics Skills. *Journal of applied behavior analysis* 42, 4 (2009), 855–860. https://doi.org/10.1901/jaba.2009.42-855

[2] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2018. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. In *arXiv preprint arXiv:1812.08008*. https://arxiv.org/abs/1812.08008

[3] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '17)*. 1302–1310. https://doi.org/10.1109/CVPR.2017.143

[4] Caroline Chan, Shiry Ginosar, Tinghui Zhou, and Alexei A. Efros. 2018. Everybody Dance Now. *CoRR* abs/1808.07371 (2018). arXiv:1808.07371 http://arxiv.org/abs/1808.07371

[5] Florian Daiber, Felix Kosmalla, and Antonio Krüger. 2013. BouldAR: Using Augmented Reality to Support Collaborative Boulder Training. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 949–954. https://doi.org/10.1145/2468356.2468526

[6] Keith Grochow, Steven L. Martin, Aaron Hertzmann, and Zoran Popović. 2004. Style-based Inverse Kinematics. In *ACM SIGGRAPH 2004 Papers (SIGGRAPH '04)*. ACM, New York, NY, USA, 522–531. https://doi.org/10.1145/1186562.1015755

[7] Raine Kajastila, Leo Holsti, and Perttu Hämäläinen. 2016. The Augmented Climbing Wall: High-Exertion Proximity Interaction on a Wall-Sized Interactive Surface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 758–769. https://doi.org/10.1145/2858036.2858450

[8] Felix Kosmalla, Florian Daiber, and Antonio Krüger. 2015. ClimbSense: Automatic Climbing Route Recognition Using Wrist-worn Inertia Measurement Units. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2033–2042. https://doi.org/10.1145/2702123.2702311

[9] Felix Kosmalla, Florian Daiber, Frederik Wiehr, and Antonio Krüger. 2017. ClimbVis: Investigating In-situ Visualizations for Understanding Climbing Movements by Demonstration. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 270–279. https://doi.org/10.1145/3132272.3134119

[10] Cassim Ladha, Nils Y. Hammerla, Patrick Olivier, and Thomas Plötz. 2013. ClimbAX: Skill Assessment for Climbing Enthusiasts. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. ACM, New York, NY, USA, 235–244. https://doi.org/10.1145/2493432.2493492

[11] B. Mackenzie. 2001. Skill Classification. Retrieved January 6, 2019 from https://www.brianmac.co.uk/continuum.htm

[12] Eleonora Mencarini, Chiara Leonardi, Antonella De Angeli, and Massimo Zancanaro. 2016. Design Opportunities for Wearable Devices in Learning to Climb. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. ACM, New York, NY, USA, Article 48, 10 pages. https://doi.org/10.1145/2971485.2971509

[13] Roberta Rikli and Gregg Smith. 1980. Videotape feedback effects on tennis serving form. *Perceptual and Motor Skills* 50 (06 1980), 895–901. https://doi.org/10.2466/pms.1980.50.3.895

[14] Frederik Wiehr, Felix Kosmalla, Florian Daiber, and Antonio Krüger. 2016. betaCube: Enhancing Training for Climbing by a Self-Calibrating Camera-Projection Unit. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1998–2004. https://doi.org/10.1145/2851581.2892393