Cocktail Pose Recommendation and Evaluation System Based on Camera and Machine Learning Methods

基於相機與機器學習方法的調酒姿態建議與評價系統建構

組別： B375　指導教授：黃朝宗　組員：張智鈞、戴瑞傑

Abstract

HomeCourt, Swing Vision and similar software implement artificial intelligence technology to provide real-time analysis of games and training videos through the iPhone camera. Building on this concept, we plan to develop a system specifically designed to provide posture training for bartenders. The goal is to train high-level shaking techniques while ensuring consistent cocktail quality.

This project utilizes a camera to capture footage, combined with machine learning to create a scoring and recommendation system, enabling users to improve their shaking techniques.

Overall System

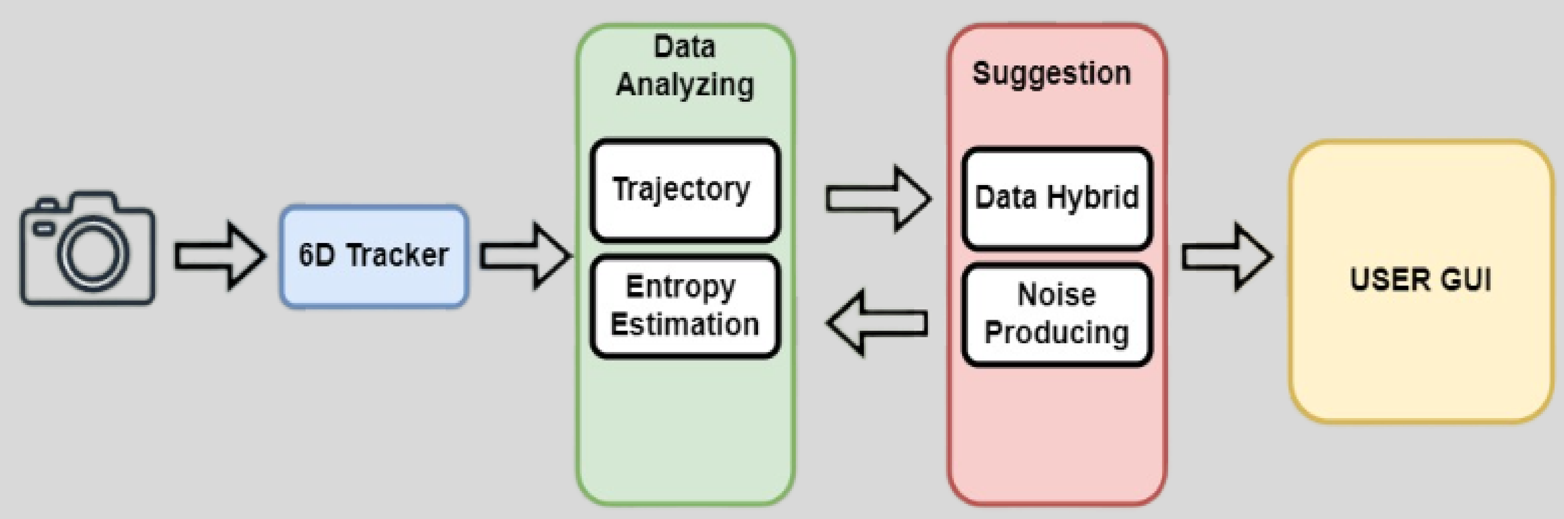

Fig.1 ： Overall system Block Diagram

6D Tracking

Using ArUco Markers to find shaker position and orientation. Regression compressing and zero-padding are used for data preprocessing.

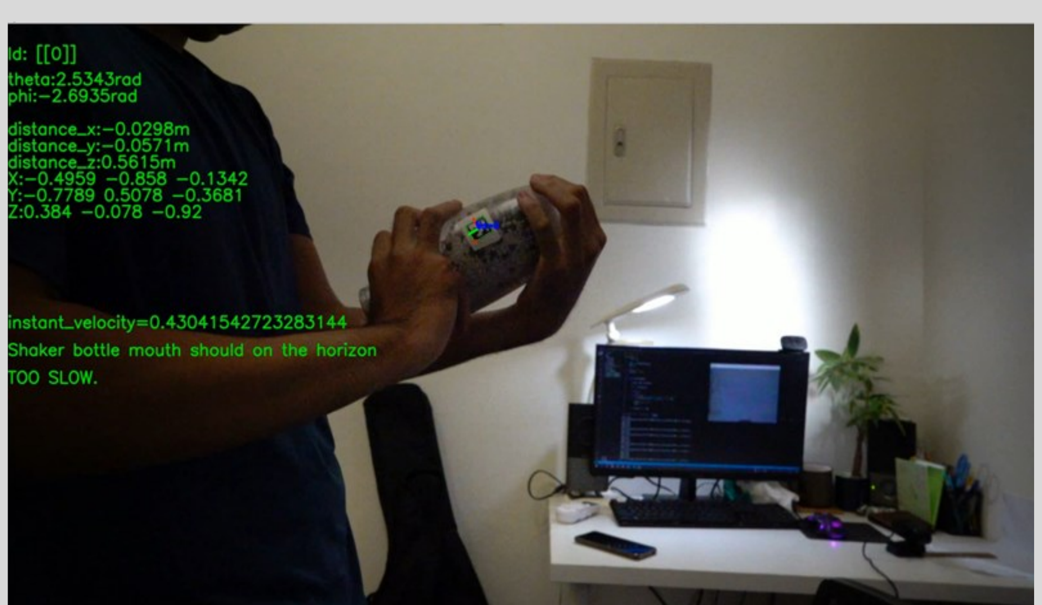

Fig.2 6D tracker with ArUco Marker

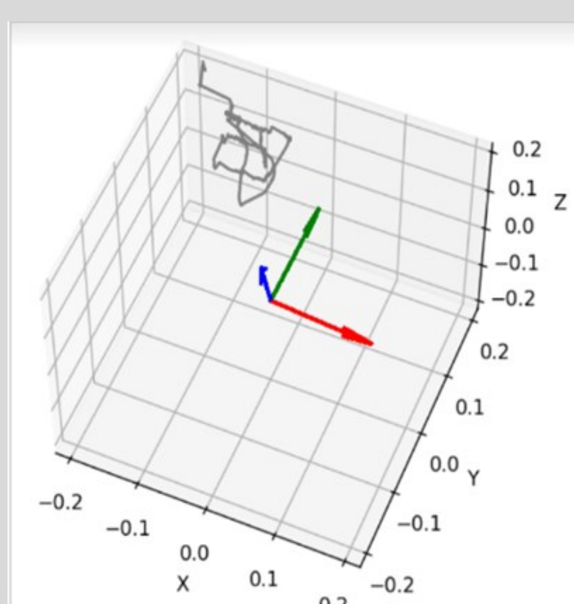

Fig.3 Trajectory validation

User Interface

The user interface we developed enables users to operate the system easily. Quantified speed and angle data help users improve their skills.

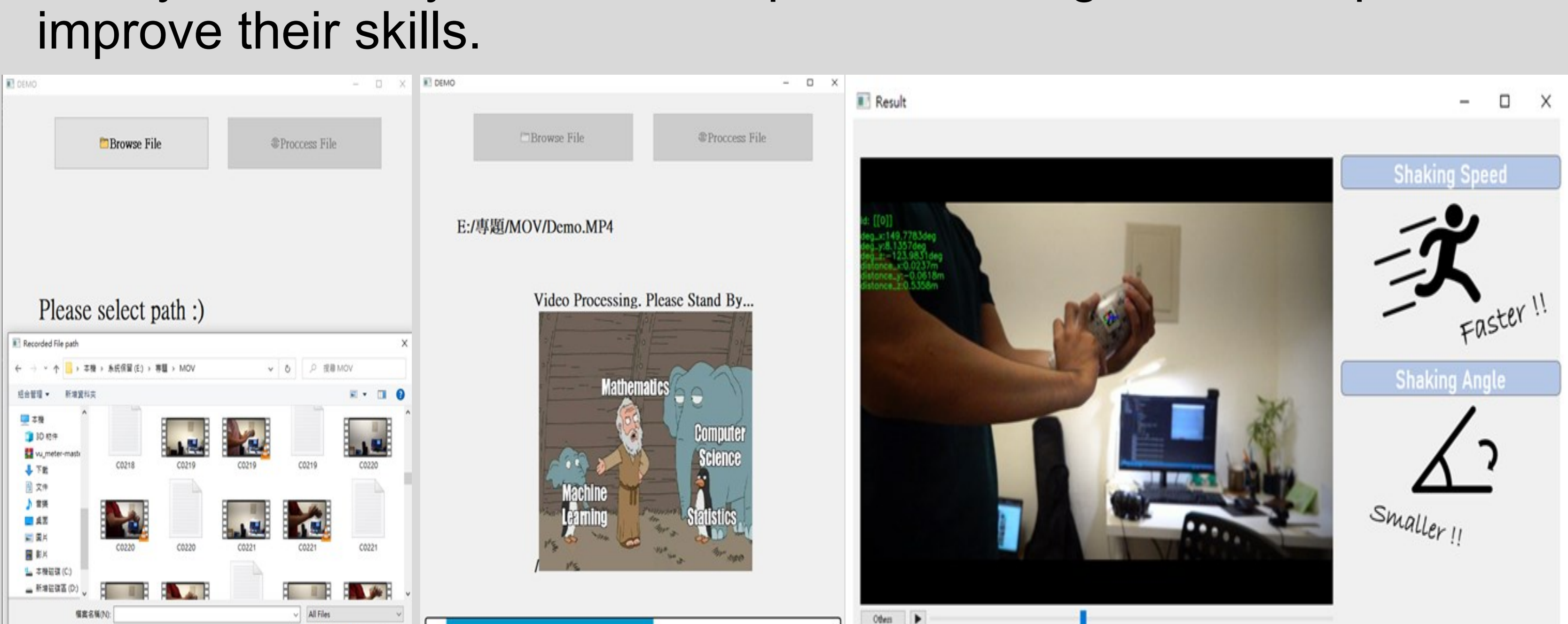

Fig.4 ： User interface

Entropy

We quantify the chaos by calculating the Discrete Information Entropy, measuring the disorder or uncertainty in a system.

$$H(X) = -\sum_{k} p_k \, log_2(p_k)$$

H(X) represents the discrete information entropy of the random variable *X*, where p_k is the probability of event k occurring.

Local entropy are measured and the total degree of chaos is quantified for the whole image.

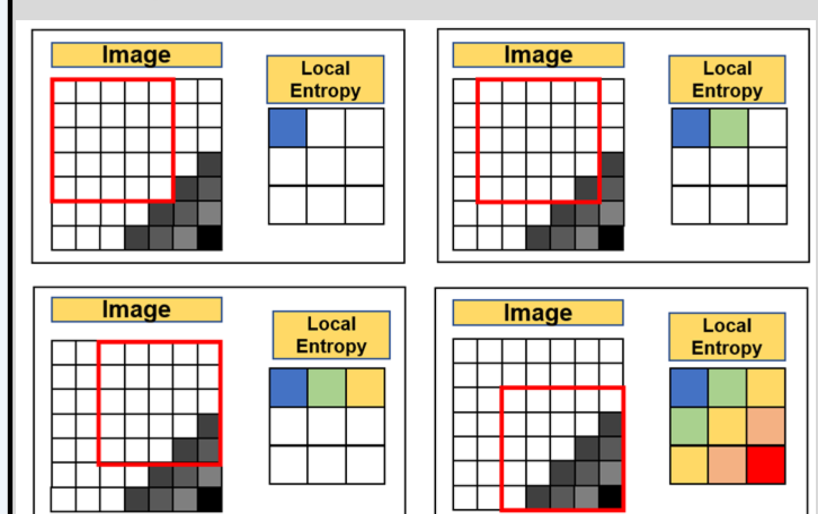

Fig.5 ： Local Entropy Algorithm

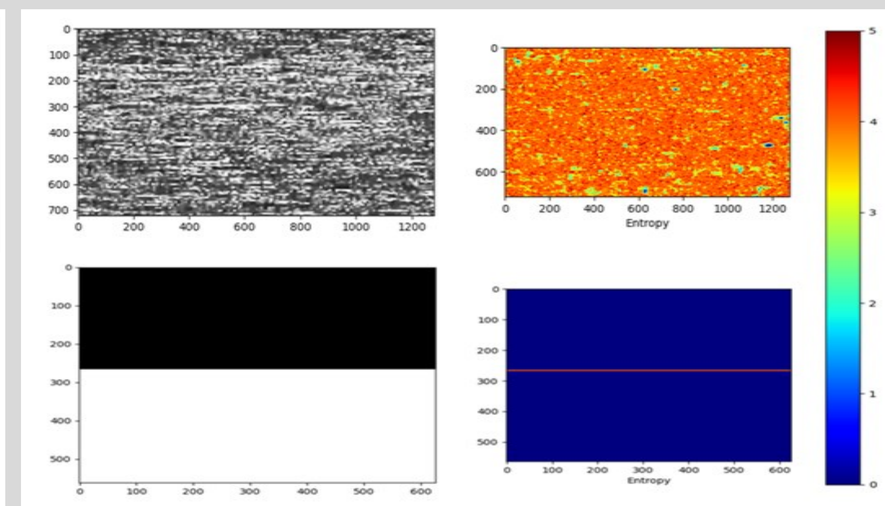

Fig.6 ： High chaos vs. Low chaos

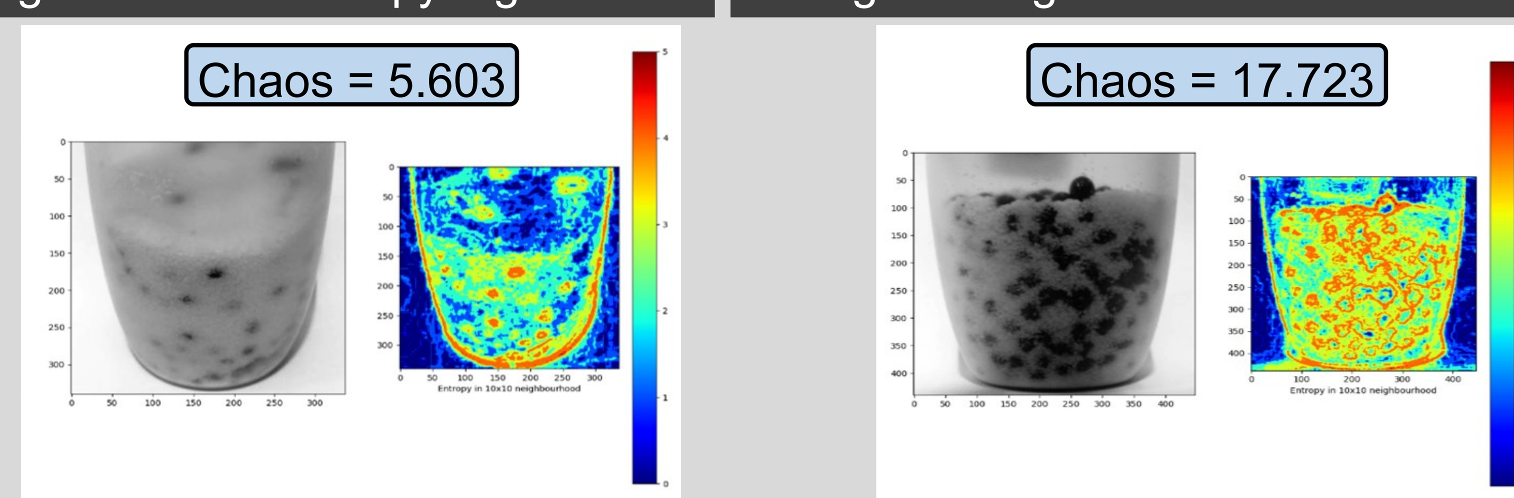

Fig.7 ： Algorithm implementation on shaker

Machine Learning Model

Evaluation system utilizes 3 different model, generating optimal virtual path for comparison to user original path.

(1) Gated Recurrent Unit (GRU)
(2) Long Short-Term Memory (LSTM)
(3) Modified Transformer

Model	Avg. MSE	Performance
Mean Model	198.268	1
GRU	146.264	1.2622
LSTM	162.996	1.1779
Transformer	225.715	0.8616

Chart 1 ： Model Performance

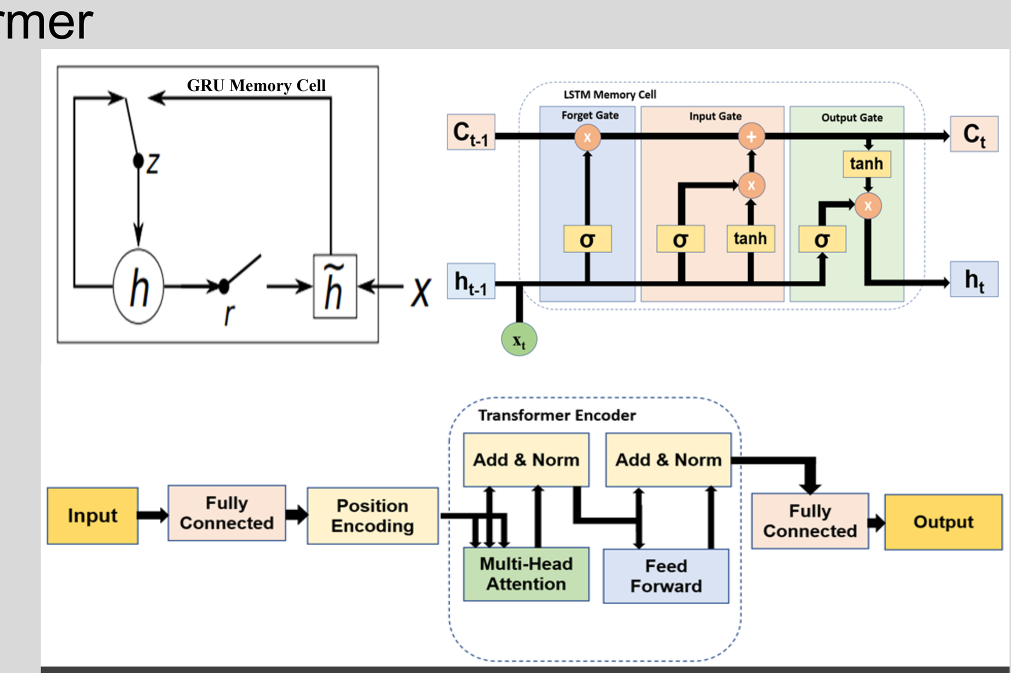

Fig.8 ： GRU, LSTM & Transformer structure

Conclusion

We have completed an initial version of a recommendation system designed to assist bartenders in posture training. There are numerous possibilities for optimization within this system, as there hasn't been much research in this area in the past. The topic may have the potential for further development. Additionally, related technologies such as entropy calculation and path tracking can find applications in various fields.

Reference

[1] Cho, Kyunghyun & Merrienboer, Bart & Gulcehre, Caglar & Bougares, Fethi & Schwenk, Holger & Bengio, Y.. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. 10.3115/v1/D14-1179.

[2] Hochreiter, Sepp & Schmidhuber, Jürgen. (1997). Long Short-term Memory. Neural computation. 9. 1735-80. 10.1162/neco.1997.9.8.1735.

[3] Vaswani, Ashish & Shazeer, Noam & Parmar, Niki & Uszkoreit, Jakob & Jones, Llion & Gomez, Aidan & Kaiser, Lukasz & Polosukhin, Illia. (2017). Attention Is All You Need.

[4] Ahmadyan, Adel & Zhang, Liangkai & Wei, Jianing & Ablavatski, Artsiom & Grundmann, Matthias. (2020). Objectron: A Large Scale Dataset of Object-Centric Videos in the Wild with Pose Annotations.

[5] 張宸。〈基於深度學習之投籃姿勢修正建議研究〉。碩士論文。淡江大學，台灣，2020。

[6] 夏至賢、陳昭和。〈籃球員自主投籃姿勢辨識系統之研究〉。《中正嶺學報》48.1 (2019): 1-11。