On-The-Go Photographer
隨行攝影師

組別：B260　　組員：陳均豪、施品瑄　　指導教授：黃朝宗

Introduction

Composition is a crucial step when taking a picture. A good composition makes the image harmonic and thus obtains more range for post-processing. On the other hand, a bad composition makes the image unbalanced and makes it hard to use the post-processing process to achieve the expected effect.

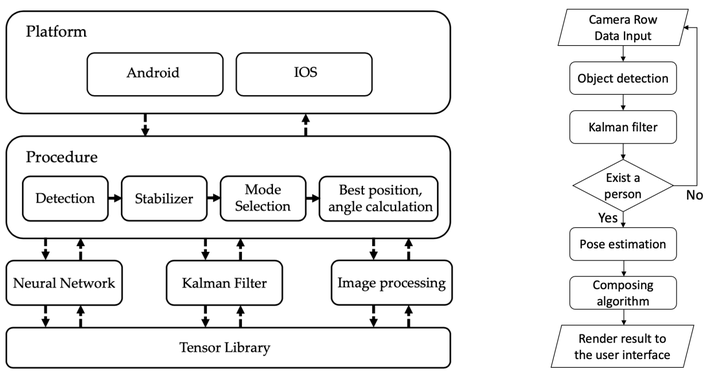
We intend to develop a camera APP to help the user compose the picture. Using the techniques such as object detection, pose estimation, etc. Analyzing the photographing environment and the character posture in real time. With the help of the algorithm, we designed to render the recommended composition in the form of an arrow or color on the user interface. Assisting the user to improve their photo composition's quality.

Problem Formulation

We know that the sense of beauty is subjective. Therefore, we need to make some assumptions to quantize this problem. Here, we assume that the rule of thirds is the golden discipline. To further reduce the complexity, we focus the main subject on a single person.

System Design

System architecture and flow chart



(a)　(b)

Figure 1: (a) System architecture (b) System flow chart

Object Detection

We need to perform object detection to identify and locate the person in the frame. We choose [2] as our detector. [2] is an FCOS-style one-stage anchor-free object detection model. Its design is CPU friendly without sacrificing accuracy.

Pose Estimation

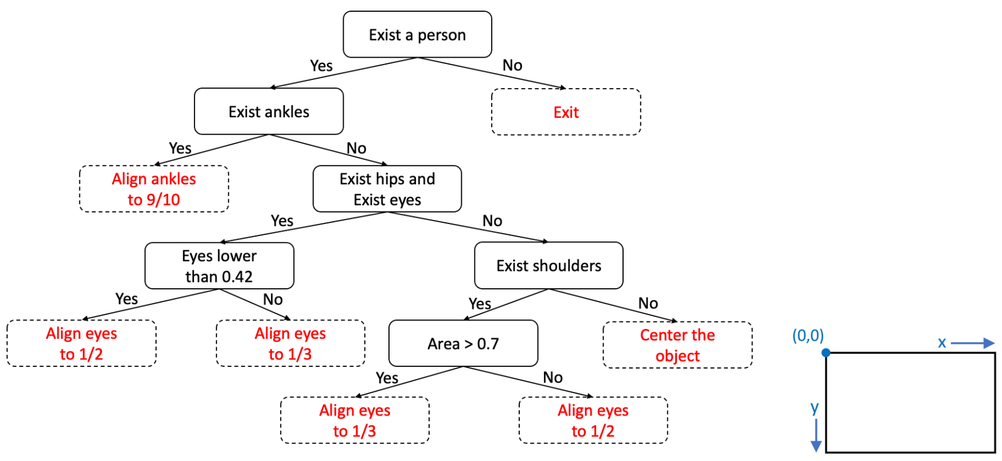
After identifying and locating the person, we need to analyze the key points of the person. By knowing the position of the key point, we can estimate the pose, providing more information to our composing algorithm. Here, we choose [3] as our pose estimator. It runs not only fast but also lightweight.

Kalman Filter

To solve the problem of unstable detection results caused by motion blur. [1] provides a solution. It is a pragmatic approach to multiple object tracking, combining the Kalman filter and the Hungarian Algorithm to achieve object tracking with high performance.

Composing Algorithm

We emphasized applying the rule of thirds on different conditions with a single person. The decision tree is the idea we used here to implement the composing algorithm. We express a person with a set of coordinates and normalize them between 0 and 1. Then, the decision tree is built to implement the composing algorithm.



(a)　(b)

Figure 2: (a) Y-axis decision tree (b) Coordinate system

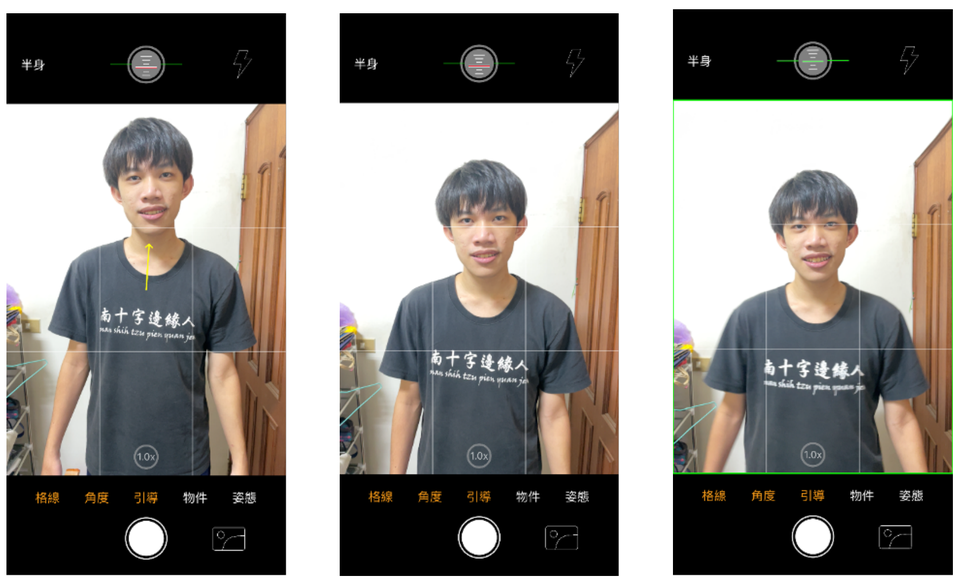
Figure 2(a) shows the y-axis decision step, and the number in the figure is a normalized y-coordinate. To achieve a specific shooting effect, the shooting angle must be considered. We use pose estimation to achieve this correction and cooperate with the built-in electronic gyroscope to complete the angle recommendation to the user.

Experimental Results

Our system can operate in different user environments and supports multiple portrait modes, such as the full-body mode in Figure 3 and the half-body mode in Figure 4. The system will display the current recommended mode on the upper-left corner of the screen to assist users in adjusting for shooting.

(a) Initial　(b) Adjust position　(c) Adjust angle

Figure 3: Full-body, Outdoor



(a) Initial　(b) Adjust position　(c) Adjust angle

Figure 4: Half-body, Indoor

The yellow arrow and the red line in the gyroscope indicate that the position and the angle could be improved. After the adjustment finish, there will display a green box on the border of the frame.

QRCode for introduction and demo video:

Conclusion

In this work, we formulated the composing problem in terms of object detection, and pose estimation. We proposed a composing algorithm based on the decision tree to achieve the rule of thirds. Experimental results show our system works well as expected that obeying the rule of thirds.

References

[1] Alex Bewley et al. "Simple online and realtime tracking". In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE, Sept. 2016. doi: 10.1109/icip.2016.7533003. url: https://doi.org/10.1109%2Ficip.2016. 7533003.

[2] RangiLyu. NanoDet-Plus: Super fast and high accuracy lightweight anchor-free object detection model. https://github.com/RangiLyu/nanodet. 2021.

[3] Bin Xiao, Haiping Wu, and Yichen Wei. Simple Baselines for Human Pose Estimation and Tracking. 2018. doi: 10.48550/ARXIV.1804.06208. url: https://arxiv.org/abs/1804.06208.