How to Get the Best Group Photo

Changyeon Park
Dept. of Electrical and Computer Engineering
Seoul National University
blackco@snu.ac.kr

Chaehyun Jeong
Dept. of Computer Science and Engineering
Seoul National University
chjeong9727@snu.ac.kr

Yuri Jeon
Dept. of Computer Science and Engineering
Seoul National University
slsl3650@snu.ac.kr

Daniel Chew
Dept. of Computer Science and Engineering
National University of Singapore,
daniel.chew@u.nus.edu

1. Overview

When we take group photos, we often see someone who blinked his/her eyes, someone who is blurred because they move at that moment or someone who didn't do the promised motion (e.g. making a V sign). These types of people spoil a nice group photo. However, it is bothersome to find the best group photo by checking such people manually among the series of photos.

The purpose of our project is to find the best group photo among these series of photos. We will first detect the faces in the group photo, and then exclude the photos which has blinking or blurred image. Finally, we will check that all people in the photo has presented the same pose which was previously promised.

2. Criteria for Best Group Photo and Input Assumption

2.1. Three Criteria for Best Group Photo

- Everyone in the group photo are stationary.
- Everyone in the group photo are non-blinking.
- Everyone in the group photo take the same motion.

2.2. Input Assumption

Assume that enough input images are given and the images are sequential group photos.

3. Paper Survey

3.1. Optical Flow

Optical flow is a classical computer vision method for detecting movements. Open source in OpenCV [1] will be referenced. This source implements dense optical flow using the Gunnar Farneback's algorithm to find the velocity vectors of each pixel in series of images.

3.2. Face Detection

There are many face detection algorithms. Open source in OpenCV [2] will be referenced. This source uses Haar Feature-based Cascade Classifiers to detect faces in an image and outputs bounding boxes around individual faces.

3.3. Eye Blink Detection

Eye blink detection is a series of detections. One implementation is eye blink detection by Soukupova [3]. It first detects facial landmarks and detects if the eye is closed or open by SVM linear classifier that is trained on blinking and non-blinking patterns.

3.4. Pose Estimation

Pose Estimation is a general problem in computer vision where we detect the position and orientation of an object. It detects keypoint locations that describe the object. The typical expected output of pose estimation is simple skeleton of human body with connected joints. One of the most common models is the Multi-Person Pose Estimation model proposed by Perceptual Computing Lab at Carnegie Mellon University [4]. This model produces 2D locations of keypoints for each person in the image. It exploits two branch multi-stage CNN - one is for predicting confidence maps of body part locations and the other is for measuring degree of association between parts.

4. Key Features for the Experiment

4.1. Detect Unprepared People for Photos

People who are ready for photo taking stay still and do not move for a set period until the photo is taken. On the other hand, people who are not ready for photo taking tend to move a lot. (e.g. fixing their hair) By using optical flow method, we will measure each pixel's velocity. If the maximum velocity is greater than the threshold, it will be considered as an image containing people who are not ready.

4.2. Detect Faces

Using real time face detectors will make eye blink detection process easier as there are multiple people in

a single image.

4.3. Detect Blink of an Eye

With each person's face image, we will be able to detect blinking eyes. By calculating the feature of an eye, Eye-Aspect-Ratio (EAR), and using SVM classifier, we will be able to detect blink of an eye.

4.4. Same Motion Among All Participants

Here we will generate skeleton picture for each person in the group picture using Multi-Person Pose Estimation model which is mentioned above. The measurement of similarity between poses are still left as a challenge for us. At first trial, we will assume that the image of model pose is given as an input, so the similarity will be decided based on the model image. When the number of people whose pose is similar enough with model image equals to the whole people detected at face detection 4.2, it is plausible to determine that everyone is taking same pose. If we succeed in doing this, we will also pick one person from the group picture as a standard, rather than receiving additional input image.

5. Expected Results

Our goal is to find the best group photo among the given images. Through our application we will be able to get the result image which all people in the group photo not blinking their eyes, are not moving, and striking the same promised pose.

6. Dataset Description

- Pose estimation dataset: COCO/ MPII Human Pose Dataset [5]

7. Reference

[1] OpenCV Dense Optical Flow https://docs.opencv.org/3.4/dc/d6b/group__video__track.html#ga5d10ebbd59fe09c5f650289ec0ece5af

[2] OpenCV Face detection https://docs.opencv.org/3.4.2/d7/d8b/tutorial_py_face_detection.html

[3] Tereza Soukupova, et al. [Real-Time Eye Blink Detection using Facial Landmarks]. CVWW, 2016

[4] Zhe Cao, et al. [Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields]. CVPR, 2017.

[5] MPII Human Pose Dataset http://human-pose.mpi-inf.mpg.de/