Tracking People and Object for Autonomous Unmanned Aerial Vehicle using Face and Color Detection

Olarik Surinta

Multi-agent Intelligent Simulation Laboratory (MISL)
Faculty of Informatics, Mahasarakham University
olarik.s@msu.ac.th

Sanya Khruahong

Department of Computer Science and Information Technology
Faculty of Science, Naresuan University
sanyak@nu.ac.th

Present at 4th International Conference on Digital Arts, Media and Technology (ICDAMT2019) - January 30 – February 2, 2019

Outline

- Introduction
- Contribution
- Proposed People and Objects Tracking Algorithm
- Experimental and Result
- Conclusion

Introduction

- Recently, the tracking of people and objects detection are seen as the crucial factors for **surveillance in the smart-city**, and can be used to **increase safety** for the digital society.
- CCTV can support to roles of the guard for the surveillance system on every situation.
- However, it my be difficult for using CCTV to monitor on the **blind spot**. It needs some technology to support its.

Introduction

- We think - **Unmanned Aerial Vehicle (UAV),** called ***drone*** can solve the problem.
- **Why** - It can control by user or program for **an auto-flying** in the mission.

Introduction

- The developer can communicate by using a smartphone or computer via Wi-Fi signal that broadcasts from the AR.Drone.
- The camera on the UAV can be recording video with a resolution of 720p.

AR.Drone 2.0

Related works

- **The ViolaJones face detector** (Viola & Jones, 2004), called **Haar-cascade classifier,** is the well-known method that first proposed in the object, face detection, and pedestrain.
  - Right now, **Haar-cascade classifier** becomes a standard technique for face detection.
- **Haar-cascade classifier** computes feature vector based on the Haar feature. It calculates from the rectangle detector or sub-window.
  - The detector scans through the image with the size of *24×24 pixel resolution*. The size of the detector will increase by 1.2 times.
  - Then, the set of the feature vector is given to the AdaBoost classifier which is the weak classifier.
- This approach can process on real-time and get high precision.

Related works

- **Dalal and Triggs (2005)** proposed a feature extraction method called ***histograms of oriented gradients (HOG)*** and first designed for human detection.
- **Feature extraction:**
  - HOG is the robust method that calculate the gradient orientation of shape/edge image.
  - Then, the orientation can lead to create the histogram.
- **Classification:**
  - Then, the histogram is fed to learn with the support vector machine (SVM) using the linear kernel.
- **Right now:** many research uses the HOG-SVM method for face detection as well.

Related works

- King (2015) proposed **max-margin object detection (MMOD)** for finding the object.
- The non-overlapping sub-window image slide through image and uses the window scoring function to find the fewest possible detection mistakes.
- As for the feature extraction method, the spatial pyramid bag-of-visual-words model combined with the HOG descriptor are used.
- The MMOD method is performed well on the face detection dataset (FDDB)

Related works

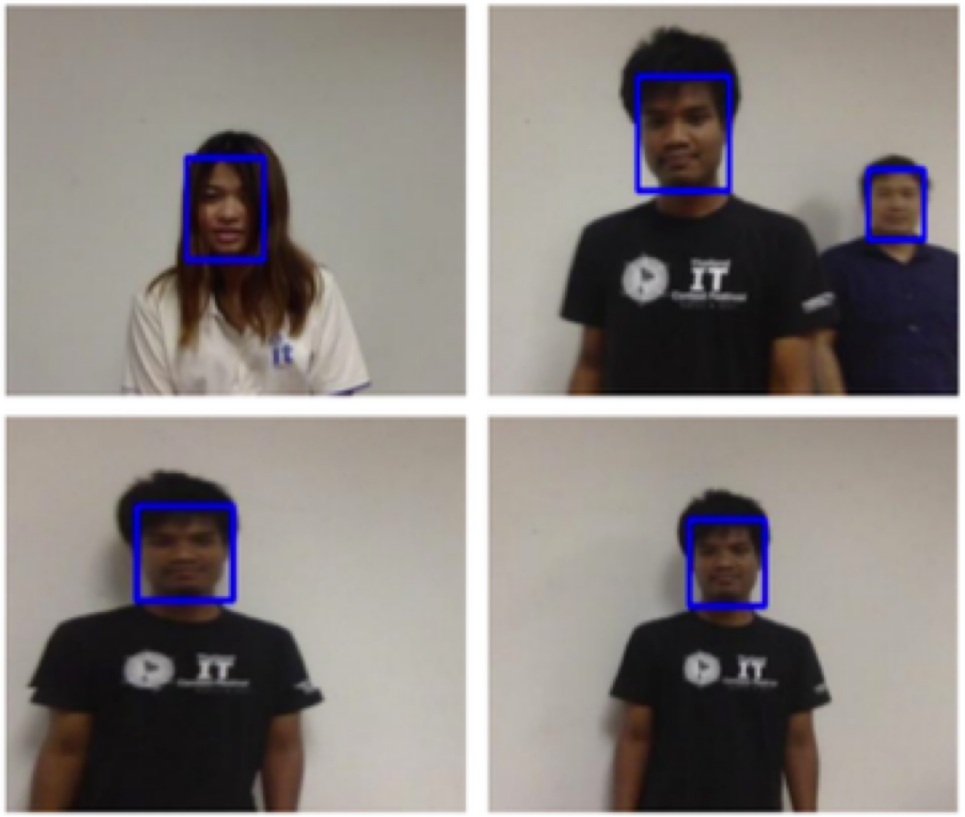
- **Binangkit and Widyantoro** (2016), a color image is converted into the HSV color space and extracted area of the traffic light color such as red, yellow, and green colors as the region of interest (ROI).
- The pixels of traffic light colors are represented as the set of feature vectors and then the feature vector send to the machine learning to create a model and classify.
- Also **Park et al.** (2017), the Haar-like feature and AdaBoost classifier was applied and reported to have high accuracy.

Contribution

- The aim of the study is a **detection development of UAV for object detection and tracking people.**
- The contribution has to approaches
  - First, **two well-known techniques of face detection** including Haar-cascade classifier and max-margin object detection with CNN based features (MMOD-CNN) are compared.
  - Second, our research focused on **the color detection of the objects**.
- These two approaches can lead to tracking the people and objects based on face and color.

Proposed People and Objects Tracking Algorithm

Haar-cascade classifier method



CNN-MMOD method

Proposed People and Objects Tracking Algorithm

- **Color Detection**
  - To detect candidate regions of interest (ROI) from the image.
  - The color image changes the color space from the RGB to the HSV color space.

Proposed People and Objects Tracking Algorithm

- **Tracking Algorithm**
  - The conception of our approach is using image processing and machine learning to control the UAV automatically by using face and object detection methods.

$$i = \frac{(x_2 - x_1)}{2}$$

$$j = \frac{(y_2 - y_1)}{2}$$

x1 and x2 are pixel position on x-axis and y1 and y2 are pixel position on y-axis.

Proposed People and Objects Tracking Algorithm

- **An illustration of finding the region of interest (ROI).**

(a) (b) (c)

(a) The original image presented with blue color
(b) converted to the HSV color space, then presented only the H value. The algorithm filtered the image containing the blue color and represented as white color on the black background.
(c) The ROI of the blue color image.

Proposed People and Objects Tracking Algorithm

- The algorithm is called "***turnDrone***" use to $C_{i,j}$ and $C_{h',w'}$ for turning the UAV, and the algorithm as shown on *Algorithm 1.*

Algorithm 1 turnDrone

Input: Set $f(h, w)$ to initial the window size, $C_{i,j}$ is the centroid of the ROI and $C_{h',w'}$ is centroid of the window

while $C_{h',w'} \neq C_{i,j}$ **do**
  if $c_i < c_{h'}$ **then**
    move drone down, update $C_{i,j}$
  else
    move drone up, update $C_{i,j}$
  end if
  if $c_i < c_{w'}$ **then**
    turn drone right, update $C_{i,j}$
  else
    turn drone left, update $C_{i,j}$
  end if
end while

Proposed People and Objects Tracking Algorithm

- The Algorithm 2 is called **movingDrone**, this algorithm used to support flying which compares the size of the ROI.

Algorithm 2 movingDrone

Input: Set T to initial acceptance ratio

Sum pixel of object: $Z_0 = \sum o(i, j)$

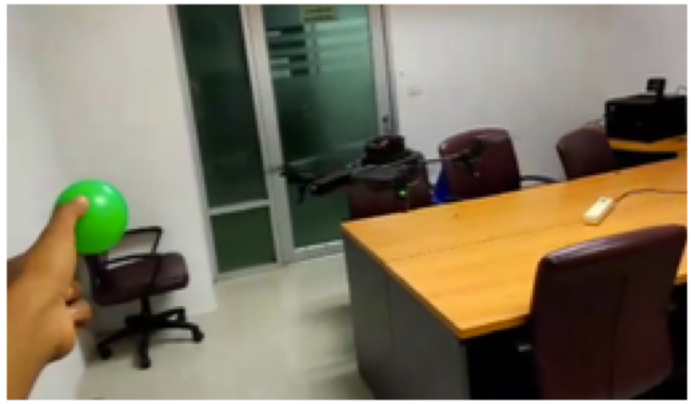
while $Z_o \neq T$ **do**
  if $Z_o > T$ **then**
    move drone backward, update Z_o
  else
    move drone forward, update Z_o
  end if
end while

Experimental and Results

- **First,** we used a web camera on a laptop computer to experiment with algorithms in the indoor environment.
  - The two face detection techniques performed the high precision of the face detection.
- **Second,** we performed face detection techniques using the frontal camera on the UAV.
  - Using images from UAV, we found that the result was not accurate detection, because of having some problem such as
    - The quality of images that obtained from the UAV
    - Sometimes lost the Wi-Fi- signal.

Experimental and Results

Experimental and Results



Experimental and Results

TABLE I: The range of HSV color space used in this research.

Color	RGB value	HSV Lower Range	HSV Upper Range
Orange	40, 99, 225	0, 100, 100	20, 255, 255
Blue	142, 52, 18	102, 100, 100	122, 255, 255
Green	65, 178, 141	30, 100, 100	50, 255, 255

(a)

The original image presented with orange color.

(b)

The binary image that containing both the orange color and skin

Found the error occurs when finding the ROI from the orange color image

Experimental and Results

- **Face detection experiment:**
  - The image from the web camera can be used for face detection and obtained high accuracy.
  - However, one important factor for testing of the face detection is **the UAV model**.
  - We used AR.Drone 2.0 model which the camera cannot change the angle view. So, the appropriate height to use the UAV for face detection is between 150 and 190 centimeters.

Experimental and Results

- **color detection experiment:**
  - The result found that green and blue color obtained the better result than the orange color.
  - Because the orange color is closely to skin human.
- In addition, the face detection needs more computation time than the color detection and needs longer time to track people.
  - Because the UAV has to move and turn during tracking

Conclusion

- In this paper, we proposed the approaches to tracking people and object by using the auto-flying control of the UAV.
  - The Haar-cascade classifier and MMOD-CNN methods are used for face detection.
  - For object detection, the HSV color space is used to find the ROI.
- The result showed that the color detection outperforms the face detection due to computation time.

THANK YOU!

Thank you for your kind attention