Proposal Presentation – Research Topics

CIS 465 Topics on Computer Vision

1. Face Recognition (Up to 3 students)

Background: Among the different biometric techniques, face recognition may not be most reliable and efficient. However, one key advantage is that it does not require the cooperation of the test subject to work. Properly designed systems installed in airports, multiplexes, and other public places can identify individuals among the crowd, without passers-by even being aware of the system. Other biometrics like fingerprints, iris scans, and speech recognition cannot perform this kind of mass identification. However, questions have been raised on the effectiveness of face recognition software in cases of railway and airport security?

Problem: In this project, we will explore using face recognition in real-world scenarios, and recognize faces in real-time through a web-camera. The identity of the face is expected as the output of the system.

Proposed solutions:

(1) Use web-camera to record the videos
(2) Detect the faces in the videos
(3) Extract facial features from the images (LBP, SIFT, VGG-Face, etc.)
(4) Use the classifier (SVM, kNN, etc.) to find the identity

Requirement:

(1) Any state of the art work? Is there any good library (Python, MATLAB, Java)? Please do a quick survey and list them
(2) What are the limitations of the current solutions? How would you solve it?
(3) Could you use both tracking and detection results to improve the quality of detected faces?
(4) Can you make an end-to-end system including *user registration* and *intruder's recognition*?

2. Human Re-Identification (Up to 3 students)

Background: In the modern computer vision community, the task of person re-ID shares similar insights with the old times. In video surveillance, when being presented with a person-of-interest (query), person re-ID tells whether this person has been observed in another place (time) by another camera. The emergence of this task can be attributed to 1) the increasing demand of public safety and 2) the widespread large camera? networks in theme parks, university campuses and streets, etc. Both causes make it extremely expensive to rely solely on brute-force human labor to accurately and efficiently spot a person-of-interest or to track a person across cameras.

Problem: In this project, we will explore the problem of person reidentification using surveillance videos of pedestrians. Assume several videos of each person are recorded, and your task is to use re-ID algorithm to retrieve the related pedestrian from an image database based on the current query pedestrian image detected from the surveillance video.

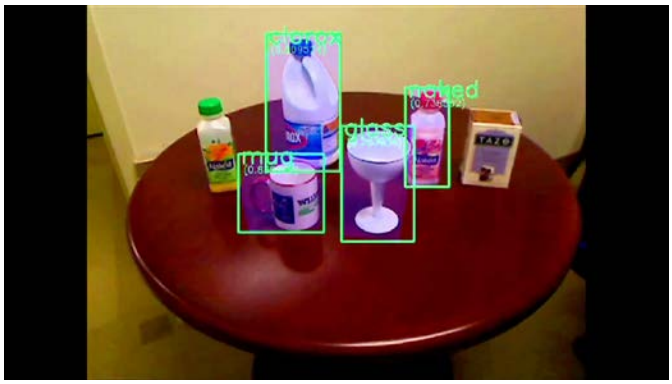
Proposed solutions:

(1) Use web-camera to record the videos to simulate the surveillance environment
(2) Track and detect the pedestrian in the videos
(3) Extract human body features from the images (LBP, SIFT, ResNet, Inception, etc.) within the bounding-box
(4) Use the certain distance metrics to retrieve the related pedestrian images.

Requirement:

(1) Any state of the art work? Is there any good library (Python, MATLAB, Java)? Please do a quick survey and list them
(2) What are the limitations of the current solutions? How would you solve it?
(3) Could you use both tracking and detection results to improve the quality of detected human body?
(4) Can you make an end-to-end system including both *user registration* and *human re-ID*?

3. Object Recognition (Up to 3 students)



Background: Object recognition is technology in the field of computer vision for finding and identifying objects in an image or video sequence. Humans recognize a multitude of objects in images with little effort, despite the fact that the image of the objects may vary somewhat in different viewpoints, in many different sizes and scales or even when they are translated or rotated. Objects can even be recognized when they are partially obstructed from view. This task is still a challenge for computer vision systems. Many approaches to the task have been implemented over multiple decades.

Problem definition: In this problem, we will design an object recognition algorithm that can recognize different/multiple objects captured by the web-cam. Given an arbitrary object in front of the camera, the system will detect and identify the category of the objects. If the object has never been registered before, the system should be able to handle this case by outputting *unknown* labels.

Proposed solutions:

(5) Use web-camera to record the videos to simulate the application environment
(6) Track and detect the objects (Faster R-CNN) in the videos
(7) Extract objects features from the images (LBP, SIFT, ResNet, Inception, etc.) within the bounding-box
(8) Use the classifier (SVM, kNN, etc.) to find the category of the objects.

Requirement:

(5) Any state of the art work? Is there any good library (Python, MATLAB, Java)? Please do a quick survey and list them
(6) What are the limitations of the current solutions? How would you solve it?
(7) There are many good object detectors using deep learning that you can take advantage of. Which one do you prefer?
(8) Can you make an end-to-end system including both *object registration* and *object recognition*?

4. Oculus Rift + Recognition Tasks (Up to 6 students, 30% additional credits)

Background: The Oculus Rift is a virtual reality head-mounted display. Software, most notably video games, must be custom programmed to use the Rift. While it is mostly employed in the video games for better experience, some recent released software such as *Virtual Desktop* has casted light on much more general using case. With *Virtual Desktop*, everything runs on the computer can be easily rendered on the VR glass in a 360 way, e.g., 360 photos, videos, multiple displays. It may offer an intuitive and efficient manner for multi-task users, especially those visual data driven tasks.

Problem definition: In this problem, we will study using the VR devices, such as Oculus Rift to render and visualize multiple interfaces of video surveillances in an efficient manner to minimize the efforts of technician in charge of video screening. Several real-time surveillance applications such as face recognition, human re-ID at different places will be simultaneously shown on the VR devices, where users can easily get access to each of them.

Proposed solutions:

(1) A few team members will work on recognition tasks like other teams
(2) Figure out the basic operations of Oculus Rift
(3) Deploy the applications and render them on Oculus through *Virtual Desktop*
(4) Collect the user responses to optimize the user experiences

Requirement:

(1) Same requirement for other recognition tasks, e.g., face/object recognition, human re-ID
(2) Excellent user experience by optimizing the: (i) surveillance APP interface, (ii) layout of multiple interfaces, (iii) design for basic user interface in 360 environment.