PLTW COMPUTER SCIENCE

Activity 3.2.2

Image Data

Introduction

People with sight obtain a huge percentage of their knowledge about the world from their vision. Using artificial intelligence to understand what objects are being seen and where they are, where they will be in the future if they are moving, and how to integrate vision with robotic motion – these are problems at the edge of computer science, areas of active research. Face recognition is among the most difficult of the problems.

Image Courtesy Microsoft ©2012

Are you ready for the day when you will live seamlessly with the computational power and data storage to identify any person you will ever see?

Are you ready to be “remembered”?

Materials

- FaceL Facile Face Labeling software from http://sourceforge.net/projects/pyvision/files/FaceL/
- Webcam
- Optional: wigs, glasses, hats, work lights

Procedure

In this activity you will be using face recognition software to recognize people and emotions.

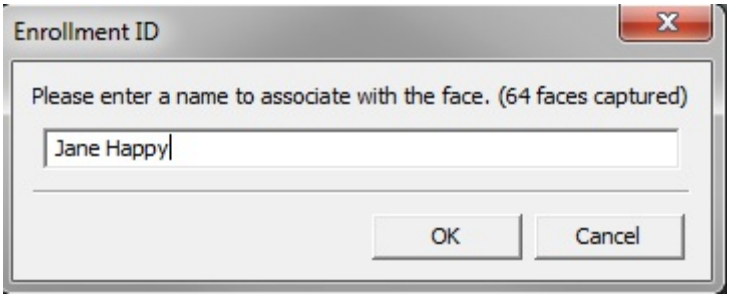
1. Form pairs as directed by your teacher. Meet or greet each other to practice professional skills. Set team norms.
2. Discuss the following article with your partner. Describe some of the challenges of face recognition. http://www.pbs.org/wgbh/nova/next/tech/the-limits-of-facial-recognition/
3. Launch the Face-L program. Video from your camera should be visible in the program, with a box around each face.
4. Change the **Enrollment Count** to 128. This will determine how many face images per identity are used to train the algorithm.
5. In many data analysis tasks, you want to determine one variable from other variables. Here we want to recognize a person’s identity from their picture. For machine learning, you need a training set of data in which you give the computer the answer. You might also have a validation set in which you know the answer so that you can see how well your program does

in predicting it. You will provide training sets with 128 images in each of four categories:

- Person #1, angry
- Person #1, surprised
- Person #2, angry
- Person #2, surprised

You could use any two emotions: sad, excited, confused, frustrated, and so on. Provide each training set by clicking on a person's face in the camera image. That person should look at the camera and think about something that makes them either angry or surprised. the person can experiment with nodding his or head head in a slight circle to provide a variety of viewing angles in the training set.

After each set has been collected, provide the person's name and emotion.



6. After all four sets of images have been captured, select the **Train Labeler** button. The software will take two or three minutes to create an algorithm for recognizing these faces.
7. Experiment with the software.
   - As you read, face recognition is difficult when a face is viewed from the side, in profile. How many degrees can you turn and still be recognized?
   - Does changing glasses, hats, hoods, or your hair confuse the face labeler?
   - Try drawing a face on paper. Does it recognize this as a face?
   - Does the computer appear to be better at recognizing a person or at recognizing emotions? Why do you think this might be?
   - Try getting closer or further from the camera. How does this affect the success rate of the face labeler?
   - Try turning off the lights in the room or using additional lighting. Under what circumstances does this undermine the face labeler's algorithm?
8. This software separates the tasks of deciding that there are faces in an image from deciding whose faces they are.
   - Under what circumstances will the labeler not even notice that there is a face in an image?
   - Why do you suppose these two tasks are separated in the software's code?
9. The software uses the Python Imaging Library (PIL) to calculate a "face chip" for each face in each image. A face chip is 60 x 100 pixels with the eyes in a standard position. Each face is compared to the training set and scored for similarity. Face recognition is much more difficult in live video than in video that has been stored. Estimate how many instructions the processor must execute between video frames (0.03 seconds) to perform live face recognition. Explain your estimate.

Conclusion

1. As you have learned, human brains are much better than computers in many ways for extracting knowledge from data. **Strong artificial intelligence** refers to artificial intelligence with generalized learning ability exceeding the computational power of the human brain. Simulating all electrical acitivity in the human brain is estimated to require 10^{16} calculations/second and 10^{13} bits of storage. It is predicted to be achieved near the year 2035. Strong AI is the opposite of **narrow AI**, in which learning occurs only to accomplish a narrow task. Is face recognition strong or narrow AI? Explain why you think so.
2. Describe some applications you can envision for software that can recognize people's faces and people's emotions.