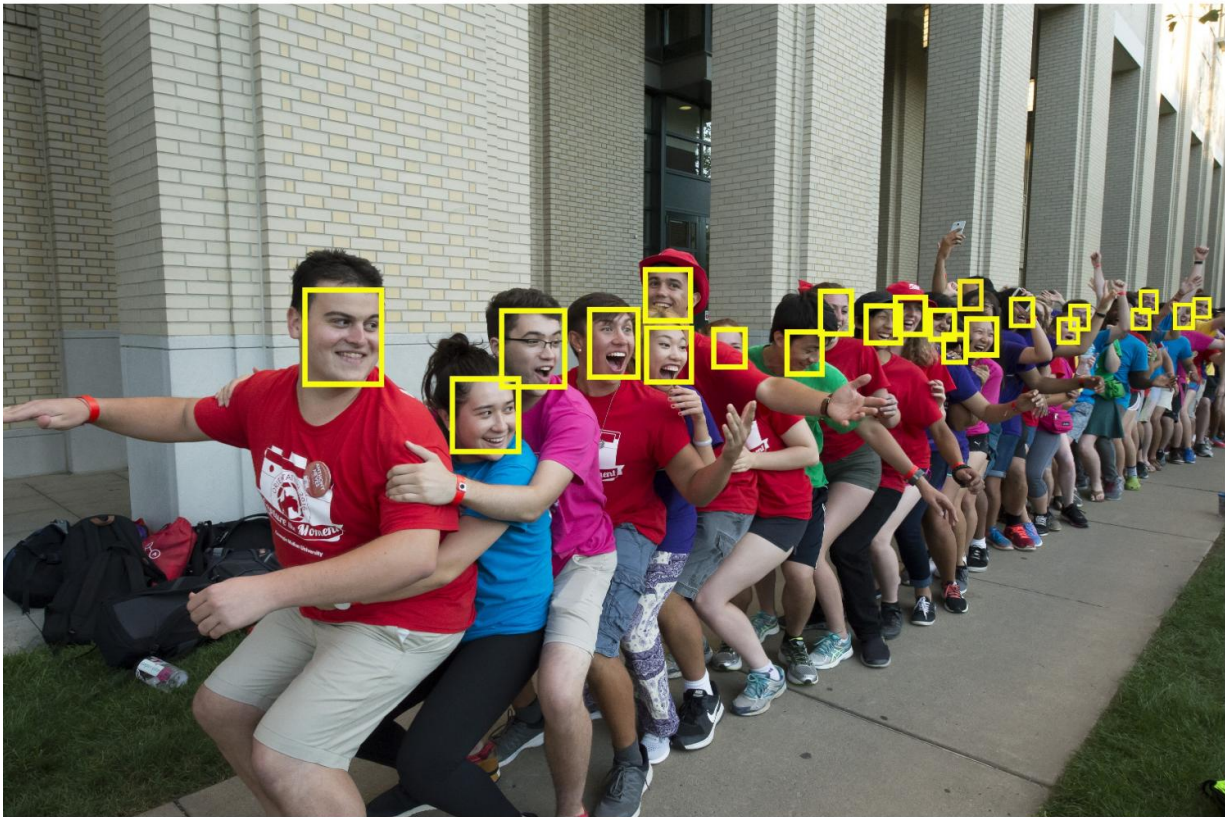


Finding faces in a crowd: Context is key when looking for small things in images

March 30 2017, by Byron Spice



An automated face detection method developed at Carnegie Mellon University enables computers to recognize faces in images at a variety of scales, including tiny faces composed of just a handful of pixels. Credit: Carnegie Mellon University

Spotting a face in a crowd, or recognizing any small or distant object within a large image, is a major challenge for computer vision systems. The trick to finding tiny objects, say researchers at Carnegie Mellon University, is to look for larger things associated with them.

An improved method for coding that crucial context from an image has enabled Deva Ramanan, associate professor of robotics, and Peiyun Hu, a Ph.D. student in robotics, to demonstrate a significant advance in detecting tiny faces.

When applied to benchmarked datasets of faces, their method reduced error by a factor of two, and 81 percent of the faces found using their methods proved to be actual faces, compared with 29 to 64 percent for prior methods.

"It's like spotting a toothpick in someone's hand," Ramanan said. "The toothpick is easier to see when you have hints that someone might be using a toothpick. For that, the orientation of the fingers and the motion and position of the hand are major clues."

Similarly, to find a face that may be only a few pixels in size, it helps to first look for a body within the larger image, or to realize an image contains a crowd of people.

Spotting tiny faces could have applications such as doing headcounts to calculate the size of crowds. Detecting small items in general will become increasingly important as self-driving cars move at faster speeds and must monitor and evaluate traffic conditions in the distance.

The researchers will present their findings at CVPR 2017, the Computer Vision and Pattern Recognition conference, July 21-26 in Honolulu. Their research paper is available online.

The idea that context can help [object](#) detection is nothing new, Ramanan said. Until recently, however, it had been difficult to illustrate this intuition on practical systems. That's because encoding context usually has involved "high-dimensional descriptors," which encompass a lot of information but are cumbersome to work with.

The method that he and Hu developed uses "foveal descriptors" to encode context in a way similar to how human vision is structured. Just as the center of the human field of vision is focused on the retina's fovea, where visual acuity is highest, the foveal descriptor provides sharp detail for a small patch of the image, with the surrounding area shown as more of a blur.

By blurring the peripheral image, the foveal descriptor provides enough context to be helpful in understanding the patch shown in high focus, but not so much that the computer becomes overwhelmed. This allows Hu and Ramanan's system to make use of pixels that are relatively far away from the patch when deciding if it contains a tiny face.

Similarly, simply increasing the resolution of an image may not be a solution to finding tiny objects. The high resolution creates a "Where's Waldo" problem—there are plenty of pixels of the objects, but they get lost in an ocean of pixels. In this case, context can be useful to focus a system's attention on those areas most likely to contain a face.

In addition to contextual reasoning, Ramanan and Hu improved the ability to detect tiny objects by training separate detectors for different scales of objects. A detector that is looking for a face just a few pixels high will be baffled if it encounters a nose several times that size, they noted.

Provided by Carnegie Mellon University

Citation: Finding faces in a crowd: Context is key when looking for small things in images (2017, March 30) retrieved 6 May 2024 from <https://phys.org/news/2017-03-crowd-context-key-small-images.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.