The World Seen Through Image Processing: A Technology for UAVs

By Alex Daniels, ECE '16

Image processing is used in many different platforms and all these platforms tend to focus on one thing: allowing humans to see things they are unable to see with their own eyes. One of the technologies that incorporate image processing is Unmanned Aerial Vehicles (UAVs).

What exactly image processing?

Imagine you're walking down the street and see a group of your friends waiting for the bus. How do you know those are your friends and not someone else? It might be because they have a nametag or a very obvious sign that follows them saying "My name is Kevin" but odds are it might be because you've seen them before and you remember them. It's very easy to for us humans to look at an image of a person and then point them out in other images or scenes. We learn how a person looks and can recognize them in various settings. Recognition isn't something we think about too deeply since this comes naturally to us. But for computers, this ability is common. How can we tell a computer that one image of a dog facing left and an inverted image of the dog is the same dog? How can computers gain the ability of recognition? One means is image processing.

Image processing takes an image of one form (whether it's a digital or physical image) and converts it into a different form. Transforming an image, such as a rotation, on a computer is one example of image processing, since the normal image is converted to another form. Usually, these conversions are for the purpose of gaining a better interpretation of the image. When a person rotates an image it is usually because the rotated version is more appealing; the dimensions of the people and objects in the image are more attractive. This idea of interpretation allows computers to gain the ability of recognition through the use of various algorithms.

Different interpretations exist for different situations. One area of interest in image processing is robotics, specifically how can we give such a technology the ability to navigate on its own as if it had its own set of eyes. With the use of a camera and a set of algorithms,

Figure 1: Example of template matching to find coins

we can collect images and convert them into an interpretation the robot can understand. One way to teach robots how to interpret these images is through the use of *computer vision*. Computer vision is defined as a discipline that studies how to reconstruct, interpret, and understand a 3D scene from its 2D images (Fisher, Dawson-Howe, Fitzgibbon, Robertson, Trucco, & Wiley, 2005).

The use of computer vision in computers includes object recognition, identification, or facial recognition, and detection. It works by getting many images from the camera and stitching them together using various transformations to create an object that it can access at any time. This would be equivalent to looking at a person's facial features to determine their identity.

A popular way to find an object within a larger image is through template tracking. The image of the desired object is considered the template and the large image is scanned for a match in pixels. The algorithm takes the template and pans across the original image until a match is found. Figure 1 shows an example of template matching being used to find an object, in this case coins.

A method of teaching identity is through feature tracking algorithms. Feature-based tracking algorithms detect objects within an image by finding corners, also known as key points using an algorithm called Harris-Stephens Corner Detector. For a computer, a corner has the most noticeable change in pixel intensity since it has a significant change in all directions it is viewed from. The image is then marked with key points. If a computer has a collection of images, it could determine which images are the same simply by comparing key points.

How are these interpretations used?

The interpretation a computer gives of data depends heavily on what this interpretation is used for. For example, an ultrasound scan sends high frequency sound waves that can go through soft tissue and fluids yet bounces off denser parts of the body. The machine detects the bouncing back signals and uses an algorithm, or set of calculations, that determines how much sound bounced back. It then displays an image where the more sound that bounces back, the grayer the image displayed is. From the machine's perspective, it only knows that more sounds means a grayer image. However, the interpretation is given to humans who use their interpretation abilities to look at this image and determine what the image depicts. In some cases, a human may be able to use the computer's interpretation. In these cases, the computer system will need to convert its own data interpretation into a platform that another system can understand.

Conclusion

The Gold Team needs to select an algorithm which detects obstacles using a camera mounted on a UAV. To accomplish this, the team utilizes feature based algorithms based on pixel grayness to see how close an object is to the UAV. Figure 3a shows an image seen by the drone's camera displayed for a human to see while Figure 3b shows the image as a gray map that the drone software needs to interpret. Through the combination

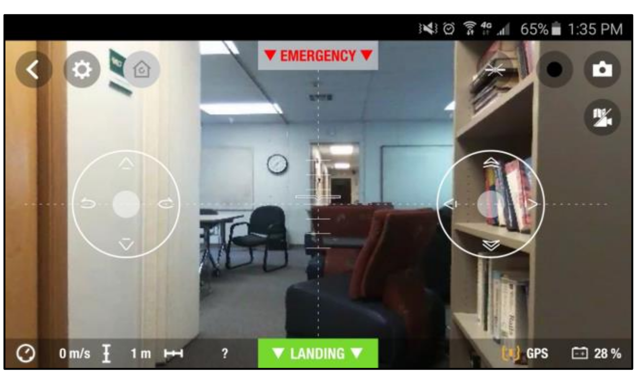

Figure 2a: Image as seen by the drone camera

Figure 2b: Image as interpreted by the drone software

of image processing tools with various proximity sensors, the drone can fly autonomously and stop itself before running into objects.

References

Fisher R., Dawson-Howe K., Fitzgibbon A., Robertson C., Trucco E., & Wiley. (2005). *Dictionary of Computer Vision and Image Processing* (2nd ed.).

Fu L., Hsiao P., Lu C. (2010). Multilayered Image Processing for Multiscale Harris Corner Detection in Digital Realization. *IEEE Transactions on Industrial Electronics*, *57*(5), 1799-1805. doi: 10.1109/TIE.2010.2040556

OpenCV-Python Tutorials. (n.d.) Template Matching. *OpenCV-Python Tutorials 1 documentation* (Revision 43532856). Retrieved from http://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_imgproc/py_template_matching/py_template_matching.html

Riccabon, M. (2014). Basics, principles, techniques and modern methods in paediatric ultrasonography. *European Journal of Radiology*, *83*(9), 1487-1494. doi:10.1016/j.ejrad.2014.04.032