DETERMINING AN INDIVIDUAL'S 3-DIMENSIONAL BODY SHAPE FROM TWO 2-D PHOTOGRAPHIC IMAGES

David Evans and Eleanor Watson
Poikos Ltd

The overall aim is to use two digital photographs of a person to compute the person's 3-D shape. The photographs are taken with 'ordinary' equipment; a monocular lens in a smartphone, or a web-camera in a lap-top computer would be typical. There is no specification on the brand of equipment.

The individual is requested to wear close-fitting clothing and to adopt particular poses for the two photographs, at a distance of about 3m from the camera. One image is face-on to the camera, and the other is side-on to the camera. Poikos ask the person his/her height for a sense of scale.

Poikos have broken the problem of deriving the 3-d shape from the 2-d photographs into a series of smaller problems. The first of which is to find the edge of the person in an image. That is, they wish to separate the person, "foreground", from the rest of the image, "background". This is achieved by comparing the image against a pre-calculated probability map for the pose, combined with colour information from the image and edge detection. Once the 2D silhouette for each input image is identified, it is compared against known 3D shapes using a similar system to 'nearest match' to convert the silhouettes into a three dimensional scan. Once a good fit is found, refinements are made to tune the shape to the subject.

What is the 'prior' information?

We should assume that the image contains one, and only one, person who is cooperating by standing in roughly the right pose, at roughly the right distance from the camera.

Poikos have a database of manually checked images. From these they have derived a kind of template which shows how likely it is, at any point in the image, for a pixel to be part of the person's body (and, therefore, not part of the background). The pixels that are highly likely to be foreground are used to determine the colour of the clothes that the person is wearing. That information is used, in turn, to help find the edges of the person.

Main difficulties

Essentially, Poikos have no control over the person's clothing, the background, the lighting, the precise distance from the camera, the angle of the camera or the accuracy of the height measurement. Indeed, people will, in general, claim that their height is closer to a perceived ideal than is actually the case. Is there a more reliable means of estimating a person's height? Although one could ask the subject to hold a 'standard object' (such as a piece of A4 paper), what if that object is not readily available? Is a 'markerless' form of scaling possible?

Any method should not be too computationally intensive: it would be desirable to get a 3-d shape estimate within seconds. The method should also be sufficiently accurate to estimate a person's waist-line to within a centimetre of the real measurement.

When presented with an image, Poikos find the head in the image first, use the head to determine the individual's skin tone, and then, use that information to locate the hands in the image. The hand location can then be used to infer other 'summary statistics' about the body shape. Could this approach be too wasteful of information, or are we retaining sufficient information at each stage of this type of process?

Study Group Problem

In February 2012, Poikos previously presented this problem at a workshop in OCIAM and received several interesting suggestions for future development. Based on these suggestions, Poikos would like to invite the Study Group to challenge their existing approach by investigating methods that combine the segmentation and 3D reconstruction (stereology) into a single process.

In addition, Poikos would also be interested to learn of methods that can be used to reduce errors associated with camera defects such as:

- Misaligned CCDs – false perspective-like distortions caused by a misalignment of the photoreceptors.
- Barrel distortion – The current approach makes a rough estimate of barrel distortion, but Poikos would like to look at fine-tuning this to allow for detection of more subtle effects.

In each of these cases, the functional form of the image distortion is known, but Poikos would like to learn of techniques to identify the unknown parameters that define the distortion in a specific image.

Furthermore, Poikos would like to hear ideas for means of determining scale without the need for standardised objects to be present in the image.