MEMORY-BASED FINE SCALE ORIENTATION DETECTION

Paul Fitzpatrick

MIT AI Lab
Cambridge, MA 02139, USA
paulfitz@ai.mit.edu

ABSTRACT

Orientation information is present in images at all scales. It is typically detected using quadrature filters applied to windows of many sizes. This paper presents a complementary approach suitable for use with window sizes that are too small to support such filters. The appearance of oriented features at fine scales can be characterized empirically by collecting labelled examples of the appearance of edges at all orientations. Since the window sizes of interest are small, it is practical to sample the space of all possible appearances quite densely. In this work, the necessary experiments are performed automatically by a robotic system using active segmentation. This data is then used to perform orientation detection with a very small window size (4 by 4 pixels) using a simple interpolating look-up table. This provides access to orientation information contained at the finest scale visible, which would be drowned out in larger window sizes.

1. INTRODUCTION

Orientation is an important visual cue for many purposes, such as object segmentation, recognition, and tracking. It is associated with neighborhoods rather than individual points in an image, and so is inherently scale dependent. At very fine scales, relatively few pixels are available from which to judge orientation. Lines and edges at such scales are extremely pixelated and rough. Orientation filters derived from analytic considerations, with parameters chosen assuming smooth, ideal straight lines or edges (for example, [1]) are more suited to larger neighborhoods with more redundant information. For fine scales, an empirical approach seems more promising, particularly given that when the number of pixels involved is low, it is practical to sample the space of all possible appearances of these pixels quite densely.

This paper is an exploration of how edges in "natural" images appear when viewed through an extremely small window (4 by 4 pixels). This window size is chosen to be large enough to be interesting, but small enough for the complete range of possible appearances to be easily visualized. Even at this scale, manual data collection and labelling would be extremely tedious, so a robotic system [2] was employed to automatically compile a database of the appearance of oriented features (Section 2). These features were extracted by sampling image patches along object boundaries, which were in turn determined using active segmentation [3]. The resulting "catalog" of edge appearances proved remarkably diverse, although the most frequent appearances were indeed the "ideal" straight, noise-free edge (Section 3). Finally, it is a simple matter to take this catalog of appearances and use it as a memory-based image processing filter (Section 4).

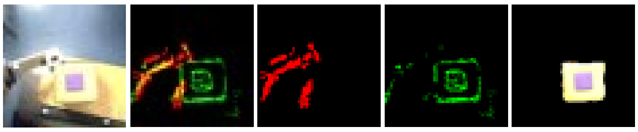

Fig. 1. Active segmentation. The robot arm is deliberately driven to collide with an object. The apparent motion after contact, when masked by the motion before contact, identifies a seed foreground (object) region. Such motion will generally contain fragments of the arm and environmental motion that escaped masking. Motion present before contact is used to identify background (non-object) regions. An optimal object region is computed from the foreground and background information using graph cuts [3].

2. DATA COLLECTION

A robot equipped with an arm and an active vision head was given a simple "poking" behavior, whereby it selected objects in its environment, and tapped them lightly while fixating them [4]. The motion signature generated by the impact of the arm with a rigid object greatly simplifies segmenting that object from its background, and obtaining a reasonable estimate of its boundary (see Figure 1). Once this boundary is known, the appearance of the visual edge between the object and the background can be sampled along

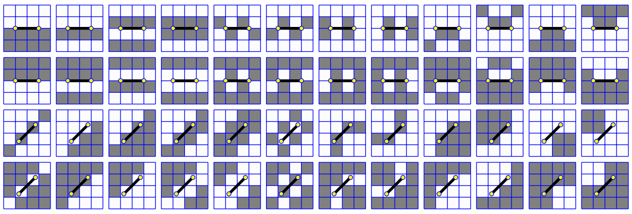

Fig. 2. Edges have diverse appearances. This figure shows the orientations assigned to a test suite prepared by hand. Each 4×4 grid is a single test edge patch, and the dark line centered in the grid is the orientation that patch was observed to have in the training data. The oriented features represented include edges, thin lines, thick lines, zig-zags, corners etc. It is difficult to imagine a set of conventional filters that could respond correctly to the full range of features seen here – all of which appeared multiple times in object boundaries in real images.

it. These samples are labelled with the orientation of the boundary in their neighborhood (estimated using a simple discrete derivative of position along the boundary). The samples are assumed to contain two components that are distinguished by their luminance. The pixels of each sample are quantized into binary values corresponding to above average and below average luminance. Quantization is necessary to keep the space of possible appearances from exploding in size. The binary quantization gives a very manageable 65536 possible appearances. About 500 object boundaries were recorded and sampled. 49616 of the possible appearances (76%) were in fact observed; the remaining 24% were all within a Hamming distance of one of an observed appearance. The orientations of these unobserved appearances was estimated using interpolation, by averaging the orientation of all the appearances within a Hamming distance of one. If the same appearance was observed multiple times, the orientations associated with these observations are averaged.

3. VISUALIZATION

In this section, the data collected is examined to determine how practical fine scale orientation detection could be. Figure 2 shows that although the data collection procedure operates on views of simple physical edges, the appearance of these edges can be quite complex. Nevertheless, the most common appearances observed are ideal, noise-free edges, as Figure 3 shows. The first four appearances shown (top row, left) make up 7.6% of all observed appearances by themselves. Line-like edges are less common, but do occur, which means that it is perfectly possible for the surfaces on either side of an edge to be more like each other than they are like the edge itself. This was completely serendipitous – it was anticipated that obtaining and automatically labelling such examples would be very difficult. Figure 4 shows the most frequently occurring image appearances with a particular orientation. Here it is clearer that the most frequent patches are generally "ideal" forms of the edges, followed by very many variations on those themes with distracting noise. Amidst the edge-like patterns are examples of a line with single-pixel thickness, and a pair of such lines running parallel. It is encouraging that examples of such appearances can be collected without difficulty and united with more classical edge patches of the same orientation.

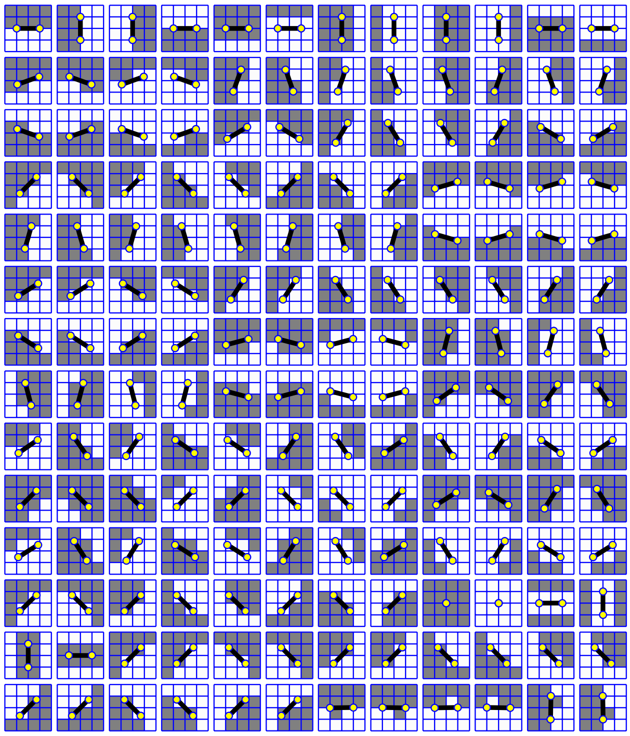

Fig. 3. The most frequently observed edge appearances. All patches observed are replicated for all 90° rotations, mirror flips, and inversion of foreground/background. The most frequent (top) are simple straight edges. The line in the center of each patch shows the orientation associated with that patch. After the straight edges, the completely empty patch is common (produced in saturated regions), followed by a tube-like feature (third-last row) where the boundary is visually distinct to either side of the edge. This is followed corner-like features and many thousands of variations on the themes already seen.

Figure 5 systematically explores the effect of adding noise to an "ideal" edge. The resilience of the orientation measure is encouraging, although a small number of gaps in coverage are revealed, suggesting that further data should be collected.

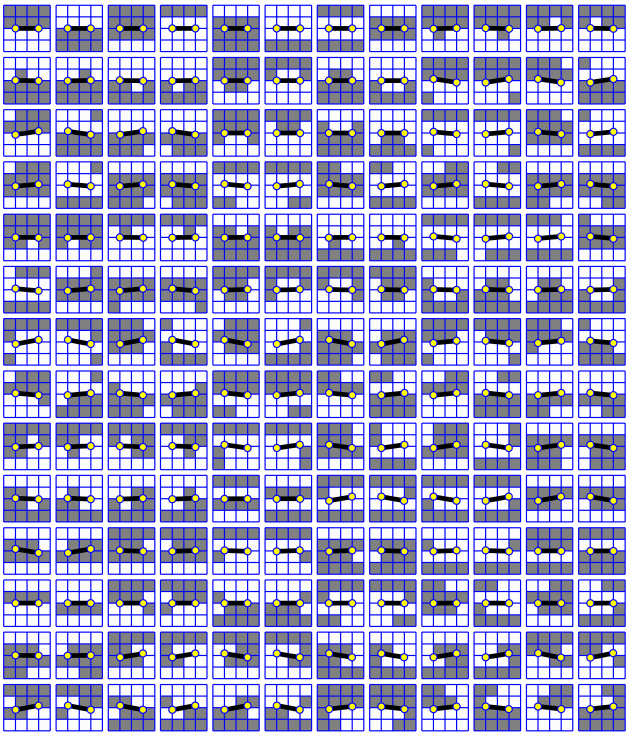

Fig. 4. The most frequently observed appearances whose orientation is within a few degrees of the horizontal. There is a clear orientation assigned to many patches that deviate a great deal from "ideal" edges/lines, showing a robustness that is examined systematically in Figure 2.

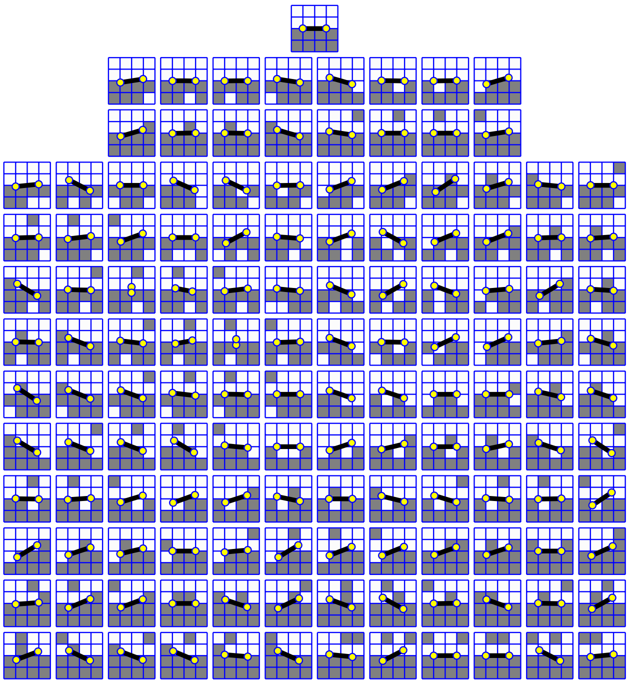

Fig. 5. The top row shows an "ideal" view of a horizontal edge. The next two rows show patches with a single pixel perturbation. The remainder of the image shows 2-pixel perturbations (25% of the pixel weight of the edge). The behavior of the orientation estimate is very intuitive, except in two cases where there was insufficient training data (a single sample).

4. FILTER IMPLEMENTATION

It is simple to use the data we have collected to filter an image for fine scale orientation features. A 4×4 window is moved across the image, sampling it as described earlier in Section 2. Each sample is used as an index into a table mapping appearance to orientation. Figure 6 shows the orientations measured for a 64×64 image consisting of a circle and square. This is based on an example in [5]. The detector gives good results for solid edges with arbitrary contrast, and various kinds of lines. The response to edges is diffuse by design – during data collection, samples are taken both along the boundary and slightly to either side of it, and treated identically. If a sharper response is desired, these side-samples could be dropped, or their offset from the boundary could be recorded. Figure 7 shows the filter operating on a very small image of a cube. Each visible edge of the cube is clearly and faithfully represented in the output. Typical filters with a window size of 8×8 and up would have little chance with such an image.

5. DISCUSSION AND CONCLUSIONS

Real images are more complex than we intuit. While this work has shown that idealized edges are empirically well grounded, in that they occur more frequently than other variants, it also shows that many more exotic forms do occur and can profitably be modeled. At fine scales, where the number of pixels used to compute orientation is low, a practical approach is to sample the appearance of edges empirically. Many elaborations on the implementation described in this paper are possible. For example, when quantizing the luminance values in a neighborhood, each possible setting could be tried for those pixels that are close to the threshold, and then the final orientation could interpolate appropriately.

With the large cache size of modern processors, this memory-based approach to orientation detection can facilitate extremely rapid orientation detection, which is important for real-time vision systems [7].

While no difficulty is anticipated sharpening the spatial response of the filter by associating offsets with appearances, this remains to be confirmed. This is important for certain applications such as edge detection.

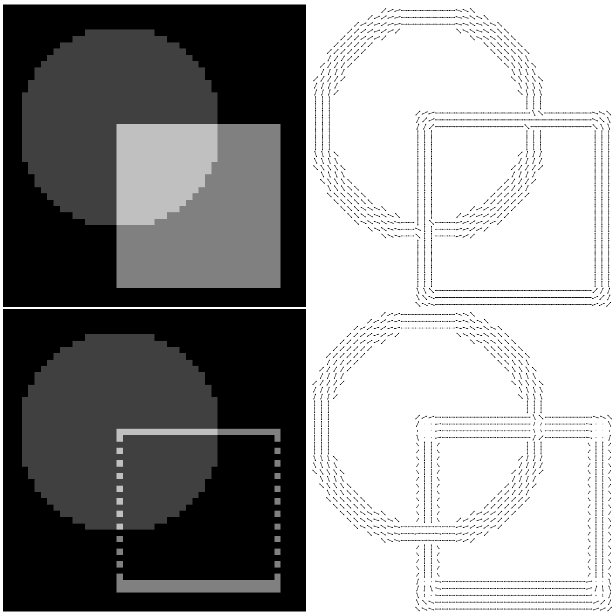

Fig. 6. A pair of simple test images (left), based on [5]. The oriented lines shown on the right represent the orientation assigned to each pixel in the corresponding image.

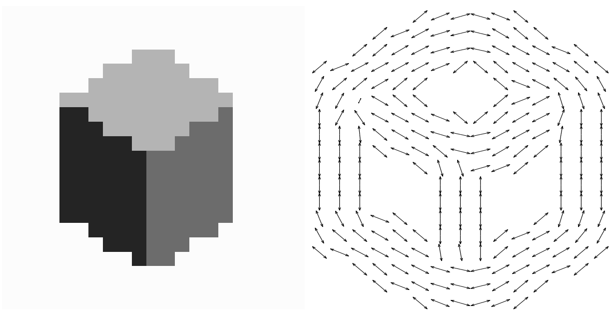

Fig. 7. A synthetic image of a cube. It is a scaled down version of that used in [6]. The length of each edge (8 pixels) is chosen to coincide with the smallest window size used in that paper. The results here are dense and well-behaved.

6. ACKNOWLEDGEMENTS

Funds for this project were provided by DARPA as part of the "Natural Tasking of Robots Based on Human Interaction Cues" project under contract number DABT 63-00-C-10102, and by the Nippon Telegraph and Telephone Corporation as part of the NTT/MIT Collaboration Agreement.

7. REFERENCES

[1] J. Chen, Y. Sato, and S. Tamura, "Orientation space filtering for multiple orientation line segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000.

[2] R. A. Brooks, C. Breazeal, M. Marjanovic, and B. Scassellati, "The Cog project: Building a humanoid robot," *Lecture Notes in Computer Science*, vol. 1562, pp. 52–87, 1999.

[3] P. Fitzpatrick, "First contact: Segmenting unfamiliar objects by poking them," submitted to IEEE Conference on Computer Vision and Pattern Recognition, 2003.

[4] P. Fitpatrick and G. Metta, "Towards manipulation-driven vision," in *IEEE/RSJ Conference on Intelligent Robots and Systems*, 2002.

[5] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.

[6] T. C. Folsom and R. B. Pinter, "Primitive features by steering, quadrature, and scale," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1161–1173, 1998.

[7] T. Kubota and C. O. Alford, "Computation of orientational filters for real-time computer vision problems i: Implementation and methodology," *Real-Time Imaging*, vol. 1, no. 4, pp. 261–281, October 1995.