Saliency Network

May 19, 2008

The problem of finding salient curves in images has a number of applications in computer vision. The saliency network [1] leads to a particularly elegant computational procedure for finding salient curves. Below we give a simplified description of the approach.

Let P be a set of points in the plane. Typically P would be the pixels in an image. We assume that there is a fixed number of oriented segments connecting each point in P to nearby points as illustrated in Figure 2. The set of segments coming out of a point p corresponds to different orientations that a curve can follow as it passes through the point. The set of all possible oriented segments is denoted by S. A curve is represented by a sequence of adjacent segments $(s_1, \ldots, s_n)$ as shown in Figure 3. We can think of P and S as a directed graph. Curves are paths in this graph.

The saliency of a curve is a score that increases with the curve length and decreases with its total curvature and "fragmentation".

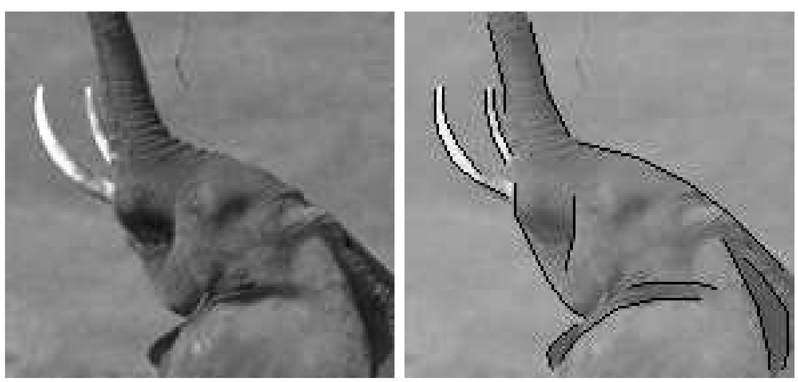

Figure 1: Salient curves found in the elephant image.

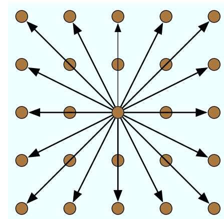

Figure 2: Example where there are 16 oriented segments leaving each point. The set S is the union of all oriented segments.

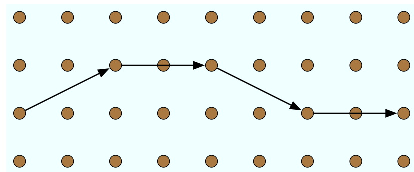

Figure 3: Curves are sequences of adjacent segments. The picture shows a curve formed by 4 segments.

Let $\sigma(s)$ be a measure of the local saliency of a segment $s \in S$. Intuitively $\sigma(s)$ should be high (positive) if there is evidence for a curve going throught s, and low (negative) otherwise.

If we are looking for curves in grayscale images the value of $\sigma(s)$ could depend on the magnitude and direction of the image gradient under s. There is evidence for a curve going through s if the gradient at each pixel under s has high magnitude and is nearly perpendicular to s.

Let $\psi(s, t)$ be a measure of the orientation difference between segments s and t. For example, we could define $\psi(s, t) = (\theta(s) - \theta(t))^2$.

The saliency of a curve $(s_1, \ldots, s_n)$ can be defined in terms of the local saliency of each segment in the curve and the orientation differences between adjacent segments,

$$\sigma(s_1, \ldots, s_n) = \sum_{i=1}^{n} \sigma(s_i) - \sum_{i=1}^{n-1} \psi(s_i, s_{i+1}).$$

Note how this measure increases with the length of a curve and deacreases with total curvature and fragmentation (here "gaps" in the curve can be defined in terms of segments with negative local saliency).

Let $\Phi_k(s)$ be the score of the most salient curve of length k starting at segment s. Note that a curve of length k starting at s is defined by s and a curve of length $k-1$ starting at a segment t that is adjacent to s. In fact, we can show that

$$\Phi_k(s) = \sigma(s) + \max_t \Phi_{k-1}(t) - \psi(s,t),$$

where the maximization is over segments t that are adjacent to s. For the base case we have $\Phi_1(s) = \sigma(s)$.

We can define the k-th saliency map as an image where the value of each pixel is the score of the most salient curve of length k leaving that pixel,

$$M_k(p) = \max_s \Phi_k(s).$$

Here the maximization is over segments leaving p. Note how the first saliency map $M_1(p)$ is similar to a local edge detector because the value of a pixel is defined by the local saliency of segments leaving that pixel. As k increases the k-th saliency map agregates information over longer and longer curves. Eventually $M_k(p)$ is high only at pixels that are in a salient curve.

References

[1] A. Shashua and S. Ullman. Structural saliency: The detection of globally salient structures using a locally connected network. In *ICCV*, 1988.