CS664 Lecture #2 | Machine Vision | 09/06/2001

Scribe: Chee Yong Lee | Lecturer: Ramin Zabih

- In the last lecture, we examine the bone fracture example. One student was asking about the need to do image restoration in this case. Why can't we just examine the x-ray intensity directly?

- We can appreciate the need for these image processing requirements from another application of motion detection. Given two images I(t) and I(t+1). These two images are roughly the same except for a slight change in the position of the moving object. In this application, we are interested in solving the *Correspondence Problem* of *what went where*. Such a problem can be viewed from the perspective of a vector field. The exploration of motion detection obtains its motivation from disciplines like medical science, where computation of the ejection fraction of the heart forms the basis for detecting certain heart malfunction.

- How are heart images obtained? The use of dye that has the ability to stop x-rays is one way to obtain x-rays diagrams of the human interior heart chamber tissues. X-ray angiography is another possible alternative where the camera is positioned in a similar manner as the x-ray set-up. MR (magnetic resonance) imaging is by far the safest way due to its low ionization effect on the human protein cells.

- Studies have been performed to show that people that have been subjected to 200 years of background radiation do not suffer side effects from excessive radioactivity exposure.

- For any scalars α, β, and images I, I', the operation B on images is a linear operation iff

  $$B(\alpha I + \beta I')[x] = \alpha B(I)[x] + \beta B(I')[x]$$

  B is shift invariant if for all Δ we have $SHIFT_{\Delta}(B(I))[x] = B(SHIFT_{\Delta}(I))[x]$.

- Convolution is an LSI, but any LSI is convolution

- Definition of unit impulse (discrete delta function) and UIR (Unit Impulse Response)

- Any input is the sum of shifted and scaled unit impulses

- Therefore, an arbitrary LSI system performs convolution with its UIR

- Review of image restoration in terms a *labeling problem*, find $f : P \mapsto Ł$, where Ł here is intensities.

- We want to give a labeling which is fairly smooth (say, piecewise constant). But that is not the only constraint.

- There is some notion that we ought to give pixels values that are close to the ones we observed. We can formalize this using a *cost function*.
- There is a cost to assign a pixel p a label $f(p)$, we will write this as $C(p, f(p))$.
- Typically, the cost function is 0 at the observed data and rises symetrically. Call the input data I; a standard cost function might be

  $$C(p, f(p)) = (f(p) - I(p))^2$$

  or

  $$C(p, f(p)) = |f(p) - I(p)|$$
- The cost function is related (we'll do this rigorously in a few weeks, after we briefly review probability) to the noise function of the camera. Simple example: suppose we knew the camera always added 1 to the true value (which is hardly noise).
- Common noise functions: gaussian (good camera, bad camera); salt and pepper; contamination model
- Stereo and motion overview; why they are similar (smoothing cost function for a given label)
- It is possible to view local filtering operations as doing *fitting*. In a fitting problem, we have some noisy data and we want to fit a model to it. The model has one or more parameters, i.e. free variables. (In a statistical setting, fitting is closely related to parameter estimation.)
- Standard example is fitting a line to data. There is a *residual* for each point, namely the difference between where it is (in the observed data) and where we expect it to be (in our model). Associated with a residual is the same kind of cost function we discussed in the context
- To view local filtering as fitting, we look at each pixel p, and assume that all the pixels in the window $W(p)$ will get the same value. So now we're looking for the best value.
- The best label i for p will be the one that minimizes some kind of cost, such as

  $$\sum_{p' \in W(p)} C(p', i)$$

  If we define $C(p', i) = (i - I(p'))^2$ (quadratic cost function), we get something we've already seen, namely the average.
- This is because given $\{x_i\}$ the z that minimizes $\sum_i (z - x_i)^2$ is the average of the x_i's.
- What about when the cost function is $\sum_i |z - x_i|$?

- In this case, given the set $\{x_i\}$, the z that minimizes $\sum_i |z - x_i|$ is the median of the x_i's. The median for this L1 fit will be the data point where both the horizontal and vertical co-ordinates are the median values of all the horizontal and vertical values of $\{x_i\}$ respectively.
- So we can view box filtering as fitting a constant line to each $W(p)$ using least squares, and median filtering as the same using least absolute values.
- What about non-uniform local averaging (i.e., with a Gaussian?) Weighted least squares.
- Robust statistics is a branch in statistics that deals with the situation where there is a mixture of both outliers (bad data points) and inliners (good data points).
- One method in robust statistics tries to get rid of the problem introduced by the squaring in LS fitting. This sort of approach is called M-estimation which utilizes the maximum likelihood of these data samples to fit the points.
- Influence function : The derivative of a quadratic cost function like the one in LS fitting is linear. This implies that bad data points that is far away from the actual line spoils the fitting by a great deal. Take for example Micheal Jordon's income totally offset the mean income of the Geology graduates of UNC 80's. The idea is to have a influence function that has a cutoff point, where the points beyond that are equally bad in regardless of their distance.
- Breakdown point is the fraction of bad data that can be introduced without effect to the fitting algorithm. M-estimation has a lousy breakdown point.
- There is a natural limit for this breakdown point. There exist an algorithm that achieves a result that is at least 50 percent of the optimal solution. Called the least median squared method, we can visualize this method as a way of finding the thinnest ruler with the smallest width w, such that it covers more than half of the available data.