In this exercise you will write a program to find an object of interest in an image. Your program should receive a sequence of images of the same scene with different viewing angles and scales. For example:

Your program should open image $I_1$ and prompt the user to select a region of interest $R$ using the mouse. It will then find and mark the region that is most similar to $R$ in each of the images $I_2, \ldots, I_n$.

As metric of similarity use:

1. The pixel-to-pixel difference in the regions.
2. The difference of the histograms of gradient of the regions. To obtain better results, you may want to partition the regions into subregions and use the concatenated histograms of the subregions as features.
3. Extra credit: SIFT descriptors, as explained in the textbook.

Your code should work at different scales. As we have discussed, this is attained by upscaling and downscaling images $I_2, \ldots, I_n$, while keeping $R$ constant.