Course notes, CS664, 8/31/04

- ADMINISTRIVIA: MEng project possibilities

STEREO

- Other vision problems often seem to have a more rigorous definition, but this is deceiving. Let us look at a few that seem like this (stereo, motion, SFM).

- Stereo problem definition: depth from multiple cameras (same time, or moving in a stationary scene).

- With 2 aligned cameras, disparity is inversely proportional to depth (try it and see).

- Human stereopsis is funny; 10% of the populace doesn't have it, the cameras are highly verged, it's really only used within arms length. But when it's useful, it's quite powerful.

- Stereo also points out yet another reason why vision is hard; the problem is inherently ambiguous.

- The world is 3D, the image is 2D.

- The fancy term for this is "ill-posed in the sense of Hadamard".

- In fact, you can't really tell **anything** about the 3D world from a 2D image. You could always be looking at a picture.

- More realistically, certain phenomena are always ambiguous and hard to interpret. Example: low-texture (note that both low- and high-texture cause problems!)

- In a certain sense, vision algorithms make up a plausible answer about the 3D world (this hints at a Bayesian approach, which is in fact quite popular).

- Stereo is also made difficult by subtle changes in appearance between the two cameras.

- Stereo algorithms work by solving a "what went where" problem; for each pixel in the left image, find where it went in the right image. Called the **(visual) correspondence** problem.

- But the left and right cameras have slightly different views, so the same "scene element" will look different.

- To understand this, let's think about what a pixel intensity actually **is**.

- To a first approximation, it measures the number of photons in a certain frequency range that travel along a set of light rays (the rays are different for each pixel) in a certain time interval. A measurement parameterized by frequency range, rays, time range.

IMAGE FORMATION

- This leads us to a short non-mathematical discussion of image formation. We will do some math related to this topic later on, when we talk about geometry and geometric transformations.

- PINHOLE CAMERA MODEL DISCUSSION

- What you see is a combination of the surface elements, the lighting and the camera (and their relative positions). To see something else, change one or more of these.

- Most computer vision uses this camera model, or something similar. More complex models handle lens distortion and various artifacts that arise in practice. More simple models make the math that follows tractable.

- Example simple model: orthographic projection. Like most simple models, it's reasonable when there isn't much depth variation relative to how far away the camera is (think of a telephoto lens).

- There is some interest in vision in building radically new types of cameras, for instance that can see a full 180 degrees. We will talk about this a bit at the end of the course.

- But most of the interesting work on non-traditional cameras comes from medical imaging

- X rays are actually fairly similar, but the imaging plane is behind the object, and there is a (penetrating) light source.

- X ray intensity is integral over time and rays of tissues' ability to absorb radiation.

- Note that photons are just photons (light, X-rays, AM radio, etc). "Bad" radiation is technically *ionizing*, which means that it breaks chemical bonds, which in turn can lead to side-effects.

- There is a great deal of study of how much ionizing radiation is safe. Largely looking at statistics of the general population, and of people exposed to significant radiation (mostly Hiroshima survivors, certain miners, workers in the nuclear industry).

- Bear in mind that the planet Earth has background radiation, so there is no way to totally avoid it. Super-polar airline flight is about 1 chest X-ray, which is on the order of a few days of background. There is no statistical evidence that airline crews are adversely affected by this.

- CT scan is about 1 year. Some studies show no measurable effect for about 200 years of excess radiation. But clearly less ionizing radiation is better.

- The only think you can do to make an X-ray image tell you anything else is to give the patient something radio-opaque (such as barium or iodide), either orally or IV.

- If you give contrast IV, you can subtract the "before" from the "after" (this use to be done in analog, with film negatives, believe it or not!)

- This is actually an important research topic; for instance, what happens if the patient moves between before and after? Also, the contrast agents themselves are not so great. Plus you're essentially taking an X-ray "movie" (this is called *flouroscopy*).

- CT is just a variant of X-ray, where the emitter and detector are mounted on a spinning cylinder with the patient inside. The patient is then slid though the cylinder. There is an elegant piece of math (Radon transform) that computes the image from these "intensities".

- US and (especially!) MR are extremely different. We won't talk about US at all, but will cover MR in some depth, because it's very exciting (and my research area).

- Sneak preview: MR uses no ionizing radiation (it's in the AM radio band). MR can tell you a wide variety of tissue properties (presence of fat, or water, or hydrogen atoms). MR, however, is slow and always has a time-versus-accuracy tradeoff. Motion artifacts are the major limitation, and hence the major research opportunity.