

Instructions: Answer all 4 questions in the bluebook provided. 75 min closed book/notes.

1. A stereoscopic camera system has a 10cm baseline. When the system is at rest, the two optical axes are parallel, and are perpendicular to the baseline. When operating, each camera can independently rotate its optical axis up to 30 degrees in any direction from its rest position.

(a) Let M be the midpoint of the baseline, and X the closest point to M which can be fixated. How far away from M is X ?

X must be on the ray beginning at M parallel to the rest optical axes. By right triangles, the distance between X and M is $5 \tan 60^\circ = 5\sqrt{3} = 8.66$ cm.

(b) What set of scene locations are fixated with zero disparity? Assume the intrinsic camera models of the two cameras are identical.

Fixation means the point is imaged at the principal point of both cameras. With equal intrinsic models, the (u,v) coordinates of both principal points will be the same, ie zero disparity. So any point that can be fixated will have zero disparity. This is the set of all scene points whose view angles with the rest axis for each eye do not exceed 30° .

2. Specify 3-4 ways in which the designs of digital camera and the human eye are very different. For each of the differences you cite, justify why camera manufacturers have chosen the solutions we see in their cameras today, rather than using the solutions chosen by nature. In each case, also state why nature designed the human eye the way it did, rather than using the camera manufacturer's design choices.

A. Camera has uniform distribution of photoreceptors vs. foveal distribution in human eye. Easier to build, also the resulting photograph when seen by a human would appear blurred everywhere but the middle if the camera had a foveal distribution. Human eye needs to acquire high resolution and low resolution data simultaneously for vision multitasking.

B. Camera has a single type of photodetector, human eye has two (rods and cones). Cameras don't need scotopic vision since a flash is far less expensive to add on. Human eye has external illumination capability.

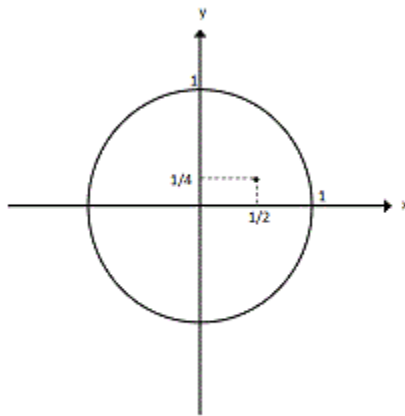
C. Camera has a flat photodetector surface, human retinal is spread over a hemisphere. Flat is less expensive to build, though it leads to smaller field of view, and to distortion of the image. Spherical eyeball is easiest geometry to maintain constant for natural building materials.

D. Camera has no pre-processing on photodetector chip, human eye has five layers of neurons doing much pre-processing. Multilayer logic on a chip is very expensive, cameras designed to do it off-chip, using a specialized image processor chip. Human eye needs to preprocess or data flow rate to the brain would require much larger optic nerve and optic centers in brain.

E. Camera has variable optical zoom, human eye doesn't. Involves multiple lenses with exact geometric relationships maintained as they move. This mechanical precision is infeasible for natural eyes.

Many others could be mentioned... use of a liquid medium in the eye, use of interchangeable filters and lenses in cameras, use of a flexible material for the lens of the eye, different spectral sensitivities for semiconductors vs. dyes in human photodetectors, etc.

3. Consider an orthographic image of a unit Lambertian sphere centered at $(x,y,z)=(0,0,5)$ in camera coordinates. The scene is lit by a single light source very far from the origin, so all source light rays are parallel. The brightest point on the image is at the point in the image plane $(x,y)=(1/2,1/4)$ as shown in the diagram. The brightness there is b .



(a) Specify the 3-D vector direction of the source light rays.

Sphere: $x^2 + y^2 + (z-5)^2 = 1$ so $(z-5)^2 = 1 - 1/4 - 1/16 = 11/16$ and $z = 5 - \sqrt{11}/4$. The coordinates of the brightest point are $(x,y,z) = (1/2, 1/4, 5 - \sqrt{11}/4)$. The normal at any (x,y,z) is proportional to the gradient there, $K(2x, 2y, 2(z-5))$. For (x,y,z) to be brightest the light must be parallel to the gradient, so the light is coming from $(x,y,z) = K(2(1/2), 2(1/4), 2(-\sqrt{11}/4)) = K(1, 1/2, -\sqrt{11}/4)$. Let $K = -1$ (inward pointing normal) and get direction of light $= (-1, -1/2, \sqrt{11}/4)$ (any K ok).

(b) What will the brightness at the point $(x,y)=(0,0)$ in the image plane be? Specify the answer in terms of b .

The light is coming from $(-1, -1/2, \sqrt{11}/4)$. The normal at that point is $(0, 0, 1)$. So $\cos(\Theta) = \text{inner product} / (\text{product of norms}) = (\sqrt{11}/4) / ((2)(1)) = \sqrt{11}/4$. Since the brightness is proportional to $\cos(\Theta)$, and $\cos(\Theta) = 1$ where the brightness is b , the brightness at $(x,y)=(0,0)$ is $b\sqrt{11}/4$.

3. A scene initially looks like Fig 1, with three visible points A, B and C. After we walk a few paces while maintaining fixation on point A, the scene now looks like Fig 2. In order to maintain fixation while we walked, our head rotated clockwise in the xz plane (z axis points into the paper).

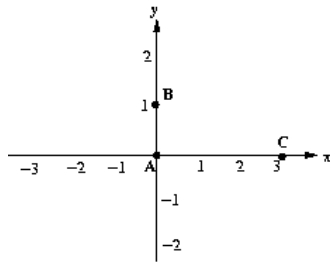


Figure 1

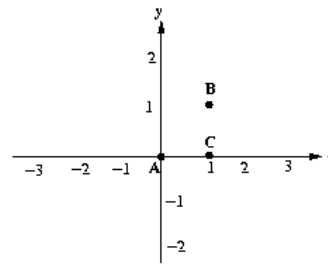
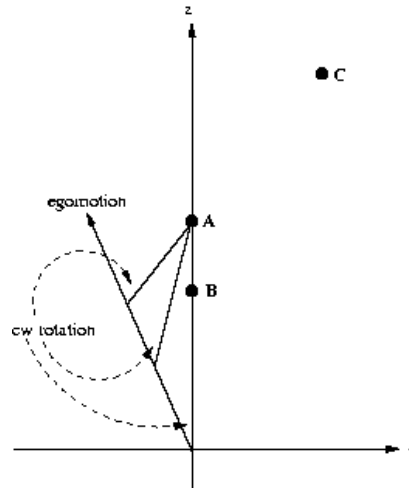


Figure 2

(a) Determine the order of distance from the camera of these three points, closest to farthest. Justify your answer.

If our head moved cw in the xz plane, our egomotion must be in the opposite direction, ie ccw in the xz plane, towards the -x direction. Since C is moving the same (-x) direction, it must be farther away than A and since B is moving in the opposite direction, B must be closer. See the sketch below, drawn in the xz plane from above.



(b) Which point is farthest from the point A? Justify your answer.

The apparent speed of movement is proportional to distance from the fixation point. So since C moves twice as far in the same time, it is twice as far from A than B is.