1. General Rules

The course paper and/or project will be the major requirement of the course. Thus, a quality job commensurate with three course credits is expected. Although size is not a measure of quality, the approximate amount of effort should be comparable to that of writing a major term paper of about 30 pages. Implicit in such a paper is library research of a dozen or so articles. Projects should show a similar amount of dedication, with the final writeup of similar size, although about half of the document can be program listings and results. If the project is done as a team (there is a limit of two people on a team, unless with special permission), the result should show proportionate effort (that is, 60 pages or its equivalent).

If you intend to submit your paper or a project to another course as well, you must obtain permission. The understanding is that such a joint investigation is for the sake of depth, and the result should show a proportionate increase in the amount of work (that is, again 60 pages). A joint effort across courses should truly be a work of art.

1.1. For a Paper

Investigate and report on the state of the art in human and/or machine visual and/or spatial perception. You may have to use libraries other than Engineering and Science. What is important in the paper is scholarly research and reflection, not breadth of coverage. You are not required to make an original contribution, although a well-written paper might see publication. You must show some evidence of an attempt at a synthesis: the field is disparate, so your work should help tie it together, at least in your own mind. Thus, the last five to seven pages or so (15-25%) should serve to future readers in the area as a personal guide. Prior course papers have included: analyses of current commercial systems (such as iris verification, fingerprint analysis, blood cell analysis, robot spacecraft), analyses of current research systems (gesture or sign language analysis, face recognition, gestural control of robots), or designs for new visual interface application areas (psychology of visual interfaces, musical instrument simulation, robotic delivery vehicles).

1.2. For a Project

Program up a small version of some visual input processor of human data. Pick any one that has appeared in the scientific literature, or invent your own. Stay small: this is difficult work, even if someone has gone before you. Ensure your completion of the project by having well-defined bail-out points; build in stages. Two dimensional data are acceptable. You can use a real-time camera (hard!) or just a real camera (not so hard) or input that has been digitized off-line (easy: for example, faces, fingerprints, photographs of parking lots, etc.) The emphasis should be on the assertion of a small symbol: a binary decision (authorization) or token from a small set (gesture class, etc.), perhaps with modifying scalar data.
(certainty factors; positions or speeds, etc.). Prior course projects have included: a visual burglar alarm (demonstrated in my own office!), a system which detected class changing times from a video of campus pedestrian traffic, and two separate systems using "visual passwords" for computer logins.

2. Proposal Structure
Since there will be a lot of work involved in an area that is not well understood, it will be important to forestall any major catastrophes early. You are required to submit a paper/project proposal, which will be critiqued and returned within a week. It should address the following checklist:

1. The phenomena or program to be investigated.
2. A description of the limits you have placed on the investigation. This is most important for programs, but even papers can get out of hand. In past experience, most initial project proposals are approximately ten times too ambitious.
3. At least two major references: either textbooks, journal articles, or (possibly) computer code. This must be a part of the proposal: start now!
4. A two-page sketch of the anticipated methods and results. For a paper, this would be usually be what sort of personal synthesis you hope to be able to get out of it. For a program, this would be program performance. You are not held to this as a promise, but it will help you decide among topics.
5. Any special help that may be required, such as exotic references or equipment. If such help is necessary but unattainable, the proposal will have to be denied.

3. Sample Topics
The following are intended only as suggestions to get you started. If you are interested in something that deals with visual (input) interfaces, please feel free to suggest it in the proposal, and something will be worked out.

3.1. For a Paper
The following domains appear to be ripe with multiple topics, with varying degrees of development and commercialization.

1. The medical domain is a rich area for visually significant human data. Blood cells, cancer cells, chromosomes, iris patterns, retinal patterns, thumb or fingerprints, palm prints, lip prints, faces, and gross body structure have all been considered and/or commercialized in systems that provide identity and/or health information via visual means.
2. There are many security and surveillance systems being developed, which include room, parking lot, freeway, and prison monitoring, sometimes using mobile robot platforms.
3. The sports world reportedly has systems to help analyze performance such as more effective use of the baseball bat, golf club, vault pole, etc. Several systems for virtual sports based on visual analysis of the athlete also exist.
4. The art world has some automated interaction systems for "performance" art.
5. Some initial attempts on coordinating facial analysis with speech understanding, and even of the reading of "mental state" (emotions, etc.), have been demonstrated.
3.2. For a Project

The following simple systems should be possible.

1. An automatic screensaver or login/logout control, based on detecting the presence of a user at the keyboard, even if only one-way: user goes away to turn screensaver on or logout, or user appears to turn screen saver off or login, or both. Possibly, the detection of head direction might also be a trigger: user looks away, etc.

2. A simple gestural password system: a gross pose or motion of the user serves as the password, which, like baseball signals, might be embedded in decoy gestures. Or, the "password" could be used to trigger some favorite application, or terminate it (a visual "here comes the boss!" switch, for example).

3. A simple activity sensor ("lots" of change over "short" intervals of time) could be used to do things like control the volume of music from the system, and which fades away as the activity calms down.

4. The gross sizing of a window keyed to the perspective effect of a looming hand: hand nearer makes window bigger, and vice versa.

5. A user-friendly reinforcement interface for the retrieval of imagery from a data base of images.

All of these require some operating system hacking, but not very much sophisticated visual processing.

If you are considering camera input, you should consider the cameras and interfaces available in the CLIC lab. There are also some low cost cameras available for purchase on the net: try starting at www.connectix.com for some ideas, or investigate the code behind Cornell University’s CUSeeMe.