Optical Flow for Obstacle Detection in Mobile Robots

Kurt Steinkraus & Leslie Pack Kaelbling

Artificial Intelligence Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: We want to let a mobile robot explore its environment, detecting and avoiding obstacles as it goes. This will allow it to create a map of that environment, and then find its way around in it. We would like it to be able to do this using visual input. It should need as little human intervention, in the way of pre-supplied maps or other input, as possible. The system should also be robust with respect to sensor noise and sloppy actuation.

Motivation: Robots currently have a hard time finding their way around on their own. This is OK if the robot is being closely watched by a person, but in some situations, this isn't practical. A mobile robot wandering the surface of Mars could not be guided remotely because of the large time-delay involved. A small army of robots moving around a factory floor, making and delivering things, would require an unreasonably large number of operators.

The reason we will focus on vision as opposed to other sensors (for example, sonar sensors), is that cameras are relatively cheap, and images from a camera give much more information than distance measurements that would be given by sonar or laser range finders. Vision provides the best balance in several factors such as sensor noise, amount of information, and cost, and the only reason it is not more widely used is because it is difficult to extract the relevant data from captured images.

Optical flow is one useful piece of information that can be extracted from sequential images captured by a mobile robot [2]. The optical flow describes how the different parts of the image are moving. As the robot advances, objects far away and toward the center of the image tend to flow more slowly than nearby objects. Calculations using the optical flow allow recovery of the distance to objects in the image. In this way, the robot can detect and avoid obstacles as it moves around.

Approach: Our approach first involved finding a stable, reliable method for calculating the optical flow. We chose to find it by matching corresponding patches of image intensity in sequential images. As an alternative to finding similiar locations, the optical flow can be found through parameterization [4]. These parameterization methods, while they give superior results when they work, tend to be either too slow for real-time mobile robots or not general enough for real-world situations.

Patch matching has the drawback of not interacting well with the discretization in images, but we have extended the within-half matching of Birchfield and Tomasi [1] to deal with two-dimensional images. We have also extended the sub-pixel technique of Nishihara [5] to work when doing within-half matching. It is therefore possible not only to know that a match occurs somewhere within half a pixel, but also to fit a parabola to the local data in order to estimate the exact sub-pixel match location.

Another drawback of patch matching is that it is only reliably correct at a small percentage of locations in images. We have developed a fast, simple method for locating reliably correct parts of the image, allowing confidence to be placed in the correct parts of the calculated optical flow.

With a reliably correct optical flow in hand, the robot can avoid obstacles by using the flow to calculate where the floor is. If the robot can find the floor around itself, it knows where it can travel and deduces that where there is no floor, there is an obstacle.

The optical flow data gives 3D data points corresponding to objects that the robot sees. The floor can be modelled in 3D space by finding a floor plane that separates the data points into those that are part of the floor and those that belong to objects. The conversion of optical flow to 3D points turns out to be unnecessary, because the movement of the floor can be parameterized in image space.

With this model of the floor's movement in hand, each part of the image can be checked for consistency with being part of the floor. Each point is examined to see if the optical flow predicted by floor model is borne out in the images. An inconsistency, i.e. the floor model's optical flow does not fit with the images captured, indicates that the point in question is not on the floor but instead is part of an obstacle.

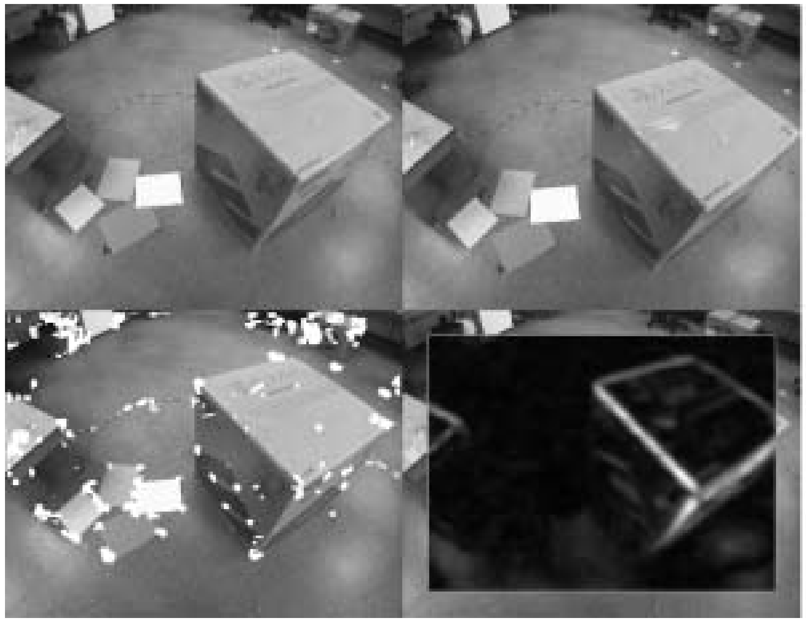

Figure 1: The top two pictures are what the robot saw before and after moving one step. The bottom-left picture shows the places in the image which have reliably correct optical flow. The bottom-right picture gives probabilities that portions of the image are part of the floor, where lighter areas are less likely to be part of the floor.

Impact: This research will give mobile robots a first, crucial processing stage in visual navigation. It will allow robots to visually explore and map their environment by combining the local geometric information about obstacles into larger and more useful topological maps. Such mobile robots will have many applications in industrial and commercial applications.

Future Work: The methods that we have developed currently work fairly well. There is work to be done, however, in generalizing and stabilizing different parts of the processing. Some of the techniques developed, for instance, require that the robot not rotate between sequentially captured images. This is quite limiting for a robot that is supposed to be wandering around irregularly-shaped environments, but this restriction can be lifted if the rotation can be subtracted out. One promising technique involves registering images based on their dominant motion, with the assumption that this dominant motion corresponds to the background, and the background motion from frame to frame is so small as to be imperceptible [3].

Another area of this research that needs attention is filling in between the reliable optical flow data points. Optical flow data (at least using patch matching) will never be able to tell a robot anything about the center of an untextured, homogenously-colored object, because of the aperture problem. Integrating the optical flow data with the processing of other image features, such as texture information or object recognition, can help fill in these gaps. This will help make robot vision more reliable, so that it can play a larger part in guiding a mobile robot.

Research Support: This work is partially funded under DARPA contract #DABT63-99-1-0012.

References:

[1] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):401–406, 1998.

[2] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[3] M. Irani, B. Rousso, and S. Peleg. Recovery of ego-motion using image stabilization. In *CVPR94*, pages 454–460, March 1994.

[4] S. Ju, M. Black, and A. Jepson. Skin and bones: multi-layer, locally affine optical flow and regularaization with transparency. In *CVPR96*, pages 307–314, 1996.

[5] H.K. Nishihara. Practical real-time imaging stereo matcher. *OptEng*, 23(5):536–545, 1984.