Safe Navigation for Autonomous Vehicles: A Purposive and Direct Solution

Gin-Shu Young\textsuperscript{1,2}, Tsai-Hong Hong\textsuperscript{1}, Martin Herman\textsuperscript{1}, Jackson C. S. Yang\textsuperscript{2}
\textsuperscript{1}National Institute of Standards and Technology (NIST), Bldg 220, Rm B124, Gaithersburg, MD 20899
\textsuperscript{2}Robotics Laboratory, Department of Mechanical Engineering, University of Maryland, College Park, MD 20742

Abstract

It is a challenge to develop autonomous vehicles capable of operating in complicated, unpredictable, and hazardous environments. To navigate autonomous vehicles safely, obstacles such as protrusions, depressions, and steep terrain must be discriminated from level terrain before any path planning and obstacle avoidance activity is undertaken. A purposive and direct solution to obstacle detection for safe navigation has been developed. The method finds obstacles in a 2-D image-based space, as opposed to a 3-D reconstructed space, using optical flow. The theory derives from new visual linear invariants based on optical flow. Employing the linear invariance property, obstacles can be directly detected by using a reference flow line obtained from measured optical flow. The main features of this approach are that (1) 2-D visual information (i.e., optical flow) is directly used to detect obstacles; no range, 3-D motion, or 3-D scene geometry is recovered; (2) the method finding protrusions and depressions is valid for the vehicle (or camera) undergoing general motion (both translation and rotation); (3) the error sources involved are reduced to a minimum, since the only information required is one component of optical flow. Experiments with both synthetic and real image data suggest that the approach is effective and robust. The method is demonstrated on both ground and air vehicles.

1 INTRODUCTION

1.1 Autonomous vehicles and safe navigation

Autonomous vehicles are vital for tasks in defense reconnaissance, planetary exploration, operation in hazardous environments, unmanned security systems and missions to be carried out in remote locations. Unlike a teleoperated vehicle which is directly controlled by a remote operator, an autonomous vehicle has independent control of its mission execution and, to some extent, can alter strategy and mission priorities.

It is a challenge to develop an autonomous vehicle capable of operating in complicated, unpredictable, and hazardous environments. One of the most important requirements for an autonomous vehicle is safe navigation, i.e., safe movement from one location to some target point under various constraints.

To navigate autonomous vehicles on the terrain safely, obstacles must be discriminated from traversable terrain before any path planning or obstacle avoidance activity is undertaken. Obstacles are defined as any region in space where a vehicle should not or cannot traverse, such as protrusions (objects lying on top of the terrain), depressions (potholes, ruts, gullies in the terrain), or steep terrain. The three types of obstacles are illustrated in figure 1. To avoid obstacles, one approach is to provide the vehicle with a complete model of the environment in advance. However, this approach is difficult to achieve in practice for the following reasons:

1. Unexpected events may occur. The environment may have been changed, and thus the predefined model may no longer be a true representation of the current environment.

2. The world model of the environment cannot be completely described especially for complicated scenes.
(3) The vehicle cannot control itself exactly as planned. Some uncertainty of the relationship between the vehicle and the environment will occur.

Rather than navigating based on a predefined model, it is more effective to equip the vehicle with sensors to detect the surrounding environment. Vision sensors, in principle, have the capacity to offer rich data for detailed interpretation of the vehicle's working environment. This provides interactive abilities for navigational tasks in dynamic environments.

1.2 Purposive active vision and direct solutions

As mentioned before, vision provides rich information about the environment. However, the cost of generating and updating a complete, detailed model of the surrounding environment based on vision information is very high. Furthermore, obstacle detection for safe navigation of autonomous vehicles requires a real-time and accurate solution. Reconstructing the environment model may be time consuming and may introduce error in the process. To solve these problems, the concept of purposive active vision was introduced by Aloimonos [1]. The purposive vision approach tries to answer the following two questions:

1. What is the visual task to be performed? and
2. How can the task be solved purposively?

Consider the task of safe navigation for example. We may ask: Are there potential obstacles (e.g., potholes or rocks) in the vehicle's path? It is obvious that the complete reconstruction of a 3-D model about irregularly shaped obstacles is not necessary. Detecting protrusions or depressions on the terrain may be sufficient. We may also ask: When will the vehicle collide with obstacles? The necessary information is the ratio of distance to speed (or time-to-collision). Recovery of both distance and speed is not necessary.

The concept of purposive vision motivated the development of direct solution methods. In the obstacle detection task described above, indirect methods always reconstruct a 3-D model before obstacles can be detected. On the other hand, direct methods detect obstacles directly from original visual inputs. Intermediate calculations (i.e., recovering distance, 3-D structure and motion from visual inputs) are not necessary. The direct approach has the following advantages:

1. Simplicity - less information is required.
2. Speed - fewer computations are done.
3. Robustness - error sources are reduced to a minimum.

1.3 The approach of this paper

In this paper, the method for obstacle detection is purposive and direct. The visual input to this method is optical flow. New visual invariants are developed as a tool for detecting obstacles directly from optical flow. These invariants involve the mapping of points that lie on any straight line segment in 3-D space into an image-based space, i.e., a space whose coordinate axes represent parameter values extracted from the image domain. There are certain image-based spaces where straight line segments in the image-based space only come from straight line segments in 3-D space. Such a mapping is described as invariant for linear relationships, or simply linearly invariant, since linear relationships are always preserved. For example, it is shown that a straight line segment $P_1P_2$ in 3-D space (see figure 2) is always a straight line segment $\bar{P}_1\bar{P}_2$ in the image-based space (see figure 3) whose coordinates are $x, y$ where $x$ is the image position along the image line $y$ and $y$ is the $y$ component of optical flow. A point $\bar{P}_5$ in 3-D space (figure 2) lies on the extended line segment $\bar{P}_1\bar{P}_2$ such that the point $P_5$ (figure 3) corresponding to $\bar{P}_5$ must lie on the extended line segment $\bar{P}_1\bar{P}_2$. It is demonstrated in Section 4 that this type of visual invariant allows us to detect obstacles using optical flow. For example, a portion of the terrain with a protrusion and depression, which is visible from an image line, is shown in figure 4. The visual information for five points $(P_1, P_2, P_3, P_4, P_5)$ lying on the terrain is represented in the image-based space $\gamma$ vs. $x$ (see figure 5). The reference flow line (see figure 5) obtained from $F_1$ and $F_2$ corresponds to a reference space line in 3-D space (see figure 4). The deviation between the measured value $\hat{y}$ ($F_1, F_2, F_3, F_4$, and $F_5$) and the $\hat{y}$ value of
the reference flow line for each image position \( x \) is calculated. In figure 6, the positive deviation corresponds to
protrusion while the negative one corresponds to depression. The features of this method are that:

(1) 2-D visual information (i.e., optical flow) is directly used to detect obstacles: no range, 3-D motion, or 3-D
scene geometry is recovered;

(2) No information about the pose of the camera relative to the ground is required;

(3) No terrain model is required, making the method useful for unknown environments;

(4) No information about the vehicle (or camera) motion is required, eliminating the need for extra sensors;

and

(5) Arbitrary camera motion (both translation and rotation) is allowed, making the method completely general.

1.4 Outline

Section 2 discusses previous work. New visual linear invariants are introduced in section 3. A purposive and
direct solution to obstacle detection for safe navigation is described in section 4 and a detailed algorithm for
obstacle detection is given in section 5. Following that, experimental results using both synthetic and real image
data are presented. Finally, conclusions are presented in section 7.

2 PREVIOUS WORK

A number of obstacle detection methods have been developed in the past (e.g., [4] [5] [6] [7] [8] [10] [12] [13] [16]
[17] [18] [19] [20]). The methods may be characterized by the following:

(1) Depth or range information extracted from optical flow, stereo, or laser range finders is often employed to
detect obstacles.

(2) With approaches in which obstacles are detected directly from optical flow, the observer's motion is often
limited to translational motion.

(3) A priori knowledge such as coordinate transformations, sensor motion, model optical flow fields, or road
models (or maps) is often required.

(4) Terrain slope, a useful feature for determining traversability, is not usually calculated.

Time-to-collision or flow divergence without 3-D reconstruction may be very useful for obstacle avoidance [2] [14].
For terrain navigation, however, the portion of the terrain nearest the observer would have small time-to-collision
values although it is not an obstacle region to be avoided.

3 VISUAL INVARIANTS

Visual invariants are unchanged visual properties. Several new visual invariants for linear relationships are
developed as tools for obstacle detection [22]. They are mappings with the property that a straight line segment
in the appropriate image-based space only comes from a straight line segment in 3-D space. The visual invariant
here is a straight line segment.

The visual invariants are based on optical flow \( \vec{F} \), which can be expressed as:

\[
\vec{F}(x, y, t) = (\dot{x}(x, y, t), \dot{y}(x, y, t)) = F(x, y, t)\vec{u}_F
\]

where \((x, y)\) is the image position, \( t \) is the instance of time, \( \dot{x} \) and \( \dot{y} \) are the components of optical flow, and \( F \) and
\( \vec{u}_F \) are the magnitude and unit directional vector of optical flow, respectively. The relationship between optical
flow and 3-D structure and motion are described as:

\[
\dot{x} = \frac{1}{Z_e}(-T_x + xT_z) + (xy_0 - (1 + x^2)\omega_y + y\omega_z)
\]

\[
\dot{y} = \frac{1}{Z_e}(-T_y + yT_z) + ((1 + y^2)\omega_x - x\omega_y - x\omega_z)
\]

where \( Z_e \) is the depth of the object in the environment relative to the camera, and \((T_x, T_y, T_z)\) and \((\omega_x, \omega_y, \omega_z)\)
are the translational and rotational motion of the environment coordinate system relative to the camera.
Note that for each instance of time, \((T_X, T_Y, T_Z)\) and \((\omega_X, \omega_Y, \omega_Z)\) are constants for all points in the stationary environment. Consider first the arbitrary line segment \(AB\) in space and its projection onto the image plane, image line segment \(\overline{ab}\) (figure 7). Note that \(AB\) may have any arbitrary motion (both translation and rotation) relative to the camera. Choose the image axis \(x\) to be parallel to line segment \(\overline{ab}\).

The following linear relationship [22] can be obtained from equation (3) for all image points lying on line \(\overline{ab}\) (i.e., \(y=\text{constant}\)) that arise from points in the scene lying on line \(\overline{AB}\):

\[
y = a_1 + a_2 x
\]  

(4)

For each instance of time, the values \(a_1\) and \(a_2\) are constants for all points on line \(\overline{AB}\). Equation (4) represents a line in the \(y\) vs. \(x\) image-based space corresponding to a line in 3-D space. This is a visual linear invariant. However, in general, no linear relationship exists in the \(z\) vs. \(x\) or \(F\) vs. \(x\) image-based space corresponding to a line in 3-D space. Note that, in principle, the line (or two coefficients \(a_1\) and \(a_2\)) in the \(y\) vs. \(x\) image-based space can be estimated from two points (say, \((x_1, y_1)\) and \((x_2, y_2)\)). This means that specific knowledge about the transformation matrix and camera motion is not required.

4 OBSTACLE DETECTION FOR SAFE NAVIGATION

In this section, we show how to detect protrusions and depressions using a purposive and direct approach (i.e., directly from optical flow without 3-D reconstruction). Detection of steep terrain was shown in [21]. Some examples are given to demonstrate the detection of obstacles by employing the properties of visual linear invariants. As described in Section 3, a mapping is a linear invariant for the \(y\) vs. \(x\) image-based space. This invariant mapping is useful for obstacle detection. Interestingly, the mapping to the \(F\) vs. \(x\) image-based space, which is full flow magnitude \(F\) as a function of position \(x\) on the image line under consideration, is not a linear invariant. Figures 8 and 9 show the results of a simulation over two different types of terrain. The simulation involves 5% noise added to synthetically generated optical flow under general camera motion. Although no obvious differences can be observed between figures 8 and 9, in figure 8 the terrain is flat and without obstacles, while in figure 9 the terrain has two obstacles, one protrusion and one depression. We see that detecting obstacles can be very difficult in an image-based space that does not have the linear invariant property.

On the other hand, when an image-based space with the property of linear invariance is used, obstacle detection becomes easy and straightforward. Consider the mapping to the \(y\) vs. \(x\) image-based space, in which we consider only one component, \(\hat{y}\), of optical flow as a function of \(x\). This mapping is a linear invariant (Equation 4) under general camera motion. Figures 10 and 11 show the results of the same simulation, using the same camera motion and terrain, as in figures 8 and 9, respectively. Figure 10 appears to coincide with a reference flow line, implying that it is obtained from a terrain which is flat (at least in one dimension). Points that do not lie on the reference line result from protrusions or depressions on the terrain. Thus in figure 11, there are two obstacles, one a protrusion and the other a depression. Figure 11 can further be mapped into figure 12 where the reference line is horizontal. Figure 12 plots the difference between the actually measured value \(y\) and the \(\hat{y}\) value of the reference line for each \(x\) value. A protrusion or a depression can easily be detected in figure 12.

5 ALGORITHM

A detailed algorithm for detecting obstacles such as protrusions and depressions is given as the following:

**Step 1:** Selection of an arbitrary straight line in the image.
This line intersects an image feature of interest, e.g., a potential obstacle. The chosen image line need not correspond to a linear feature in the scene.

**Step 2:** Estimation of the reference flow line for the chosen image line.
The reference flow line in the image-based space, say \(\hat{y}\) vs. \(x\) for an image line \(y\), corresponds to a reference space line in 3-D space. The latter reference line and the camera focal point define a plane in space (figure 4). The intersection of this plane with objects and terrain in the environment defines a set of curves lying in the plane which are visible in the camera. Deviations between these curves and the 3-D reference line are represented in the
image-based space (figure 6). The 3-D reference line can be arbitrarily chosen but it should probably correspond to a surface in the environment such that deviations from this surface represent protrusions and depressions. On a road, for example, the 3-D reference line should probably lie on the road surface. For a vertical image line, the reference flow line can be obtained from the measured optical flow located in the lower image positions which corresponds to regions on the ground surface near the vehicle. For any arbitrary image line, the reference flow line can be obtained by fitting a straight line from the measured optical flow selected at some image positions which correspond to the reference region in the scene. We want the 3-D reference line to lie in the surface represented by the reference region. Note that only the component of optical flow normal to the image line is used. In principle, only two points are required to estimate the reference flow line.

Step 3: Computation of the deviation.
The deviation between the reference line obtained in step 2 and the measured flow at all image positions lying on that image line is computed.

Step 4: Representation of the regions of obstacles.
The computed deviation in step 3 is used to detect obstacles. If the difference is larger than some threshold value, the observed point in space is considered to be an obstacle. Its corresponding image position is marked as an obstacle to be avoided for safe navigation.

To detect obstacles easily, an image-based space with the property of visual linear invariance should be used. For arbitrary camera motion, the proper image-based space should be \( y \) vs. \( x \) (or \( z \) vs. \( y \)) for any image line \( y=\text{constant} \) (or \( z=\text{constant} \)). In this method, only one component of optical flow is needed. Therefore, normal flow, the component of optical flow along the intensity gradient direction, could be used. Information such as specific knowledge of vehicle (or camera) motion, or knowledge of the coordinate transformation between the camera and the ground is not required. Therefore, the method reduces error sources to a minimum since it employs minimum information. The approach is simple because obstacles are detected directly in the image-based space, without performing 3-D reconstruction. Without any assumption of a terrain model, this method can be used for ground vehicles navigating in man-made roadways or natural outdoor terrain. This method can also be used for air vehicles undergoing general six-degree-of-freedom motion while landing on either known or unknown terrain.

6 EXPERIMENTS AND RESULTS

In this section, we present the results of three experiments demonstrating the simplicity and usefulness of visual linear invariants applied to obstacle detection. These experiments include ground as well as air vehicles. Both synthetic and real indoor and outdoor scenes are used. To extract optical flow from real imagery, we have used two different algorithms. Only the components of flow normal to the selected image lines are used.

Experiment 1

This experiment involves detecting hazardous regions for air vehicles, such as mountains, using the realistic-looking synthetic data of a 3-D natural outdoor scene. The Yosemite Fly-Through sequence used in this experiment was obtained from Barron et al [3], courtesy of Lynn Quam. An image frame of the sequence and its optical flow field are shown in figures 13 and 14, respectively. This sequence represents a complex test case [3]. In the upper right, the clouds translate to the right with a velocity of about one pixel/frame, while velocities in the lower left are about 4 pixels/frame. This sequence is challenging because of the range of velocities and the occluding edges between the mountains and the horizon. In fact, it is too difficult to detect obstacles for safe navigation directly from the optical flow field shown in figure 14. However by employing the property of visual invariance, hazardous regions such as mountains can easily be detected.

Here, we consider only column number 270 and row number 210 labelled in figure 13. Both true optical flows and those extracted from the imagery are used as inputs to our system. In this way, we can compare our results to ground truth. The optical flow extracted from the imagery, shown in figure 15, was obtained using the first-order, local differential method of Lucas and Kanade implemented by Barron et al [3]. The average angular error of optical flow is 5.20 degrees with a standard deviation of 9.45 degrees and a measurement density of 35.1% [3]. Note that the 10% bound on acceptable velocity errors for ego-motion and structure from 2-D motion correspond
to angular errors of roughly 2.5 degrees [3].

Column no. 270:

First, the true optical flow was used as input. We define the reference flow line that corresponds to the valley region in the lower right portion of the image at column no. 270 (figure 13). The reference flow line for column no. 270 is therefore obtained by fitting the data points near the lower right image. The deviation from the reference flow line is shown in figure 16. The hazardous mountains can be easily detected. The big gap in the deviation is due to the occluding edge between the mountains and the sky. Next, the optical flow extracted from imagery was used as input. To reduce the effects of highly noisy data, a median filter is used. The reference flow line is obtained by fitting a straight line to the extracted flow near the lower right portion of the image. The result is shown in figure 17. The mountains can still be easily detected. The region marked as "no information" is due to the lack of optical flow in the sky.

Row no. 210:

The procedure for detecting hazardous regions at row no. 210 is similar to the procedure used for column no. 270. Using the true and measured optical flow as inputs, the results are shown in figures 18 and 19 respectively. Some of the mountain region in figure 19 can still be detected even though there are few optical flow values available.

Experiment 2

This experiment involves detecting one main obstacle such as a stationary car on a road way in a real outdoor scene. A camera is mounted on an Army HMMWV vehicle, which is on a road and approaching a stationary car. The vehicle (or camera) motion is unknown. An image frame of the sequence is shown in figure 20. Optical flow was obtained using the 3-D hermite polynomials method developed by Hong [9]. The components of optical flow in the x and y directions are shown in figures 21 and 22, respectively. To reduce the effects of highly noisy data, a median filter is applied to the optical flow image. Figure 23 shows the result for scan line 118 using the measured \( \hat{y} \) (figure 22) as inputs. Scan line 118 is labelled in figure 20. For the full image, employing multiple scan lines, the results superimposed on an original image are shown in figure 24. The white region denotes the detected obstacle.

Experiment 3

This experiment involves detecting multiple obstacles in a real outdoor scene such as a parking lot. A camera is held by a passenger in a moving vehicle, approaching several parked cars on the right side. An image frame of the sequence is shown in figure 25. The vehicle (or camera) motion is unknown. Optical flow was obtained using the first-order, local differential method of Lucas and Kanade implemented by Barron et al [3]. Here, we only consider scan lines 130 through 133 labelled in figure 25. The \( y \) components of optical flow in the scan lines 130 and 133 are shown in figures 26 and 27, respectively. To reduce the effects of highly noisy data, a median filter may be applied to the optical flow image. Figures 28 and 29 shows the results for scan lines 130 and 133 using the measured \( \hat{y} \) in figures 26 and 27 as inputs. The protrusions denoting parked cars on the right side are easily detected. Note that the rightmost image position shows a larger deviation denoting a more dangerous obstacle. This is because the parked cars in the rightmost image positions are closest to the observer.

7 CONCLUSIONS

A purposive and direct solution to obstacle detection for autonomous vehicles has been presented. Only 2-D visual information (i.e., one component of optical flow) is required to detect protrusions and depressions. The developed method is valid for autonomous vehicles moving under arbitrary motion on unknown terrain. The results of the experiments in the paper suggest that the approach is effective and robust.
8 ACKNOWLEDGEMENTS

The authors would like to thank Dr. James Albus, Dr. David Coombs and Marilyn Nashman for important comments on this work.

References


---

**Fig. 1** Terrain with obstacles

**Fig. 2** Space line and image line

**Fig. 3** Flow line in an image-based space \( \hat{y} \) vs. \( x \)

**Fig. 4** Terrain visible from one image line
Fig. 5 Reference flow line in an image-based space $\hat{y}$ vs. $x$

Deviation

Image position $x$

Fig. 6 Obstacle detection without 3-D reconstruction

Fig. 7 Space line segment $\overline{AB}$ and image line segment $\overline{ab}$

Fig. 8 $F$ vs. $x$ without any obstacle

Fig. 9 $F$ vs. $x$ with two obstacles

Fig. 10 $\hat{y}$ vs. $x$ without any obstacle

Fig. 11 $\hat{y}$ vs. $x$ with two obstacles

Fig. 12 $\Delta \hat{y}$ vs. $x$ with two obstacles
Fig. 13 One image from the Yosemite sequence

Fig. 14 True flow field for the Yosemite sequence

Fig. 15 Optical flow field produced by Lucas and Kanade method

Fig. 16 Hazardous regions detected from the true flow at column no. 270

Fig. 17 Hazardous regions detected from the measured flow at column no. 270

Fig. 18 Hazardous regions detected from the true flow at row no. 210

Fig. 19 Hazardous regions detected from the measured flow at row no. 210
Fig. 20 One image frame of a sequence in experiment 2

Fig. 21 The x component of the measured flow in experiment 2

Fig. 22 The y component of the measured flow in experiment 2

Fig. 23 Obstacles detected at row no. 118 in experiment 2

Fig. 24 Obstacles detected for the full image in experiment 2
Fig. 25 One image frame of a sequence in experiment 3

---

Fig. 26 The $y$ component of the measured flow at scan line 130 in experiment 3

Fig. 28 Obstacles detected at scan line 130 in experiment 3

---

Fig. 27 The $y$ component of the measured flow at scan line 133 in experiment 3

Fig. 29 Obstacles detected at scan line 133 in experiment 3