Visual orientation in the sewer – adaptation to the environment

Marina Kolesnik
Fraunhofer Institute for Media Communication, Schloss Birlingshoven, D-53754 Sankt-Augustin, Germany, email: marina.kolesnik@imk.fraunhofer.de

Abstract
Most biological systems employ visually acquired information for their locomotion. In the course of evolutionary history, the visual system of organisms has evolved to be adapted to the environment. As a consequence of this adaptation, biological systems often display highly efficient visual skills. This reasoning has motivated the development of a specific visual system, which serves the purpose of navigation in an unusual environment – a sewer. The sewer environment exhibits two dominating features: restricted geometry of its inner surfaces and absolute darkness. These features are exploited by the hybrid vision system of the autonomous robot consisting of a crosshair laser projector and a camera. If a priori knowledge about the sewer geometry is taken into account, orientation of the robot can be derived from a visual analysis of a regular laser pattern projected onto the sewer surface. Because the footprint image is acquired in an entirely dark environment, the camera records a mostly dark image with the bright footprint in it. The analysis of such an image is very fast and knowledge of the robot’s instantaneous orientation derived from this analysis is enough to guide its navigation. It is concluded that proper exploitation of the environmental constraints has lead to the development of this highly efficient visual system.

1 Introduction
The classical approach towards visually guided navigation is based on the following generic plan: Get an image from CCD cameras. Perform image analysis in order to build an accurate model of the outside world. Use this model to plan the robot’s actions. Finally, these actions are executed in the real world and the cycle is repeated. The problem with this approach, however natural it may seem, is that it does not work in the real world, at least not in general. The reasons have been discussed in the literature (e.g. [1], for a more general perspective on this point [10]). In essence, it is that the robots fail to react to events in the rapidly changing surrounding in real time. One reason for this is that the classical approach is not enough adapted to the environmental conditions and does not exploit the task specific constraints.

A completely different approach – cheap vision – which is well adjusted to specific conditions of the environment where the navigation is performed, was suggested by Ian Horswill in 1992 [2]. Horswill’s robot Polly [3] was the first practical demonstration of the power of the cheap vision idea, which states that vision should not be looked at in isolation, but rather in the context of a complete agent (definition of the complete agent is introduced in [10]) and navigation scenario. Since then, the cheap vision approach has proved to be extraordinarily effective in several case studies which mimicked the principles of visually guided navigation of biological systems, such as the homing of the desert ant Cataglyphis [7], [9], and the flying bees [4]. Each vision solution of this kind embodies a navigation strategy which is only functional under the given conditions, but in these conditions it is perfectly reliable, fast and highly precise [11].

A visual system described in this work employs the principles of cheap vision for navigation inside a sewer. The sewer inspecting robot [12] has to move inside a long system of inaccessible to humans pipes in order to collect a video record of sewer conditions for a later check by experts. Operation of the robot’s visual system is built upon characteristic constraints of the sewer environment. Thus, the visual system is well adapted to navigation in the sewer where it is only operational.

A second principle exploited in this work is based on the fact that local navigation can be considered as a continuous process of finding the direction for the robot’s next move. If navigation is treated like this, the task for the robot’s visual system becomes much simpler. No longer does it need to construct a metric model of the outside world, but rather just a trivial check on the right direction for the robot’s next move. There is, however a snag in this scheme. Because the local direction must be continuously updated on the move, the visual analysis has to operate in real time. This real time requirement on the operation of vision was and is a major bottleneck in the classical approach. The cheap vision opens up an entirely new opportunity for the real time orientation update, which, in turn, brings sensor-motor coordination of the robot to a new level [10].

These points will be illustrated in this paper. We start by looking at local navigation in the sewer as the problem of finding the robot’s instantaneous orientation. This has to be solved in the presence of specific constraints imposed by the sewer environment. These constraints govern the design of the visual system adapted to the sewer environment (Section 2.1). Next, we describe the visual system setup and give a detailed account of the principles of its operation (Sections 2.2-2.3). Real time orientation within the sewer falls into two categories: orientation within a straight pipe and at a manhole (Section 3). Finally, we take a closer look at the concepts of local navigation, adaptation, and real time vision and the way all three are interconnected.

2 The visual system
In this section we describe the most characteristic features of the sewer environment in light of their
implications on the navigation task and, ultimately, on a setup of the visual system.

2.1 The sewer environment

Modern concrete sewers consist of cylindrical pipe segments having a 30 or 60cm diameter. These are joined together into the longer straight pipe portions. The straight portions intersect each other in T-, L- or X- shaped junctions called manholes. The latter are regions where humans can access the sewer from outside. Manholes are constructed out of preformed standard-shaped blocks, which are portions of vertical cylinders of about 2m diameter with the pipe entrances in perpendicular directions inside them. There are also four little stairs between the pipe entrances.

These elements of sewer construction define a dominating constraint of the sewer environment – the restricted geometry of its inner surfaces. The second constraint is obvious: the underground sewer world is absolutely dark.

2.2 The task of navigation in the sewer

Any navigation scenario is ultimately driven by two mechanisms. The first one is global navigation. It guides the agent’s motion towards the final destination and requires a capability to update the global direction at all times along the way. The objective of the second, local navigation, is mainly to find a safe, collision-free path in the close vicinity of the agent. In a typical navigation scenario the point of destination is not visible most of the way. It follows that global navigation cannot rely on visual sensing. Contrast this with local navigation, that has to detect obstacles in the “visible” proximity ahead of the agent. Visual sensing is well suited to this task.

Given the global direction updated by independent non-visual means, there remains a problem of local navigation: the robot has to accomplish a set of intermediate goals before it arrives at the final destination. The strategy for the local navigation results from a trade-off between two driving forces: the straight way towards the final destination and a safe path, which is suitable for locomotion. It follows that the best direction for the local motion is one, which is (1) feasible for the motion and (2) has the smallest deviation from the global direction. If these general considerations could be exploited by the robot, than the robot’s task would be reduced to a trivial check on the distance between the camera and the laser is about 14cm.

These considerations have lead to the idea of using a hybrid vision system that consists of two components: 1) a pen-size laser crosshair projector and 2) an optical camera (Figure 1). The laser is equipped with a special optical head to project a high quality crosshair from the laser beam. The crosshair consists of two perpendicular planar sheets spanned within an accurately known fan angle. By projecting the ideal laser generated pattern onto the sewer surface whose geometrical features are roughly known, we extract a small number of surface points that, like a condensed print, carry information about the surface geometry. This information may be enough to recover the relative orientation of robot with respect to the surface. Clearly, the dark conditions within the sewer will enhance the image of the laser footprint cast on the sewer surface. An important point of this approach is that the visual system directly exploits the environmental constraints and, in fact, benefits from them.
The camera acquires images of the pipe causing both curvature on the quadratic curves. As the laser orients along vertices direction of its beam, i.e. with the pipe and single point on the plane at infinity, will lie at infinity. Because all parallel lines intersect in a two perpendicular planar sheets central point of the laser crosshair, which is the origin of the footprint image. Let and the images of and. At some point this distance will become too small and the images of will eventually coincide. This is illustrated in Figure 3.a. Note that by including infinity in our considerations, we go beyond constructions of Euclidean geometry into an infinite world of Projective geometry. This is justified by the fact that a pinhole camera performs perspective projection mapping of 3-dimensional projective space onto a projective plane.

It follows from these geometrical considerations that the image of the laser footprint in a cylindrical pipe is uniquely connected to the instantaneous orientation of the robot head inside the pipe. Moreover, the discrepancy between the points and is a function of the robot’s deviation from the direction along the pipe axis. A critical condition that makes this scheme work is the distance separating the camera and the laser. For the camera to record the footprints as curves, it must be shifted away from the laser’s origin. Simulations made in [5] and images acquired by the robot’s camera within typical sewer pipes have proven that a sufficient distance is feasible in practice.

3 Visually guided local navigation

Typically, inner surfaces of the sewer can be classified into four different types: a) pipe; b) ridge; c) right wall; d) left wall. Each surface type is characterised by a priori knowledge of the possible configuration of surfaces, performs a quick analysis of the footprint shape to identify the type of surface it currently looks at and, in the case of a pipe, to decode its instantaneous orientation relative to the pipe axis.

Figure 2. Geometry of the crosshair footprint when projected on the surface of a cylinder.

Let us take a closer look at the geometrical rules that govern the footprint image. Let be an infinite circular cylinder modelling a sewer pipe (Figure 2). Let be the central point of the laser crosshair, which is the origin of the two perpendicular planar sheets and, generated by the laser. and intersect in a line which is the laser beam. When the crosshair is projected onto the pipe surface, each planar sheet slices through in a quadratic curve (a segment of an ellipse). Let and be the two ellipse segments, cut by the planar sheets and, respectively. Let and be the vertices of and, respectively.

Orientation of the laser within the pipe is related to the direction of its beam, i.e. with . As the laser orients along the pipe and becomes more parallel to the axis of , the vertices and move further away from . At some point, when is parallel to the axis of , both and will lie at infinity. Because all parallel lines intersect in a single point on the plane at infinity, and the axis of will meet at this point.

Consider a camera located at a point distinct from . The camera acquires images of and, which are two quadratic curves. Given a certain distance between and, the vertices and are viewed as two points of maximum curvature on the quadratic curves. As the laser orients along the pipe causing both and to slip away closer to infinity, the camera records a smaller distance between and. At some point this distance will become too small and the images of and will eventually coincide. This is illustrated in Figure 3.a. Note that by including infinity in our considerations, we go beyond constructions of Euclidean geometry into an infinite world of Projective geometry. This is justified by the fact that a pinhole camera performs perspective projection mapping of 3-dimensional projective space onto a projective plane.

Figure 3. Four typical footprint images classified as different surface types. (a) Pipe. The points of maximum curvature on the horizontal and vertical parts of the laser footprint coincide indicating the robot’s heading is well aligned along the pipe axis. (b) Ridge – this shape occurs when the laser footprint is cast onto the stair inside manhole. (c) Right wall. (d) Left wall. These footprints occur when the laser is projected onto the sewer surface located either on the left (d) or right (c) hand side from the robot’s head segment.

The visual analysis of the specifically shaped footprints, whose formal description is given in [6], consists of several fairly simple processing steps such as: segmenting the footprint from the background; approximating the horizontal and vertical parts of the footprint by a cubic polynomial function, finding the points of extremas on these functions; defining the type of the underlying surface
by analysing relative configuration of the extremal points. Because the camera records a mostly dark image with a bright laser footprint, extraction of the footprint points is simple and robust. Because the footprint points constitute just a tiny fraction of all image pixels, their processing is extremely fast. In fact it takes 50msec for analysis of a single footprint image with the size 320x240 pixels by a Pentium, 166MHz processor.

When the robot moves along a straight pipe, it continuously projects the crosshair pattern, acquires the image of the laser footprint, makes analysis of the shape of the footprint in the image, identifies its instantaneous orientation and corrects its heading. At a manhole, the robot finds the pipe emanating from the manhole in the direction of interest, orients towards the pipe axis and moves in. Indeed, the visual orientation in the sewer by the laser footprint analysis proved to be extraordinarily effective and reliable. We had the astonishing experience that instructions obtained from the visual module corrected the robot’s heading with an accuracy to 1 degree.

4 Discussion and Summary

The navigation of the sewer robot by the hybrid visual system illustrates two important points that we would like to discuss here.

The first point is that for any visual system to be effective it must be adapted well to the environment of its operation. If environmental constraints are exploited by the visual system, the visual analysis often becomes remarkably simple, hence the name - cheap vision. But adaptation is incomplete if the visual system is not adapted to the task. Here comes the second point to make: all that is needed for successful navigation is the right direction to move. Consequently, the role of vision while directing the local navigation is just to find continuously the right direction where to move next. On the one hand this point of view simplifies the task of vision, but on the other - it imposes a tough condition on the speed of visual analysis that must update the instantaneous orientation in real time. A consequence of these two points is that if vision indeed exploits the constraints, it becomes quick enough to make an instantaneous decision on the direction of motion. The two points start to work for each other making real time visual analysis possible and bringing an important concept of sensor-motor coordination into reality.

These points, however obvious they may seem, have been given little attention by the computer vision community particularly in those classical computer vision applications where structure is extracted from images and is used for navigation purposes. At the same time most, if not all, biological systems that navigate by vision, do so without use of precise metric information. Tests with human observers have reported poor performance when judgments of metric properties are to be made [8]. This, however, hardly limits our excellent navigation skills. How does it come about? Clearly, adaptation of biological vision to the task and the environment is a decisive factor in all instances of navigation. The visual system of the sewer robot, which in some ways is deeply adapted to the sewer environment, illustrates our point and shows that visual navigation should not necessarily be a complicated process.

One may argue that the way we treat the visually driven local navigation is not generic. There may be cases when cheap vision will not properly find the right direction for motion and navigation will fail. Indeed, our approach is not more universal than any other operational robotic or biological system on Earth. What we do point out is that not much of visual data is needed in order to extract and update the right direction for motion. This is illustrated by the hybrid visual system, which one should see not only as a cute engineering solution to a particular navigation scenario, but rather as a demonstration of the generic principles of cheap vision. The success and easiness of the visual orientation did not come by chance, but resulted from a thoughtful use of ideas on adaptation of the visual system to the specific environmental conditions.

References