

Geoffrey Taylor
Lindsay Kleeman

Department of Electrical and Computer Systems Engineering
Monash University, Clayton 3800
Victoria, Australia
geoffrey.taylor@eng.monash.edu.au
lindsay.kleeman@eng.monash.edu.au

Stereoscopic Light Stripe Scanning: Interference Rejection, Error Minimization and Calibration

Abstract

This paper addresses the problem of rejecting interference due to secondary specular reflections, cross-talk and other mechanisms in an active light stripe scanner for robotic applications. Conventional scanning methods control the environment to ensure the brightness of the stripe exceeds that of all other features. However, this assumption is likely to be violated for a robot operating in an uncontrolled environment. Robust scanning methods already exist, but suffer from problems including assumed scene structure, acquisition delay, lack of error recovery, and incorrect modeling of measurement noise. We propose a robust technique that overcomes these problems, using two cameras and knowledge of the light plane orientation to disambiguate the primary reflection from spurious measurements. Unlike other robust techniques, our validation and reconstruction algorithms are optimal with respect to sensor noise. Furthermore, we propose a procedure to calibrate the system using measurements of an arbitrary non-planar target, providing robust validation independently of ranging accuracy. Finally, our robust techniques allow the sensor to operate in ambient indoor light, allowing color and range to be implicitly registered. An experimental scanner demonstrates the effectiveness of the proposed techniques. Source code and sample data are provided in the multimedia extensions.

KEY WORDS—depth perception, light stripe, active stereo, specular reflections, robustness, calibration, interference rejection, optimal reconstruction, color registration

1. Introduction

Light stripe ranging is an active, triangulation-based technique for non-contact surface measurement that has been studied for several decades (Shirai and Suwa 1971; Agin and Binford 1973). By projecting a known feature onto the measured

surface, active scanners provide a more robust solution to the measurement problem than passive ranging techniques. Reviews of light stripe scanning and related range sensing methods can be found in Hebert (2000), Bastuscheck (1989), and Besl (1988). Range sensing is an important component of many robotic applications, and light stripe ranging has been applied to a variety of robotic tasks including navigation (Nygards, Högström, and Wernersson 2000; Aldon and LeBris 1994), obstacle detection (Haverinen and Rönning 1998), object recognition for grasping (Alshawish and Allen 1995; Rao et al. 1989), and visual servoing (Khadraoui et al. 1996).

The drawback of conventional single-camera light stripe range sensors is that favorable lighting conditions and surface reflectance properties are required, so that the stripe can be identified as the brightest feature in the captured image. However, the range sensor presented in this paper is intended for use on a humanoid robot operating in an uncontrolled domestic environment (Taylor and Kleeman 2001, 2002). Under these conditions, various noise mechanisms interfere with the sensor to defeat conventional stripe detection techniques: smooth surfaces cause secondary reflections, edges and textures may have a stripe-like appearance, and cross-talk can arise when multiple robots scan the same environment. The motivation for this work was to develop a robust light stripe sensor suitable for operation in these noisy conditions.

A number of techniques for improving the robustness of light stripe scanners have been proposed in other work, using both stereo and single-camera configurations. Magee, Weniger, and Franke (1994) develop a scanner for industrial inspection using stereo measurements of a single stripe. Spurious reflections are eliminated by combining stereo fields via a minimum intensity operation. This technique depends heavily on user intervention and a priori knowledge of the scanned target. Trucco and Fisher (1994) also use stereo cameras to measure a laser stripe, and treat the system as two independent single-camera sensors. Robustness is achieved by imposing a number of consistency checks to validate the range data, the

most significant of which requires independent single-camera reconstructions to agree within a threshold distance. Another constraint requires valid scanlines to contain only a single stripe candidate, but a method for error recovery in the case of multiple candidates is not proposed. Thus, secondary reflections cause both the true and noisy measurements to be rejected.

Nakano, Watanabe, and Kanno (1988) develop a similar method to reject false data by requiring consensus between independent scanners, but using two laser stripes and only a single camera. In addition to robust ranging, this configuration provides direct measurement of the surface normal. The disadvantage of this approach is that each image only recovers a single range point at the intersection of the two stripes, resulting in a significant acquisition delay for the complete image.

Other robust scanning techniques have been proposed using single-camera, single-stripe configurations. Nygard and Wernersson (1994) identify specular reflections by moving the scanner relative to the scene and analyzing the motion of reconstructed range data. In Haverinen and Röning (1998), periodic intensity modulation distinguishes the stripe from random noise. Both of these methods require data to be associated between multiple images, which is prone to error. Furthermore, intensity modulation does not disambiguate secondary reflections, which vary in unison with the true stripe. Alternatively, Clark, Trucco, and Cheung (1995) use linearly polarized light to reject secondary reflections from metallic surfaces, based on the observation that polarized light changes phase with each specular reflection. However, the complicated acquisition process requires multiple measurements through different polarizing filters.

Unlike the above robust techniques, the method proposed in this paper uniformly rejects interference due to secondary reflections, cross-talk, background features, and other noise mechanisms. The proposed algorithm solves the association problem to allow the true stripe to be disambiguated from a set of noisy candidates. Thus, our method offers a significant improvement over previous techniques for error detection, by providing both noise rejection and recovery of valid measurements. The depth data reconstructed by our stereo scanner is optimal with respect to sensor noise (unlike the stereo techniques in Nakano, Watanabe, and Kanno 1988 and Trucco and Fisher 1994), and fuses stereo measurements with the light plane parameters. Thus, our sensor provides greater precision than a single-camera configuration, as demonstrated experimentally. Furthermore, we develop an image-based calibration procedure using a scan of a non-planar but otherwise arbitrary surface. Our calibration procedure provides robust validation independently of ranging accuracy. Finally, the ability of our sensor to operate in ambient indoor light allows surface color and range to be captured in the same cameras. Since the intended application is to support object modeling and tracking (Taylor and Kleeman 2001), this implicit registration of



Fig. 1. Stereoscopic light stripe scanner hardware.

color and range is highly desirable. The system described in this work was first proposed in Taylor, Kleeman, and Wernersson (2002), and is extended here by providing a more rigorous treatment with respect to measurement noise.

In the following section we briefly describe the hardware and operation of our stereo stripe scanner. In Section 3 we develop the theoretical framework for optimal noise rejection and reconstruction. Image-based calibration of the system is described in Section 4. Implementation details are discussed in Section 5, and in Section 6 we present experimental results to compare the performance of our system with other stripe scanning techniques. The multimedia extensions to this paper provide three-dimensional (3D) models of scans produced by our sensor, along with source code and raw stereo image sequences to demonstrate the proposed algorithms.

2. Overview

2.1. Basic Operation

Figure 1 shows the components of the experimental stereoscopic light stripe scanner, which is mounted on a pan/tilt active head. A vertical light plane is generated by a laser diode module with a cylindrical lens, and is scanned across a scene by rotating the laser about a vertical axis. The angle of rotation is measured by an optical encoder connected to the output shaft via a toothed belt. PAL color cameras capture stereo images of the stripe at 384×288 pixel (half-PAL) resolution.

Captured images are processed to extract all possible locations of the light stripe on each scanline. Scanlines typically contain multiple candidates, with the true stripe (primary reflection) hidden amongst spurious measurements. The true stripe is identified on each scanline by searching for the pair of stereo measurements most likely corresponding to the projection of a 3D point on the light plane, given the current encoder measurement. The validation process is detailed in

Section 3. Finally, the validated measurements are used to reconstruct the 3D profile of the illuminated surface.

As the stripe is scanned across the scene, the laser profiles are assembled into an array of 3D points, which is referred to as a "range map". A color image is captured and registered with the range map at the completion of a scan. Captured frames are processed at PAL frame rate (25 Hz) on the 2.2 GHz dual Xeon host PC. Motor control and optical encoder measurements are implemented on a PIC microcontroller, which communicates with the host PC via an RS-232 serial link.

2.2. Coordinate Frames and Notation

The following sections adopt the convention of representing 3D points in upper case, such as \mathbf{X} , and two-dimensional (2D) points on the image plane using lower case, such as \mathbf{x} , with all vectors in homogeneous coordinates. Coordinate frames are specified in superscript, such as ${}^A\mathbf{X}$, and the homogeneous transformation matrix ${}^B\mathbf{H}_A$ transforms points from frame A to frame B as ${}^B\mathbf{X} = {}^B\mathbf{H}_A {}^A\mathbf{X}$. Figure 2 shows the relevant coordinate frames for the experimental scanner: L and R denote the left and right camera frames, P is rigidly attached to the light plane, and W is the world frame.

The cameras are modeled by the 3×4 projection matrices ${}^L\mathbf{P}$ and ${}^R\mathbf{P}$, which project points from the world frame onto the image planes according to the homogeneous transformation ${}^L,R\mathbf{x} = {}^L,R\mathbf{P}\mathbf{X}$. The cameras are arranged in a rectilinear stereo configuration (parallel axes and coplanar image planes), with optical centers at $\mathbf{C}_{L,R} = (\mp b, 0, 0, 1)^T$ (taking the negative sign for L and positive for R) in the world frame. In practice, the cameras can verge about the y -axis (violating the conditions of rectilinear stereo) but rectilinear measurements are recovered by applying "projective rectification" to captured frames (see Hartley 1999). Thus, without loss of generality, the analytical models consider only the rectilinear arrangement shown in Figure 2.

2.3. System Model

A theoretical model of the experimental scanner is now developed to identify points on the light plane. In frame P , which is rigidly attached to the light plane, points ${}^P\mathbf{X}$ on the plane may be represented by the plane equation ${}^P\boldsymbol{\Omega}^T {}^P\mathbf{X} = 0$, where ${}^P\boldsymbol{\Omega}$ represents the parameters of the light plane. Now, P is defined such that the light plane is approximately vertical and parallel to the z -axis, which allows ${}^P\boldsymbol{\Omega}$ to be expressed as

$${}^P\boldsymbol{\Omega} = (1, B_0, 0, D_0)^T \quad (1)$$

where B_0 is related to the small angle between the plane and the y -axis ($B_0 \ll 1$ for an approximately vertical plane), and D_0 is the distance of the plane from the origin. During a scan, frame P rotates about its y -axis with angle θ_y , where $\theta_y \triangleq 0$ when the light plane is parallel to the optical axes of the

cameras. The rotation axis intersects the xz -plane of the world frame at $(X_0, 0, Z_0)^T$, and the direction of the rotation axis relative to the y -axis of the world frame is defined by the small fixed angles θ_x and θ_z (see Figure 2). The scan angle θ_y is linearly related to the measured optical encoder value e via two additional parameters m and k :

$$\theta_y = me + k. \quad (2)$$

Now let ${}^W\mathbf{H}_P$ represent the homogeneous coordinate transformation from P to W . If $\mathbf{R}_x(\theta_x)$, $\mathbf{R}_y(\theta_y)$ and $\mathbf{R}_z(\theta_z)$ represent the homogeneous transformations for rotations about the x -, y -, and z -axis, and $\mathbf{T}(X_0, 0, Z_0)$ represents the transformation for translation by $(X_0, 0, Z_0)^T$, ${}^W\mathbf{H}_P$ can be expressed as

$${}^W\mathbf{H}_P = \mathbf{T}(X_0, 0, Z_0) \cdot \mathbf{R}_z(\theta_z) \cdot \mathbf{R}_x(\theta_x) \cdot \mathbf{R}_y(\theta_y). \quad (3)$$

It is straightforward to show that if ${}^W\mathbf{H}_P$ is the coordinate transformation from P to W , the plane parameters transform from P to W as

$${}^W\boldsymbol{\Omega} = ({}^W\mathbf{H}_P)^{-T} \cdot {}^P\boldsymbol{\Omega}. \quad (4)$$

Finally, combining eqs. (3) and (4) and making the simplifying assumptions $B_0, \theta_x, \theta_z \ll 1$ to eliminate insignificant terms, the laser plane parameters can be expressed in the world frame by the approximate model

$${}^W\boldsymbol{\Omega} \approx \begin{pmatrix} \cos \theta_y \\ \theta_x \sin \theta_y + \theta_z \cos \theta_y + B_0 \\ -\sin \theta_y \\ -X_0 \cos \theta_y + Z_0 \sin \theta_y + D_0 \end{pmatrix} \quad (5)$$

Now, using eqs. (2) and (5), points in the world frame can be identified as coincident with the light plane (given the system parameters and current encoder measurement) using the plane equation:

$${}^W\boldsymbol{\Omega}^T (B_0, D_0, X_0, Z_0, \theta_x, \theta_y, \theta_z) {}^W\mathbf{X} = 0. \quad (6)$$

3. Robust Stripe Scanning

3.1. Problem Statement

As discussed earlier, the problems associated with light stripe scanners result from ambiguity in identifying the primary reflection. The following sections describe an optimal strategy to resolve this ambiguity and robustly identify the stripe in the presence of secondary reflections, cross-talk and other sources of noise by exploiting redundancy.

Figure 3 shows a simplified plan view of the scanner to demonstrate the issues involved in robust stripe detection. The surface at \mathbf{X} causes a primary reflection that is measured (using a noisy process) at ${}^L\mathbf{x}$ and ${}^R\mathbf{x}$ on the stereo image planes. However, a secondary specular reflection causes another stripe to appear at \mathbf{X}' , which is measured on the right image plane at