Fourier Techniques and Monocular Vision for Simplistic and Low-Cost Visual Odometry In Mobile Robots

Ricardo Ramirez

Advisor: Dr. Randy C. Hoover

This work was made possible by the National Science Foundation REU Site: Bringing Us Together, Improving Communications and Lives (Award Number EEC-1359476).

Introduction

What Is Visual Odometry?

- Visual Odometry(VO) is the process of determining the location and motion of a robot motion by analyzing camera images as it moves through an environment.
- VO is widely used in robotics to facilitate localization, mapping, and navigation.

How it works:

- As a robot moves across an environment, video is captured by two(stereo vision) or a single camera(monocular vision).
- **Image registration** techniques are then used to calculate linear or rotational translation between consecutive frames.
  - This is the process of matching two images given translation, rotation, and scale differences between the two.
  - Prominent methods use feature matching, optical flow (measuring pixel based motion), or Fourier transform techniques.

Motivation

The development of accurate localization systems for mobile robots typically involves the use of several sophisticated and high-cost sensors (e.g., multiple cameras, GPS, IMUs). However, when robot design has strict cost or size constraints, use of multiple, high-cost sensors may not be feasible.

Typically, less sensing equipment leads to less accuracy in odometry and localization.

This motivates investigations/explorations to create comparably accurate VO system using simplistic and low-cost sensors (e.g., webcams)

Objective

Creating a Low-Cost approach to VO:

- Low end webcams offer good resolutions and speeds at a low cost.
- When used with a downward-facing camera, Fourier Techniques may be a good option in the localization process.
- This approach is attempted using a **Logitech Quickcam Fusion** webcam mounted on an **iRobot Create**.
- The aim is to attempt to accurately measure translative offset between consecutive frames to measure linear and rotational distance as the robot navigates in a controlled environment.

The iRobot Create and the Logitech Quickcam Fusion are the simplistic and low cost components used for this VO system.

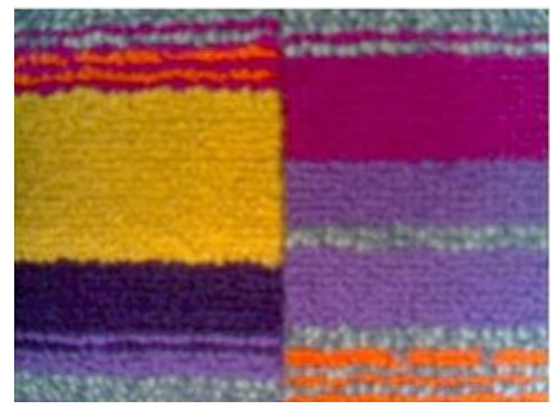
Procedure

Calculating translation between two frames using *Phase Correlation*:

Phase correlation is an approach which relies on a frequency-domain representation of the image, usually calculated by fast Fourier transforms.

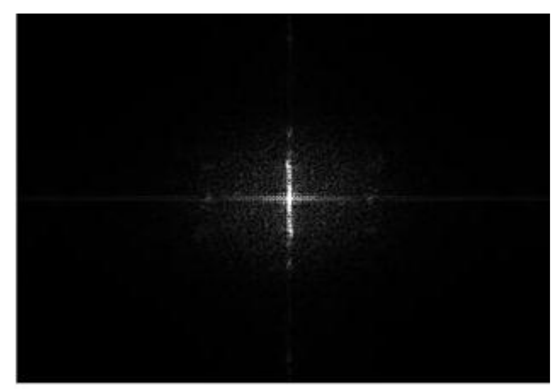
The Method:

1. Given 2 images g_a and g_b

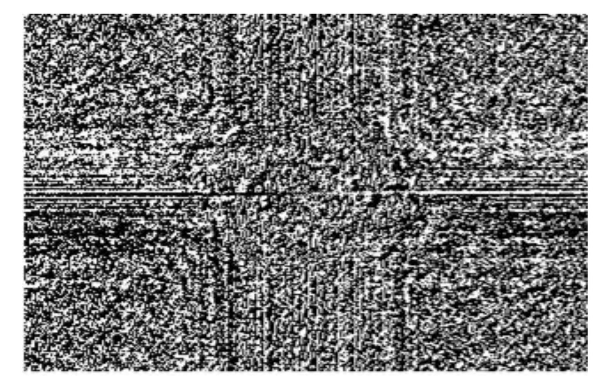


2. Calculate the discrete 2D Fourier Transform of both images: $G_a = F\{g_a\}, G_b = F\{g_b\}$

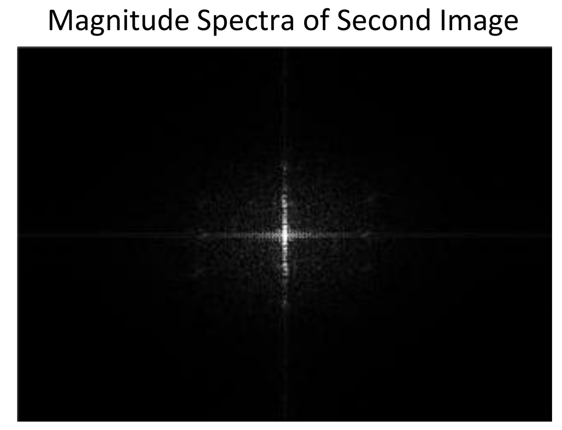
Magnitude Spectra of First Image



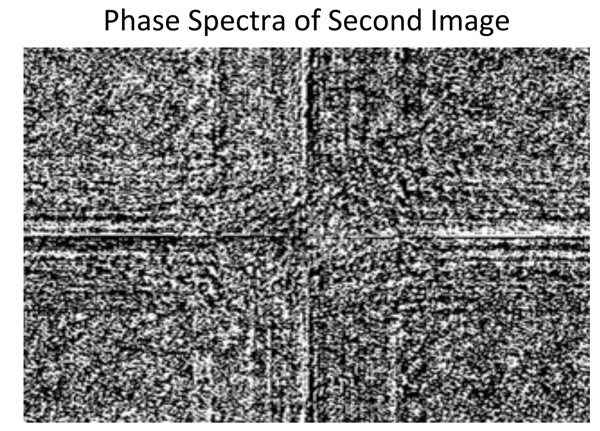
Phase Spectra of First Image



Magnitude Spectra of Second Image



Phase Spectra of Second Image



3. Calculate the cross-power spectrum by multiplying the first Fourier transform and the complex conjugate of the second and normalizing the product elementwise.

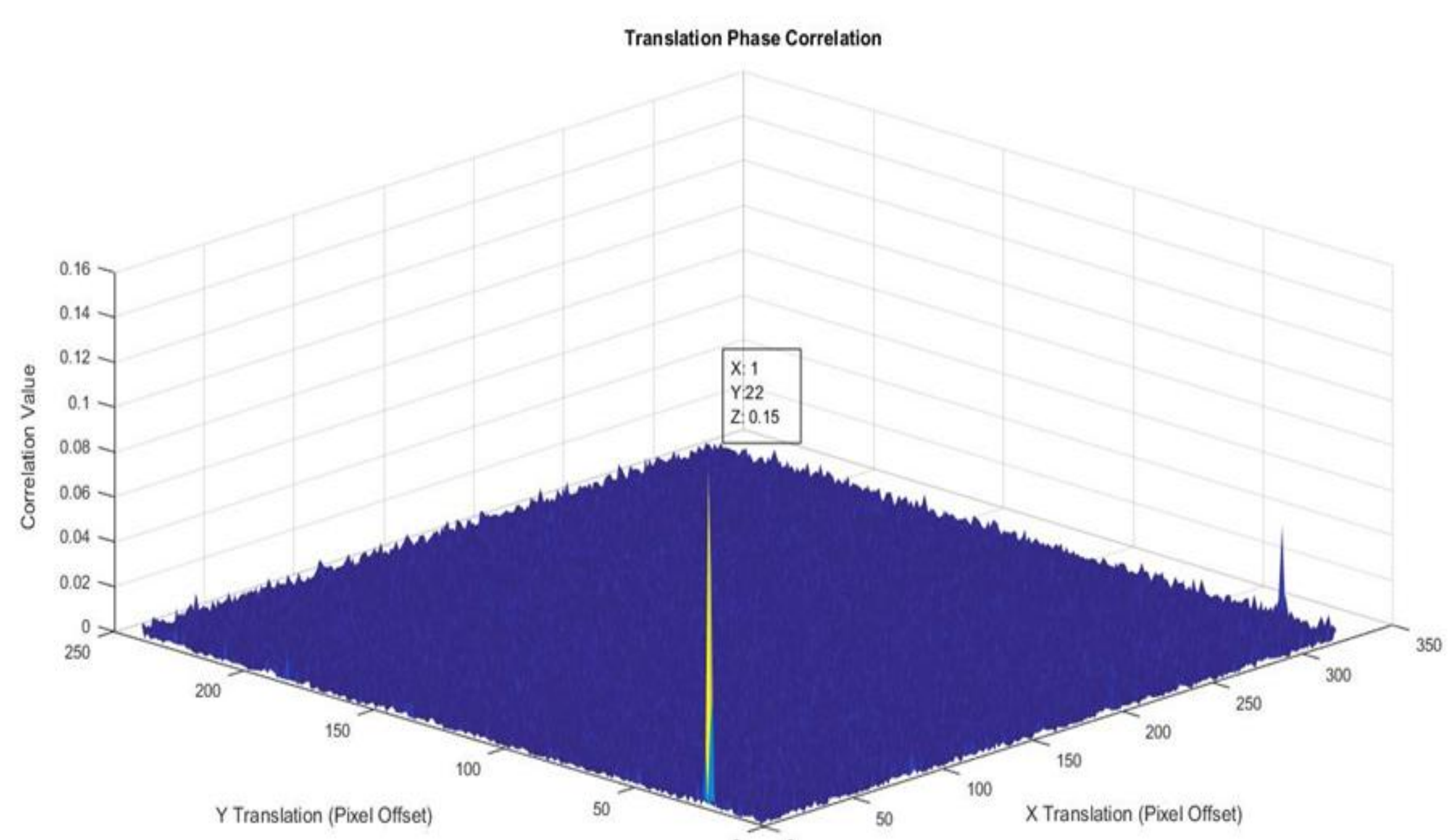
$$R = \frac{G_a \circ G_b^*}{|G_a \circ G_b^*|}$$

4. Apply the inverse Fourier transform to obtain the normalized cross-correlation.

$$r = F^{-1}\{R\}$$

5. Determine the location of the peak in **r**.

$$(\Delta x, \Delta y) = argmax_{(x,y)}\{r\}$$



The phase correlation is shown as a peak.

Trial Results

Performance

- To measure performance of this VO approach, linear displacement and a rotational displacement trials were performed over several surfaces.
  - Both linear and rotational trials were performed using images recorded with the robot and camera.
  - MATLAB's Image Acquisition Toolbox software package was used to capture and process the images.

Examples of Surfaces that were imaged during operation of VO system.

Linear Translation (Driving Forward):

VO linear translational displacement computations were compared to the iRobot's internal wheel odometer as well as physical measurements taken with a yardstick.

Throughout trial period, camera parameters and robot speed were varied to analyze differences between visual odometry and wheel odometry(WO).

Exhaustive trials proved the following parameters resulted in the best performance.

Frame rate:	**15fps**
Resolution:	**320x480**
Robot Speed:	**~7 cm/s**

When 50, 100, and 200 frames were registered with these parameters, the difference in translation distance between VO and WO was under 8%. Outside of these parameters, difference increased above 8%.

Rotational Translation (Rotating in Place):

Computing rotational translation posed problems with this particular VO approach. Trials consited of left and right rotations of the robot, while summing the 2D change in pixels as the robot rotated 45 or 90 degrees. Exhaustive trials showed no consitency in total pixel displacement despite continued change in camera parameters and rotational speed.

Discussion

This VO method shows limited success in computing translational displacement for the purpose of mobile robot localization. It was observed that a consistency of under 8% difference was maintained between WO and VO throughout forward movement trials when ideal parameters were set. It is possible that the VO method is accurate and that the constant difference in odometry is due to an unexplored camera setting.