

ESTIMATION OF TERRAIN GRADIENT CONDITIONS & OBSTACLE DETECTION USING A MONOCULAR VISION-BASED SYSTEM

Final Demonstration

Connor Luke Goddard (clg11)

Department of Computer Science, Aberystwyth University

BACKGROUND

BACKGROUND

The ability for an autonomous robot to navigate from one location to another in a manner that is both **safe**, yet **objective** is vital to the **survivability** of the machine, and the **success of the mission** it is undertaking.

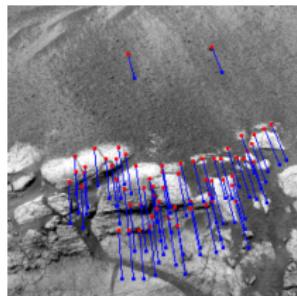


Figure 1: Tracked features from MER [?].

Vision based obstacle detection has enjoyed a high level of research focus over recent years. Less work has concentrated on providing a means of reliably detecting changes in terrain slope through the exploitation of observed changes in motion.

RELATED WORK

- [1] *A Robust Visual Odometry and Precipice Detection System Using Consumer-grade Monocular Vision*, J. Campbell et al., IEEE, 2005
- [2] *Obstacle Detection using Optical Flow*, T. Low and G. Wyeth, School of Information Technology and Electrical Engineering, University of Queensland, 2011
- [3] *Appearance-based Obstacle Detection with Monocular Color Vision*, I. Ulrich and I. Nourbakhsh, AAAI, 2000



Figure 2: Left: Precipice detection [1], Right: Obstacle colour detection [3]

PROJECT OVERVIEW

Focus

Investigation into a system capable of utilising a single, forward-facing colour camera to provide an estimation into current terrain gradient conditions, obstacle location & characteristics and robot ego-motion.

Potential Metrics For Observation

- Changes in terrain gradient (i.e. slopes).
- Location & characteristics of positive and negative obstacles (i.e. rock or pit respectively).
- Current robot speed of travel.
- Changes in robot orientation.

WORKING HYPOTHESIS

WORKING HYPOTHESIS

*“From images captured using a **non-calibrated** monocular camera system, analysis of optical flow vectors extracted using **localised appearance-based** dense optical flow techniques can provide certain estimates into the current condition of terrain gradient, the location and characteristics of obstacles, and robot ego-motion.”*

MOTION PARALLAX

Approach focussed on the effects of **motion parallax**.

i.e. Objects/features that are at a greater distance from the camera appear to move less from frame-to-frame than those that are closer.

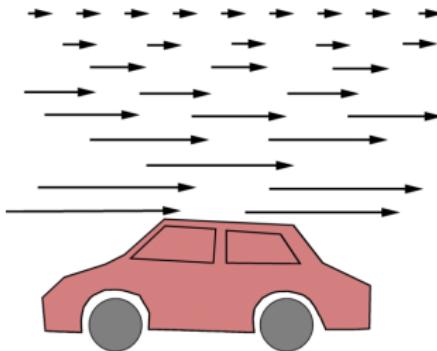


Figure 3: Typical example of motion parallax. Courtesy:
<http://www.infovis.net>.

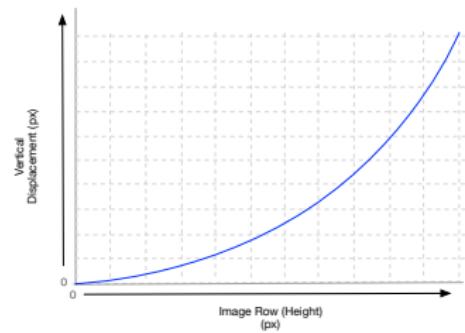
Key Aspects

1. The exclusive use of appearance-based template matching techniques to provide a localised variant of dense optical flow analysis over the use of sparse optical flow techniques relying upon a high number of available image features.
2. The use of a formalised model to represent detected changes in vertical displacement as part of efforts to estimate changes in terrain gradient in addition to the location characteristics of potential obstacles.

VERTICAL DISPLACEMENT MODEL

Provides a mechanism by which to evaluate the prediction:

"Features captured towards the bottom of an image will show greater displacement between subsequent frames, than those features captured towards the top of the image."



— Expected general trend between image row and vertical displacement demonstrated.

Figure 4: Example of “perfect” vertical displacement model.

INFERENCE OF TERRAIN SLOPE & POTENTIAL OBSTACLES

From observing discrepancies between the current model and “baseline” model, it should be possible to infer the presence of potential obstacles:

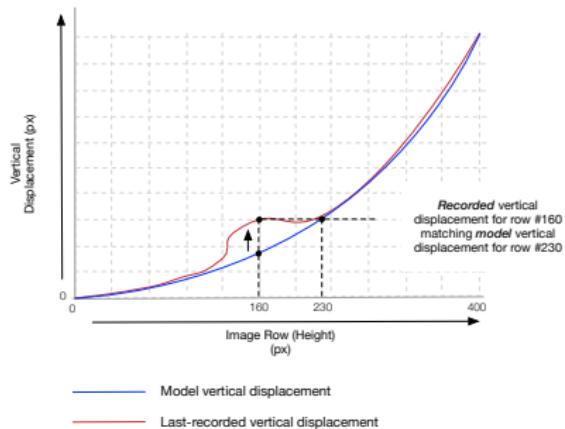


Figure 5: Vertical displacement model indicating the potential presence of a positive obstacle.

INFERENCE OF TERRAIN SLOPE & POTENTIAL OBSTACLES

Differences in model discrepancy behaviour should provide enough information to establish between a potential obstacle and a change in terrain slope:

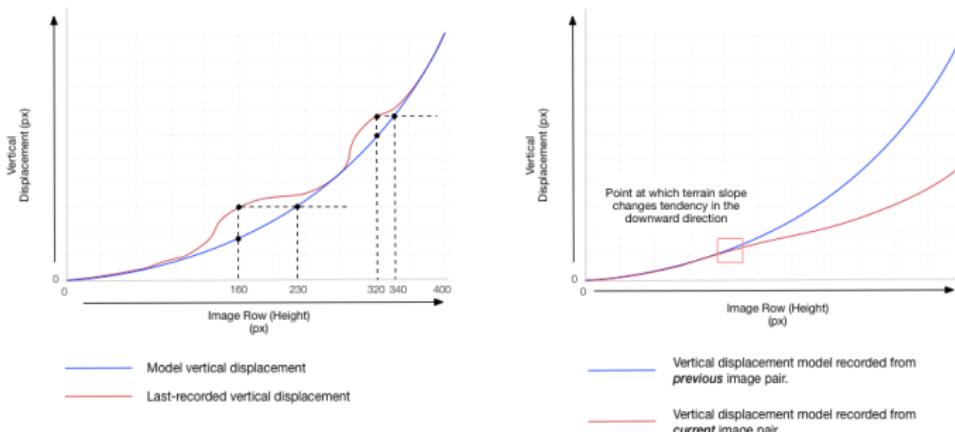


Figure 6: Example of observed differences between detection of obstacles and change in terrain slope.

ROBOT ROTATION

Two types of observed motion:

- Translation (*Vertical* component of optical flow vectors)
- Rotation (*Horizontal* component of optical flow vectors)

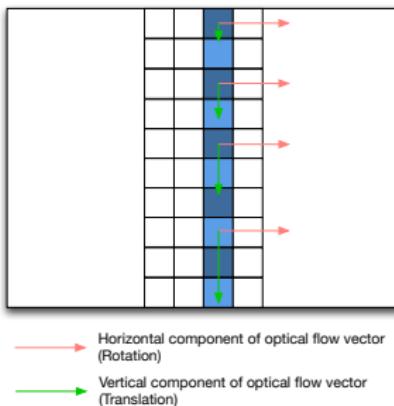


Figure 7: Diagram demonstrating the expected difference in behaviour between the separated vertical and horizontal components of observed optical flow vectors calculated from *forward-turning motion*.

ROBOT SPEED

By recording multiple “calibration” models at different speeds, it should be possible to later infer the current speed using the standard *speed-distance-time* equation.

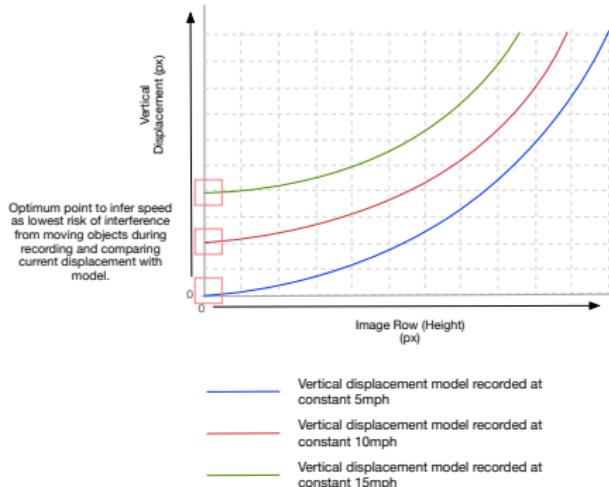


Figure 8: Diagram indicating examples of alternative vertical displacement models calculated while travelling at different constant speeds.

Primary Aims

1. Establish which out the following appearance-based template matching similarity measures:

- Euclidean Distance
- Correlation Coefficient
- Histogram Correlation
- Histogram Chi-Square

best supports the predicted positive correlation between the vertical position of features within an image, and the level of vertical displacement demonstrated.

2. Use the generated model of vertical displacement to identify potential obstacles and changes in terrain gradient based upon the predicted ‘comparative trend’ behaviour.

INVESTIGATION AIMS

Secondary Aims

1. Further extend the capabilities of the vertical displacement model to provide estimates into the current speed of travel demonstrated by the robot.
2. Add support for providing an estimation into the change in orientation.

EXPERIMENT METHODS

GATHERING DATASETS

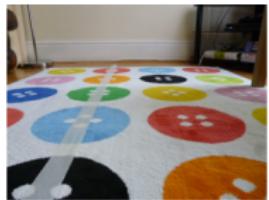
No existing imagery datasets were found to be suitable, therefore new sets had to be collected.



Figure 9: Camera-rig setup used to capture the experiment datasets from front and back profiles.

GATHERING DATASETS

Images taken at a number of locations, but **four** were eventually selected based upon their variety in terms of **terrain texture** and **lighting conditions**.



Site 1: Living Room Rug



Site 2: Brick Paved Road



Site 3: Asian Rug



Site 4: Slate Tile Footpath

Figure 10: Examples of the first frame captured from each of the four location datasets.

EXPERIMENT METHODS

Over time, **investigation focus changed** from exploring a series of reasonably “generalised” research aims, to instead concentrating almost exclusively on one small, but very important aspect;

Ensuring that accurate results for appearance-based motion displacement could continue to be achieved, across terrain lacking any significant features and/or demonstrating highly repetitive patterns.

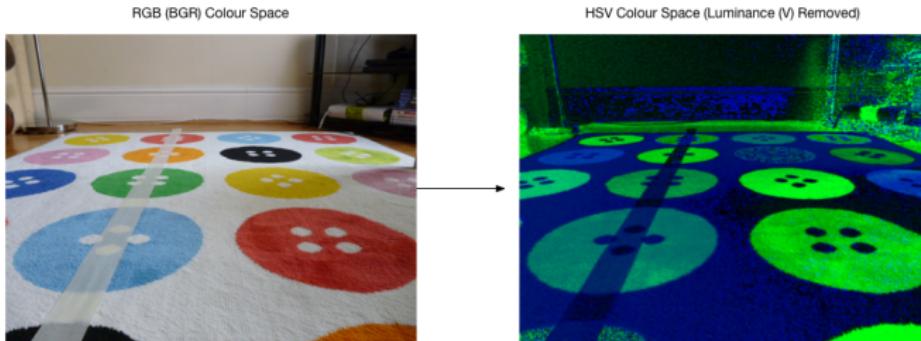
Three Experiments

1. Template Matching - Multiple Small Patches
2. Template Matching - Full-width Patches | Non-Scaled
3. Template Matching - Full-width Patches | Geometrically Scaled

EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage One

Import two consecutive images, and convert from RGB (BGR in OpenCV) colour space to HSV.

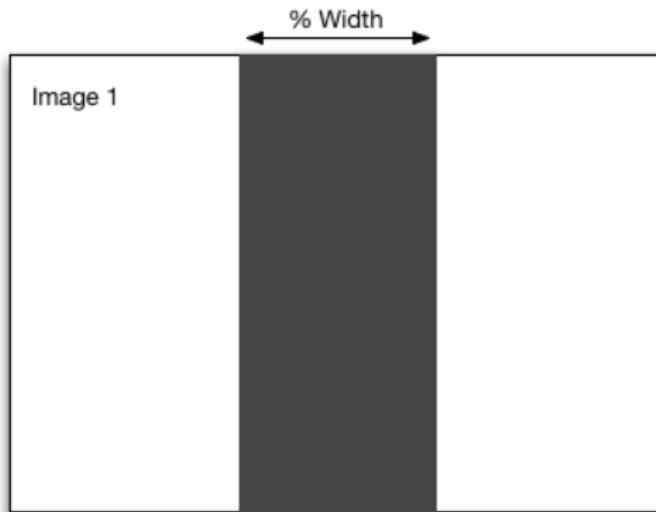


The 'V' channel is then removed in order to improve robustness to lighting changes between frames.

EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage Two

Extract a percentage-width region of interest (ROI) from centre of first image.



Why do we need to extract a ROI?

Focus-of-expansion: Objects in images do not actually move in 1-dimension (i.e. straight down the image).

This effect is minimised towards the centre of the image.

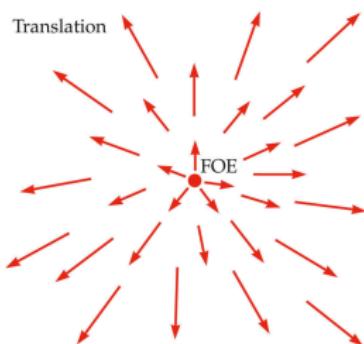


Figure 11: Courtesy: J.Pillow, University of Texas

EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage Three

Extract patches of a fixed size around each pixel within the extracted ROI.

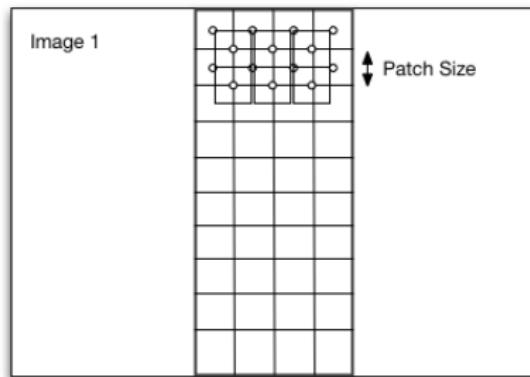


Figure 12: Simplified example of patch extraction within ROI.

EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage Four

For each patch extracted from *image one*, move down through a localised search window (column) in *image two* searching for the best match against the template patch.

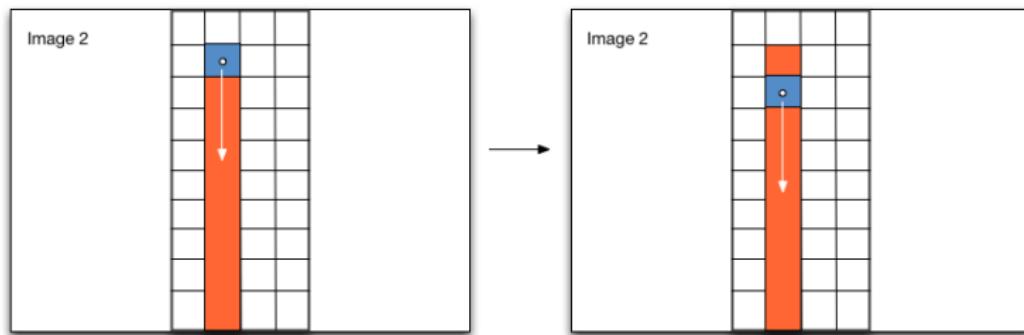
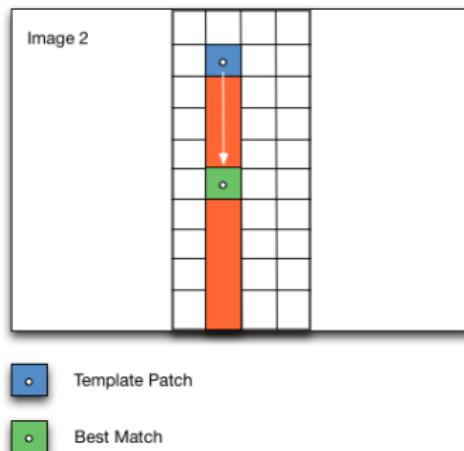


Figure 13: Example of “best match” search within local column.

EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage Five

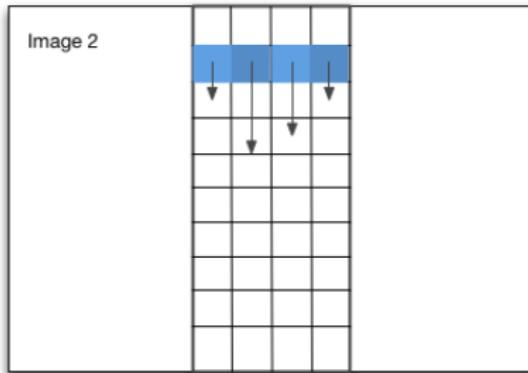
Identify the patch within the localised search window that provides the “best match” via correlation-based matching (e.g. Euclidean Distance, SSD or Correlation coefficient).



EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage Six

Average all measured displacements for each pixel along a given row.

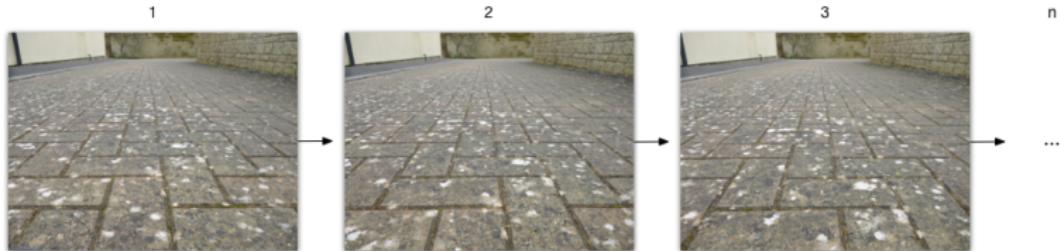


Outliers are removed by ignoring any displacements that lie outside of $(2 \times \text{Standard Deviation})$ of the mean.

EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Repeat Stages 1-6

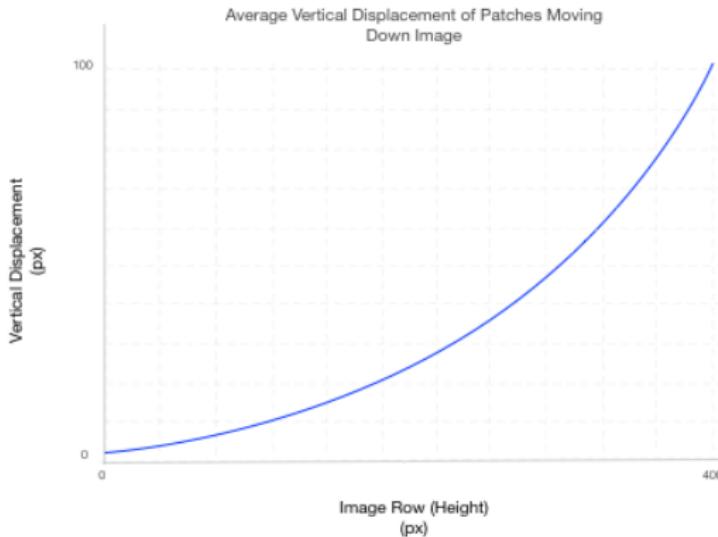
Repeat stages 1-6 for an entire collection of “calibration” images taken of *flat, unobstructed* terrain.



EXPERIMENT ONE - MULTIPLE SMALL PATCHES

Stage Seven

Plot the *average displacement* for each ROI row, calculated from the displacements recorded over all calibration images.



EXPERIMENT TWO - FULL-WIDTH PATCHES (NON-SCALED)

Focus

Establishing if adopting *larger* patches in fewer numbers provided better appearance-based matching accuracy than using many more, but critically much *smaller* overlapping patches.

While the underlying method remained the same as in the first experiment, some key changes were made:

- Extracting a central region of interest relative to the **ground plane**, as opposed to extracting a fixed-width column from the image plane.
- Moving away from multiple overlapping small patches, in favour of adopting a single, **full-width** patch to represent a single row in the image.

EXHAUSTIVE VS. NON-EXHAUSTIVE SEARCH

Issue discovered regarding the use of **non-exhaustive** searching causing erroneous results. All tests in experiments two and three were conducted using both exhaustive and non-exhaustive approaches.

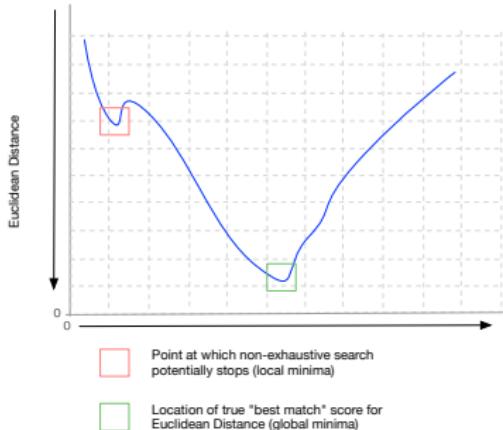


Figure 14: Example of potential early termination of non-exhaustive search due to entering local minima caused by noise, rather than global minima which represents to true best match.

PERSPECTIVE DISTORTION CALIBRATION TOOL

To extract a region-of-interest along the *ground plane*, the system would need to take account of *perspective distortion*. Therefore, a simple **calibration tool** was implemented, enabling users to define the region of interest within an image.



Figure 15: Results of the verification test for the approach towards geometric scaling of template patch pixel coordinates.

EXPERIMENT THREE - FULL-WIDTH PATCHES (SCALED)

Focus

Add **geometric scaling** of template patch in order to account for objects appearing larger as they approach the camera.

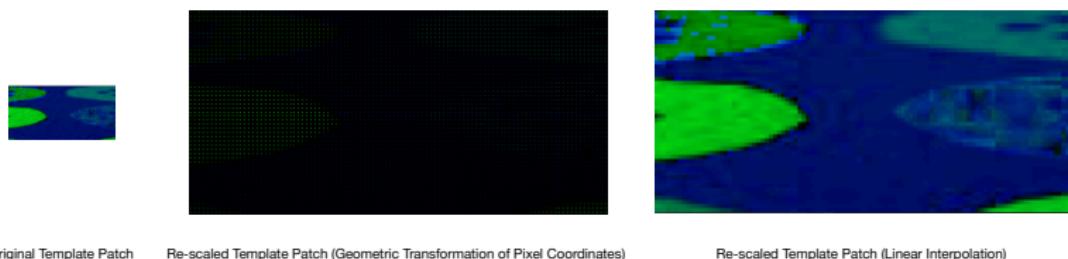


Figure 16: Comparison between the various approaches of performing scaling of the original template patch image. In the case of linear interpolation, additional information has been added to image as a result of “estimating” the colour that lies between two original pixels.

EXPERIMENT THREE - FULL-WIDTH PATCHES (SCALED)

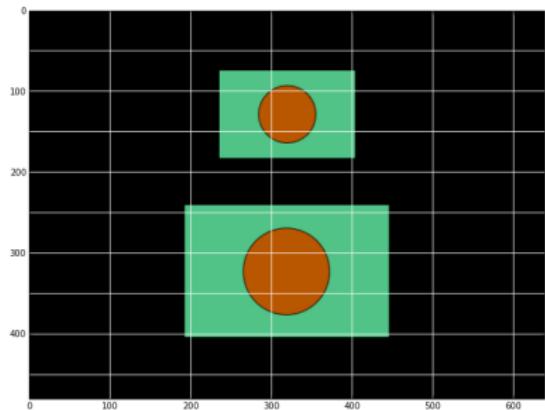
6 Stage Process

1. Obtain the width and height of the current template patch.
2. Obtain the calibrated width of the search window with respect to the current image row.
3. Calculate the independent scale factor for the width and height between the template patch and the search window.
4. Treating each pixel as a 2D coordinate in geometric space, scale the position of each pixel coordinate within the template patch relative to the position of centre coordinate.
5. For each scaled pixel coordinate, extract the value of the pixel at the corresponding position with the search window and add to a temporary “image” data structure.
6. Perform template matching between the original template patch, and the new “extracted” temporary image.

EXPERIMENT THREE - FULL-WIDTH PATCHES (SCALED)

Testing of Scaling Approach

A simple test was devised to confirm geometric scaling worked as expected under “perfect” conditions.



Euclidean Distance Scores - Scaled Template Matching

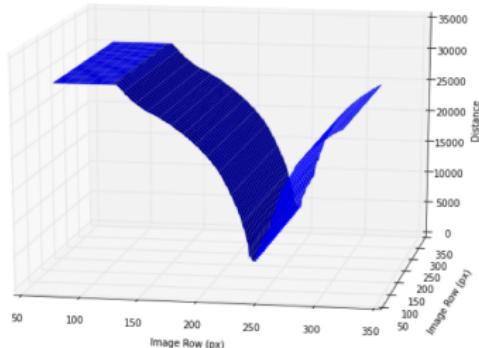


Figure 17: Results of the verification test for the approach towards geometric scaling of template patch pixel coordinates.

INVESTIGATION RESULTS

RESULTS

Generally “mixed” results at this stage.

From the results obtained, we have learned:

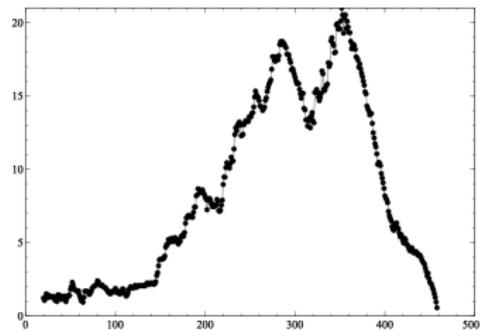
1. It is possible to potentially establish a relationship between row height in an image, and average downwards pixel displacement.
2. The current approach to appearance-based tracking using multiple “small” patches **does not work well** for many typical terrain conditions (i.e. outside).

RESULTS: EXAMPLE 1

Input Collection:

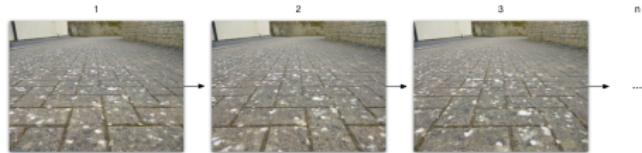


Result:

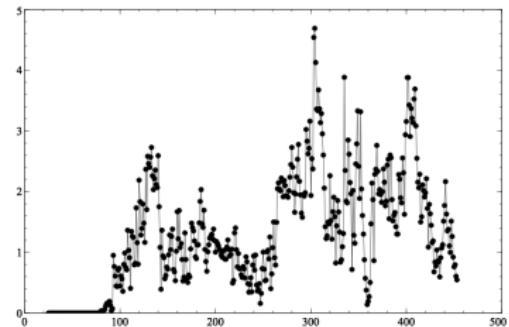


RESULTS: EXAMPLE 2

Input Collection:



Result:



EVALUATION & CONCLUSION

FUTURE WORK

FUTURE WORK

SUMMARY

QUESTIONS?

PROJECT BLOG

SLIDE DESIGN: MATTHIAS VOGELGESANG - (GITHUB)