Tackling the Scale Factor Issue in a Monocular Visual Odometry Using a 3D City Model

Paul Verlaine Gakne and Kyle O’Keefe

Position, Location And Navigation (PLAN) Group
Department of Geomatics Engineering
University of Calgary, Canada

ITSNT 2018, Toulouse, France
13th – 16th November
Contents

• Introduction
  • Various existing solutions
  • Scale Factor (SF) ambiguity and solutions

• Objectives

• Methodology
  • Skyline-based positioning (concept, matching, similarity metric)

• Slant distance-based SF computation

• Experiment and results

• Conclusions
Introduction

- In open sky environment with good satellite visibility, GNSS provide seamless and accurate solutions.

- However, in urban canyons:
  - poor GNSS availability
  - corrupted GNSS signals

- In such areas, signals are classified in three groups:
  - Direct
  - Shadowed
  - Blocked
Various Existing Methods

- Camera/Vision
- Lidar
- Ultrasonic
- Infrared
- Mobile Signals (e.g., LTE)
- Radar
- Barometer
- Odometer
- GNSS
- Gyroscopes
- IMES
- Ultra-Wideband (UWB)
- GNSS Repeaters
- Bluetooth
- WiFi
- Accelerometers
- TV
- Magnetometers

3 DBM
With a monocular system, overall scale cannot be directly measured.
Scale Factor (SF) Ambiguity – 2/2

![Graph showing the Scale Factor (SF) Ambiguity](image)

Legend:
- **Reference**
- **VO**
- **Start** (green square)
- **End** (red square)
Solutions to the SF Ambiguity

- INS can help to obtain scale, but INS error will lead to scale error, which leads to error in aiding INS, which leads to scale error.

- There must be some absolute knowledge of scale somewhere in the system in order to avoid scale uncertainty:
  - Baseline between stereo cameras
  - Feature locations/relative separations are known
  - Depth from defocus/coded apertures
  - In an aircraft, knowing altitude and ground height
Objectives of the Research

• Overcome the scale factor issue observed in a monocular VO

  • Combine the information obtained from a 3DBM by matching synthesized 3DBM images and camera images and also with VO

  • Calculate the scale factor from the 3DBM to increase the translation magnitude accuracy for monocular VO
Input Images

Features Detection, Extraction and Matching

Motion Estimation

\[(r_c, t_c) = \arg\min_{R_c, T_c} \sum_{i=0}^{M-1} \omega_i \left\| (R_i p_i + T_c) - p_i \right\|^2 \]
Methodology – 3/3

• Extract “ideal” skylines from a 3DBM

• Obtain “observed” skyline images from an upward facing camera

• Rectify and segment the observed images

• Match the segmented “observed” and “ideal” skylines
Skyline-based positioning – 1/2

• Ideal skylines (georeferenced images) generated from a 3D city model at any location of the user.
Skyline-based positioning – 2/2

• Skyline parametrized as function of azimuth. For each azimuth, the highest elevation angle amongst all the obstructing surfaces was recorded with the corresponding height

\[ \zeta_p = \left\{ (\varepsilon_{p,j}, h_{p,j}) \right\}, \; j = 0..N - 1 \]

• \( \zeta_p \) is the ideal DB used for comparison with the observed images
• Example of a 3DBM-based synthesized image at a given location:
Skyline-based positioning – 2/2

Observed Camera Image

3DBM Synthesized Image

1288

728
Skyline Matching

- At this stage we have 2 types of binary images: from the 3DBM and the observed image.

- Similarity metric used is the cross-correlation coefficient:

\[ CC_{I_{b(p)}}(u,v) = \frac{1}{n} \sum_{u,v} \left( [I_b]_{\text{cam}}(u,v) \circ [I_b]_{3D(p)}(u,v) \right) \]

- User position: position in the DB corresponding to the best match.
Similarity Metric
Scale Factor Computation

- After the system initialization, the scale factor can be computed:

\[ \hat{s}_k = \sqrt{(\hat{x}_k - \hat{x}_{k-1})^2 + (\hat{y}_k - \hat{y}_{k-1})^2 + (\hat{z}_k - \hat{z}_{k-1})^2} \]

- When the 3DBM solution is available:

\[ s_g = \frac{d_s}{\|P_{3DMB} - P_{VO}\|} \]

Where the slant distance is defined as:

\[ d_s = \frac{h}{\sin \varepsilon} \]
• Experiment conducted in downtown Calgary
Experiment and Results – 2/3

![Graph showing scale error over time for VO and VO-aided 3DBM]
Conclusions

- Scale factor recovery for monocular visual odometry was presented
  - skyline-based positioning using a narrow FOV visual spectrum camera
  - Slant distance obtained from a 3DBM used to recover the SF and aid the VO

- Results showed the effectiveness of the proposed method
  - 90% of the time improvement compare to VO-only
THANK YOU!

pvgakne@ucalgary.ca

plan.geomatics.ucalgary.ca