literature research

Wolfgang Baumgartner

September 22, 2017

Contents

wes

1 Chapter 1

1.1 Read Literature

- An Evaluation of the Suitability of FPGAs for Embedded Vision Systems, 2005
- Embedded vision system for real-time object tracking using an asynchronous transient vision sensor, 2006
- Stereo Vision System on Programmable Chip (SVSoC) for Small Robot Navigation, 2006
- A 128x128 120dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor, feb 2008
- A Tele-operated Gesture Recognition Mobile Robot using a Stereo Vision, 2008
- Smart cameras, 2010
- Dynamic stereo vision system for real-time tracking, 2010
- Accurate hardware-based stereo vision, 11.08.2010
- PIXHAWK: A micro aerial vehicle design for autonomous flight using onboard computer vision, 23.02.2012
- Low-cost FPGA stereo vision system for real time disparity maps calculation, 27.02.2012
- Parallelization of Belief Propagation on Cell Processors for Stereo Vision, June 2012
- An Overview to Visual Odometry and Visual SLAM: Applications to mobile robotics, 13.11.2015
- Precise localisation of a UAV using visual odometry, December 2015
- Potential-path approach for UAV with movable camera on-board, January 2017

1.2 Summaries

1.2.1 Dynamic stereo vision system for real-time tracking

- paper describes an embedded system with two special vision sensors and a DSP
- objective is the tracking of objects in address-even representation
- intensity changes over time are registered as events
- events are grouped together with a bounding box to recognize moving objects
- recognized objects can be tracked
- use of AER reduces necessary computation power
- DVS very expensive
- use the event stream coming from the sensor, rectify pixels, put them in a time frame and match corresponding pixels

1.2.2 Embedded vision system for real-time object tracking using an asynchronous transient vision sensor

- uses DVS for monitoring of traffic and persons tracking
- uses event(pixel detecting intensity change) clusters to track objects
- every new event is added to a cluster and influences cluster center
- path of cluster is path of cluster center

1.2.3 A 128x128 120dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor

- this paper describes a sensor that asynchronously detects intensity changes and sends the address of the corresponding pixel
- this includes circuit design, testing, previous work
- seems hard to replicate this exact idea on an fpga
- this is the DVS sensor used in the above paper

1.2.4 Smart cameras

- a book about all kinds of image sensors, image processing systems
- interesting chapter about embedded stereo vision
- extensively describes algorithm in 1.2.1

1.2.5 PIXHAWK: A micro aerial vehicle design for autonomous flight using onboard computer vision

- paper extensively describes a lightweight drone with a unusually powerfull computation platform
- drone capable of using complex computer vision algorithms normally reserved for heavier drones
- improves navigation by combining inertia measurment unit with computer vision data
- objective is to build an autonomous drone that does all calculations online
- article not wholly finished as extensively describes all aspects of design

1.2.6 Parallelization of Belief Propagation on Cell Processors for Stereo Vision

- paper explains complicated algorithm that minimizes an energy function to find matching pixels between the two images of a camera pair
- algorithm is rather complicated
- algorithm is very accurate but needs a lot of resources
- shows that the algorithm can be used on a multicore system using parallelization to achieve about 6 fps
- necessary resources not available on ramstix, therefore not relevant

1.2.7 A Tele-operated Gesture Recognition Mobile Robot using a Stereo Vision

- bad English, not very nice to read
- service robot that recognizes face and hands and hand gestures
- does computations on board and on a server
- mention an FPGA but not clear if it is on the server or on the robot
- main topic is an algorithm for face recognition and hand gesture recognition

1.2.8 Low-cost FPGA stereo vision system for real time disparity maps calculation

- describes a complete system that performs stereo vision
- implementation on a XlinixVirtex-4 XC4VLX60 FPGA chip
- consists of image acquisition, image rectification, disparity computation and validation
- reaches over 300 fps for a resolution of 640x480
- claims to perform well while costing 1/10 of comparable systems

1.2.9 Accurate hardware-based stereo vision

- article about a new algorithm for real-time systems to be implemented on FPGAs or ASICs
- uses sum of absolute differences (SAD) in combination with a modified census transform (MCT)
- these two are combined to get rather high accuracy for an embedded system while reaching a framrate of 60 fps
- article complicated, explains used algorithm in detail
- write an IP core for FPGAs that can be adapted to the size of the FPGA, this way it also fits on smaller FPGAs

1.2.10 An Evaluation of the Suitability of FPGAs for Embedded Vision Systems

- looks at advantages and disadvantages of using FPGAs in an embedded computer vision context
- FPGAs are very powerful when one can use parallelism
- Seems to be the case for many computer vision algorithms
- biggest disadvantage is the need to use fixed-point arithmetic
- floating-point arithmetic is very costly
- easy to make and change a design

1.2.11 Stereo Vision System on Programmable Chip (SVSoC) for Small Robot Navigation

- describes a stereo vision system that recognizes obstacles and plans a possible path
- uses three cameras together with a board that combines FPGA and DSP
- stereo vision seems to be done only on FPGA
- allows a hexapod drone to avoid obstacles

1.2.12 Potential-path approach for UAV with movable camera on-board

- RAM report of Elisa Rimondi about a camera that stays focused on a high-feature view to improve localisation of drones
- extensive with a lot of background information
- stereo visual odometry search term for more suitable literature
- explains different position systems such as GPS or odometry

  [rro] How does visual odometry work?

  [rro] What does stereo vision add to visual odometry and when is it necessary?
- SLAM is the process of generating a map and simultaneously the position of a drone in it while both depend on each other

  [rro] How important is SLAM for Aeroworks? Seems that map creation is not necessary
- active vision means that a drone actively chooses a path with a lot of features to reduce localisation uncertainty

  [rro] active vision insofar relevant as it doesn't change the planned path, i.e. camera orientation changes to search high-feature areas
- uses Robot Operating System (ROS)
- ROS is a framework that basically provides a way of connecting different kinds of software elements as nodes

- it provides topics that all nodes can publish to and read messages from
- describes different forms of feature recognition

  [rro] What criteria are important for the choice of a feature detector?
- every pixel in image gets an energy and the camera points at location with lowest energy to keep camera as stable as possible
- extensive description of system architecture with all ros-nodes, coordinate systems and transformations
- improves localisation by focussing camera on features
- only achieves 15Hz because of ROS latency

  [rro] How can frequency be improved?
- proposes using ROS2 in combination with global shutter high framerate camera with large field of view

  [rro] Can stereo vision be done with large field of view? fish-eye lens?
- would be interesting to compare Elisa's method with other methods of odometry to evaluate performance and combine localisation data with other sensor data

1.2.13 An Overview to Visual Odometry and Visual SLAM: Applications to mobile robotics

- recounts history of VO and VSLAM
- VO: match features in different frames to compute translation matrix

  [rro] How can that be implemented?

  [rro] How can VO data be combined with other sensor data to improve accuracy? (Kalman filter)
- RANSAC often used to improve feature detection

  [rro] What is RANSAC? (see further below)
- explains different kinds of SLAM
- explains different feature extraction algorithms and how to associate features in different frames

  [rro] What are good criteria to choose the most suitable feature extraction algorithm and data association?
- RANSAC is an algorithm to find outliers in measurement data
- Bundle Adjustment seeks to optimize camera pose and 3D structure parameters by minimizing a cost function

1.2.14 Precise localisation of a UAV using visual odometry

- aims at using stereo vision on a UAV in combination with IMU for localisation
- uses an AEROWORKS drone with vi-sensor
- fuses SVO measurements with IMU with a kalman filter

  [rro] How does a kalman filter work? (described in this paper)
- makes use of ROS
- chooses between different implementation options for localisation and measurement fusion

  [rro] What does noise look like in VO?
- uses an intel NUC i5 with Linux Ubuntu 14.04 and ROS Jade
- SVO is fused with IMU data to get more accurate localisation and orientation of the drone

- IMU has drift over time but SVO as well
- when SVO fails, no pose is available, therefore backup pose estimation used
- measurements from flight data
- error in some directions bigger than in others, yaw introduces error
- transformation between coordinate frames important
- motors introduce noise for IMU
- contains recommendations such as using more sensors or eliminating drift with SLAM

[rro] Other ways of reducing drift in VO?

A Appendix 1

Tip: Make a copy of this document, since this is a manual of the template. This will prevent losing the document while modifying it. (And probably breaking it.)

A.1 Required TexLive packages

The following packages are required for TeX Live 2012 and 2013 (tested under Ubuntu)

```
texlive
texlive-lang-dutch      -> Dutch hypenation
                           (old TeX Live distributions, 2013-)
texlive-lang-europeans -> Dutch hypenation (amongst others)
                           (new Tex Live distributions 2013+)
texlive-latex-extra
texlive-fonts-extra     -> fourier
texlive-humanities      -> lineno
texlive-pstricks        -> pstricks (2013+?)
                           (not really used in the template??)
```

Bibliography