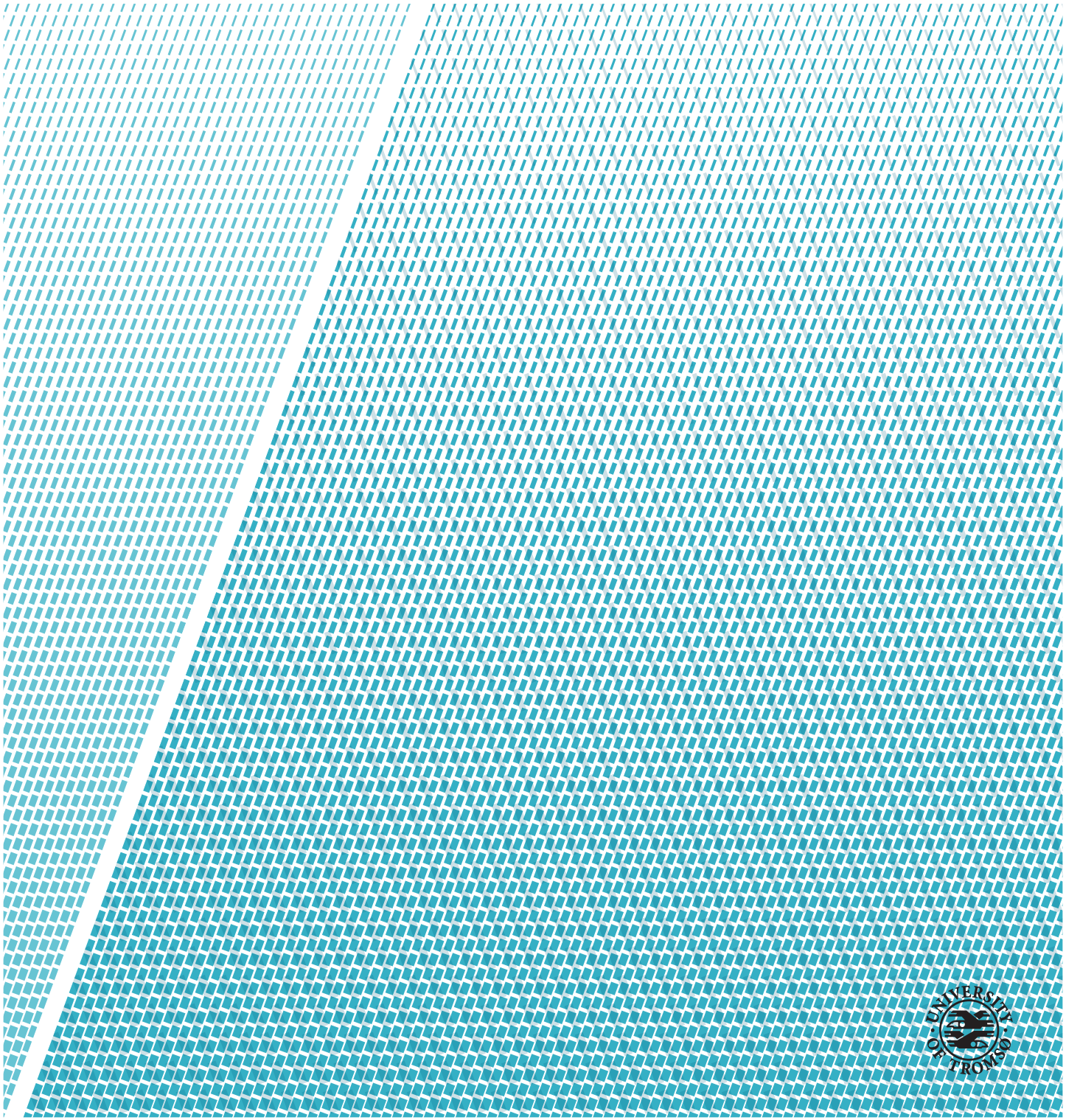
UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

Faculty of Science and Technology
Department of Computer Science

From Physical to Virtual Sensors (PVS)

Camilla Stormoen

INF-3983 Capstone Project in Computer Science . . . December 2017



Abstract

W3 Whats wrong with the word? / motivation 1-3 setninger

Architecture - 1-3 setninger

Design- 1-3 setninger

Implementation - 1-3 setninger

Experiments - 1-3 setninger

Results - 1-3 setninger

Lessons learned/main conclusion - 1-3 setninger

Kutt heller etterpaa

This dissertation present/describe ...

Contents

List of Figures

1 Introduction

- Mention focus on camera-sensors/data, and not other sensors?!?[1]
- Talk a little bit about COAT in general?

This project will develop an abstraction for virtual sensors, and do a prototype of the abstraction on a set of computers with physical sensors.

The purpose is to provide for a more powerful and flexible sensor in the COAT monitoring of the arctic tundra. As such, a fox feeding station is the usage domain to be used for the prototype.

1.1 Motivation

The motivation!

- W3
- Problem definition: This project investigated ... x, with the purpose of y.

The motivation behind this project is that no single sensor may cover the sensing needs, and that sensing needs can change rapidly over time. Consequently, there is a need for sensor fusion, and allow for combining sensors at different

computers.

1.2 Contributions

What was the contribution?

1.3 Assumptions

AVGRENSE, VIKTIG!! Something about motivation and stuff

1.4 Limitations

AVGRENSE, VIKTIG!!

- Mention focus on camera-sensors/data, and not other sensors?!

1.4.1 A subsection

2

Background and Related Work

- Taking Sensor Networks from the Lab to the Jungle
- Wireless Sensor Networks for Habitat Monitoring
- Building Virtual Sensors and Actuators over Logical Neighborhoods
- A virtual sensor system for user-generated, real-time environmental data products
- Integrating modeling and smart sensors for environmental and human health
- Capability representation model for heterogeneous remote sensing sensors: Case study on soil moisture monitoring
- Dice: Monitoring Global Invariants with Wireless Sensor Networks

2.1 Something

gggg

3

Architecture

Functionalities, abstractions, tell it clean/neat.

There are 6 components in the system: physical sensors, data storage, fused data, virtual sensors and the user. However, the main components in the system are the data storage, the fused data, the virtual sensors and the user. In this chapter, we will describe the architecture of the data storage, fused data, virtual sensors and the user. The architecture of the system is presented in Figure 3.1.

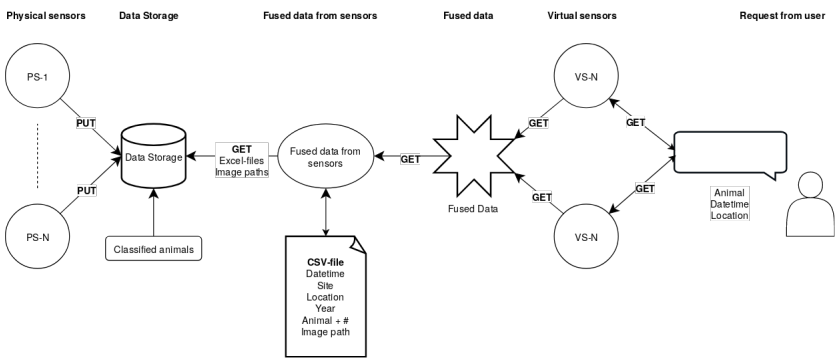


Figure 3.1: Figure showing the system architecture.

3.1 Physical Sensors and Data Storage

The physical sensors transmit their data to the data storage. The data storage consists of several set of images from different sensors and excel-sheets containing information about each picture. The physical sensors are placed at different location in the Arctic Tundra such as Nordkynn and Varanger, which includes Komag, Nyborg and Stjernevann.

3.2 Fused Data

The fused data retrieves it's data from the data storage and store the fused data into an CSV-file on demand. The fused data is the image-path from all the directories in the data storage combined with necessary information from the excel-sheet such as date-time, location, site, year, what kind of animal was in the image and also how many animals there was.

3.3 Virtual Sensors

The virtual sensors are divided such as each sensor represent an animal that scientists are interested in, e.g. one raven-sensor, one red fox-sensor and one golden eagle-sensor. The user types in what animal it wants to see, where it is and the date-time and the search is redirected to the sensor related to that specific animal. The virtual sensor receive its result from the fused data from the CSV-file.

3.4 Result from Virtual Sensor to User

A user wants to retrieve images from the data store. The user interacts with a user application. This application takes care of the communication to the virtual sensors. When a user gives a command/commands, the user application sends it to one of the virtual sensors, the virtual sensor retrieve data from the fused data as described in the section above, and deliver the response back to the user. The image(s) are displayed through an image-vision library/software.

4

Design

Client/Server, p2p, put/get, pub/sub, protokoller etc.. BESKRIV INTERAKSJONEN MELLOM ENHETENE!!

Virtual sensors probably uavhengige prosesser, ikke threads ettersom man evt vil addere flere sensorer og unnga a starte alle sensorer på nytt igjen.. Er de virtuelle sensorene servere eller client/publisher?

4.1 Data Storage

4.2 Fused Data from Sensors

Rekursiv traversering av directories og leser metadata fra bildene (ca 1,6 mill bilder)) - Lagres i en dictionary hvor datetime og sted er key og image pathen er value).

Ca 16-17000 rows i et excel-ark. Leser ut info som dyr, antall, sted, datetime, site, year og putter i en string.

Disse sammenlignes og de som matcher blir skrevet til en CSV-fil.

4.3 Virtual Sensors

4.4 User

5

Implementation

Threads, data structures, language ... Pandas (dataframe [1] [2]), CV2 (show image), exifread, Python 2.7, missing testing (CPU, memory, time?)

The system is implemented and written in Python 2.7[3] because .. (frameworks available in this language??).

To visualize/show pictures, a Python library called OpenCV [4] was implemented. To read exif/metadata from pictures, we used a Python library called exifread 2.1.2 [5].

1. https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.html
2. http://pandas.pydata.org/pandas-docs/stable/
3. https://www.python.org/
4. https://opencv-python-tutroals.readthedocs.io/en/latest/
5. https://pypi.python.org/pypi/ExifRead

6

Evaluation

metrics, define (CPU, memoury, lantecy.), benchmarks (mirko, kernel... How to measure, where done, PSEUDOCODE

Time Finding folders and metadata takes: 1:43:13.488799, Reading excel file takes: 0:00:17.413845, Comparing takes: 4:43:30.705587, Overall time is 6:27:01.608355. Med alle bilder m/metadata og hele fotoboks2011_nordkynn_nordkynn.2011.xlsx.

New time Finding folders and metadata takes: 1:46:17.406581 Reading excel file takes: 0:01:07.686779 Comparing takes: 19:03:11.177869 Overall time is 20:50:36.271415 Med alle bilder m/mETADATA og hele nordkynn og varanger

"Concurrent" python find_folder_ok_concurrent.py Finding folders and metadata takes: 1:30:50.250732 Reading excel file takes: 0:00:18.127597 Comparing takes: 4:18:33.368199 Overall time is 5:49:41.746759 – Pool(processes=4) FRA fotoboks2011_nordkynn_nordkynn.2011.xlsx

"Concurrent" Finding folders and metadata takes: 1:44:18.788013 Reading excel file takes: 0:00:17.269817 Comparing takes: 3:54:23.483186 Overall time is 5:38:59.541248 – Pool(processes=100) FRA fotoboks2011_nordkynn_nordkynn.2011.xlsx

This chapter describes the experimental setup and metrics used to evaluate the implemented system.

6.1 Experimental Setup

All experiements was done on a Lenovo ThinkCenter with an Intel® Core™ i5-6400T CPU @ 2.20GHz × 4, Intel® HD Graphics 530 (Skylake GT2), 15,6 GiB memory and 503 GB disk. It ran on Ubuntu 17.04 64-bit.

6.2 Something!?

6.3 Results

What does the result say? Each experiment, result, meaning

/7 Discussion

Idea, architecture, design, results, other solutions, "arch have scale-problem?"

7.1 Architecture

7.1.1 Other solutions

- Virtual sensors probably uavhengige prosesser, ikke threads ettersom man evt vil addere flere sensorer og unnga a starte alle sensorer på nytt igjen..

- Updates of "database" in background, not on demand -> future work

7.2 Design

7.2.1 Other solutions

- hash-map instead of multiple lists with different animals. hash on animal and/or place

- parallel/concurrent program when doing metadata-collecting and read-

ing from excel-sheet and writing to csv

-

8

Contributions

Combine with conclusion??

9

Conclusion

9.1 Contributions?

9.2 Future Work

10

Future Work?

11

Appendix?

readme, source code, dataset measurement RAW

Bibliography

[1] Robert Sedgewick *Algorithms in C - parts 1-4*. Addison-Wesley Publishing Company, 3. Edition, 1998.