Markus Stocker

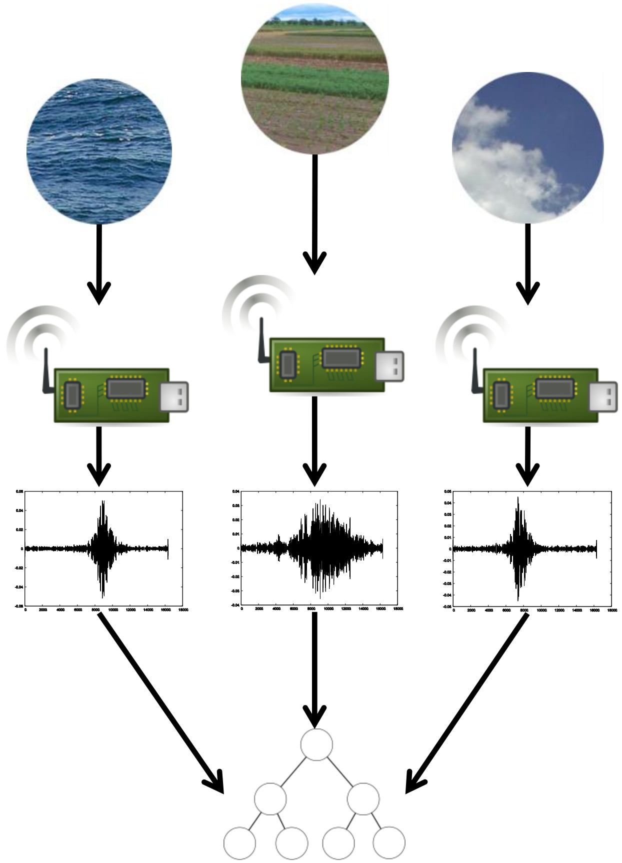
Nanjing, People's Republic of China, November 2011

Environmental monitoring

Discussed for road vehicle classification and lake monitoring

UNIVERSITY OF
EASTERN FINLAND

Introduction



- Environmental monitoring
  - “Observe and record the conditions of the natural environment” [1]
  - Involves *measurement*, i.e. the “process of empirical, objective, assignment of numbers to the properties of objects and events of the real world in such a way as to describe them” [2]
- Sensor layer
  - Distributed sensor network
- Data layer
  - Raw measurement spatiotemporal data
  - Potentially high frequency

- Ontology layer
  - Formal and explicit domain knowledge

Introduction
Overview

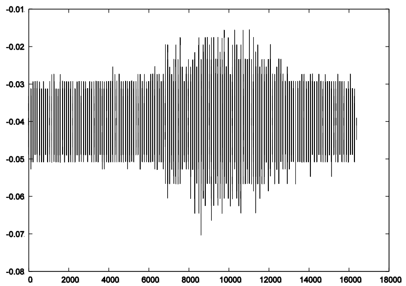
- Discuss
  - Environmental monitoring
  - Sensor layer
  - Data layer
  - Ontology layer (briefly)
- For each topic, discuss
  - Things to consider
  - Applied to road vehicles (concrete use case)
  - Applied to lake monitoring (potential use case)

Environmental monitoring
Things to consider

- Purpose of monitoring
  - Must be clearly defined
  - Determines sensor network and monitored variable(s)
- Monitored variable(s)
  - Must be clearly defined
  - Determines sensor type(s)
- Sensor type
  - Determines data type, data volume, data communication, ...
- Data type
  - Determines data parsing and processing
- Data volume
  - May be huge, especially for high frequency sensors and/or large sensor networks
- Data communication
  - TCP/IP, HTTP, proprietary protocols

Environmental monitoring
Applied to road vehicles

- Purpose of monitoring
  - Road vehicle detection and classification
- Monitored variable
  - Road pavement vibration
- Sensor types
  - Vibration sensor (accelerometer)
  - Camera sensor
- Data types
  - WAVE-encoded vibration sampling data
  - JPEG camera images
- Data volume
  - 2000 Hz sampling frequency (vibration)
  - Variable 1-3 fps at 640x480px (~300-400 Kb)
  - One month continuous monitoring: 5bn values, 1.7Tb data
- Data communication
  - TCP/IP, HTTP, wired/wireless links



Environmental monitoring
Applied to lake monitoring

- Purpose of monitoring, e.g.
  - Water quality for drinking purposes and/or bathing
  - Ecosystem health monitoring (e.g. Eutrophication)
- Monitored variables, e.g. for drinking water quality
  - Hardness, chlorine, nitrates, metals, radioactive materials, pH, ...
- Sensor type, e.g. for hardness
  - Calcium carbonate concentration (CaCO3 sensor)
- Data type
  - May write measurement data to text file (data logger)
- Data volume
  - May be little for a few low-frequency sensor, or huge
- Data communication
  - Possibly locally-wireless, remotely-wired (Internet), networks

Sensor layer
Things to consider

- Sensor network topology
  - Single sensor or sensor network
  - Communication between sensors or not
  - Data aggregation at sensor level or not
- Sensor installation
  - Proper deployment of sensor in environment
- Sensor calibration
  - Proper parameterization of sensor in environment
- Sensor durability
  - Powered by battery or connected to power network
- Sensor resistance
  - To weather conditions (e.g. temperatures and precipitation)
  - To interference of noise on signal

Sensor layer
Applied to road vehicles

- Sensor network topology
  - One vibration sensor (total three) and one camera sensor (total two)
  - Sensor data directly reported to database for persistent storage
  - Database aggregates data
- Sensor installation
  - Sensor is mounted on metal bar of about 70 cm
  - Metal bar horizontal in ground underneath road asphalt
- Sensor calibration
  - Sampling frequency and duration
  - Sensor sensitivity
- Sensor durability
  - Sensors are connected to power network
- Sensor resistance
  - Vibration sensors are resistant to weather conditions and sound

Sensor layer
Applied to lake monitoring

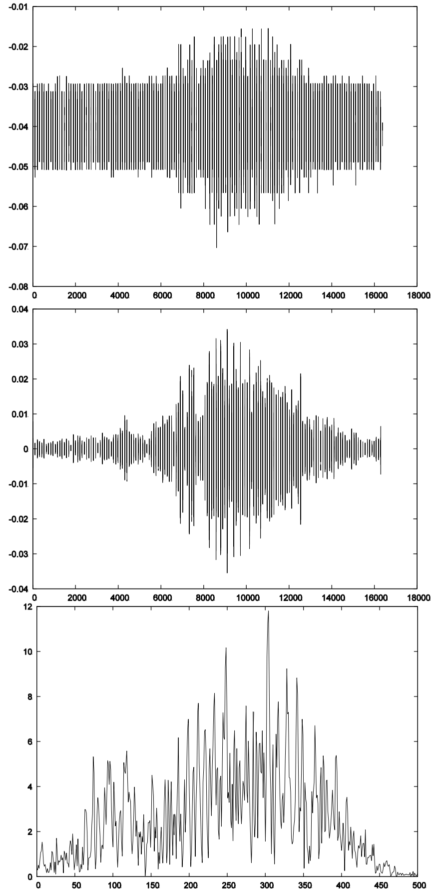
- Sensor network topology
  - Depending on monitoring purpose possibly a large network
  - Sensors may not communicate between each other
  - Data aggregation may be necessary
- Sensor installation
  - Very dependent on sensor
  - Location may be difficult to access
- Sensor calibration
  - To match different environment, e.g. solar radiation for turbidity
- Sensor durability
  - Possibly battery powered (if distributed and difficult to access)
  - Make sure battery lasts for months (affects data transfer)
- Sensor resistance
  - To water, mechanical interference

Data layer
Things to consider

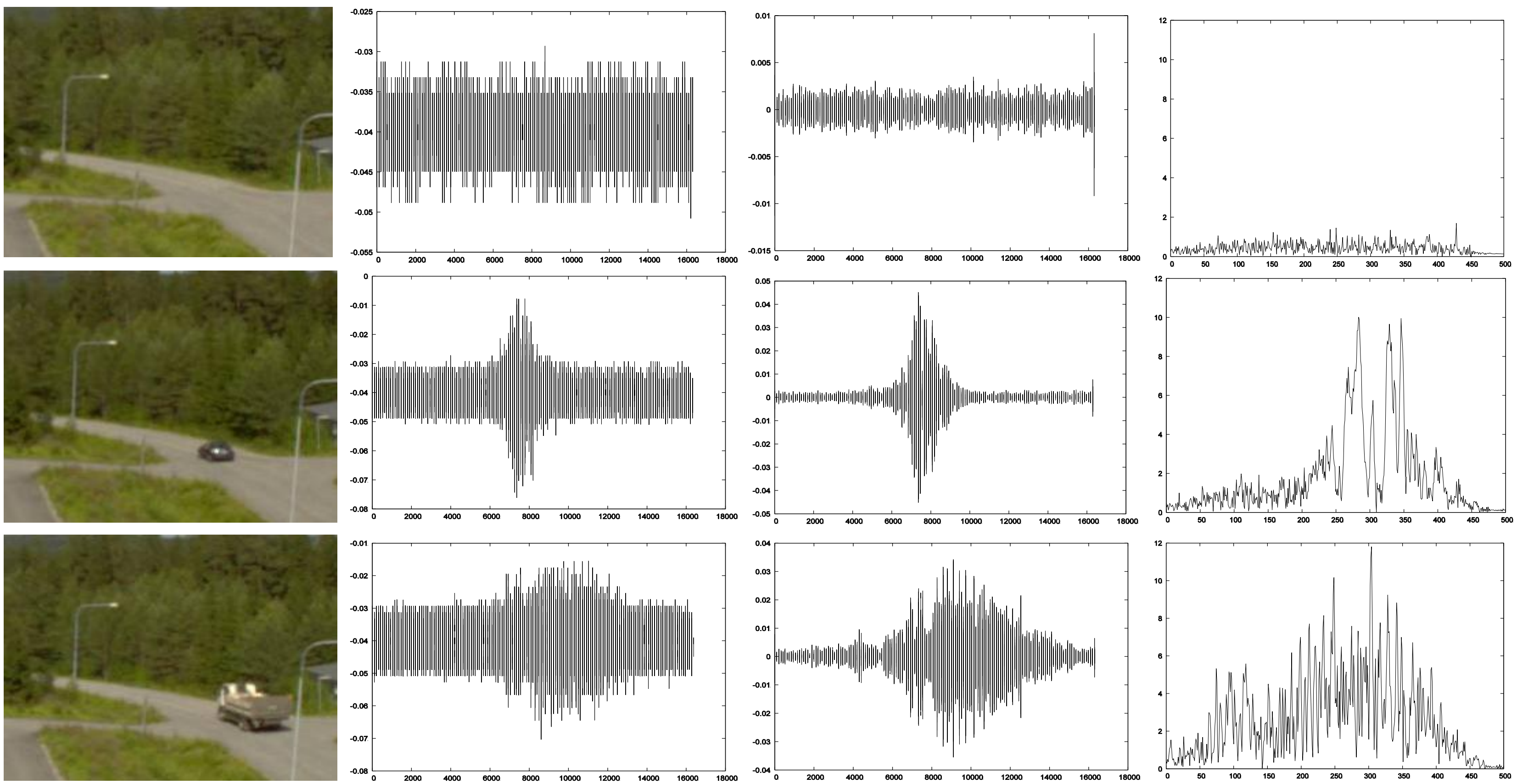
- Data access
  - How to read data from sensor
- Data parsing
  - Sensor data may be encoded (possibly binary format)
  - Parsing to extract measurement values we actually want
- Data processing
  - Aggregation, e.g. hourly average
  - Fusion, i.e. combining data from two or more sensors
  - Data quality, e.g. missing data
- Data mining
  - Discover patterns in sensor (network) data

Data layer
Applied to road vehicles (1)

- Data access
  - URL for both vibration and camera sensors
- Data parsing
  - WAVE data to extract measurement values
  - Camera images are provided as JPEG
- Data processing
  - For every window with 16,384 values (i.e. ~8s)
  - Band-pass filter, 80-130 Hz
  - Fourier transform
  - Extract frequency profile, 80-130 Hz



Data layer
Applied to road vehicles (2)



Data layer
Applied to road vehicles (3)

- Data mining
  - It is possible to classify frequency profiles and if so to what accuracy?
  - Supervised learning using Multi-Layer Perceptron (MLP)
  - Training dataset with 13 vehicle types + background
  - Total of 1911 labeled examples
    - Frequency profiles corresponding to 8s windows
- Detection task
  - Discriminate between background and not-background examples
  - About 95% classification performance for this binary case
- Classification task
  - Vehicles types grouped into four broader classes
  - About 70% classification performance

Data layer
Applied to lake monitoring

- Data access, parsing, processing
  - Highly dependent on sensor network
  - For battery-powered sensors
    - Much of data processing may be done by sensors directly
    - Power is a constraint on data transmission
- Data mining
  - Clustering, e.g.
    - Identify prototypical values (ranges) for lakes rich/poor in a nutrient
  - Classification, e.g.
    - Trophic state of lake (oligotrophic, mesotrophic, eutrophic)
    - Suitable for drinking/bathing purposes
  - Regression, e.g.
    - Early warning for Eutrophication
    - Determine potential future trends

Ontology layer
What is an ontology?

- Defined as "an explicit specification of a conceptualization" [3]
- A means to formally represent knowledge of a domain
  - Concepts of some area of interest
  - Relations that hold among concepts
- Example
  - Domain: Limnology, the "study of fresh or saline waters contained within continental boundaries" [4]
  - Concepts: Lake, Oligotrophic, Mesotrophic, Eutrophic, Nutrient
  - Relations: hasSurfaceArea, hasVolume, hasTotalNitrogen
  - Individual: Lake(*lakeSuperior*), Oligotrophic(*lakeSuperior*), hasVolume(*lakeSuperior*, "12,000 km^3")

Ontology layer
Applied to road vehicles

Schema (TBox)

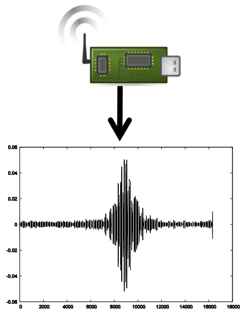
VehicleObservation $\sqsubseteq$ Observation $\sqcap$ $\forall$hasObservedObject.Vehicle

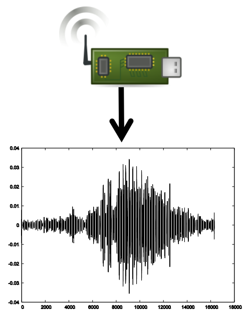
(VehicleObservation(s) are Observation(s) that observe Vehicle objects)

PersonalCar $\sqsubseteq$ Vehicle (subclass: PersonalCar is a subclass of Vehicle)

Truck $\sqsubseteq$ Vehicle

PersonalCar $\sqcap$ Truck $\sqsubseteq$ $\bot$ (disjointness: PersonalCar and Truck are disjoint classes)





Data (ABox)

Observation(o1)	Observation(o2)
PersonalCar(v1)	Truck(v2)
hasObservedObject(o1, v1)	hasObservedObject(o2, v2)

Ontology layer
Applied to lake monitoring

- Discussed in a second presentation

Ontology layer
Some implications

- Learning ontology
  - About individual observations (e.g. timestamp, vehicle type)
  - About ontology schema (e.g. vehicle classes)
- Abstraction
  - Work at knowledge level rather than data level
  - Keep knowledge and discard data (may be necessary)
- Consistency checking
  - Is my ontology consistent, schema, data w.r.t. schema?
  - Ontology reasoning service
- Query
  - Perform domain-oriented, rather than data-oriented, query

Conclusions

- Environmental monitoring
  - Purpose of monitoring — What conditions of natural environment do we want to monitor?
  - Monitored variables — What variables are affected by change in those conditions?
  - Sensor type — What technology exists to measure change in those variables?
  - Data type — What type of data does the technology deliver?
  - Data volume — How must data does the technology deliver?
  - Data communication — How can we communicate with the technology?
- Sensor layer
  - Sensor network topology — What network configuration do we need?
  - Sensor installation — How must the sensor be installed in natural environment?
  - Sensor calibration — How must the sensor be adapted to local conditions?
  - Sensor durability — For how much time is the sensor powered?
  - Sensor resistance — Do environmental factors affect the sensor?
- Data layer
  - Data access — How can data from the sensor be requested?
  - Data parsing — Is parsing to extract measurement values we need necessary?
  - Data processing — Do we need to further process data?
  - Data mining — What knowledge can we extract from data?

References

[1] Ioannis N. Athanasiadis and Pericles A. Mitkas. A methodology for developing environmental information systmes with software agents. In Advanced Agent-Based Environmental Management Systems. (Cortés, Ulises and Poch, Manel, Eds.), Birkhauser, 2009, pp. 119-138

[2] Sydenham, P.H.:: Handbook of Measurement Science: Volume 1 Theoretical Fundamentals. John Wiley & Sons (1982)

[3] Thomas R. Gruber. A translation approach to portable ontology specifications. Knowledge Acquisition 5(2), 199-220 (1993)

[4] Alexander J. Horne and Charles R. Goldman. Limnology. Second Edition, McGraw-Hill, Inc. (1994)