EUSIPCO2020
28th European Signal Processing Conference
January 18-22, 2021
Amsterdam

ULISSES AIR SENSORS FOR EVERYONE, EVERYWHERE

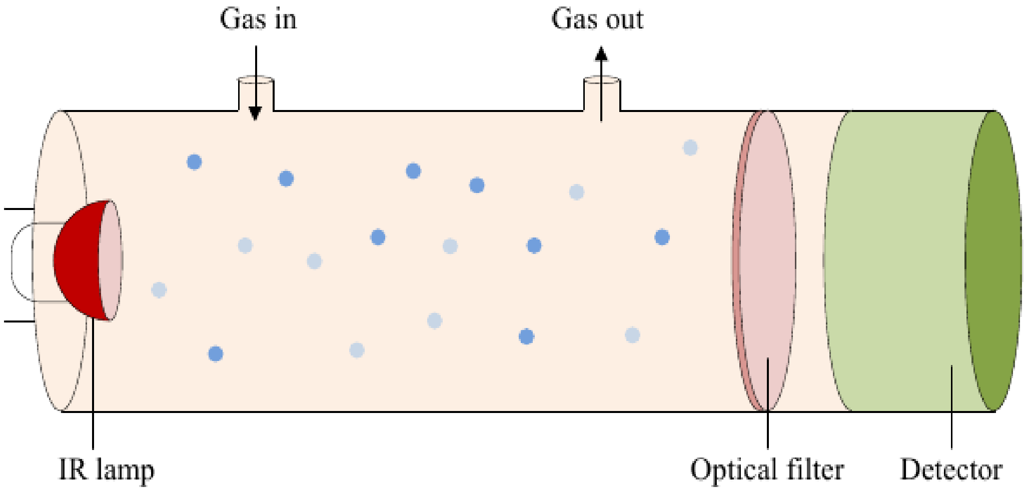
KTH ROYAL INSTITUTE OF TECHNOLOGY

Hidden Markov Model Based Data-driven Calibration of Non-dispersive Infrared Gas Sensor

Yang You and Tobias Oechtering
Division of Information Science and Engineering, EECS
KTH, Royal Institute of Technology, Sweden

NDIR CO_2 Sensor

- Sensor Mechanism:
  - Different gas molecules → Absorption of light with different wavelengths (Beer-Lambert law [1])
  - Attenuation of light intensity → Absorption of the measured CO_2 → CO_2 concentration
- Measurement process:
  - $Abs = f(zero, IR, T)$
  - $CO_2 = L(R - Abs)$
  - Abs: Absorption
  - ***IR***: Received IR light
  - ***Zero***: Calibration coefficient
  - ***R***: Reference level



[1]. D. F. Swinehart, "The beer-lambert law," Journal of Chemical Education, vol. 39, no. 7, pp. 333, 1962.

Drift Analysis of NDIR Sensor

- Zero coefficient:
  - Calibration parameter used for adjusting the sensor baseline offset
- ***Why is there a drift?***
  - IR signal varies with time, temperature and/or other factors
  - Measurement error if same zero coefficient is being used for calculating absorption.
- Sensor calibration requires adjusting the zero coefficient
  - Build a stochastic model of the 'true' zero coefficient
  - Estimate the true zero coefficient at any time given noisy observation

State Of The Art - Technology

- The traditional Automatic Baseline Correction (ABC) algorithm [1]
  - The sensor is calibrated to a fixed value which is assumed to be the fresh air CO_2 concentration
  - Fails when the sensors never get exposed to fresh air in a calibration period
    - E.g., in mega-cities

- Need to design more robust and smart self-calibration algorithms

[2]. "TN-011," Tech. Rep., SenseAir, 2000.

State Of The Art - Research

- Data-dirven modeling aims to find relationships between the system state variables (input and output) without explicit knowledge of the physical behavior of the system [3].
  - Powerful tool for smart sensor self-calibration [4-7].

- **[3]. D. Solomatine, L. M. See, and R. J. Abrahart, “Data-driven modelling: concepts, approaches and experiences,” in Practical hydroinformatics, pp. 17–30. Springer, 2009**
- **[4]. N. Roy and S. Thrun, “Online self-calibration for mobile robots,”in Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C), May 1999, pp. 2292–2297 vol.3.**
- **[5]. A. A. Boechat, U. F. Moreno, and D. Haramura Jr, “On-line calibration monitoring system based on data-driven model for oil well sensors,” IFAC Proceedings Volumes, vol. 45, no. 8, pp. 269–274, 2012.**
- **[6]. T. Wissel, B. Wagner, P. Stber, A. Schweikard, and F. Ernst, “Datadriven learning for calibrating galvanometric laser scanners,” IEEE Sensors Journal, vol. 15, no. 10, pp. 5709–5717, Oct 2015.**
- **[7]. D. Wang, J. Liu, and R. Srinivasan, “Data-driven soft sensor approach for quality prediction in a refining process,” IEEE Transactions on Industrial Informatics, vol. 6, no. 1, pp. 11–17, Feb 2010.**

State Of The Art - Research

- Machine learning approaches provide increasing levels of automation and improved accuracy by discovering and exploiting regularities in the training data.
  - Research on data-driven sensor self-calibration have combine machine learning approaches [8],[9].

- In this paper, we follow the idea of combing the concept of data-driven modeling and approaches of machine learning.

- **[8]. L. Wijeratne, D. Kiv, A. Aker, S. Talebi, and D. Lary, "Using machine learning for the calibration of airborne particulate sensors," Sensors, vol. 20, no. 1, pp. 99, 2020.**
- **[9]. W. Xie and P. Bai, "A pressure sensor calibration model based on support vector machine," in 2012 24th Chinese Control and Decision Conference (CCDC). IEEE, 2012, pp. 3239–3242.**

Modeling the Drift

- Utilize the hidden Markov model (HMM) to build a data-driven self-calibration system for NDIR sensors.

  - Build a joint probalistic model on time series of CO2 measurements, temperature and ‘true’ zero coefficients

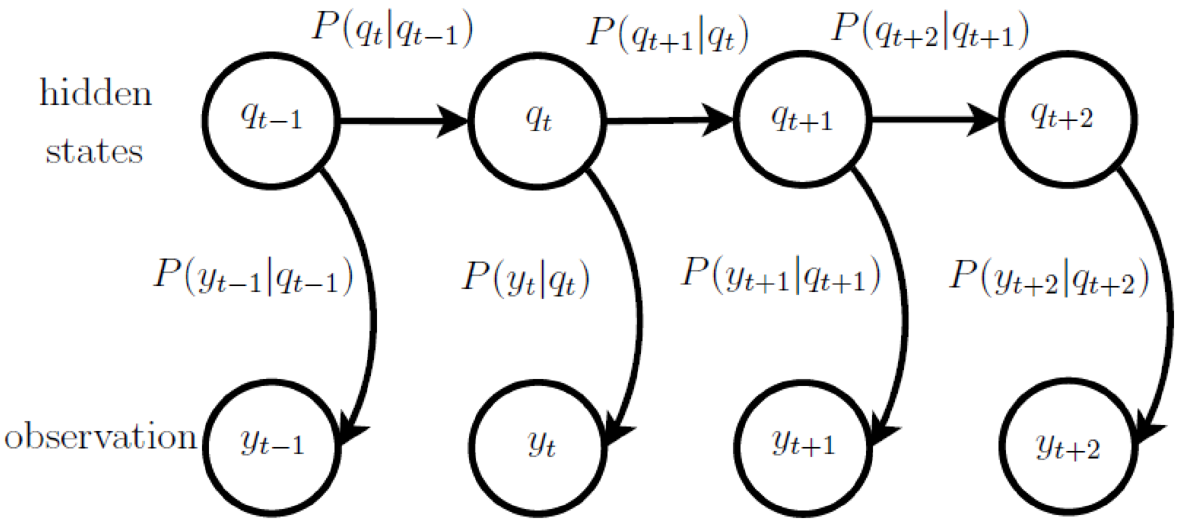
  - Learn the transition behavior of ‘true’ zero coefficient

  - Infer the ‘true’ zero coefficient at different time steps

Hidden Markov Model

- Brief introduction on HMM:
  - Probalistic model composed of observation space and hidden state space (state space should be discrete)

Main Tasks of HMM:
Learning & Inference



HMM setting for ZERO Coefficient Inference

- Implementation:
  - Model:
    - ‘True’ zero coefficients are quantized to different integer levels, i.e., ‘true’ zero coefficients within the same range will be quantized to the same level.
    - The change of temperature $\Delta T_t = T_t - T_{t-1}$ is also quantized to certain different levels
    - We set the pair of quantized ‘true’ zero coefficients and temperature change as the hidden state of HMM.
    - Observation sequence: CO_2 measurements

HMM setting for ZERO Coefficient Inference

- Notations:
  - Hidden state at time t $X_t = (ZERO_t, \Delta T_t)$
  - Observation at time t $Y_t = CO_{2\,t}$
- Definitions (with two-dimensional hidden state be $(ZERO_t, \Delta T_{t+1})$):
  - Transition probability: $P(x_{t+1}|x_t)$
  - Emission probability: $P(y_t|x_t)$
  - Prior distribution of the hidden state: $P(x_1)$
  - Posterior distribution of the hidden state: $P(x_t|y_1^T)$
- Assumption:
  - The temperature dependency of the model is time independent, i.e., $P(x_{t+1}|x_t)$ and $P(y_t|x_t)$ are time-invariant

Unsupervised learning of HMM

- Unsupervised training:
  - Exact values of true zero coefficients are always hard to know (only with some prior knowledge).
- Baum-Welch algorithm [10]:
  - Find an approximation to the maximum likelihood estimate:

$$\hat{\lambda}^* = \arg\max_{\lambda} P(y_1, y_2, ..., y_L | \lambda) \qquad \frac{\log P(y^L | \hat{\lambda}_{i+1}) - \log P(y^L | \hat{\lambda}_i)}{\log P(y^L | \hat{\lambda}_i)} < \gamma$$

  - λ denotes the underlying paramters of the HMM, i.e., transition probability, emission probability, and prior distribution.
  - $\{\hat{\lambda}^{(1)}, \hat{\lambda}^{(2)}, \dots\}$ denote the sequential parameters estimates
  - Notice that we only apply the Baum-Welch algorithm to a subsequence y^L instead of the whole sequence y^T for computational simplicity reason

[10]. C. M. Bishop, Pattern recognition and machine learning, springer, 2006.

Inference of ZERO coefficient

- Viteribi Decoding:
  - Maximum a posterior estimation:

$$\hat{x}^T = \arg\max_{x^T} P(x^T | y^T, \hat{\lambda})$$

  - Notice $X_t = (ZERO_t, \Delta T_t)$
  - The temperature change ΔT can be observed
  - Plug sequence ΔT^T into the above function and only find the best estimate for the sequence $ZERO^T$

Numerical Experiments

- Dataset:
  - Research is performed using the data provided by Senseair
  - The data is acquired from 10 sensors which are put in a station at highway E18 in Sweden
  - Main content of the dataset:
    - Raw CO_2 measurement
    - Current temperature
    - Current time
  - The sampling interval is 15mins during most time:
    - However, some samples are missing.
      - Solution: Dataset is split into 3 subsets to make the data sequences included synchronized

Numerical Experiments

- Several initial conclusions:
  - The 'true' zero coefficients in the dataset is within the range [12434,12637]
  - The CO_2 measurement in the dataset is within the range [379ppm, 536ppm]

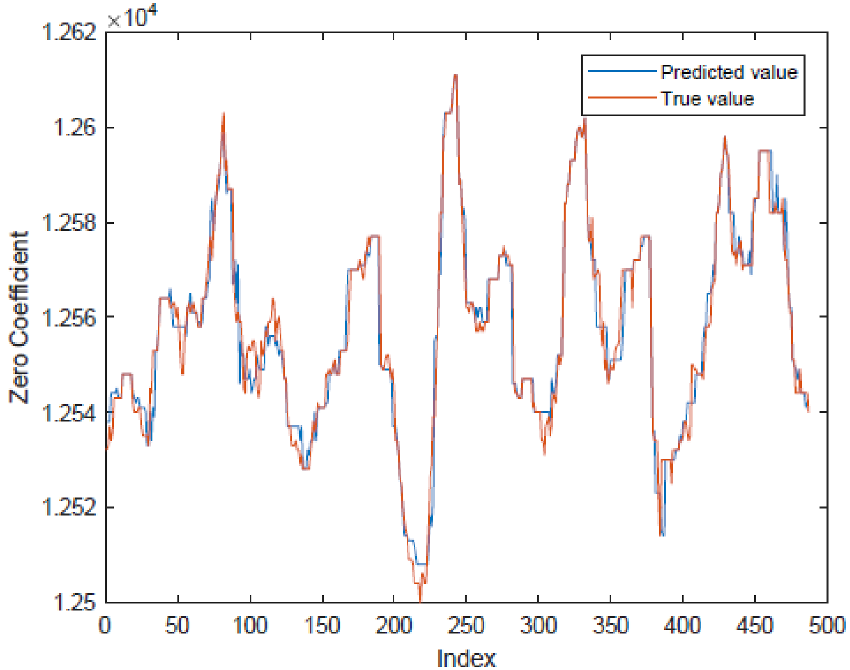
- We also provide supervised learning approach as the benchmark of our proposed unsupervised learning approach

Experiment Setting – Supervised Learning

- High-resolution quantization:
  - True zero coefficient is quantized to each integer in the range [12434,12637]
  - CO_2 measurement is also quantized to each integer in the range [379ppm, 536ppm]
  - Quantization of ΔT:

Interval	Level	Interval	Level
<-1	-2	[0.1,0.3]	0.2
[-1,-0.7]	-0.85	[0.3,0.5]	0.4
[-0.7,-0.5]	-0.6	[0.5,0.7]	0.6
[-0,5,-0.3]	-0.4	[0.7,1]	0.85
[-0.3,-0.1]	-0.2	>1	2
[-0.1,0.1]	0		

Experiment Results – Supervised learning



Results for supervised learning using the data from 2018. 04.14-05.30

Experiment Setting – Unsupervised Learning

- Low-resolution quantization to guarantee a fast convergence:
  - Quantization of true zero coefficient:

Interval	Level	Interval	Level
$[12434, 12485)$	$zero_1$	$[12555, 12565)$	$zero_8$
$[12485, 12495)$	$zero_2$	$[12565, 12575)$	$zero_9$
$[12495, 12505)$	$zero_3$	$[12575, 12637]$	$zero_{10}$
...	...		

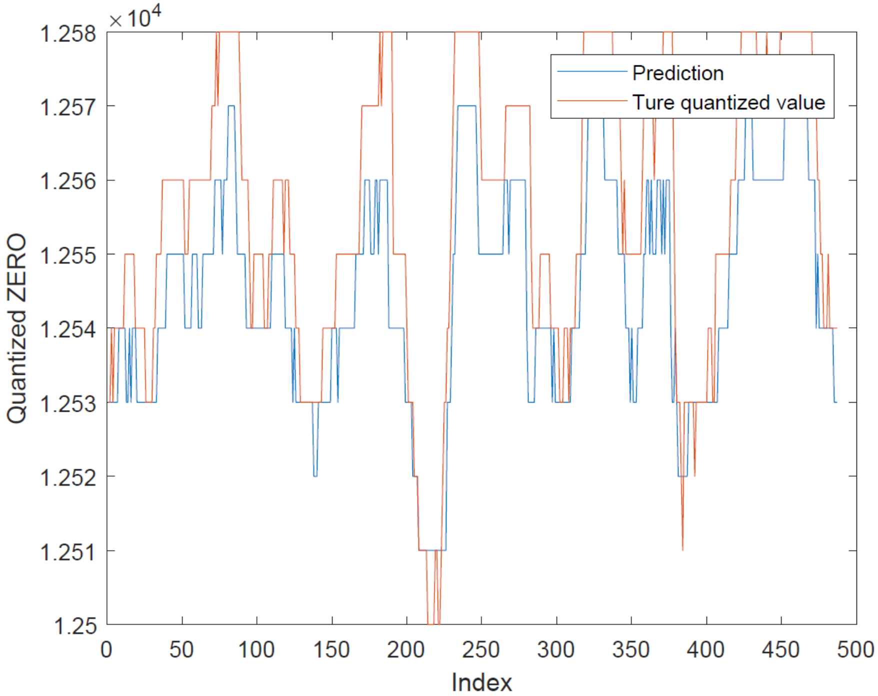
  - Quantization of CO_2 measurement:

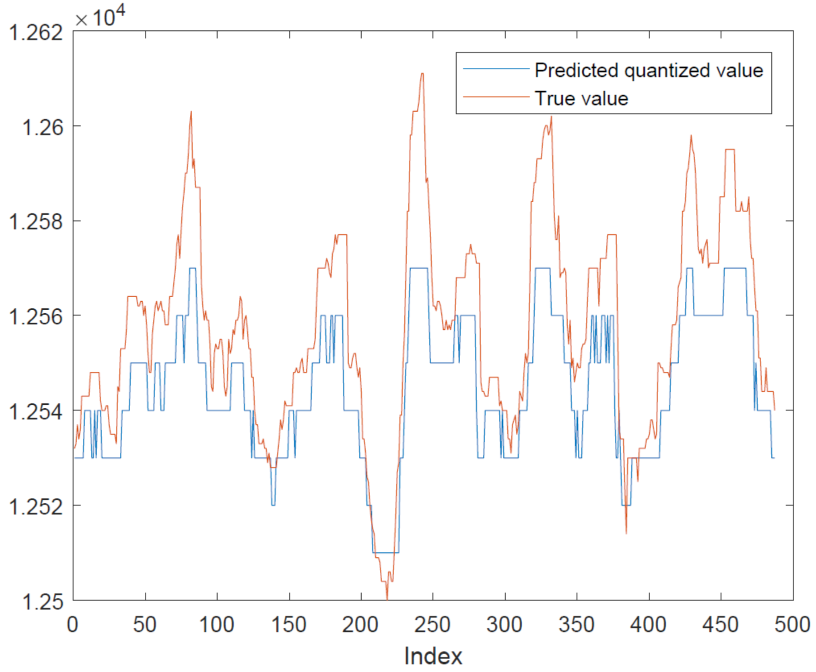
Interval	Level	Interval	Level
$[379, 410)$	Ω_1	$[470, 480)$	Ω_8
$[410, 420)$	Ω_2	$[480, 490)$	Ω_9
$[420, 430)$	Ω_3	$[490, 536]$	$\Omega_1 0$
...	...		

  - Quantization of ΔT:

Interval	Level	Interval	Level
$(-\infty, -0.2)$	$\Delta T(1)$	$(0.2, +\infty)$	$\Delta T(2)$
$[-0.2, 0.2]$	$\Delta T(3)$		

Experiment Results





Result for unsupervised learning using the data from 2018. 04.14-5.30 compared to quantized true value

Result for unsupervised learning using the data from 2018. 04.14-05.30

Conclusions

- The calibration parameter of the NDIR is temperature dependent

- The drift of the NDIR sensor measurement can be fully described by the drift of the calibration parameter, i.e., true zero coefficient

- The self-calibration problem is formulated into a statistical inference problem of the true zero coefficient, both supervised and unsupervised HMM-based approaches are developed and studied.