Georgia Tech ISyE
H. Milton Stewart School of Industrial and Systems Engineering

Machine Vision Enabled In-Process Quality Improvement in Smart Manufacturing

Jianjun Shi
The Carolyn J. Stewart Chair and Professor
Georgia Institute of Technology

contact: jianjun.shi@isye.gatech.edu
https://sites.gatech.edu/jianjun-shi/

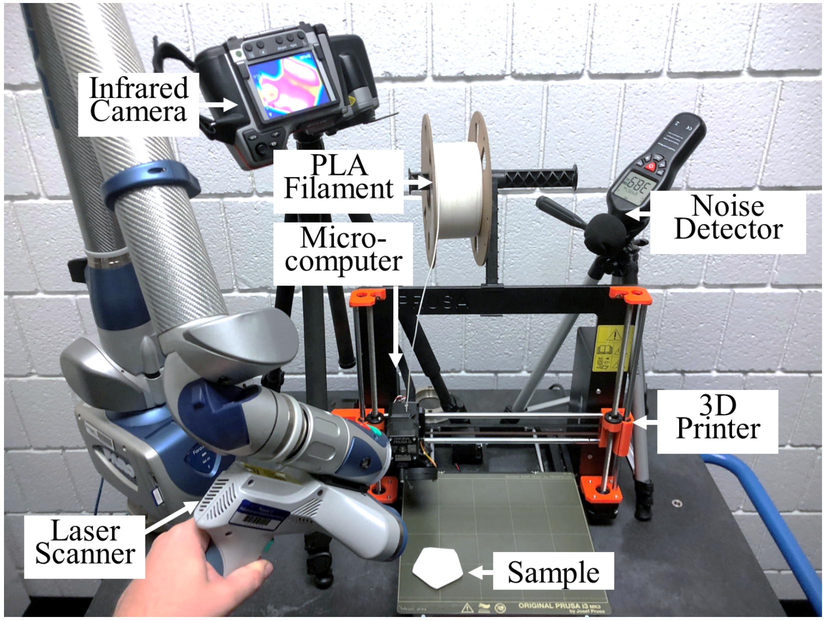
Introduction
Machine Vision Enabled In Process Quality Improvements

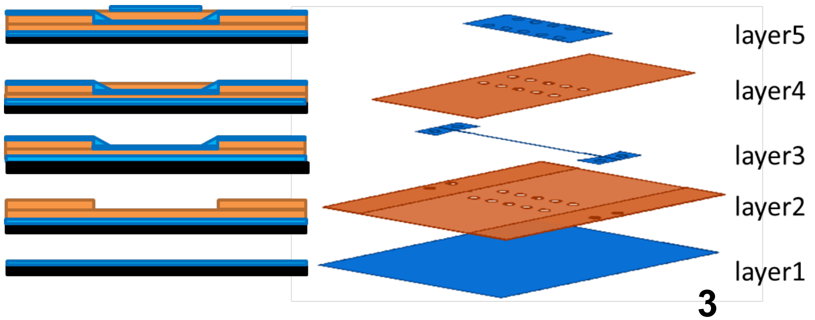
- Machine Vision (MV) has been widely used in smart manufacturing systems to monitor product quality and machine conditions
- Many R&D efforts have been done to improve the MV capabilities in terms of anomaly detection, defect detection, etc.
- Few efforts to integrate data from MV with other production data for quality and productivity improvements
- It is desirable to use MV as an essential tool to enable the "In-Process Quality Improvements (IPQI)" to achieve in-site process monitoring, root cause diagnosis, predictive control, and product defect prevention

Machine Vision Enabled IPQI:
Key Enabling Methodologies (examples)

3D Printing / Multilayer Additive Manufacturing

- Effective algorithms in anomaly detection or feature extraction using machine vision data
- Modeling of heterogeneous data (images, video signals, 3D point cloud data, functional curve data, text data, etc.) to represent relationships between production process variables and product quality variables
- Modeling, prediction, and control of 3D profile propagation based on machine vision signals in Multistage Manufacturing Process (MMP)





Outline

- **Introduction**
  - **Machine Vision Enabled In-Process Quality Improvements**

- **R&D Examples for Machine Vision Enabled IPQI**
  - **① Unsupervised Anomaly Detection with Machine Vision**
  - **② Multiple Tensor-on-Tensor Regression Model for MMP**
  - **③ Machine Vision-Based Automatic Control**
  - **④ In-situ Product Quality Prediction based on 3D Point Cloud Data**
  - **⑤ DETONATE: Nonlinear Dynamic Evolution modeling of Time-dependent 3-dimensional point cloud profiles**

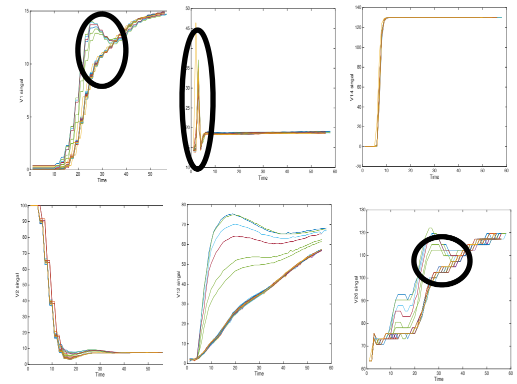
- **Summary**

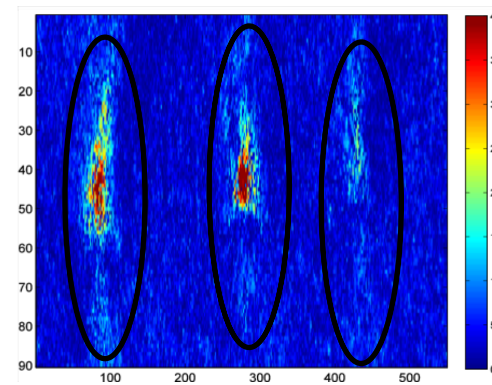
Sequential High-Dimensional Data Analysis
for Anomaly Detection and System Monitoring

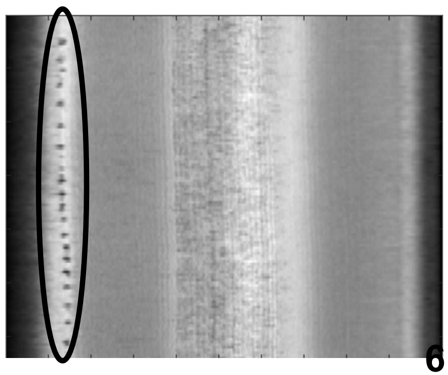
- Yan, H., Paynabar, K., Shi, J., 2017, "Anomaly Detection in Images with Smooth Background Via Smooth-Sparse Decomposition", *Technometrics,* Vol. 59, No. 1, pp102-114. https://doi.org/10.1080/00401706.2015.1102764
- Yan, H., Paynabar, K., Shi, J., 2017, "Online High-dimensional Monitoring and Diagnostics via Recursive Spatio-Temporal Smooth Sparse Decomposition", *Technometrics,* Vol. 60, No.2, pp181-197. https://doi.org/10.1080/00401706.2015.1102764

Common Characteristics and Challenges of High Dimensional (HD) Streaming Data

- **High-dimensionality**
- **High velocity**
- **Complex spatio-temporal structure**
- **Unknown anomaly occurrence, location, and shape**
- **Goal: Unsupervised feature extraction from HD data**
  - Labeling/quality measurement is expensive.
  - Process data is typically cheap.



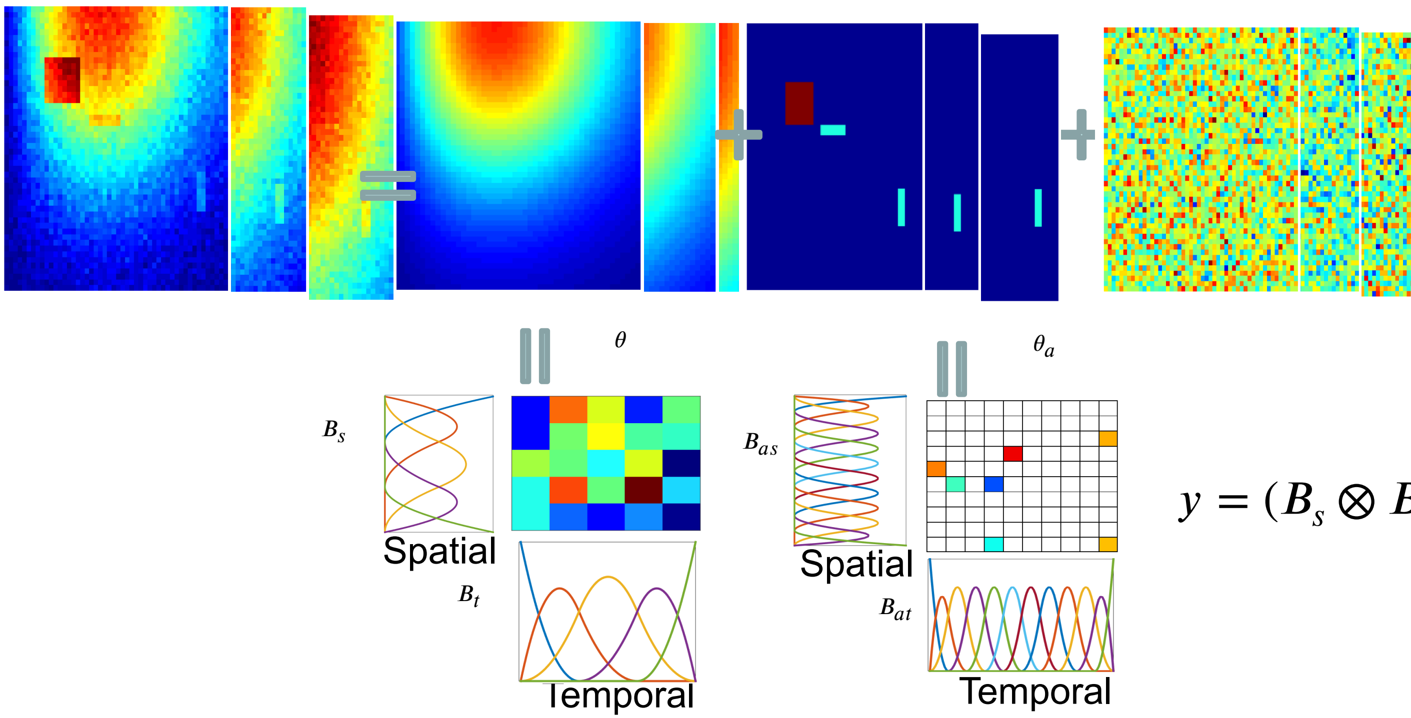




Anomaly Detection for High-dimensional Data

Profile Sequence = **Spatio-temporal Mean** + **Anomaly** + **Noise**

Smooth | Sparse | Small



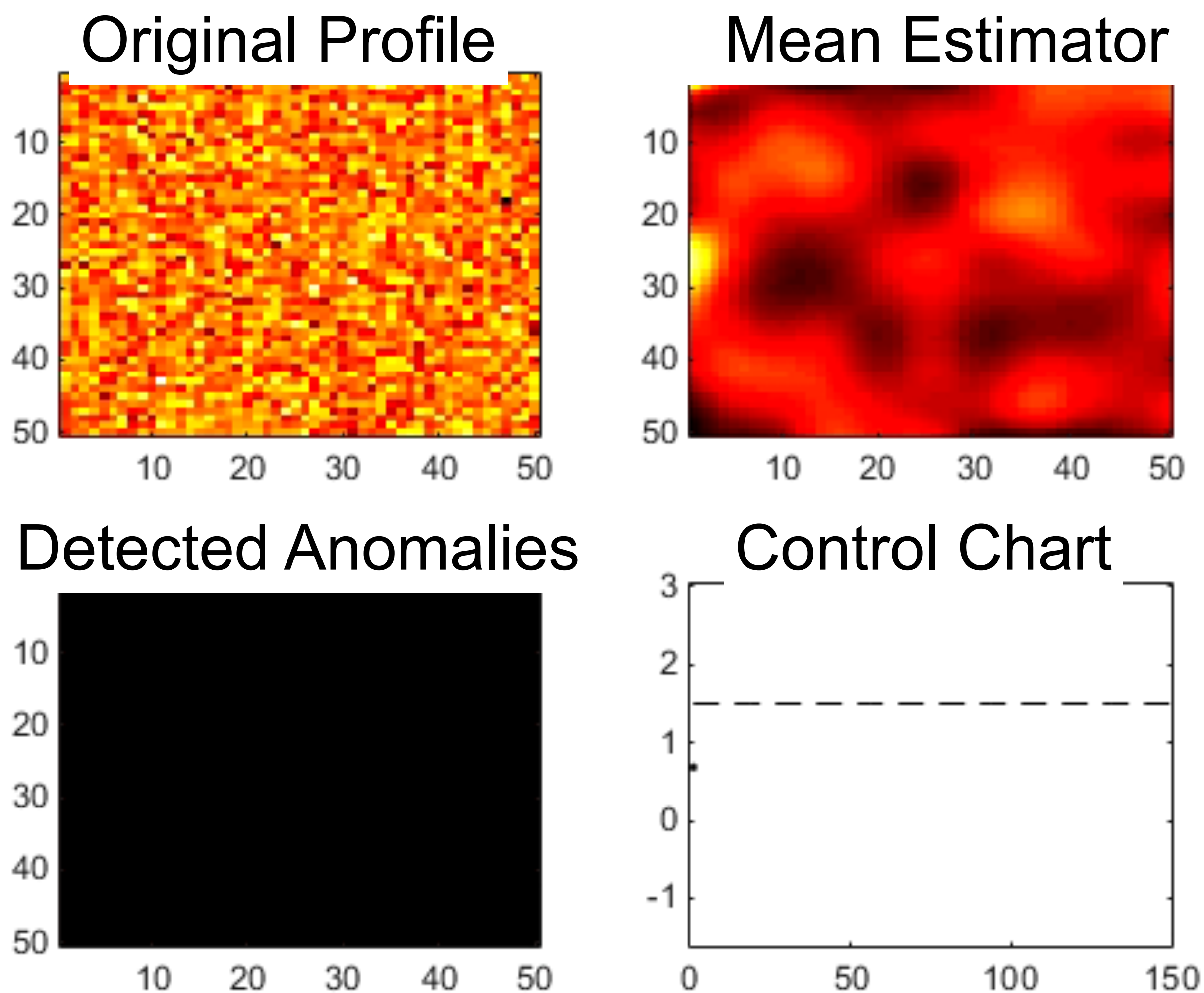
$$y = \mu + a + e$$

$$y = (B_s \otimes B_t)\theta + (B_{as} \otimes B_{at})\theta_a + e$$

$$\operatorname{argmin}_{\theta,\theta_a} \quad \lambda\theta' R\theta + \gamma\left|\theta_a\right|_1 + ||e||^2, \ s.t. \ \ y = (B_s \otimes B_t)\theta + (B_{as} \otimes B_{at})\theta_a + e$$

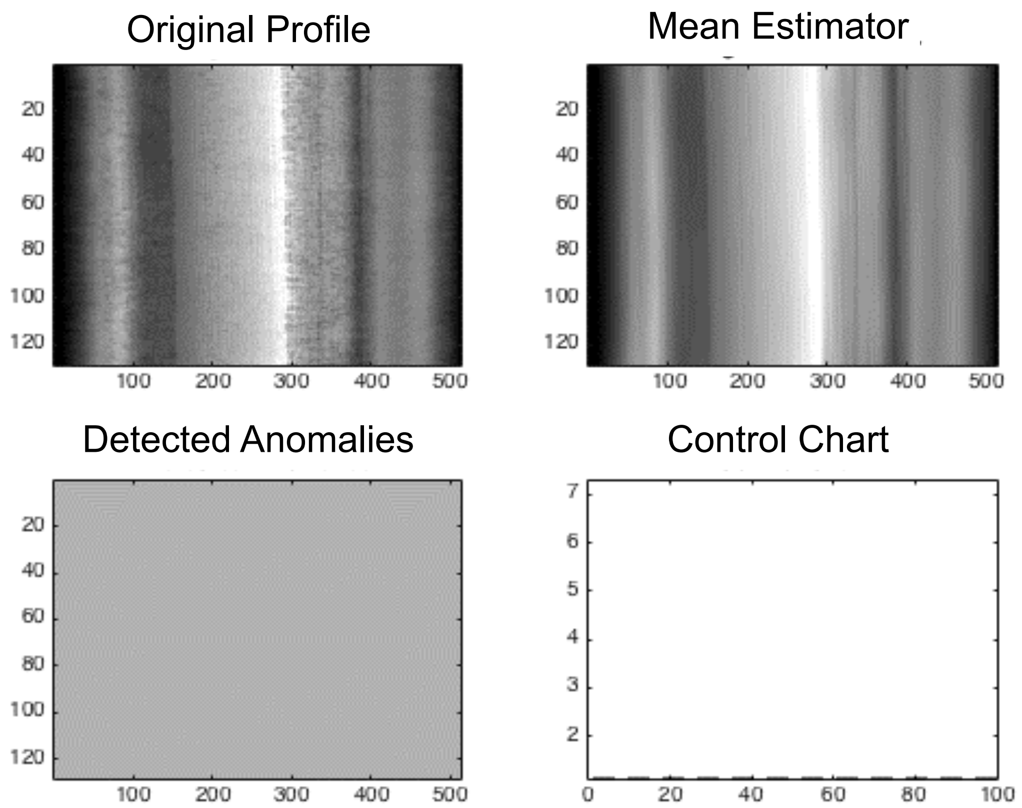
- Yan, H., Paynabar, K., Shi, J., 2018, "Online High-dimensional Monitoring and Diagnostics via Recursive Spatio-Temporal Smooth Sparse Decomposition", *Technometrics,* Vol. 60, No.2, pp181-197.

Simulation Results:
Anomaly Detection for Dynamic HD Streaming Data



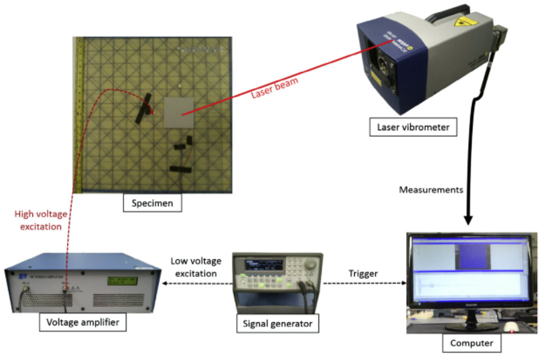
Case Study: Rolling Process Monitoring

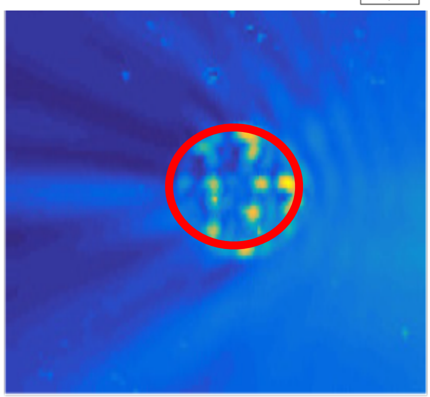
- **Data Size:** 128×512**, with 100 samples**
- **Background: smooth in** y **direction**
  - $B_x = I_x$,
  - B_y: **B-spline base with 5 knots**
- **Goal: Detect scattered surface anomaly:** $B_{ax} = I_{ax}, B_{ay} = I_{ay}$



Adaptive sampling for anomaly detection

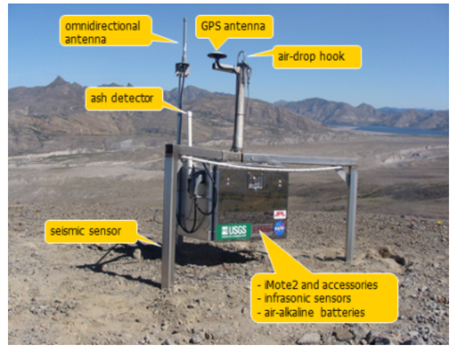
Point-based Inspection System

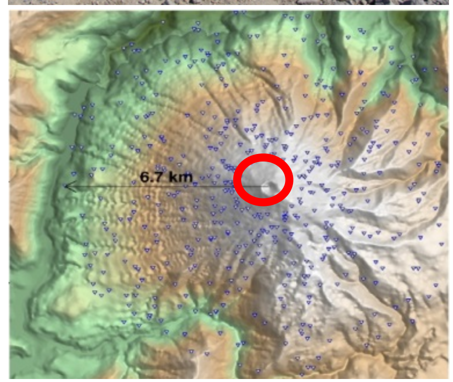




Low sampling frequency

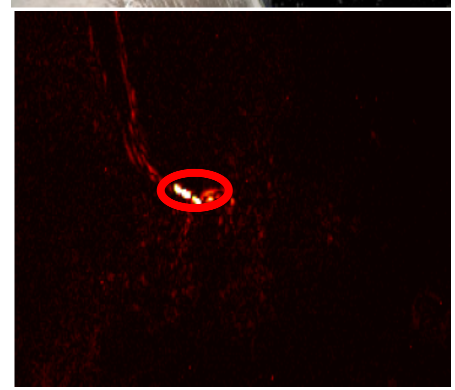
Volcano detection





Low battery or bandwidth

Solar flare detection

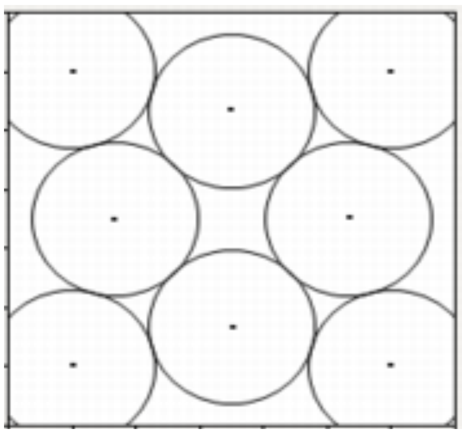


Low transmission and processing capability

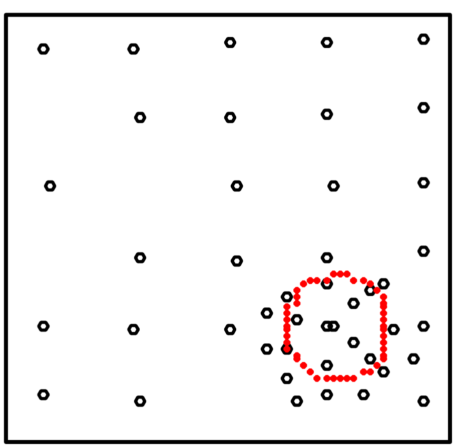
- Time-consuming for dense sampling
- Sparse anomaly: most points/sensors are irrelevant and only a subset is important.
- **Objective**: Adaptive sampling strategy to quickly locate and examine anomalies

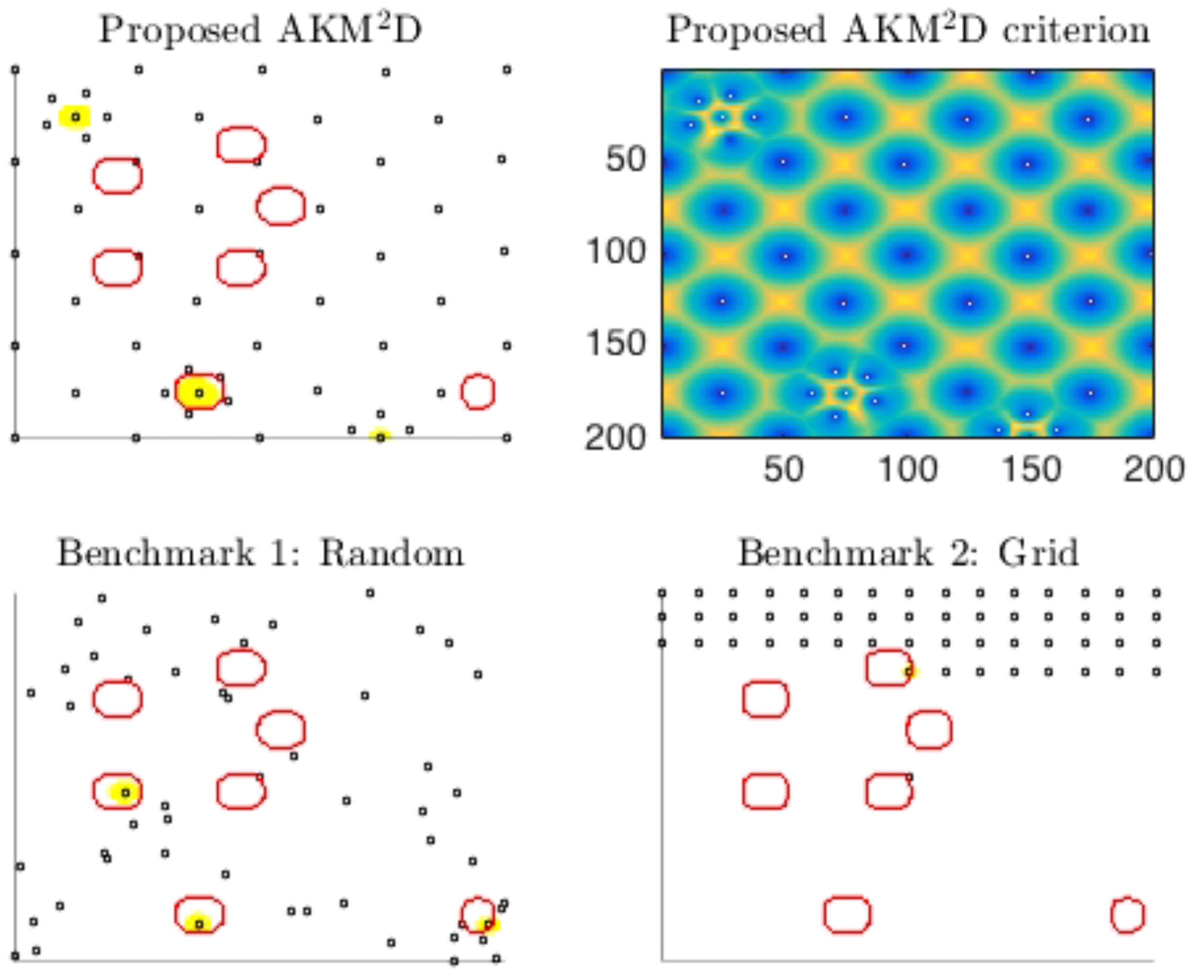
Anomaly Detection Result (250 Points)

Space Filling (Exploration)

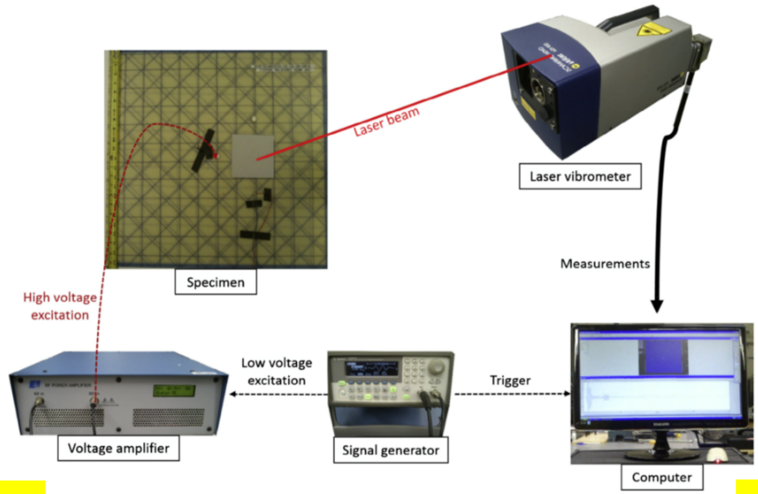


Focus Sampling (Exploitation)

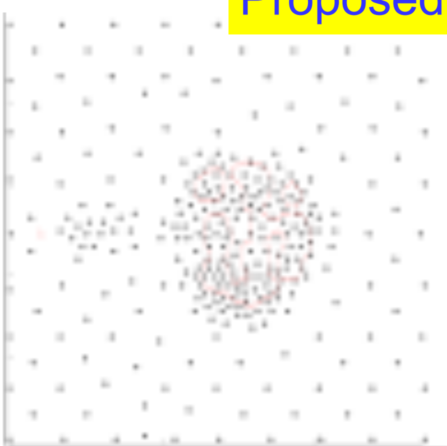


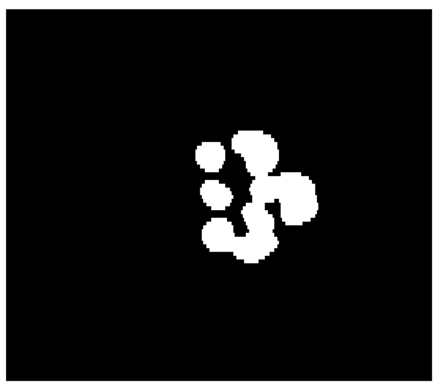


Real Case Study—Guided Wave Test

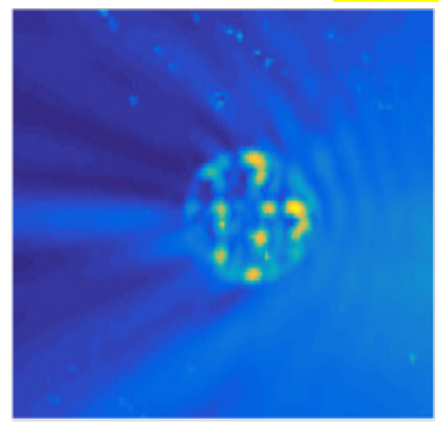


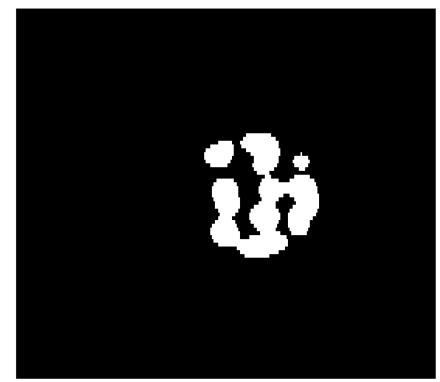
Proposed AKM²D





Full Sampling





	Proposed AKM²D		Full Sampling
# of points	300	# of points	27000
Measurement time	2mins	Measurement time	2 hours

- **Capable of detecting the anomaly with 1% of full sampled data**
- **Reduce measurement time from 2 hours to 2minutes for 1m² sample**

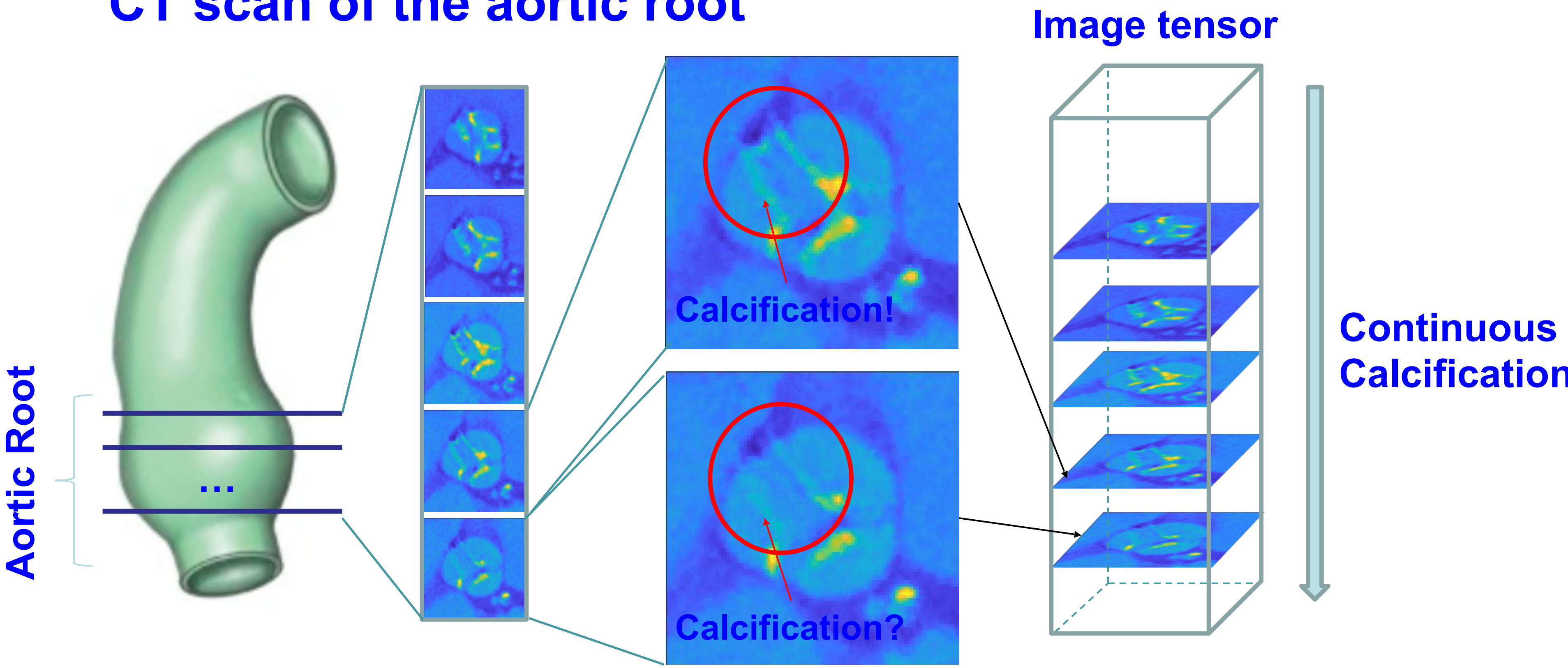
Additive Tensor Decomposition (ATD)

Robust learning for tensor data

Mou, S., Wang, A., Zhang, C., & Shi, J. (2021). Additive Tensor Decomposition Considering Structural Data Information. *IEEE Transactions on Automation Science and Engineering*, 19(4), 2904-2917.

Motivating: Unsupervised pixel-wise calcification region extraction in CT image

CT scan of the aortic root

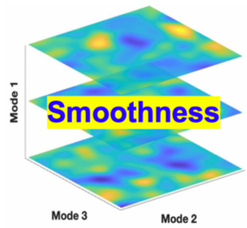


Calcification region sparse inside each slice but continuous across nearby slices

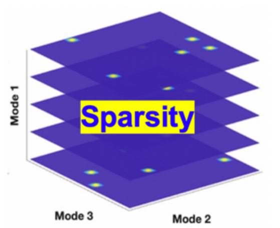
Additive Tensor Decomposition (ATD)

Robust tensor signal restoration by incorporating tensor structural priors

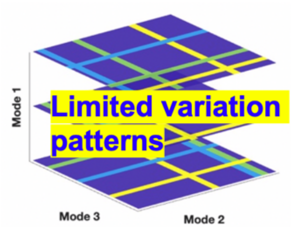
Tensor structural priors



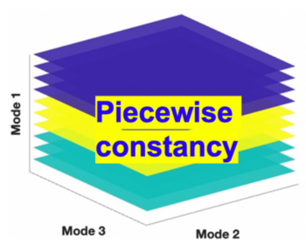
$$p(\mathcal{X}) = \sum_{s=1}^{k} \|\mathbf{D}_{d_s}^{l} \mathcal{X}_{(d_s)}\|_F^2$$



$$p(\mathcal{X}) = \sum_{i_1=1}^{I_1} \cdots \sum_{i_k=1}^{I_k} \|\mathrm{vec}(\mathcal{X}(1, \dots, k, :, \dots, :))\|_2$$

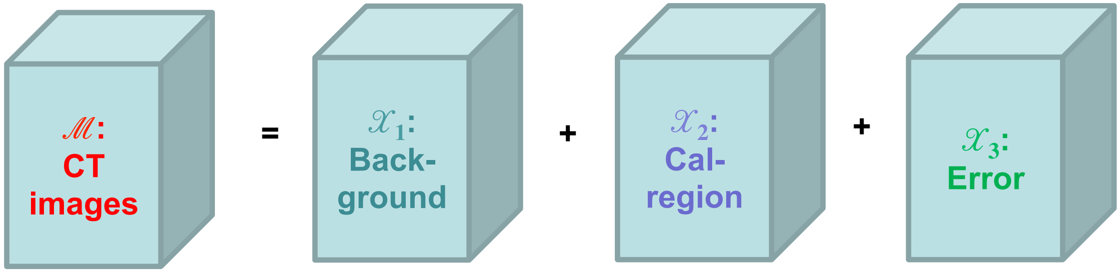


$$p(\mathcal{X}) = \sum_{i_q=1}^{I_q} \cdots \sum_{i_{q+k}=1}^{I_{q+k}} \left\| \left(\mathcal{X}(:, \dots, :, i_{q+1}, \dots, i_{q+k}, :, \dots, :)\right)_{(I_1 \dots I_q) \times (I_{q+k+1} \dots I_n)} \right\|_*$$



$$p(\mathcal{X}) = \sum_{i_{d_1}=1}^{I_{d_1}} \cdots \sum_{i_{d_k}=1}^{I_{d_k}} \sum_{i_s=1}^{I_S-1} \left\| \mathbf{vec}\left(\mathcal{X}(:, \dots, :, i_{d_1}, :, \dots, :, i_{d_k}, :, \dots, i_{d_s} + 1, \dots, :) - \mathcal{X}(:, \dots, :, i_{d_1}, :, \dots, :, i_{d_k}, :, \dots, i_{d_s}, \dots, :)\right) \right\|_2$$

Data: CT images of aortic root region



Background Smooth

$$\underset{\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3}{\text{minimize}} \lambda_{1,1} \left(\left\| \mathbf{D}_1 \mathcal{X}_{1(2)} \right\|_F^2 + \left\| \mathbf{D}_1 \mathcal{X}_{1(3)} \right\|_F^2 \right) + \lambda_{2,1} \left\| \mathbf{D}_1 \mathcal{X}_{2(1)} \right\|_F^2 + \lambda_{2,2} \left\| \mathrm{vec}\left(\mathcal{X}_2\right) \right\|_1 + \lambda_{3,1} \left\| \mathcal{X}_3 \right\|_2^2$$

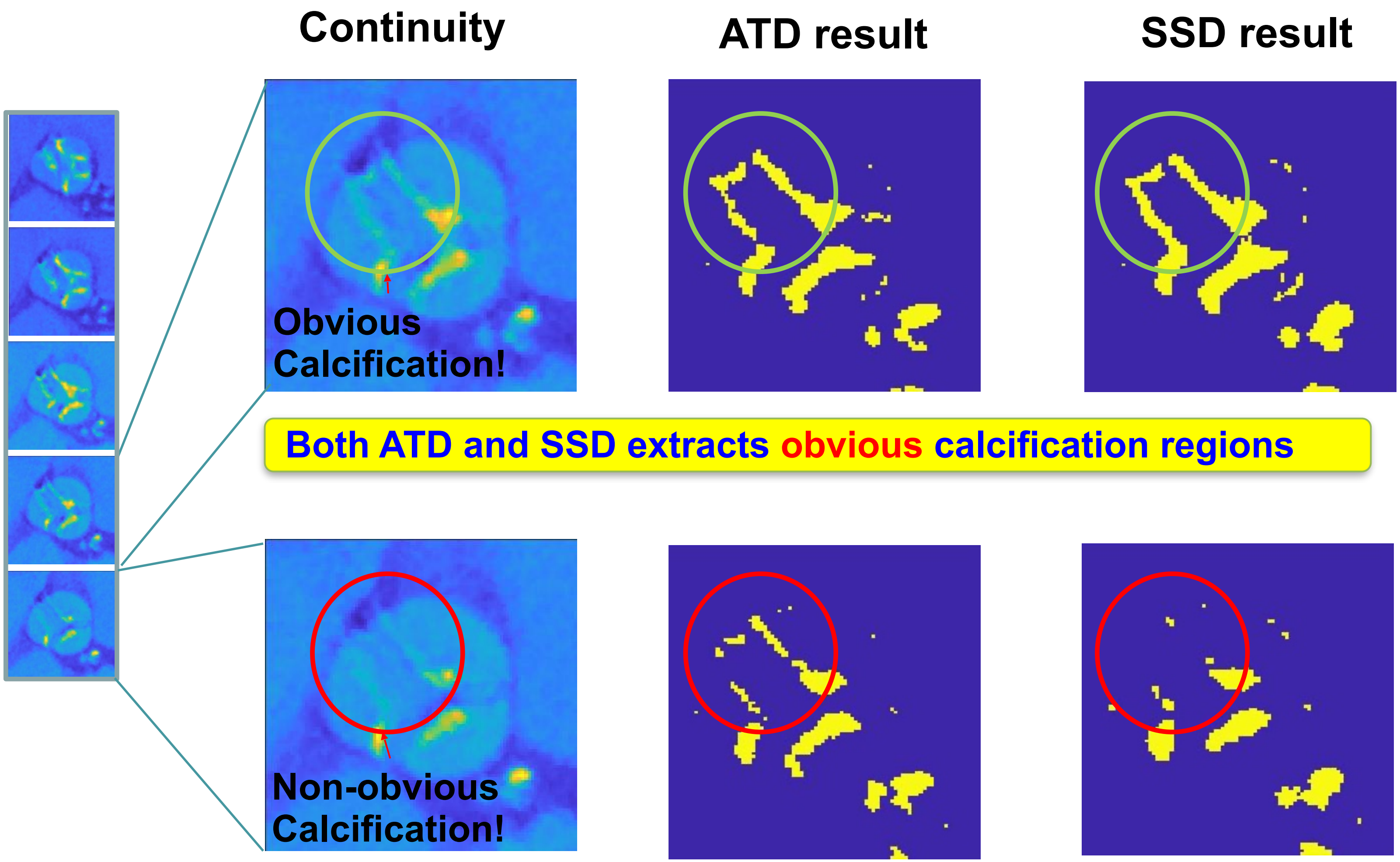
Calcium continuous across slices

Calcium sparse inside each slice

Error small

Generalize the robust learning methods to tensor data.

Results Calcification Region Extraction



Both ATD and SSD extracts obvious calcification regions

ATD outperforms SSD for non-obvious calcification regions

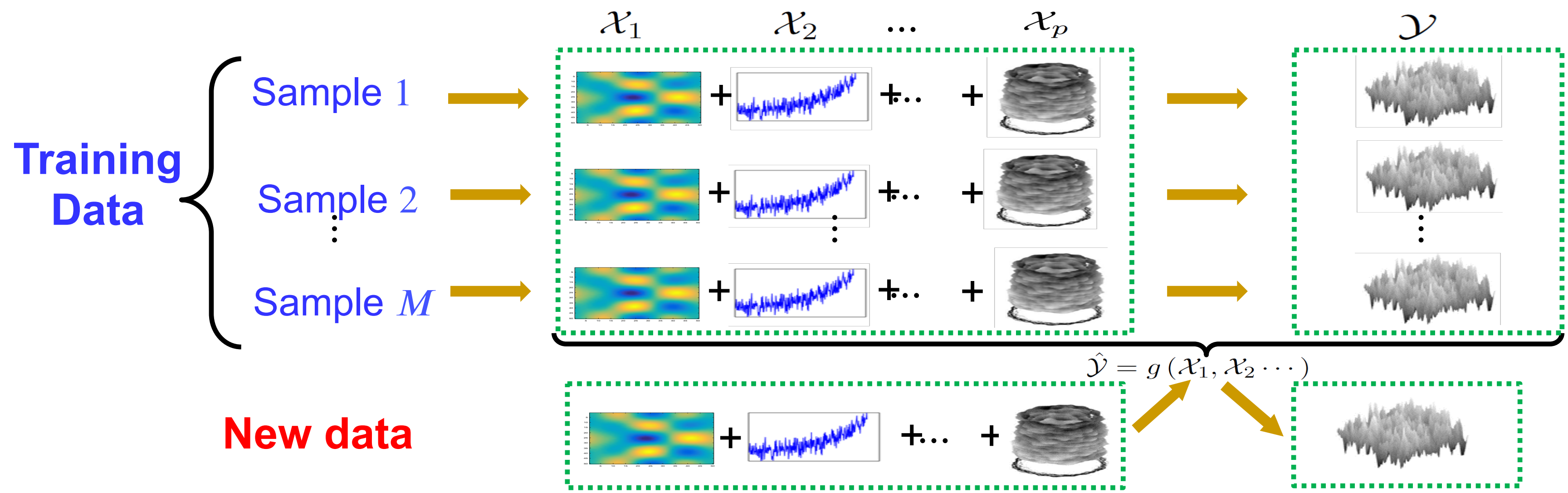
Integration of tensor structural information improves the performance for tensor signal restoration.

Mou, S., Wang, A., Zhang, C., & Shi, J. (2021). Additive Tensor Decomposition Considering Structural Data Information. *IEEE Transactions on Automation Science and Engineering*, 19(4), 2904-2917.

Multiple Tensor-on-Tensor Regression Model for Multistage Manufacturing Processes (MMP)

- Reisi Gahrooei, M., Yan, H., Paynabar, K., Shi, J., 2020, "Multiple Tensor on Tensor Regression: An approach for modeling processes with heterogeneous sources of data", *Technometrics*, *63*(2), 147-159.
- Miao, H., Wang, A., Chang, Z, and Shi, J. (2021), "Structural Tensor-on-Tensor Regression with Interaction Effects and Its Application to A Hot Rolling Process", *Journal of Quality Technology,* Vol. 54, Issue 5, p. 547-56

Multiple Tensor-on-Tensor Regression
To estimate an HD output given a set of HD inputs



Model:

$$\mathcal{Y} = \sum_{j=1}^{p} \langle \mathcal{X}_j, \mathcal{B}_j \rangle_{l_j} + \mathcal{E}$$

$$\mathcal{B}_j = \mathcal{C}_j \times_1 U_{j1} \times_2 \cdots \times_{l_j} U_{jl_j} \times_{l_j+1} V_1 \times \cdots \times_{l_j+d} V_d$$

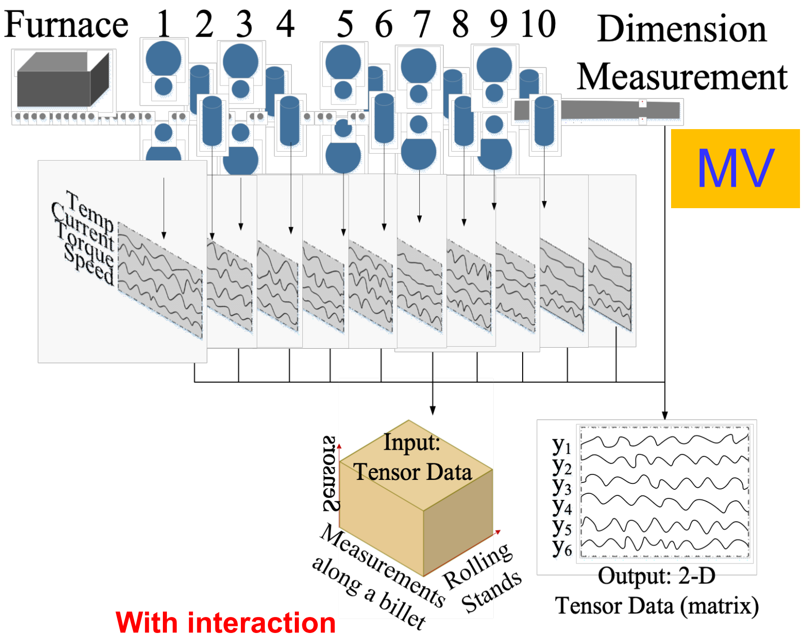
$$V_k^{\mathrm{T}} V_k = I_{\tilde{Q}_k} \qquad U_{ji}^{\mathrm{T}} U_{ji} = I_{\tilde{P}_{ji}}$$

Objective:

$$\arg\min_{\mathcal{C}_j, V, U} \left\| \mathcal{Y} - \sum_{j=1}^{p} \langle \mathcal{X}_j, \mathcal{C}_j \times_1 U_{j1} \times_2 \cdots \times_{l_j} U_{jl_j} \times_{l_j+1} V_1 \times \cdots \times_{l_j+d} V_d \rangle_{l_j} \right\|_F^2$$

Reisi Gahrooei, M., Yan, H., Paynabar, K., Shi, J., 2020, "Multiple Tensor on Tensor Regression: An approach for modeling processes with heterogeneous sources of data", *Technometrics*, *63*(2), 147-159.

Tensor-on-Tensor regression with Interaction effects Models with Application to a Hot Rolling Process



Model:

$$\mathcal{Y} = \mathcal{X}_1 * \mathcal{B}_1 + \mathcal{X}_2 * \mathcal{B}_2 + (\mathcal{X}_1 \circ \mathcal{X}_2) * \mathcal{B}_{[1,2]} + \mathcal{E}$$

Interaction effects

Challenges:

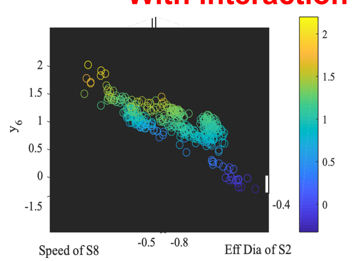
Curse of dimensionality

$\mathcal{B}_{[1,2]} \in \mathbb{R}^{K_1 \times M_1 \times K_2 \times M_2 \times D}$ is **a five-order tensor**.

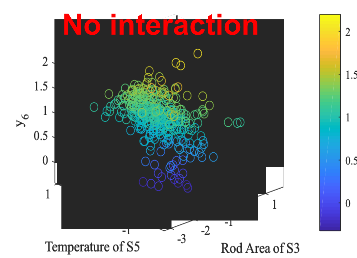
Complex structures

It is difficult to capture the structure of a high dimensional tensor.

With interaction



3D scatter plot among speed of stand 8, effective diameter of stand 2 and quality y_6



3D scatter plot among rod area of stand 3, temp of stand 5 and quality y_6

	OLSI	PLSI	CP_TOTI	Tucker_MTOTI	STOTI
MSPE	0.0976 (0.0397)	4.18e-04 (5.14e-05)	7.10e-04 (6.28e-05)	7.50e-04 (1.15e-04)	2.87e-04 (3.28e-05)

Miao, H., Wang, A., Chang, Z, and Shi, J. (2021), "Structural Tensor-on-Tensor Regression with Interaction Effects and Its Application to A Hot Rolling Process". *Journal of Quality Technology*, Vol. 54, Issue 5.

Image-Based Feedback Control Using Tensor Analysis

Zhong, Z., Paynabar, K., Shi, J., (2023). "Image Based Feedback Control using Tensor Analysis." *Technometrics,* DOI: 10.1080/00401706.2022.2157880

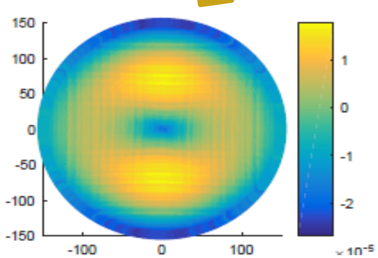
Motivation for Image-based Control

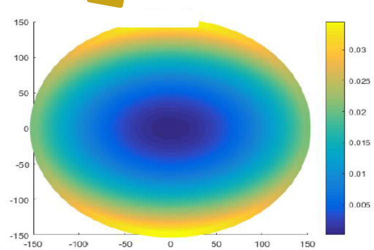
- **The product quality measures are sequential images or videos.**
- **Adjustment of process control variables will impact product quality.**

Stage position L_i, lens height $h_i, \ldots$

Stage position L_j, lens height $h_j, \ldots$

Lithography Machine





A set of control variables $\Rightarrow$
Minimize overlay output

Cutting Depth D_i, Speed v_i

Cutting Depth D_j, Speed v_j

Machining process

A set of control variables $\Rightarrow$
Minimize the deviations from nominal

Objective and Challenges

Objective:

Develop an optimal control framework for streaming image outputs by adjusting the input variables.

Challenges

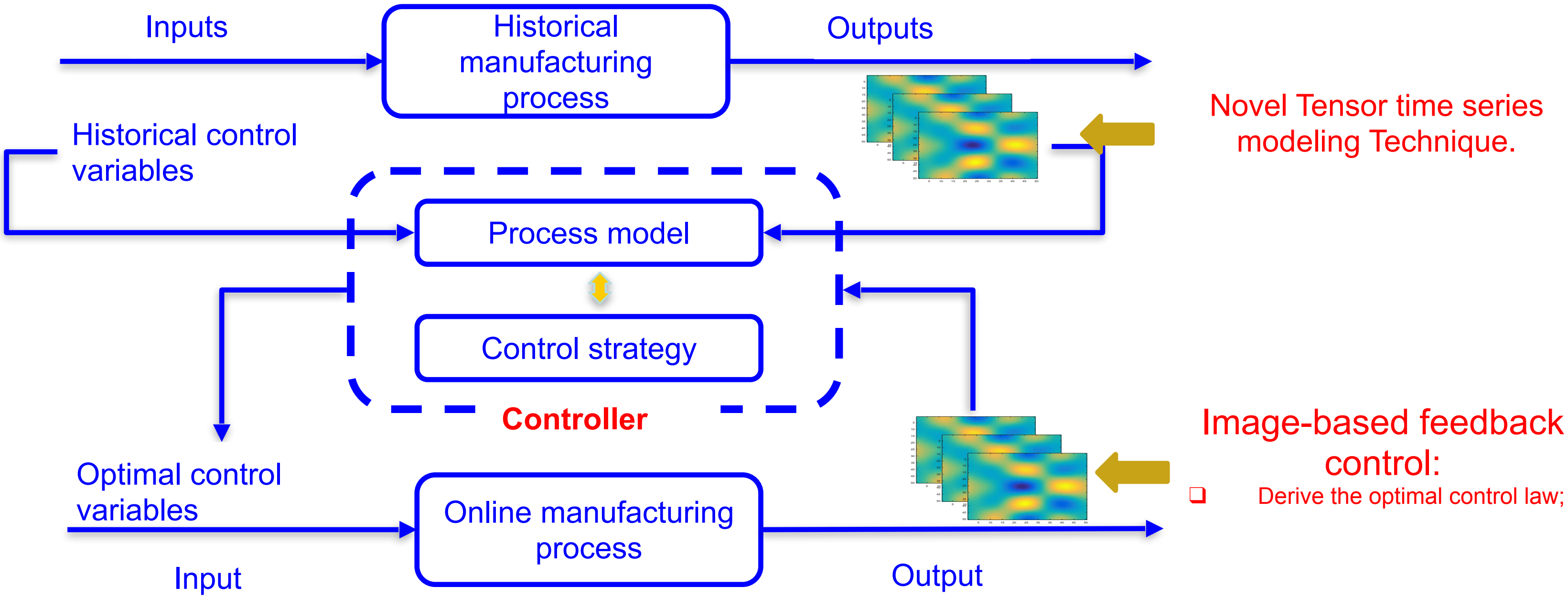
1. **High-dimensionality: How to avoid overfitting?**
2. **Spatial and temporal correlation structure: How to exploit?**
3. **Non-i.i.d Noise: How to model?**

Methodology:

Tensor-based time series modeling and control

1 2 3

Overview of Image-based Control



Process model:

High dimensional coefficients

$$\mathscr{Y}_t = \Sigma_{j=1}^{p} \mathscr{Y}_{t-j} * \mathscr{A}_j + \Sigma_{n=1}^{l} \mathscr{X}_{t-n} * \mathscr{B}_n + \delta E_t .$$

Current and previous image output

Control variables

Non-i.i.d Noise

How to learn

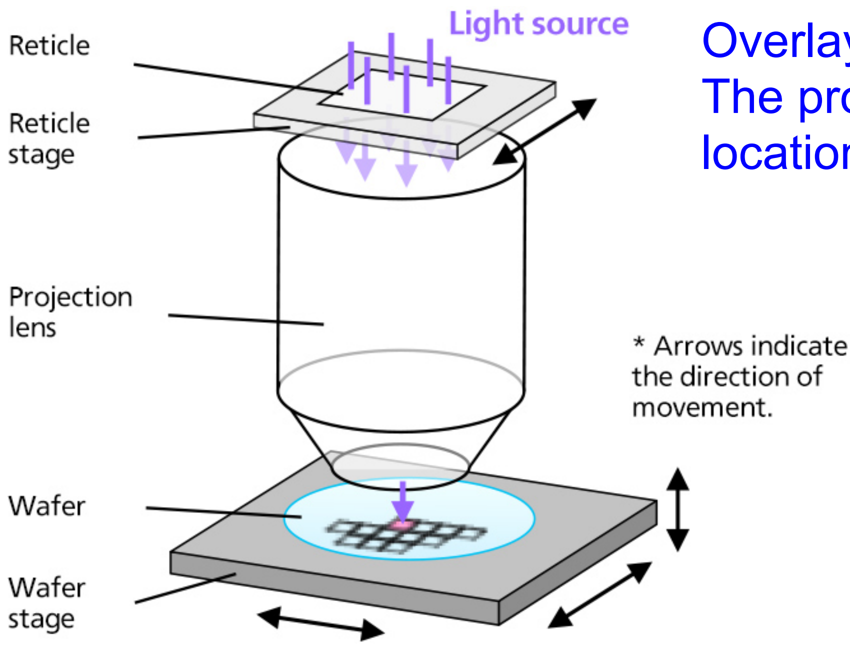
1) $\mathscr{A}$ and $\mathscr{B}$?
2) The correlation structure of $\delta\mathscr{E}$?

Control Objective function:

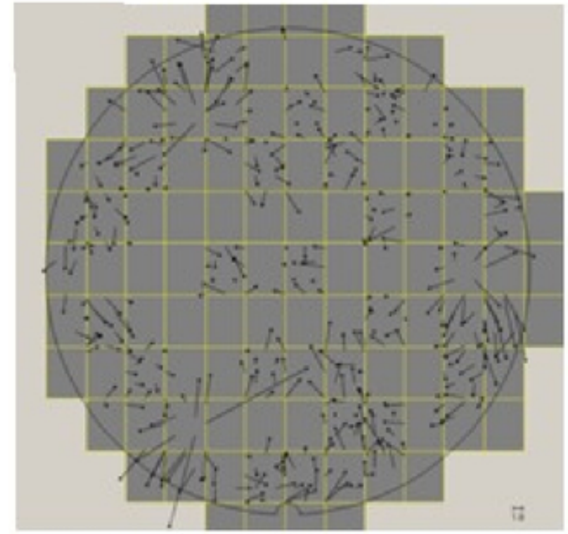
$$\mathrm{Min}_{\mathscr{X}_t} E\left(\hat{\mathscr{Y}}_{t+1|t}(\mathscr{X}_t) - T\right)^2$$

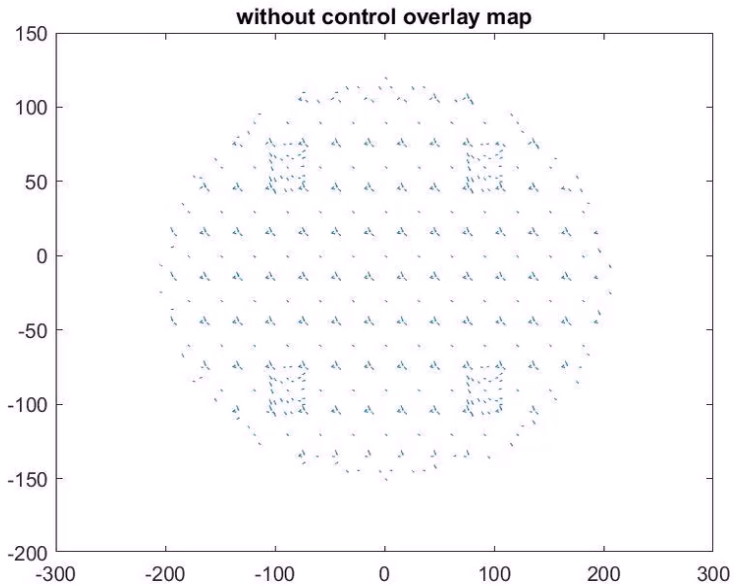
closed-form solution for control law $vec(\mathscr{X}_t)$

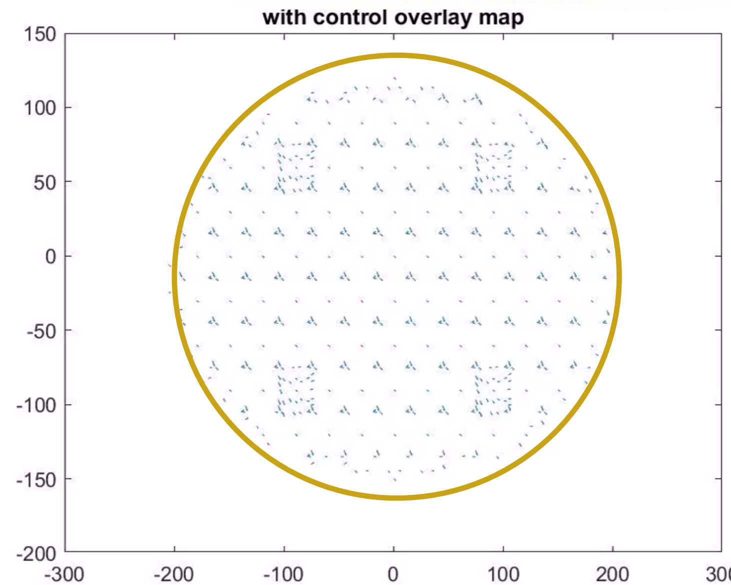
Case Study: The photolithography process



Overlay vector: The alignment error between (i) The projected pattern. (ii) The desired projected location on the wafer.





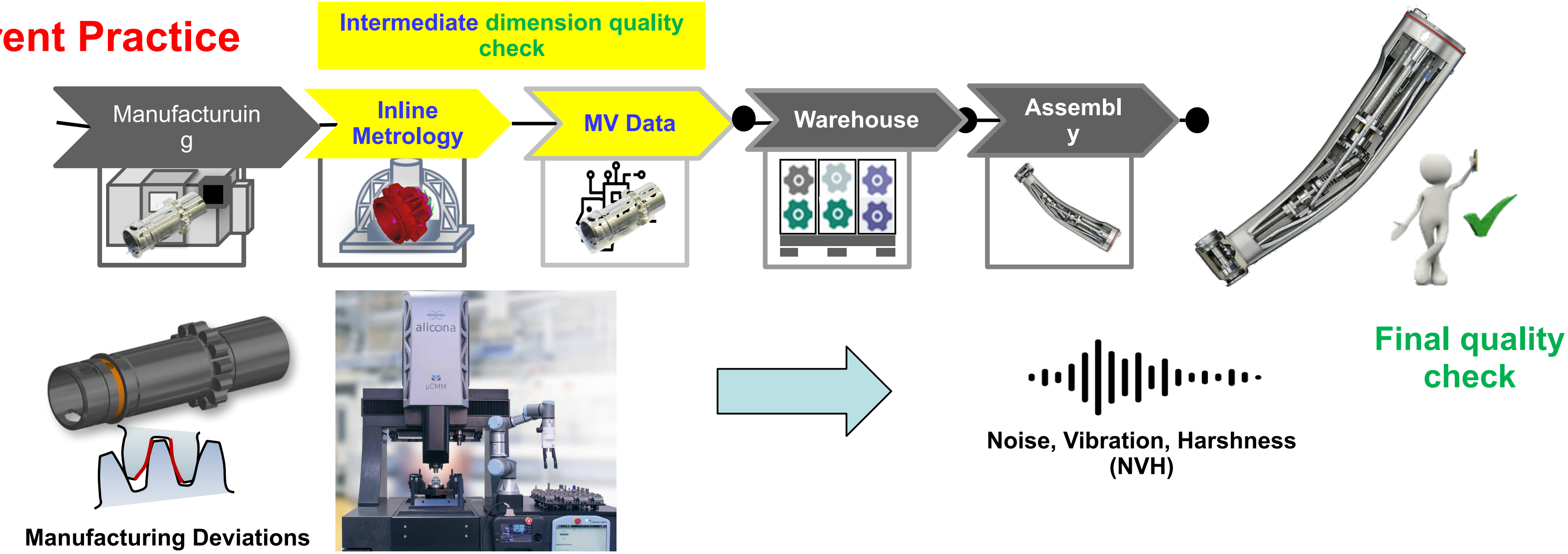


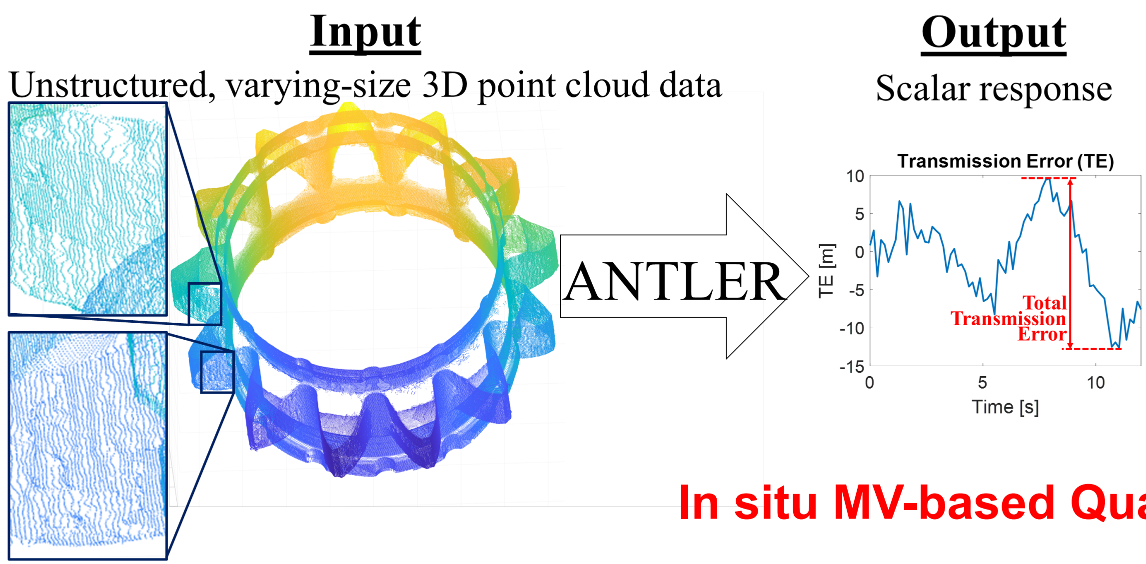
In-situ Product Quality Prediction based on 3D Point Cloud Data

Biehler, M., Yan, H., &Shi, J. (2022) "ANTLER: Bayesian Nonlinear Tensor Learning and Modeler for Unstructured, Varying-Size Point Cloud Data," in *IEEE Transactions on Automation Science and Engineering*, doi: 10.1109/TASE.2022.3230563.

Motivating Example – Medical Engineering

Current Practice

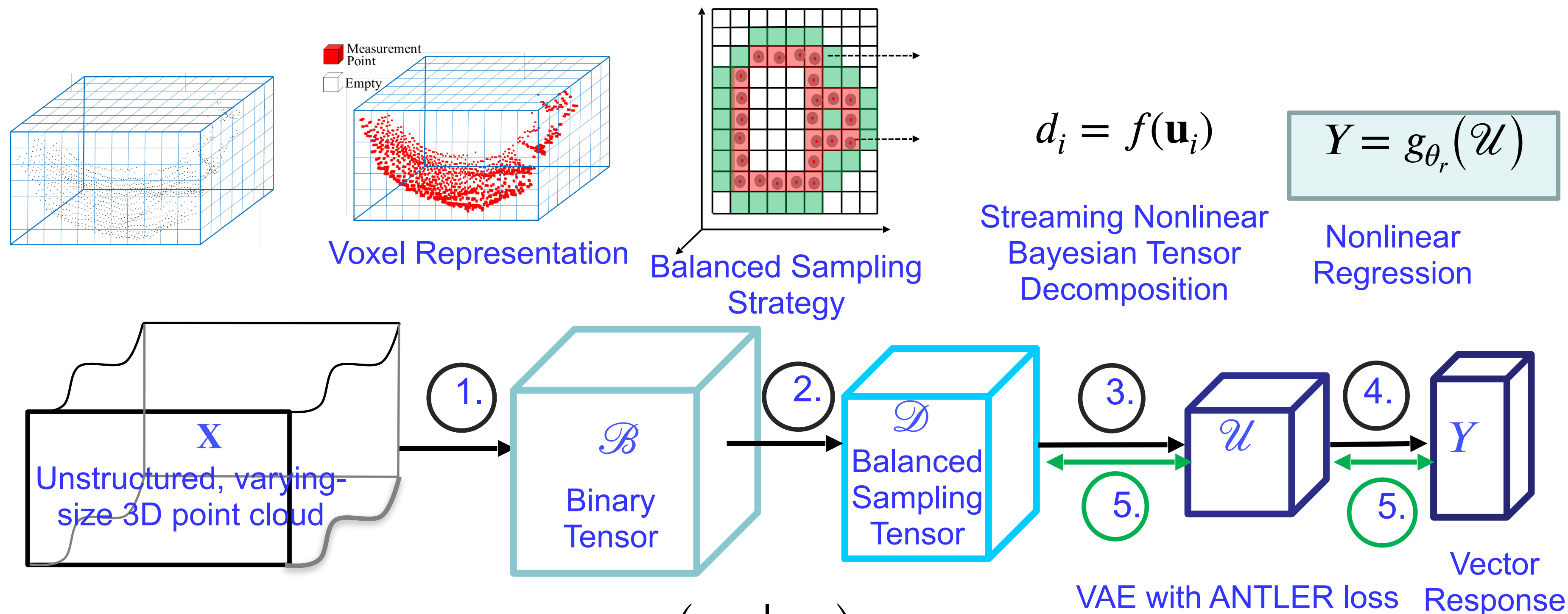




In situ MV-based Quality Prediction

Objective: Model a scalar response as a function of an unstructured, varying-size 3D point cloud

ANTLER Methodology – An Overview



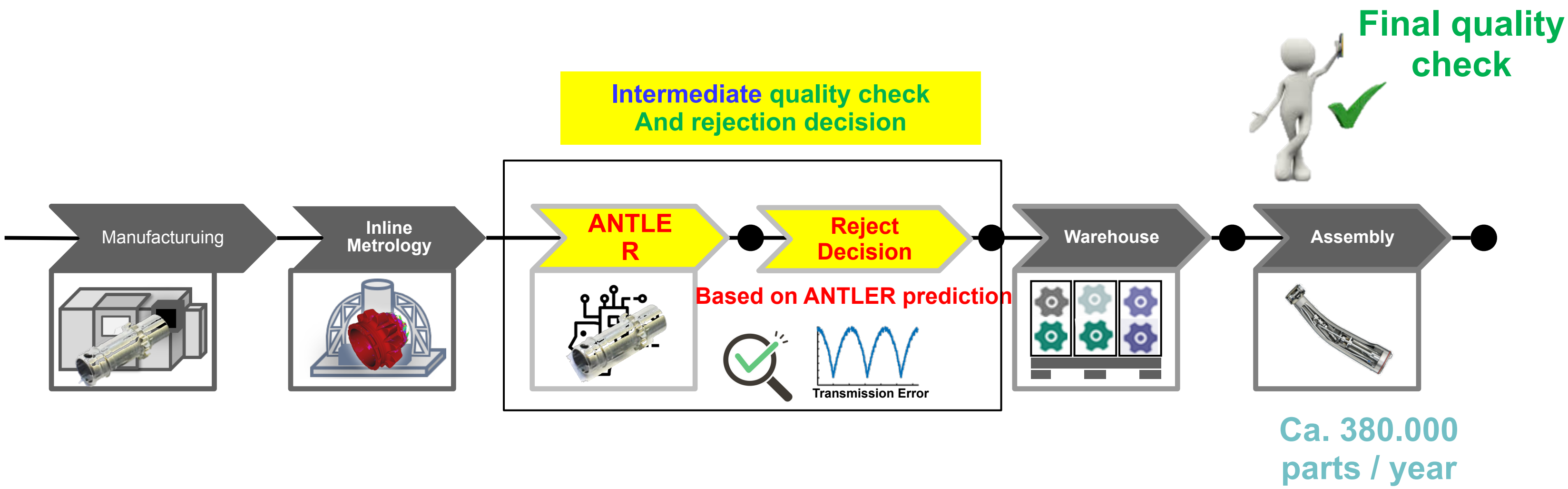
$$\mathscr{L}_{\phi,\theta,\theta_r}(\mathscr{D},Y) = \sum_{\mathscr{D}_i \in \mathscr{D}} \log \frac{1}{S}\sum_{j=1}^{S} \frac{p_\theta\left(\mathscr{D}_i \middle| \mathbf{z}_{(j)}\right)}{q_\phi\left(\mathbf{z}_{(j)} \middle| \mathscr{D}_i\right)} + \lambda_1 D_{KL}\left(q_\phi\left(\mathbf{z} \middle| \mathscr{D}_i\right) \middle\| p(\mathbf{z})\right)$$

$$+\lambda_2 \left\| \mu_z - \mu_{\mathscr{U}_{SNBTD}} \right\|_2^2 + \lambda_3 \left\| g_{\theta_r}(\mu_z) - Y \right\|_2^2$$

- Utilize **high expressivity** of VAE and **fast inference time**
- Leverage performance of **SNBTD** on **small sample sizes**
- Simultaneously optimize regression and low dimensional representation

Biehler, M., Yan, H., &Shi, J. (2022) “ANTLER: Bayesian Nonlinear Tensor Learning and Modeler for Unstructured, Varying-Size Point Cloud Data,” in *IEEE Transactions on Automation Science and Engineering*, doi: 10.1109/TASE.2022.3230563.

ANTLER – Industry Implementation and Impact



- **Introduced intermediate quality check based on ANTLER prediction results:**
  - **Decision rule:**
    - **Accept if predicted transmission error < Engineering Tolerance**

- **Since implementation (February 2022):**
  - **Reduction of total scrap cost by 61%**

Biehler, M., Yan, H., &Shi, J. (2022) "ANTLER: Bayesian Nonlinear Tensor Learning and Modeler for Unstructured, Varying-Size Point Cloud Data," in *IEEE Transactions on Automation Science and Engineering*, doi: 10.1109/TASE.2022.3230563.

DETONATE: Nonlinear Dynamic Evolution modeling of Time-dependent 3-dimensional point cloud profiles

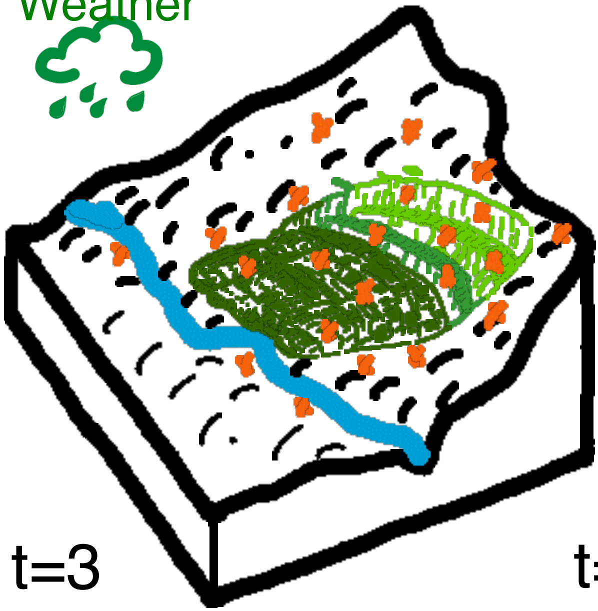
Biehler,M., Lin, D. & Shi, J. (2023) DETONATE: Nonlinear Dynamic Evolution Modeling of Time-dependent 3-dimensional Point Cloud Profiles, *IISE Transactions*, (accepted, in press) DOI: 10.1080/24725854.2023.2207615

DETONATE: Nonlinear Dynamic Evolution modeling of Time-dependent 3-dimensional point cloud profiles

Motivation: dynamically evolving 3D shapes are common

Landslide

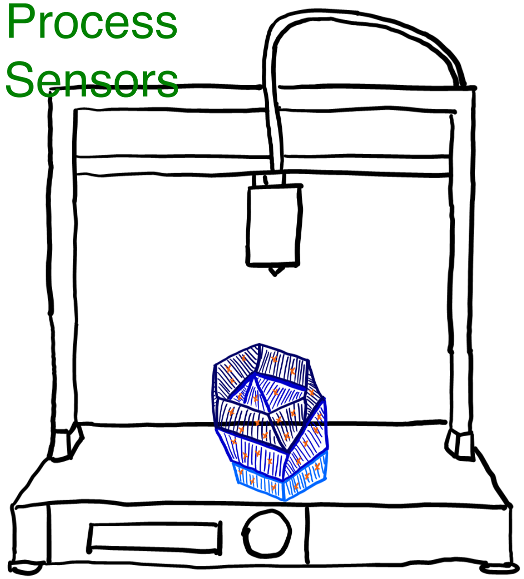
Weather



t=3

3D Printing

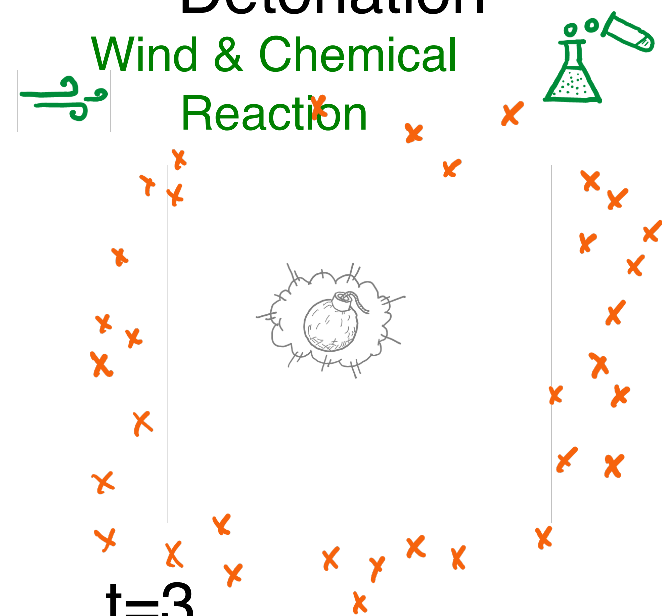
Process Sensors



t=3

Detonation

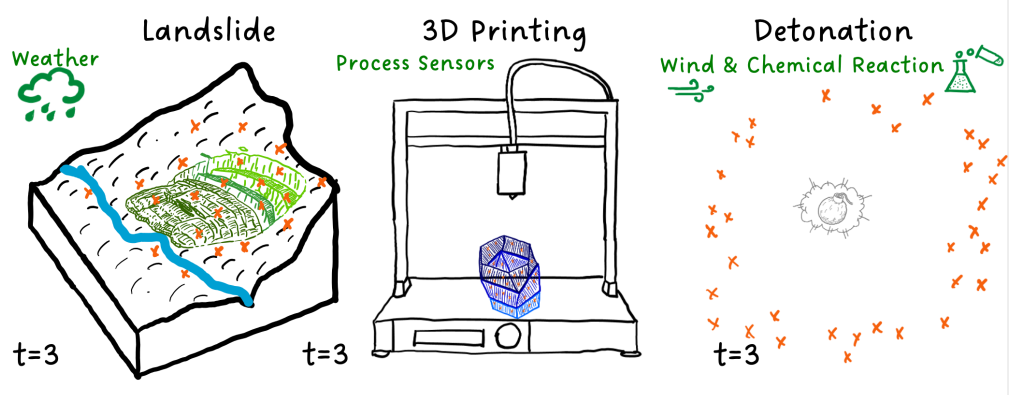
Wind & Chemical Reaction



t=3

X Unstructured Measurement Points

Objective: Modeling of dynamically evolving 3D shapes according to temporal propagation and heterogenous inputs



Common Data Characteristics:

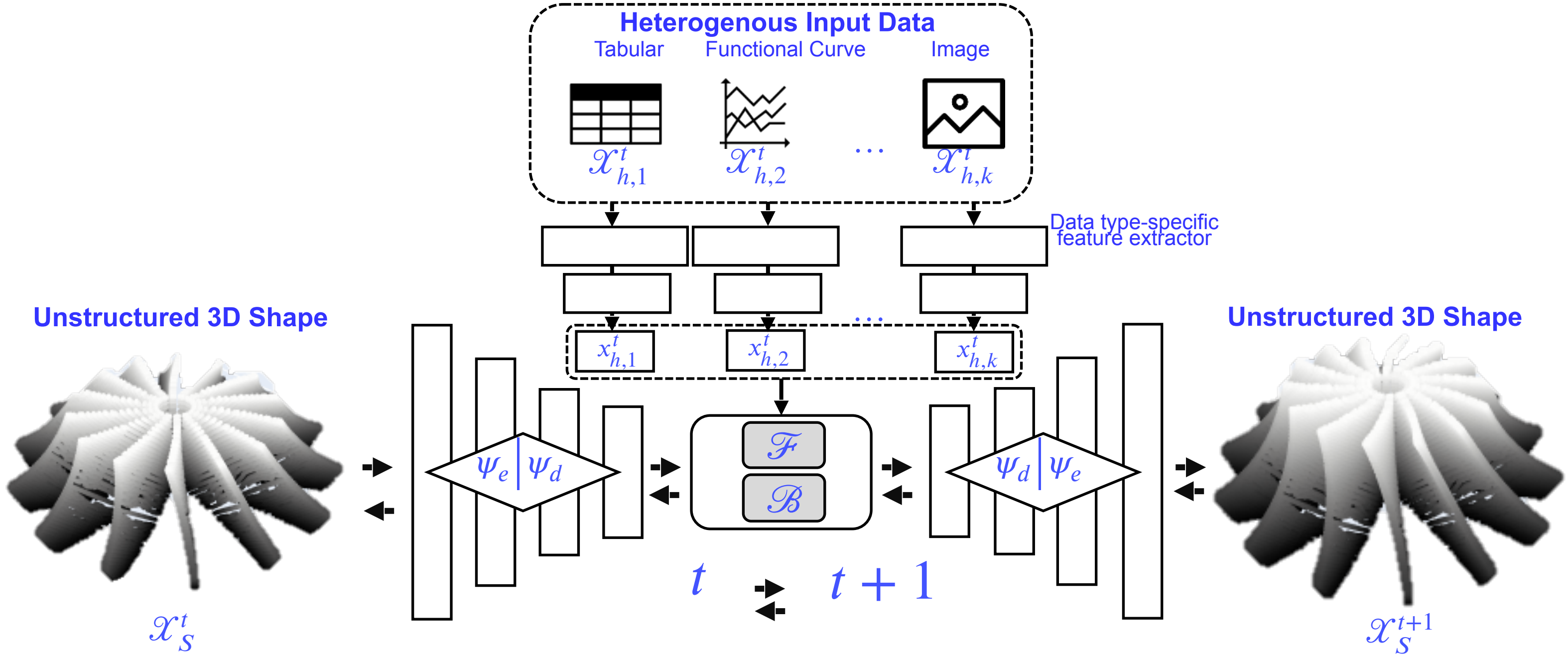
- Dynamically, nonlinear, temporally evolving 3D shape profiles
- 3D shapes are represented by unstructured 3D point clouds
- 3D shape profiles are spatially affected by heterogeneous input data
- Backward predictions are also relevant: Root Cause Analysis

Challenges

1. Evolving 3D shapes exhibit **complex spatio-temporal structure**: **How to model?**
2. **Forward and backward** predictions: **How to exploit and combine?**
3. **Unstructured** data structure of **3D point clouds: How to process?**
4. **Heterogenous** input **data: How to fuse with temporal model?**

Proposed DETONATE Methodology

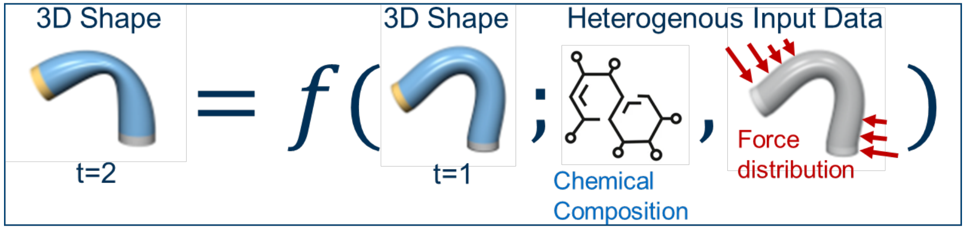
- Problem setup using Koopman Operator Theory
- Latent Encoding via 3D Autoencoder
- Forward and Backward Dynamics
- Consistent Dynamics
- Heterogenous Input Data Sources
- Unified DETONATE framework



M. Biehler, D. Lin, and J. Shi (2023) "DETONATE: Nonlinear Dynamic Evolution Modeling of Time-dependent 3-dimensional Point Cloud Profiles", IISE Transactions (accepted), http://dx.doi.org/10.13140/ RG.2.2.28631.34720/1

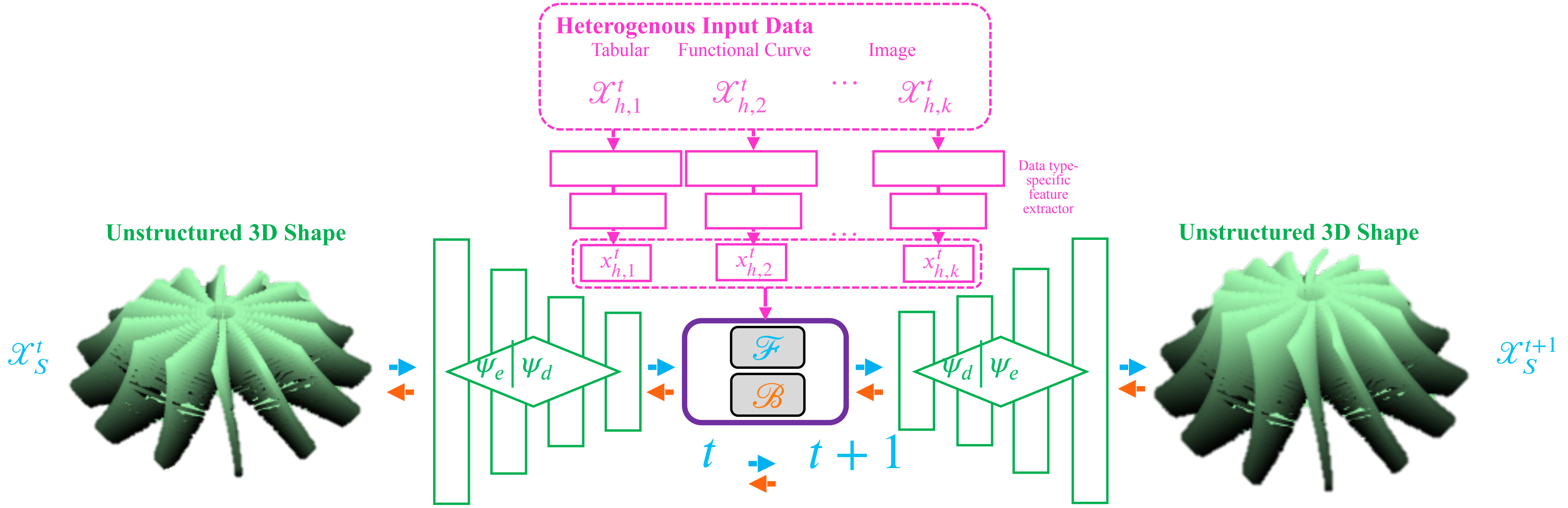
Unified DETONATE Framework

- **Goal:** $\mathscr{X}_S^{t+1} = f(\mathscr{X}_S^t, \mathscr{X}_{h,j}^t)$



- **Unified DETONATE Loss:**

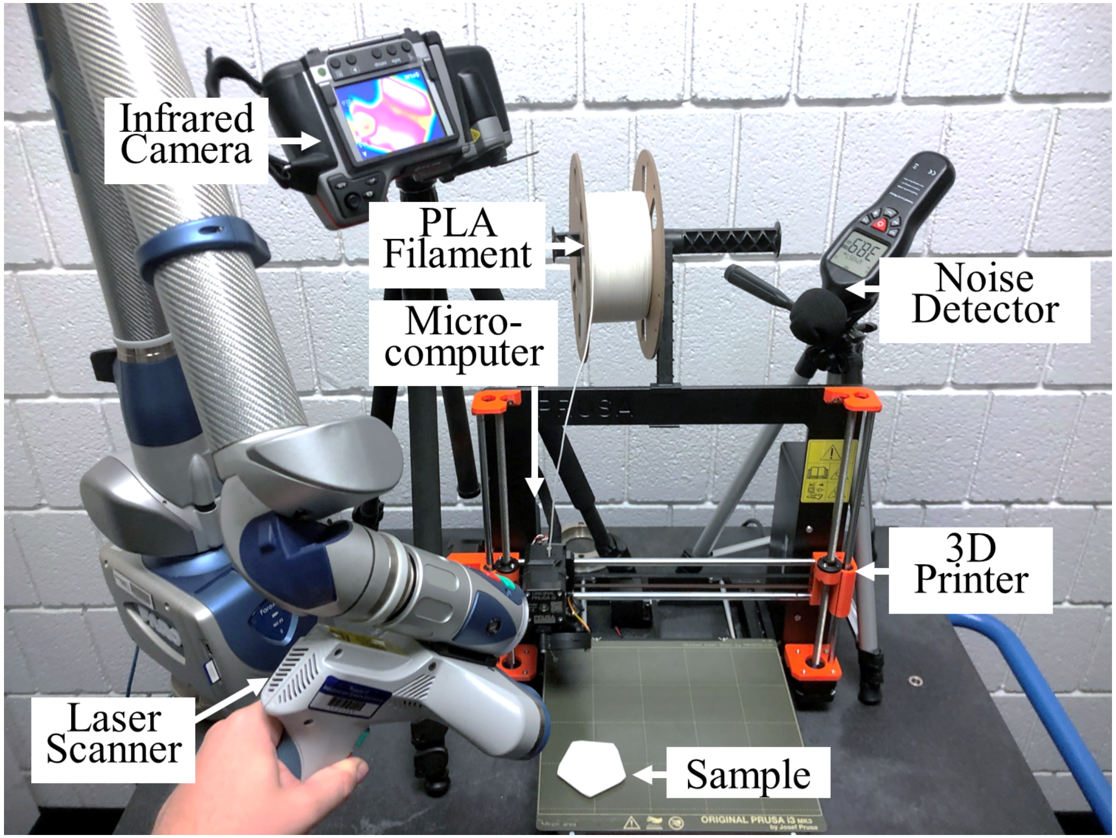
$$\mathscr{L}_{DETONATE} = \mathscr{L}_{rec} + \lambda_1 \bullet \mathscr{L}_{fwd} + \lambda_2 \bullet \mathscr{L}_{bwd} + \lambda_3 \bullet \mathscr{L}_{con} + \lambda_4 \bullet \mathscr{L}_{h}$$

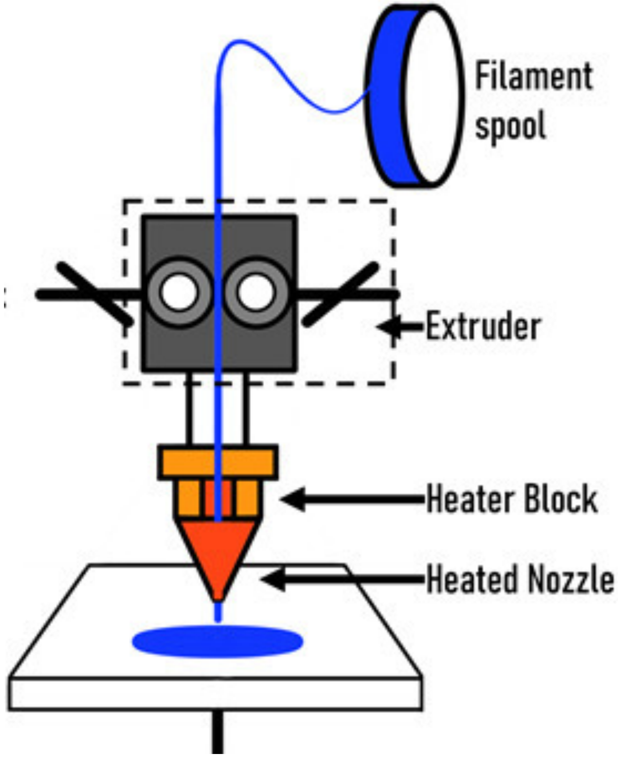


Utilize a unified Optimization Framework to combine all components to improve predictive performance

Case Study – 3D Printing Experiments

- **3D printing of PLA specimen via Fused Filament Fabrication (FFF)**



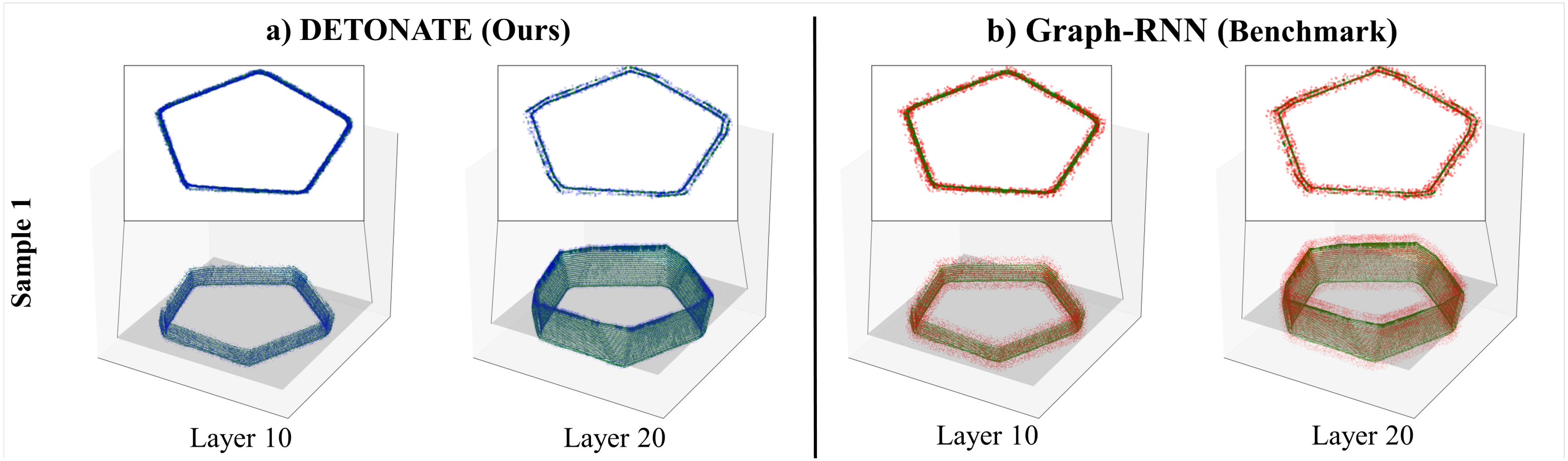


- Prusa MK3S FDM printer
- Infrared Camera FLIR T360
- Laser Scanner FARO Quantum ScanArm
- Microcomputer to log nozzle and print bed temperature
- Noise detector

Case Study – 3D Printing - Results

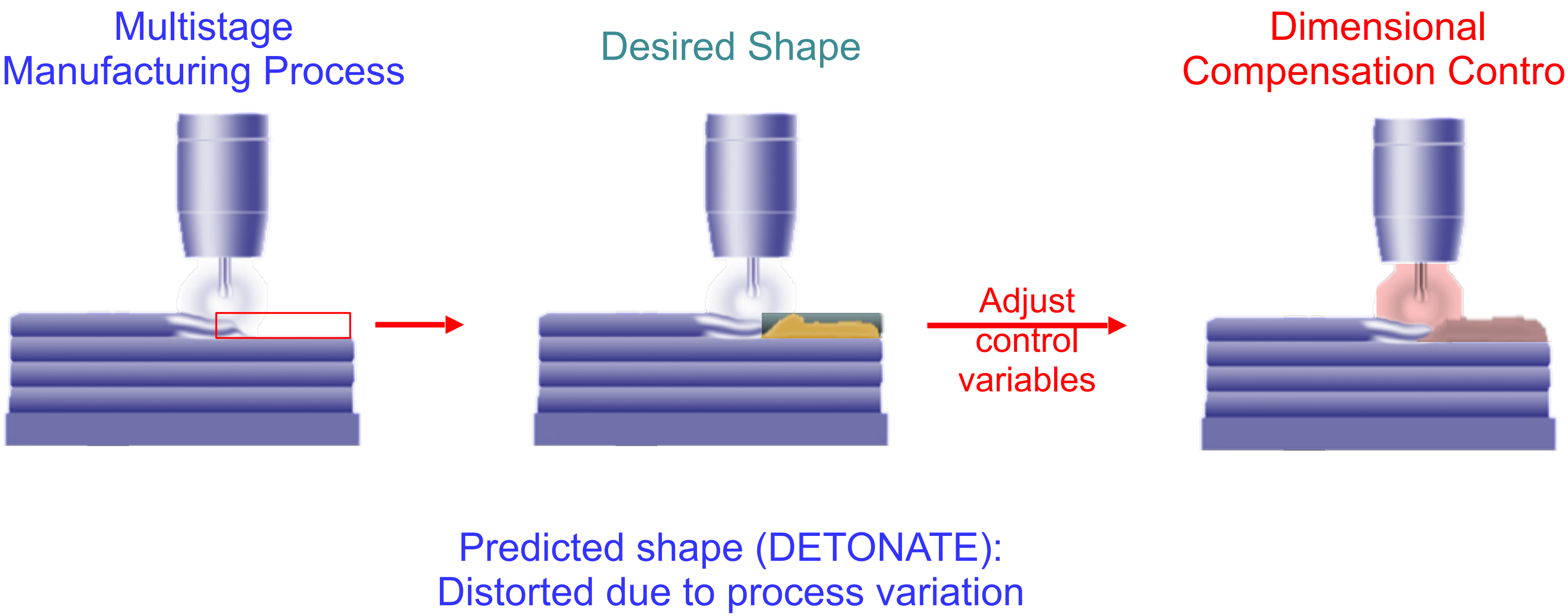
- **Temporally evolving 3D point cloud profiles:**
  - **Up- or down-sampled to a fixed-point number of** $N_p = 60{,}000$
  - **If more precision is required: Our ANTLER work on <u>varying-size</u>, unstructured point clouds can be adapted**

- **Prediction Results**

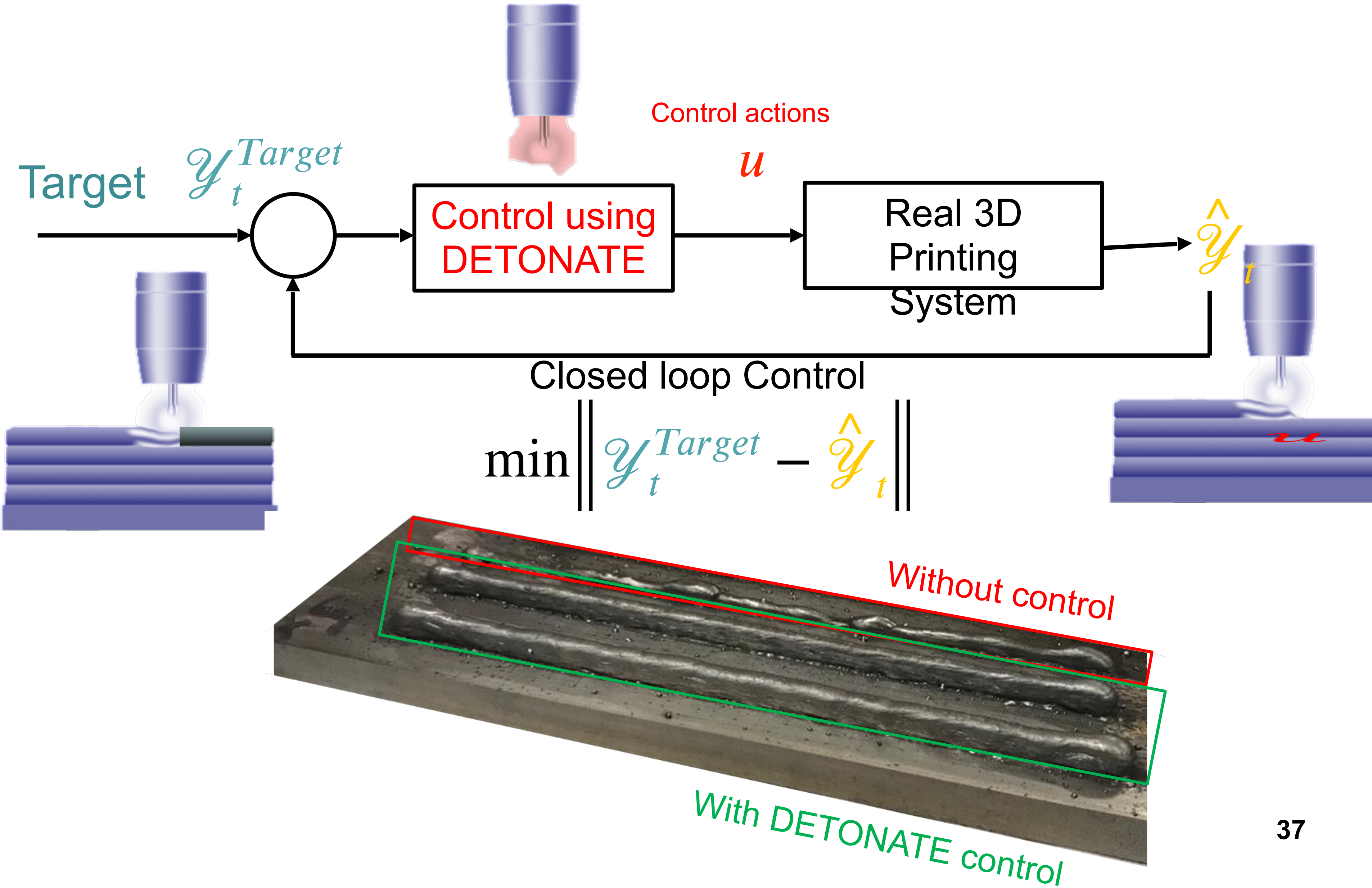


M. Biehler, D. Lin, and J. Shi (2023) "DETONATE: Nonlinear Dynamic Evolution Modeling of Time-dependent 3-dimensional Point Cloud Profiles", IISE Transactions (accepted), http://dx.doi.org/10.13140/RG.2.2.28631.34720/1

- **DETONATE enables next generation of In-Process Quality Improvement (IPQI) methodologies**
- **Compensation and Control: Concept**
  - **Adjust process variables to achieve desired 3D shape**



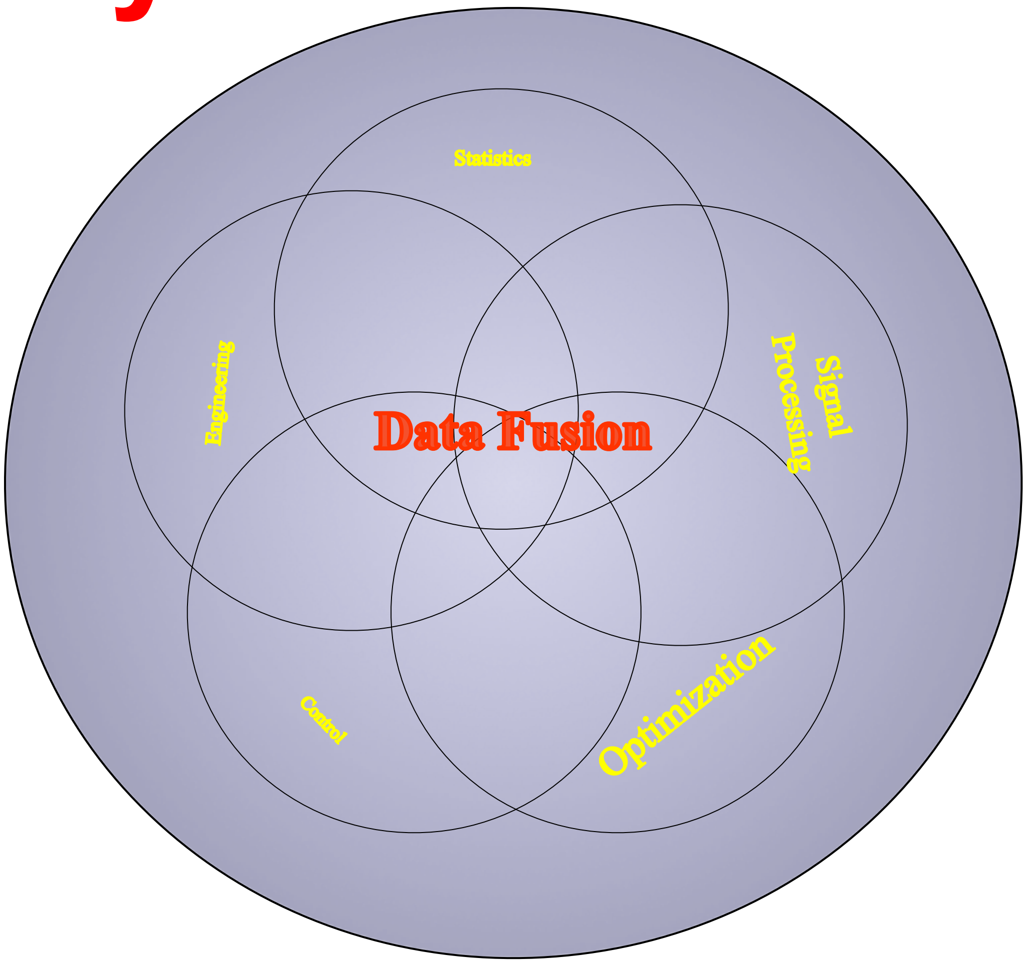
DETONATE Model-based Dimensional Compensation Control - Concepts and Experimental Results



Summary

- Machine Vision has been widely used in smart manufacturing to measure product quality and process/machine conditions, which enables the IPQI.
- Machine Learning and Data fusion are key components of R&D, which requires multidisciplinary efforts from engineering, statistics, machine learning, and control.
- Machine Vision enabled IPQI has been generated significant economic impacts in numerous manufacturing systems, and much more work needs to be done in this important area.
- Close collaborations between industry and academia are essential to move forward of machine vision enabled IPQI.

Thank you!



Statistics
Engineering
Signal Processing
Data Fusion
Control
Optimization