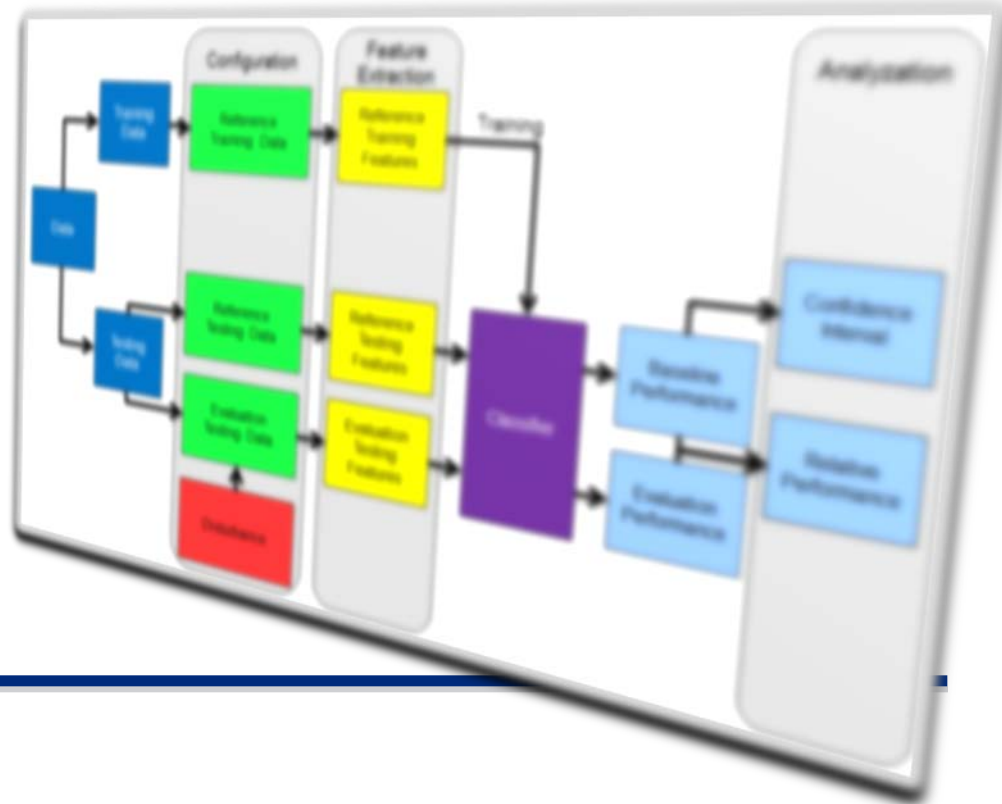


DERMALOG

On Testing the Robustness of Liveness Detection Feature Extractors for Fingerprint Live Scanners



Outline

- ▶ Why liveness detection?
- ▶ Typical Workflow
- ▶ Known challenges
- ▶ How to overcome
- ▶ An example
- ▶ Discussion & Conclusion

Who is DERMALOG?

- ▶ HQ in Hamburg, Germany
- ▶ Outpost in Kuala Lumpur, Malaysia
- ▶ Modality: fingerprints
- ▶ Main Products:
 - ▶ AFIS
 - ▶ Live scanners

Why Liveness Detection?

Spoofer

Easy &
cheap

Manuals
available

Vulnerable
systems

Risk

False
accepts

Loss of
trust

Loss of
acceptance

Why Liveness Detection?

Detect spoofing



Increase security



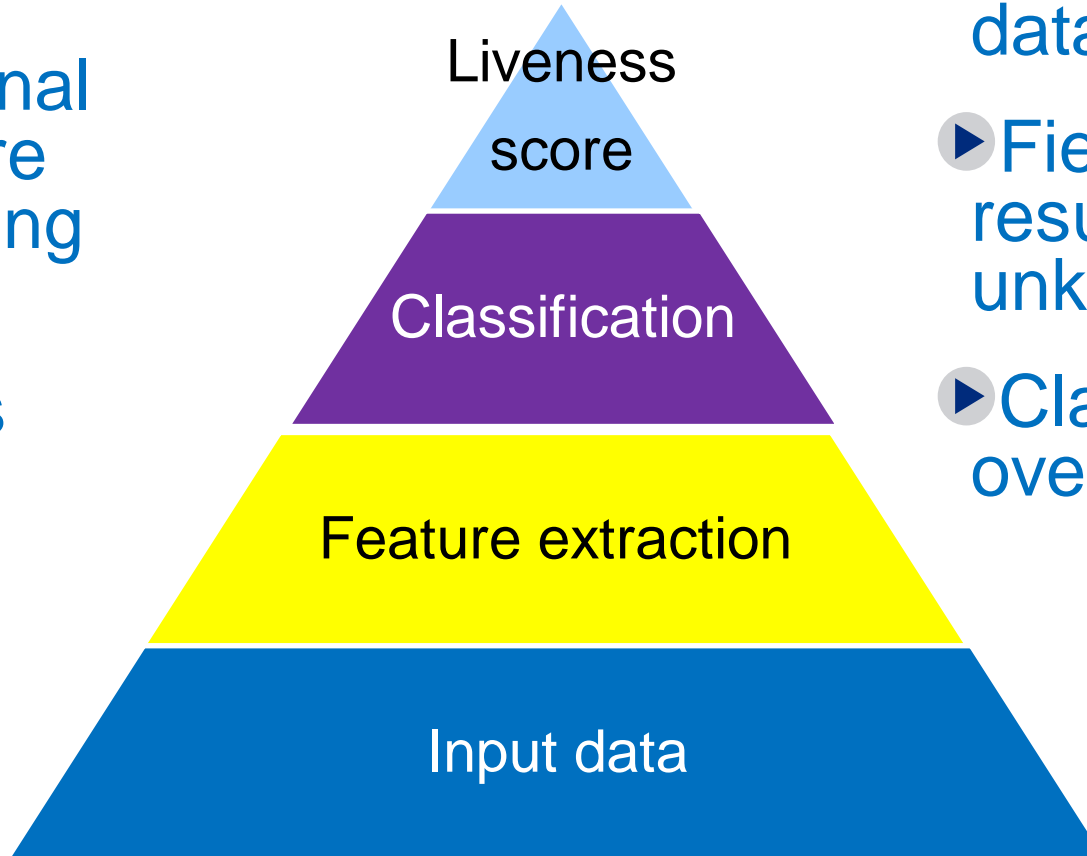
Increase customer's trust



Increase acceptance

Typical Workflow

- ▶ Grayscale images
- ▶ Additional hardware measuring
- ▶ Result: liveness score



- ▶ Lab: good results on known data
- ▶ Field: bad results on unknown data
- ▶ Classifier overtrained?

Known Challenges

External

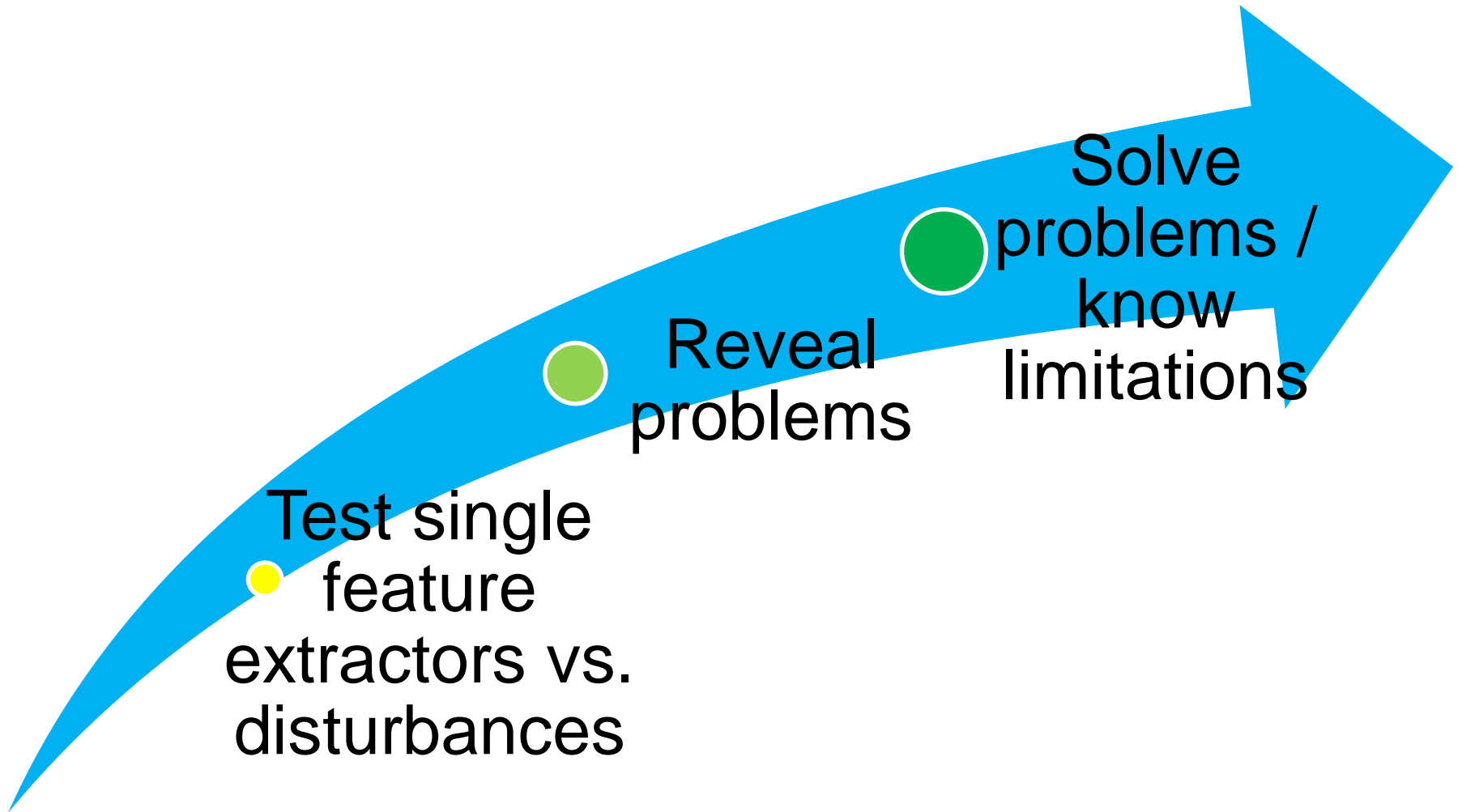
- Illumination
- Temperature
- Humidity
- Dust/Dirt
- Latents
- Population (skin etc.)
- ...

Internal

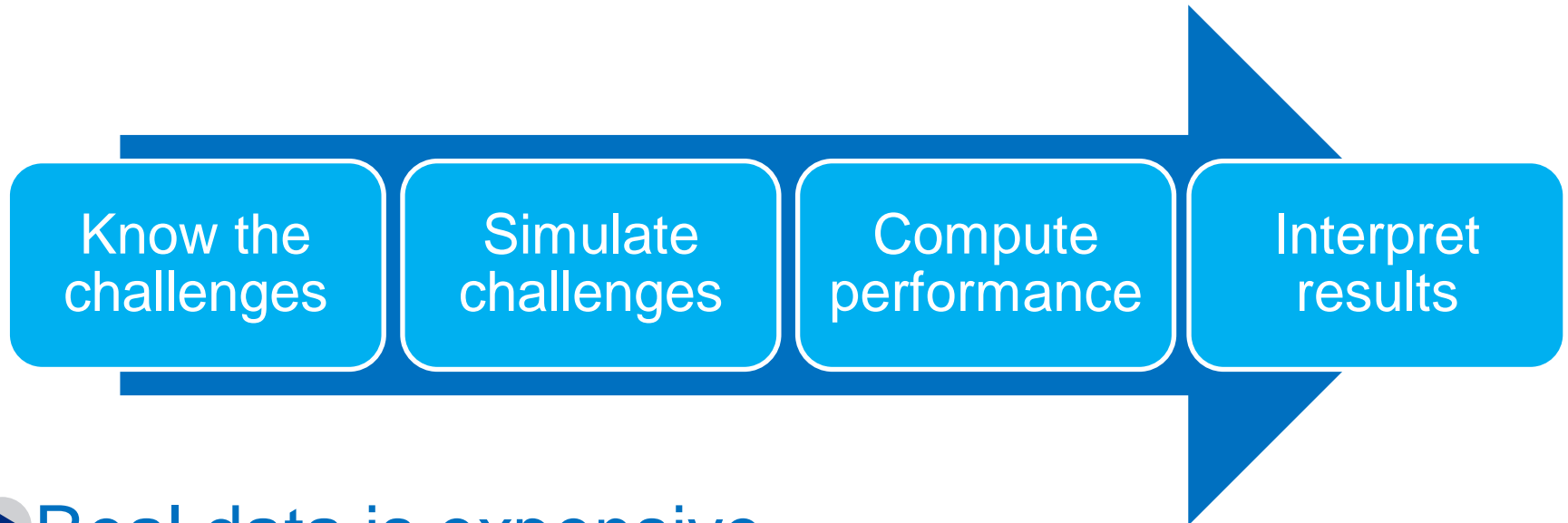
- Noise
- Defects
- Sharpness
- Fabrication tolerances
- ...

► **ISO/IEC 29197: Evaluation Methodology for Environmental Influence in Biometric Systems Performance**

Known Challenges

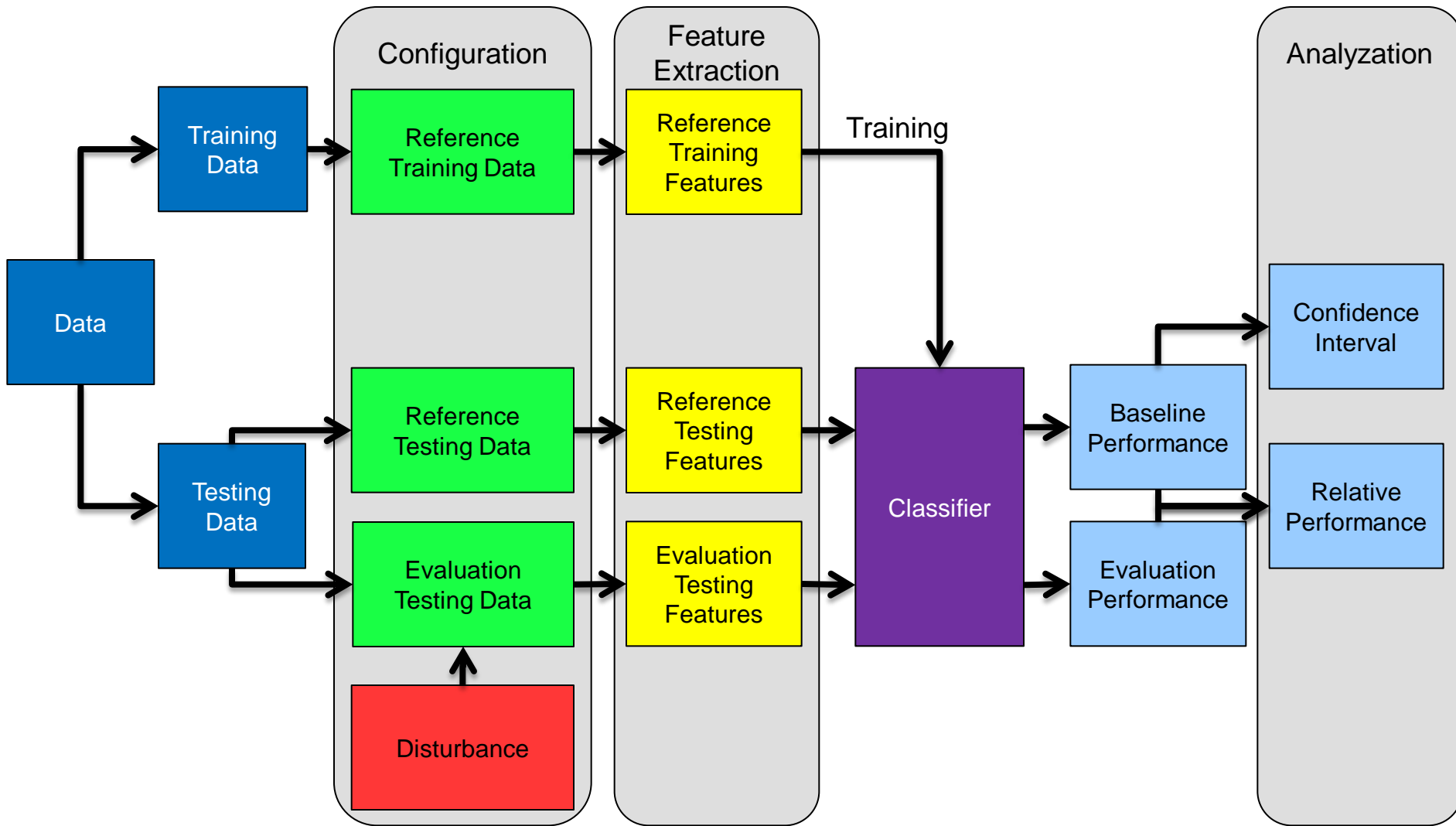


How to Test



- ▶ Real data is expensive
- ▶ Disturbances:
 - ▶ How to model them mathematically
 - ▶ What must be expected?

How to Test



Example - Data

- ▶ LivDet2009 data set

- ▶ Subset only

- ▶ Identix DFR2100

- ▶ 686 dpi, 720x720 pixels

- ▶ Cropped ROI

Training
data

750
genuine

250
gelatin

Testing
data

3,000
genuine

1,000
gelatin

Example - Tested Feature Extractors

$$\gamma_2$$

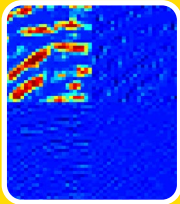
Kurtosis

- Proposed by Abhyankar et al. [2006]
- Baseline performance: EER ~ 21%

$$\mu / \sigma$$

Coefficient of variation

- Proposed by Abhyankar et al. [2006]
- Baseline performance: EER ~ 8.9%



Surface coarseness

- Proposed by Moon et al. [2008]
- Baseline performance: EER ~ 25%



Combined

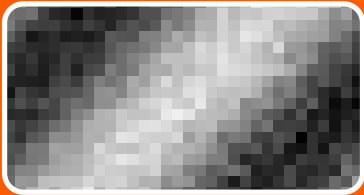
- Combination of the previous three
- Baseline performance: EER ~ 8.4%

Example - Disturbances to be Tested



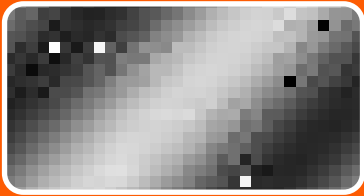
Illumination offset

- Internal/External illumination variation
- Homogenous offset



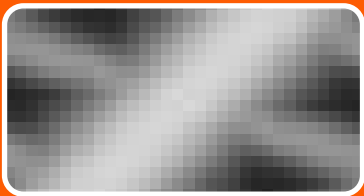
White Gaussian Noise

- Noise in components



Shot Noise

- A.k.a. “Salt and Pepper Noise”
- Defect pixels



Latents

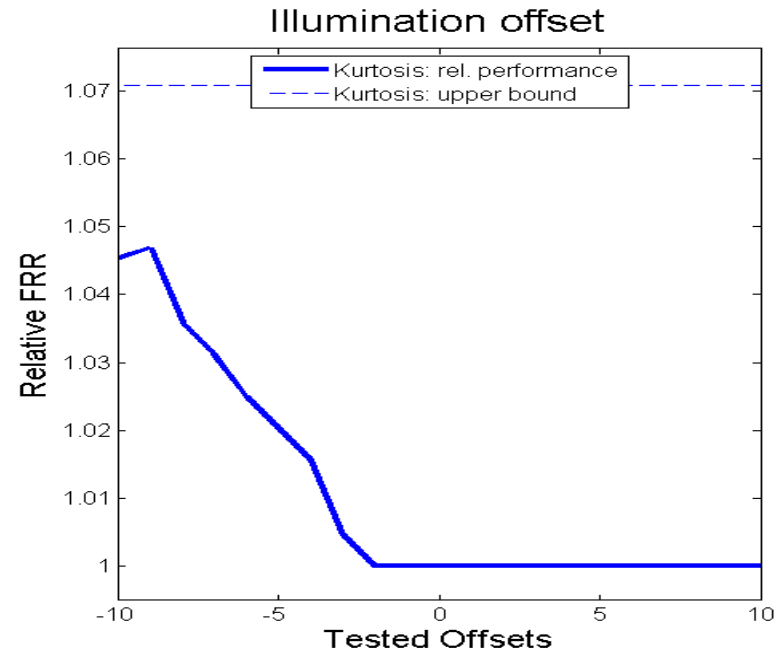
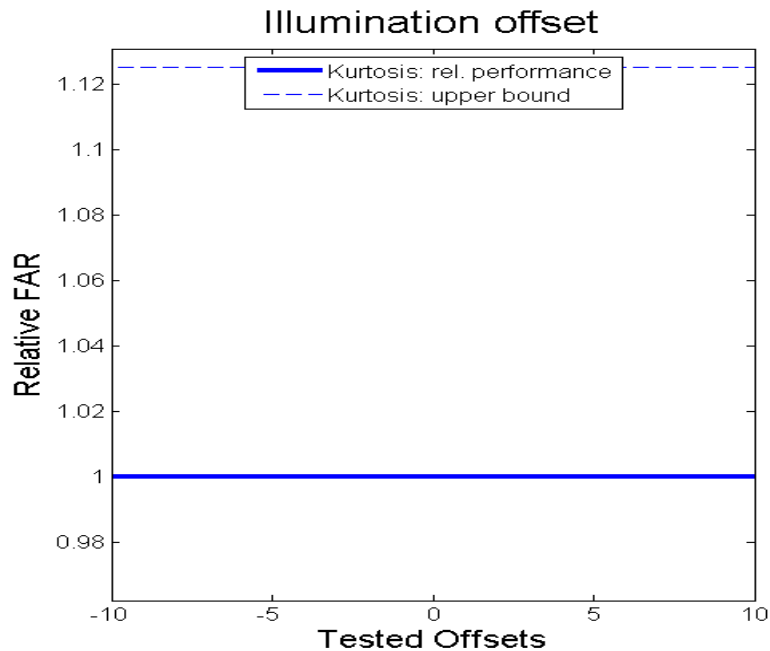
- “Dirt” from previous acquisitions
- Simulated latents

Example – Confidence Intervals

- ▶ Naive Bayes Classifier
- ▶ False suspicious presentation detection rate (false SP-DR aka FRR)
- ▶ False suspicious presentation non-detection rate (false SP-NDR aka FAR)
- ▶ Use upper bounds of Confidence Intervals (significance level $\alpha = 0.05$)
- ▶ Only relative error rates will be used

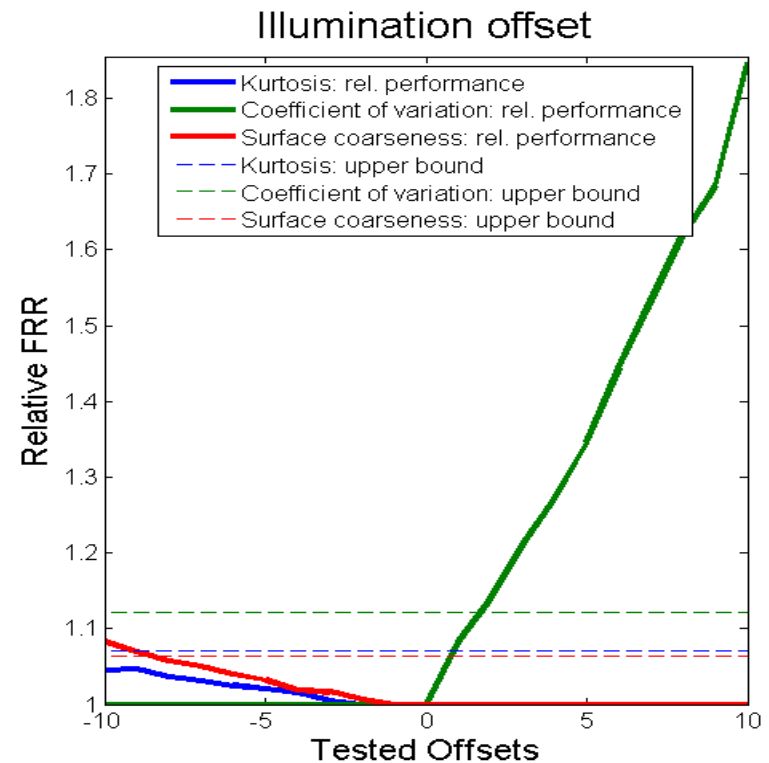
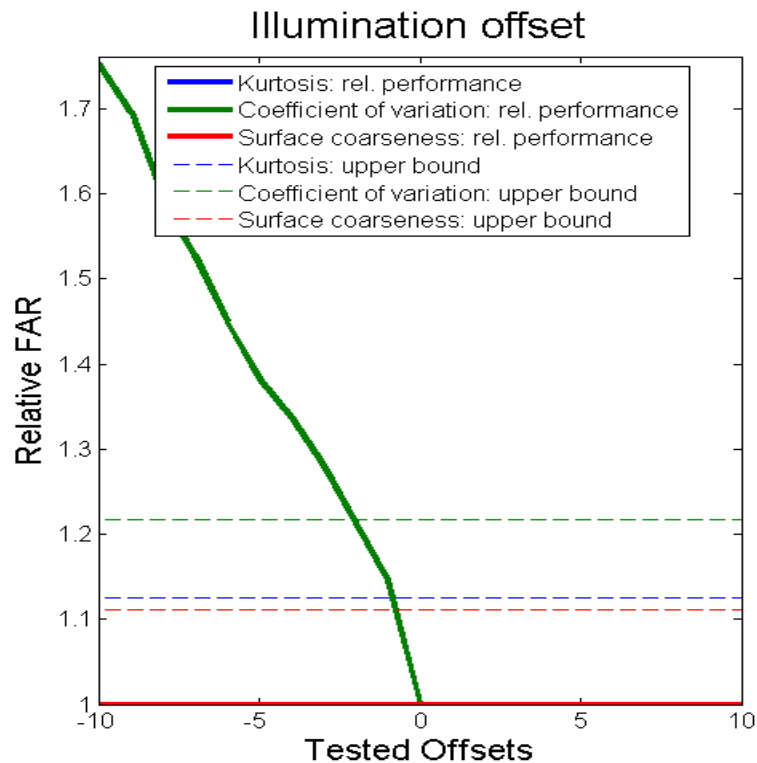
Feature Extractor	EER on Testing Set	Upper Bound for Relative FRR	Upper Bound for Relative FAR
Kurtosis	21.3%	1.07	1.12
Coefficient of variation	8.9%	1.12	1.22
Surface coarseness	25.0%	1.06	1.11
Combined	8.4%	1.13	1.23

Example – Some Results



- ▶ Baseline performance for offset 0
- ▶ Tendencies revealed in relative error rates
- ▶ Apply confidence intervals
- ▶ Crop to relative performance 1.0

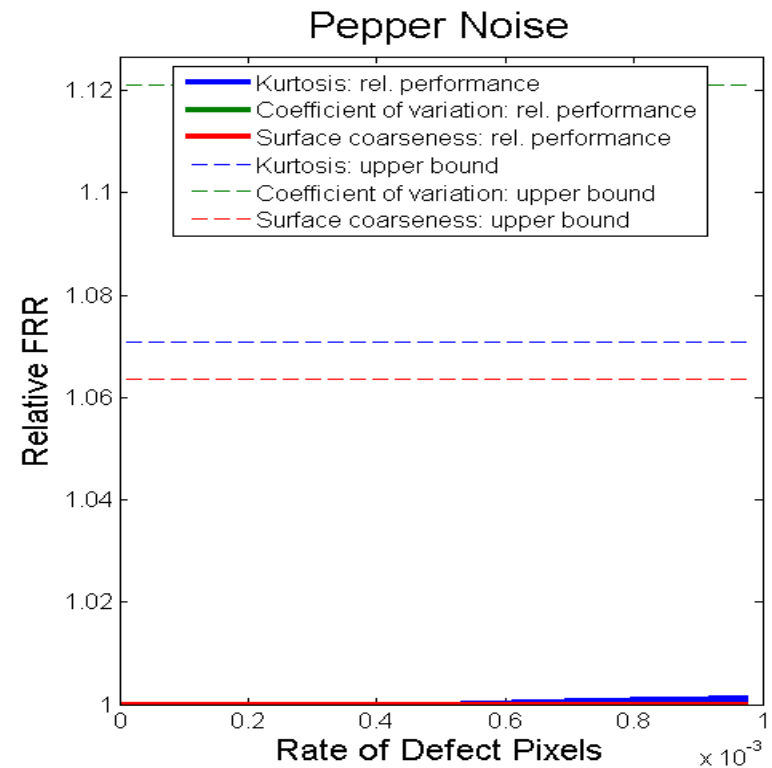
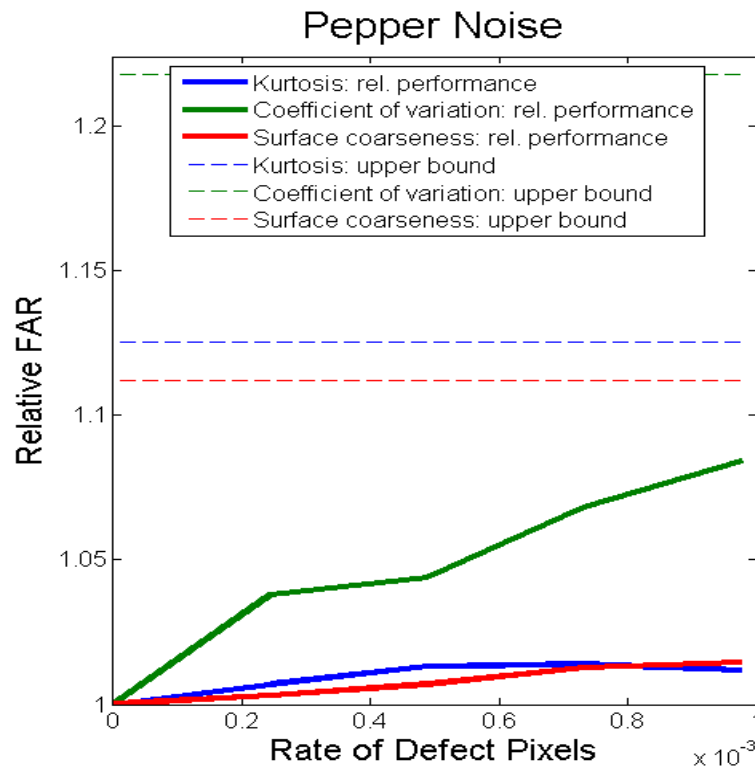
Example – Some Results



► Parameter range [-10 10]

► Strong decay for coefficient of variation

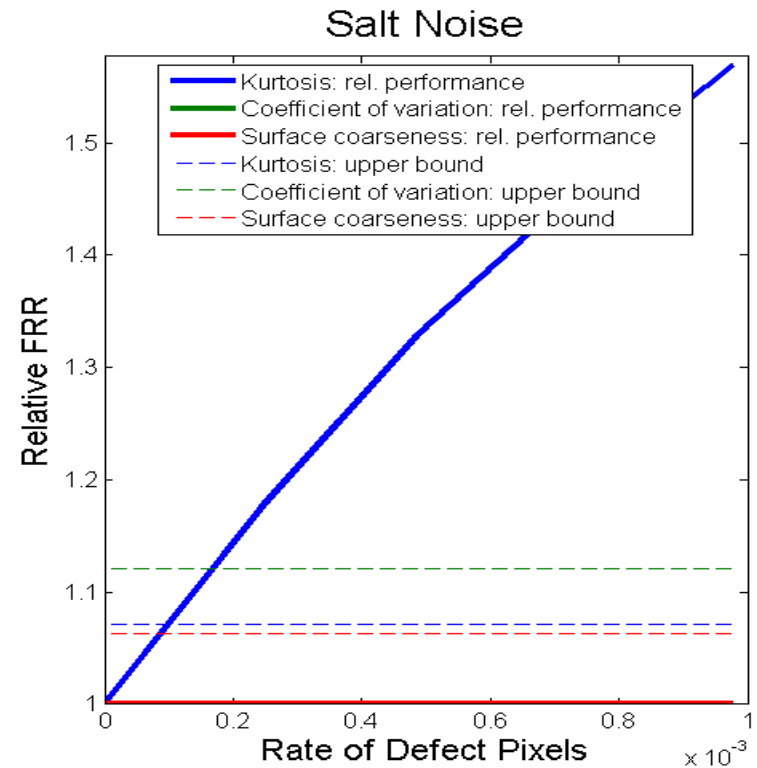
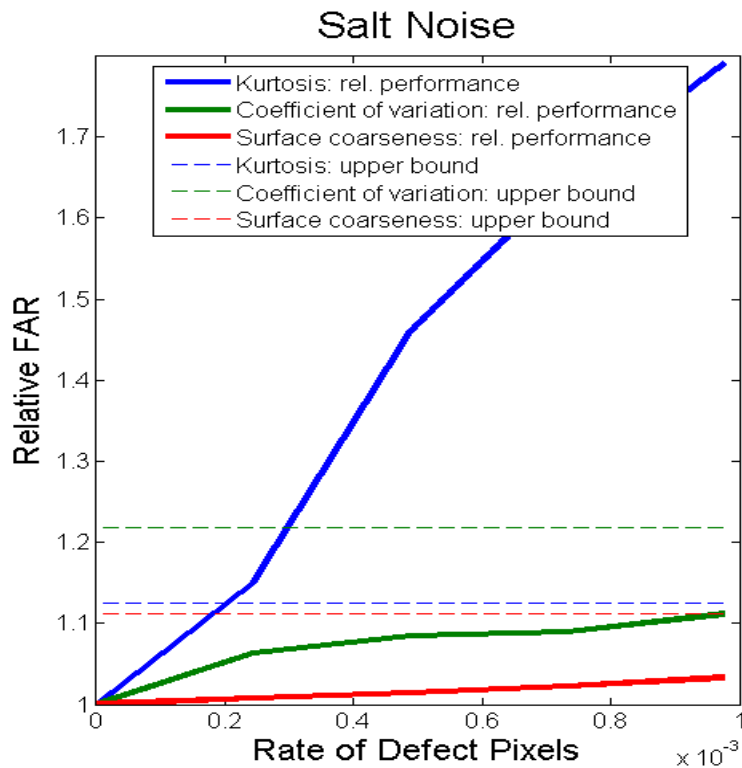
Example – Some Results



► Parameter: Rate of defect pixels

► “Pepper Noise”: All seem to be robust.

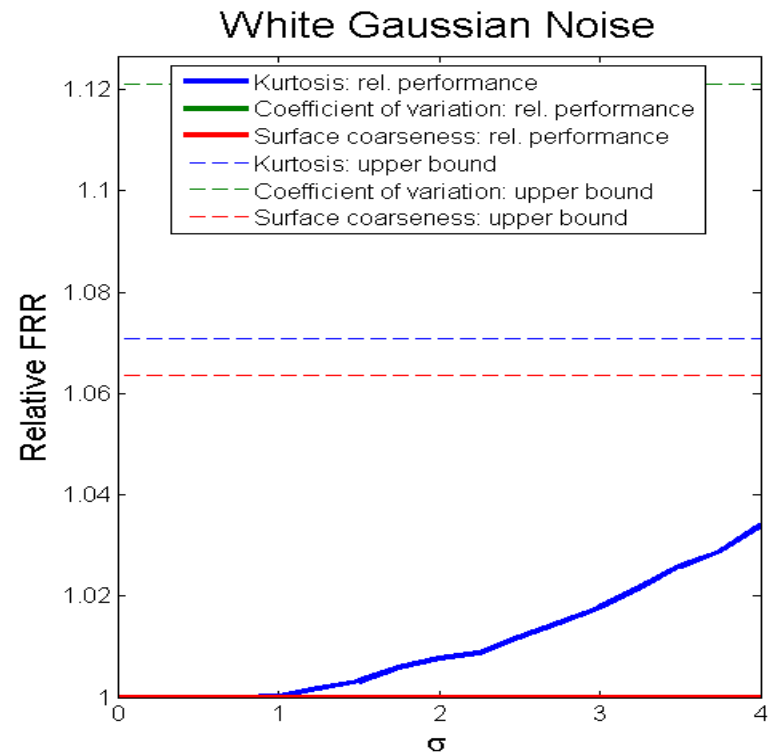
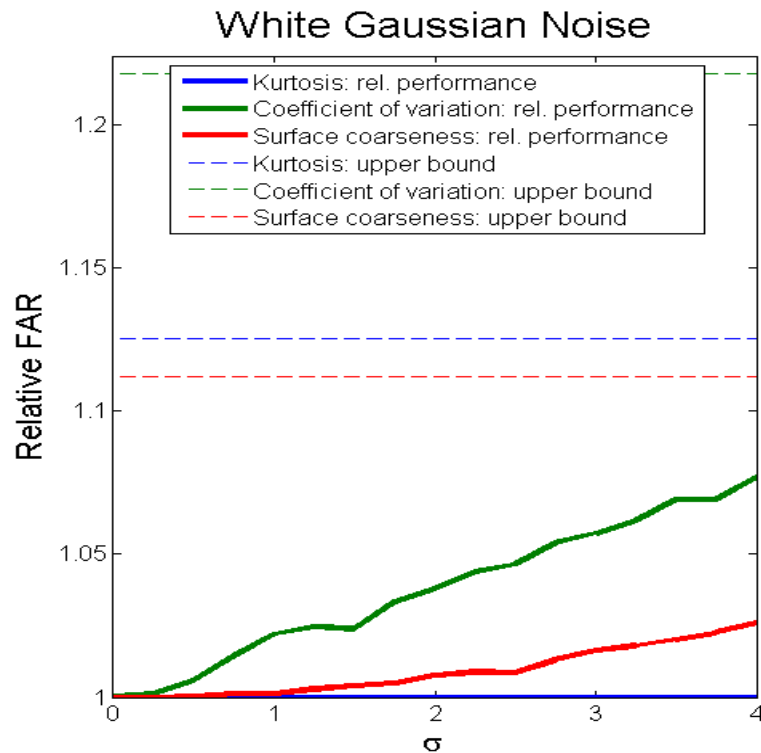
Example – Some Results



► Parameter: Rate of defect pixels

► “Salt Noise”: Kurtosis is not robust.

Example – Some Results

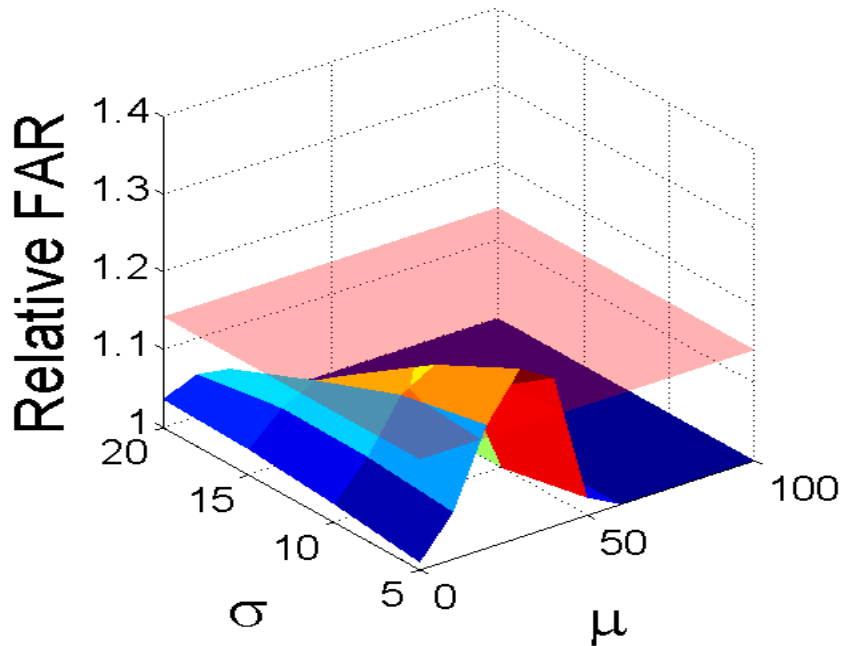


▶ Parameter range [0 4]

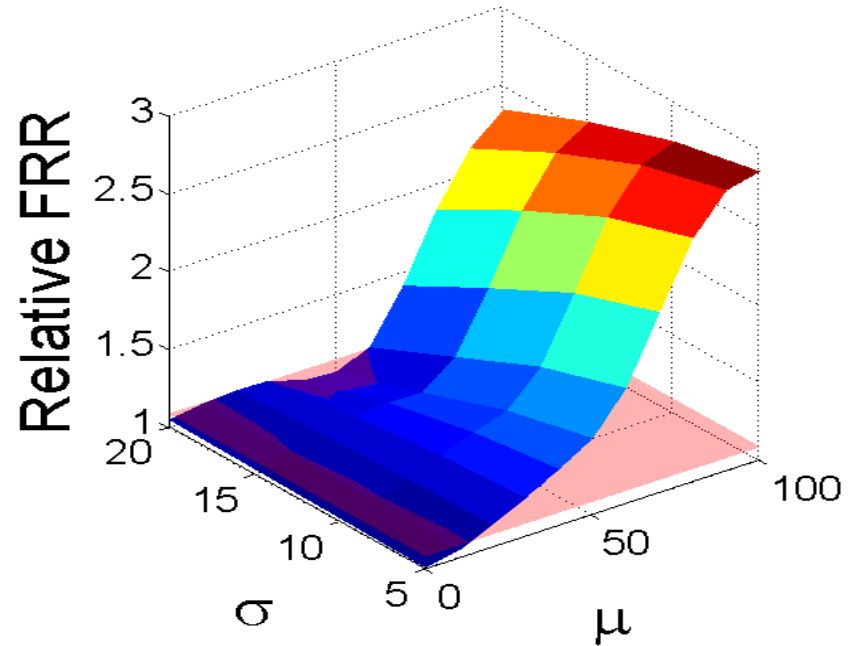
▶ All extractors seem to be robust.

Example – Some Results

Latents

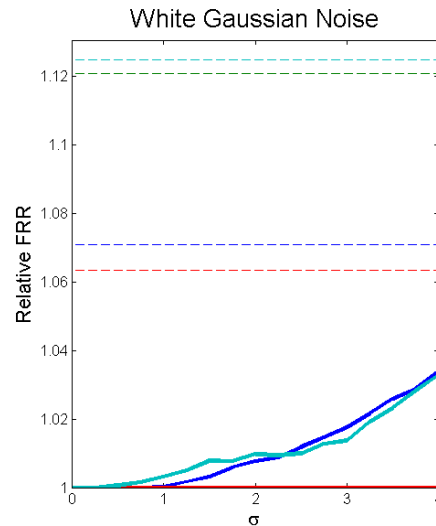
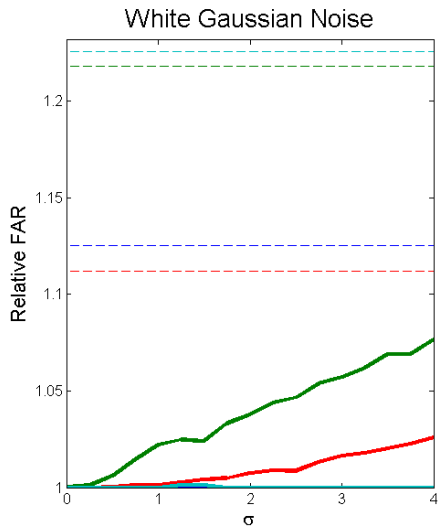
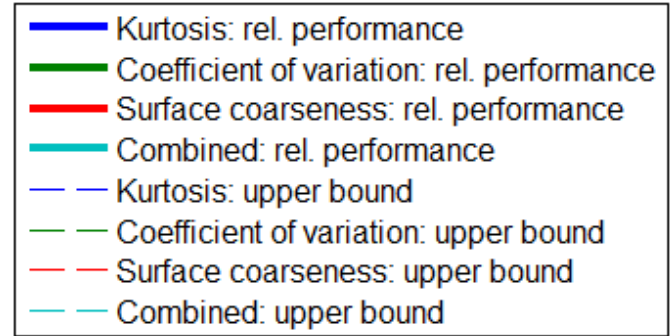
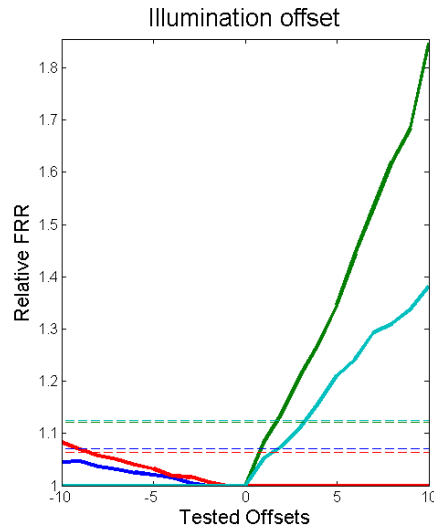
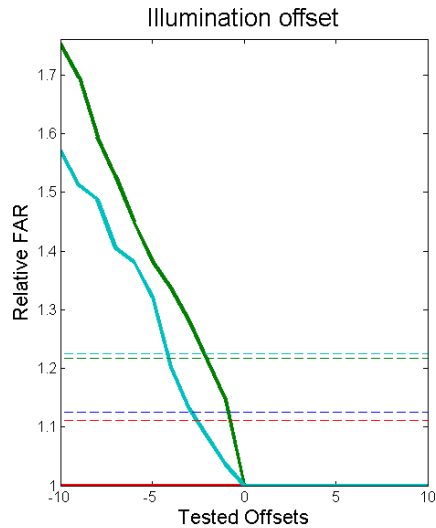


Latents



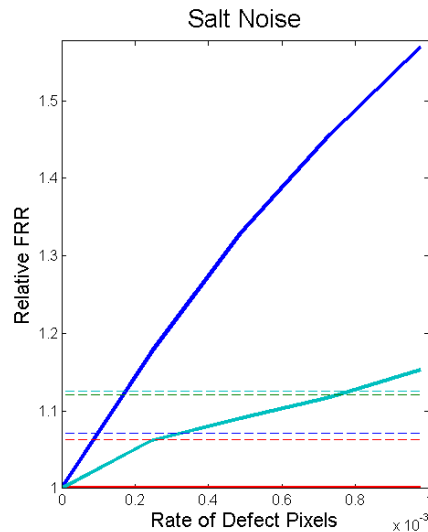
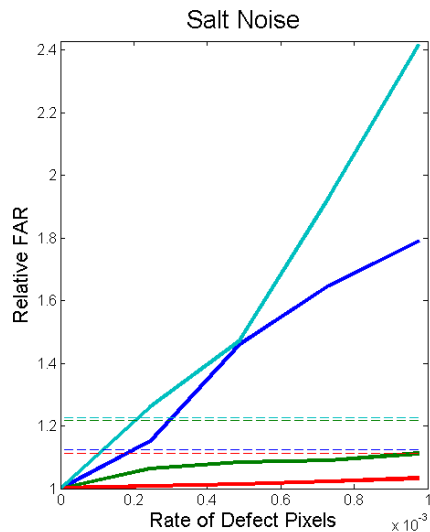
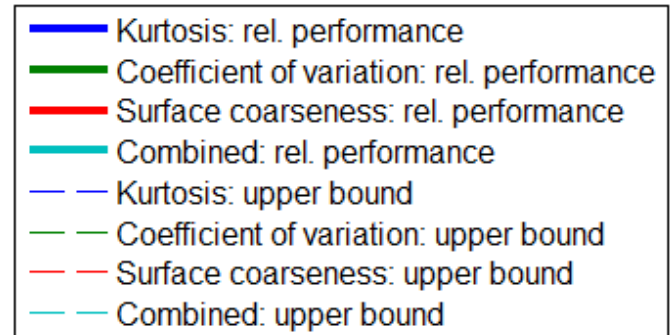
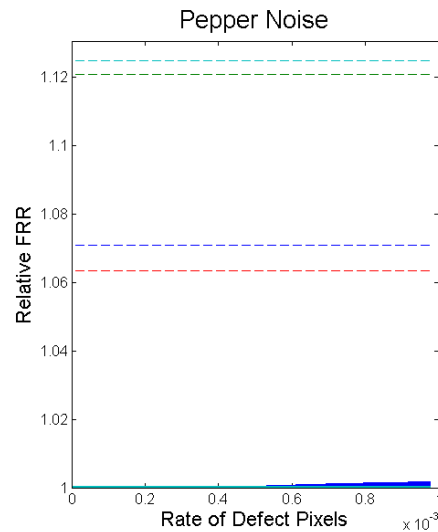
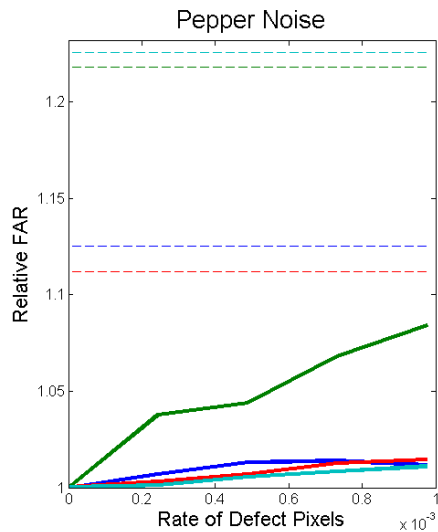
- ▶ Extractor: Kurtosis
- ▶ Parameter range: mean [0 100], deviation [0 15]
- ▶ Decay when mean gets close to common ridge mean.

Example – Some Results



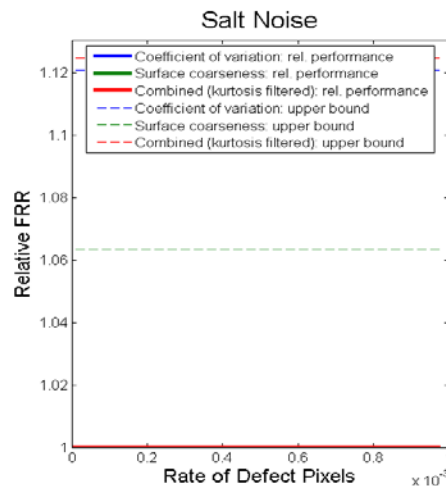
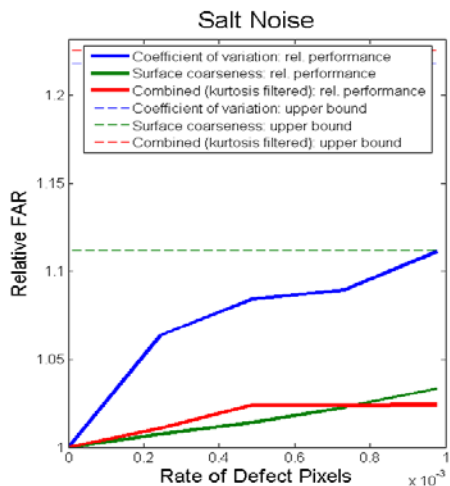
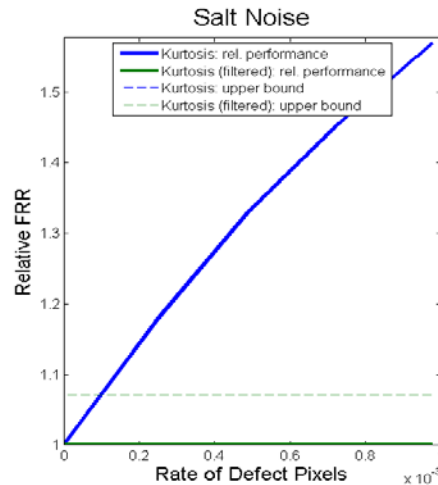
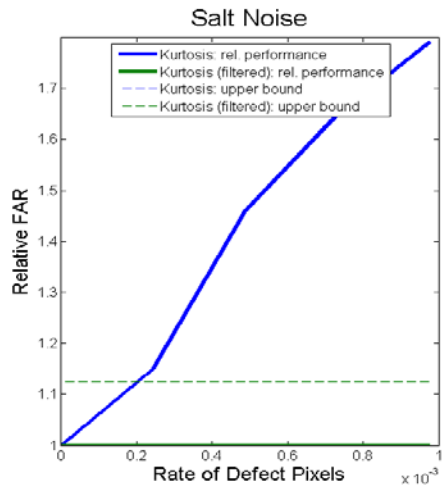
- ▶ Combination tested
- ▶ Single extractor may have impact on combined performance.

Example – Some Results



- ▶ Combination tested
- ▶ Single extractor may have impact on combined performance.

Example – Some results



- ▶ Revealed problem: Kurtosis vs. “Salt Noise”
- ▶ Solution: apply special filter
- ▶ Robustness gained
- ▶ Iterative process

Discussion

- ▶ Testing simulations...

 - ▶ shall not replace real data.

 - ▶ is cheaper.

 - ▶ is better than no testing.

- ▶ Curse of dimensionality: testing all combinations

- ▶ Classifiers influence needs inspection.

Summary

▶ Proposed method:

▶ Easy workflow

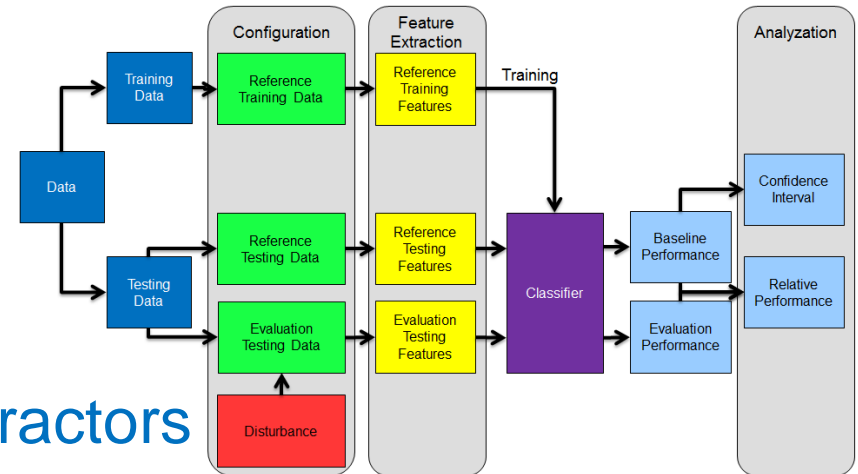
▶ Concentration on feature extractors

▶ Know and model the environment

▶ Tendencies observable in relative error rates.

▶ Not limited to fingerprint

▶ Weakest link may have serious impact



Outlook

- ▶ Revealed weaknesses in our own extractors
 - ▶ Robustness gained
 - ▶ Performance stabilized in the field
- ▶ Robustness over performance
- ▶ More disturbances to be modeled
 - ▶ More sophisticated models

On Testing the Robustness of Liveness Detection Feature Extractors for Fingerprint Live Scanners

Do you have any questions?
Thank you.

Patrick Schuch
DERMALOG Identification Systems GmbH
Hamburg, Germany
patrick.schuch@dermalog.com