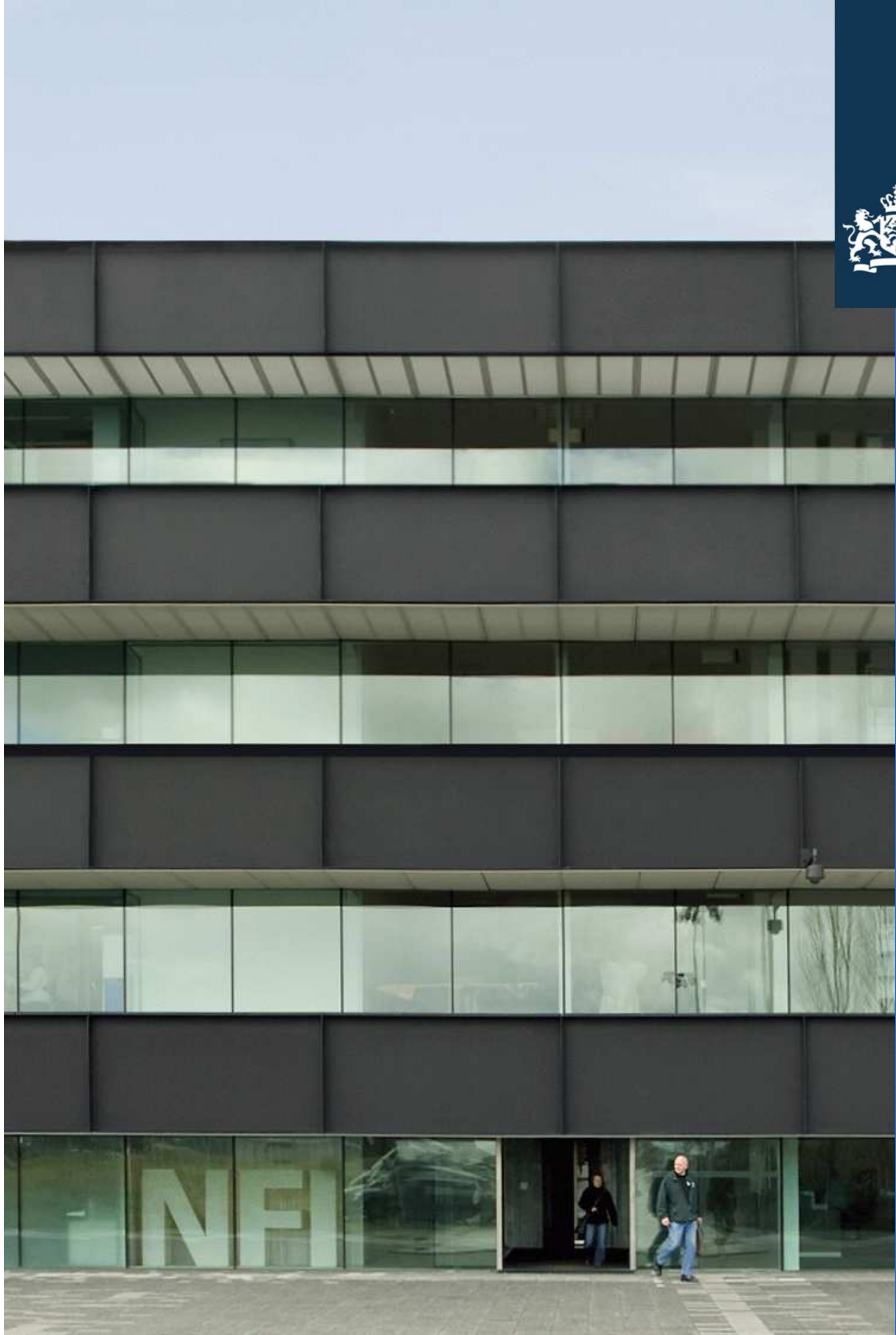


Netherlands Forensic Institute
Ministry of Security and Justice

Interlaboratory comparison for 3D analysis of bullet land engraved areas (LEA)

Martin Baiker-Sørensen[1]
Xiaoyu Alan Zheng[2]

[1] Team Firearms and Tools
Netherlands Forensic Institute

[2] National Institute of Standards and Technology

FBI, US:
Earl Gliem, Erich D. Smith, Jennifer L. Stephenson

Sensofar, Spain:
Cristina Cadevall, Neus Vintró

John Jay College of Criminal Justice, US:
Yu Chia Chen, Brady Huang, En Tni Lin, Victor Lin, Nicholas Petraco

NIST, US:
Xiaoyu Alan Zheng

Cadre Forensics, US:
Ryan Lilien, Eric Meschke, Chris Radford

Alabama Department of Forensic Sciences, US:
Adam Grooms, Ngo Phong

Laboratory Imaging, Czech Republic:
Tomáš Staněk, Martin Sábl

m.baiker@nfi.nl

Disclaimer:

a) *Neither NFI nor NIST intend to promote any particular acquisition method and/or vendor.*

b) *All measurements presented were taken between 2016 and 2020 and reflect the performance of the acquisition devices back then.*

m.baiker@nfi.nl

Oblique light enables to represent a surface as light shadow pattern

m.baiker@nfi.nl

Why 3D topography?

- *Oblique light illumination is dependent on the light angle and information is lost in shadow areas*

➢ *2D imaging does not yield true depth information*

m.baiker@nfi.nl

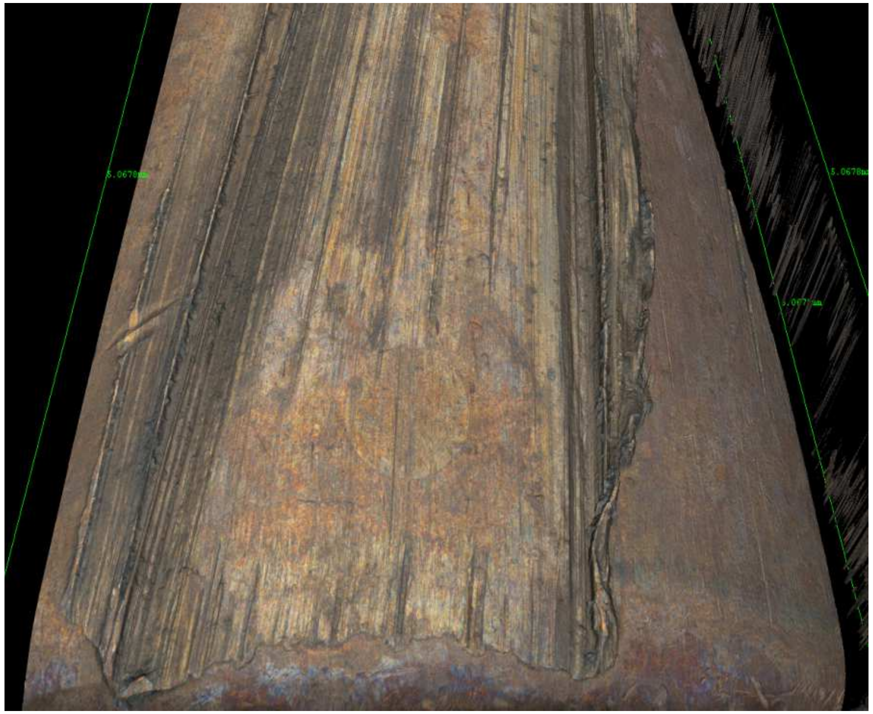
High magnification means a short depth of field

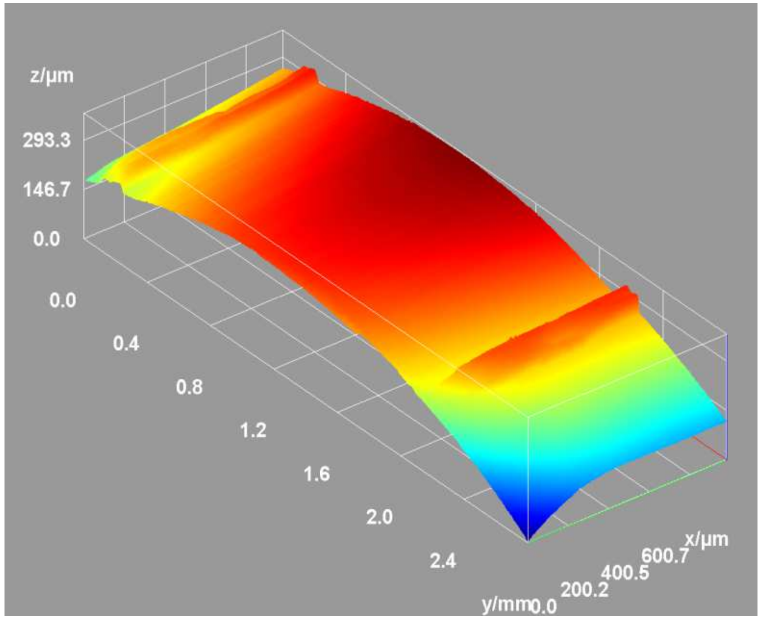
Why 3D topography?

- *Oblique light illumination is dependent on the light angle and information is lost in shadow areas*
  - ➢ *2D imaging does not yield true depth information*
- *High magnification means short depth of field and limited field size*
  - ➢ *2D does not provide high resolution images with all areas in focus*

m.baiker@nfi.nl

Surface topography yields true depth information and is more objective





[Zheng et al. 2014]

m.baiker@nfi.nl

Topographic microscopy

- *True depth information and no information loss*
- *More objective data and less variability*

Problem

- *How reproducible are the 3D surface measurements in practice?*
- *To date, limited statistical data available*

m.baiker@nfi.nl

Topographic microscopy

- *True depth information and no information loss*
- *More objective data and less variability*

Problem

- *How reproducible are the 3D surface measurements in practice?*
- *To date, limited statistical data available*

Goal

- *Study the reproducibility of results with data from different labs*

m.baiker@nfi.nl

Research questions

What are the *qualitative* and *quantitative* differences between bullet LEA data as well as known match and known non-match similarity scores if surface data of the *same* set of bullets is:

1.) Acquired at *different* labs, using different technologies, and compared with the *same* (NFI) algorithm?

2.) Acquired at *different* labs, using different technologies, and compared with *lab specific* algorithms (All methods using their own data)?

m.baiker@nfi.nl

Experimental setup to acquire bullet data

- *10 consecutively manufactured barrels of a Beretta 92F/FS*

  (Courtesy of Scott McVeigh of the Prince George's County Police Dept.)
- *3 bullets each with Remington UMC FMJ Copper Jacket, 9 mm*
- *6 LEAs were measured, at the base of the bullet*

➢ *180 surface measurements!*

m.baiker@nfi.nl

Experimental setup to acquire bullet data

- *Eight labs in the US, Netherlands, Czech republic and Spain*
- *Six manufacturers and six acquisition techniques*
- *Operators with varying experience with 3D microscopy (from student interns to experienced operators)*
- *Acquisition parameters: sampling distance (resolution of the data), mark orientation*

m.baiker@nfi.nl

Various techniques were used for surface acquisition


Alicona IFM SL & G5
Focus variation

Nanofocus µsurf
Confocal

Sensofar S Neox & Plu Neox
Confocal or Focus variation


Cadre TopMatch-3D
Photometric stereo

Laboratory Imaging LUCIA BalScan
FV + Photometric stereo

EvoFinder 4X4
FV + Photometric stereo

m.baiker@nfi.nl

Various techniques were used for surface acquisition

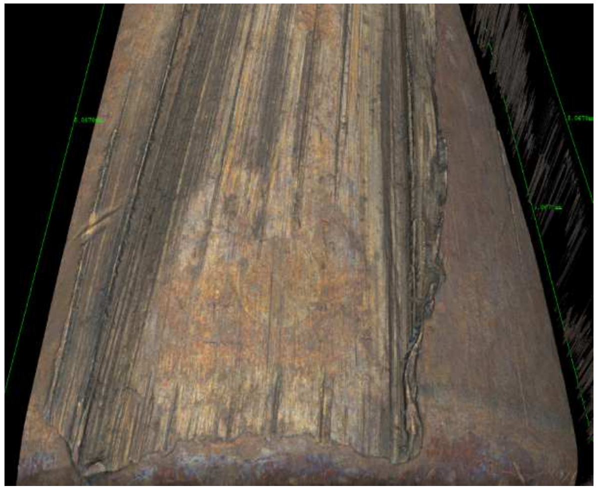
Lab	Model	Technology	Lateral sampling	Scanning time (LEA)	Operator
Lab A	Brand 1, Model I	Focus variation	501 nm	≈ 1 min	Experienced operator
Lab B	Brand 2, Model I	Confocal	645 nm	≈ 2 min	Experienced operator
Lab C1	Brand 3, Model I	Confocal	1560 nm	≈ 3.5 min	Experienced operator
Lab C2	Brand 4, Model I	Focus variation + Photometric stereo	3570 nm	≈ 3 s (10 s all lands)(*)	Experienced operator
Lab D	Brand 1, Model II	Focus variation	520 nm	≈ 6 min	Student interns
Lab E1	Brand 2, Model I	Continuous confocal	1300 nm	≈ 1.5 min	Student interns
Lab E2	Brand 2, Model I	Focus variation	1300 nm	≈ 1.5 min	Student interns
Lab F	Brand 2, Model II	Confocal	830 nm	≈ 2.5 min	Experienced operator
Lab G	Brand 5, Model I	Focus variation + Photometric stereo	3070 nm	≈ 18 s (210 s all lands)(*)	Experienced operator
Lab H	Brand 6, Model I	Photometric stereo	1440 nm	≈ 1 min	Student interns

(*) Systems acquire all six lands in one measurement

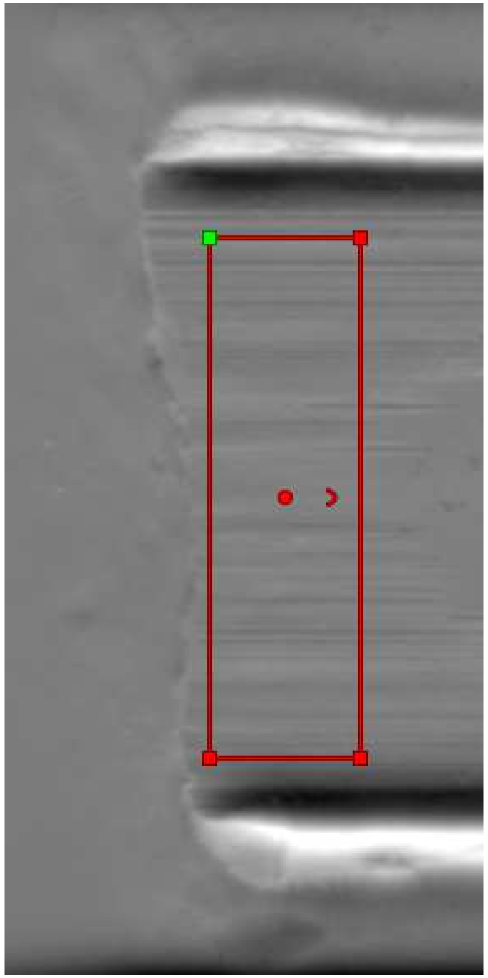
m.baiker@nfi.nl

Data post processing and comparison

- *At NFI, using Scratch*
- *Manual cropping of Region of Interest (ROI)*
- *All ROIs had equal size*



LEA surface rendering

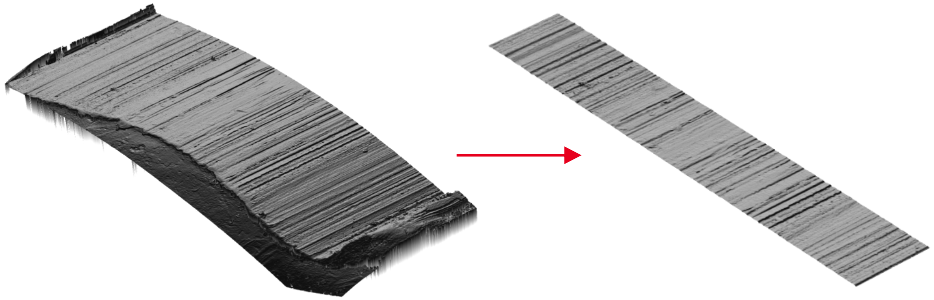


LEA height map

m.baiker@nfi.nl

Data post processing and comparison

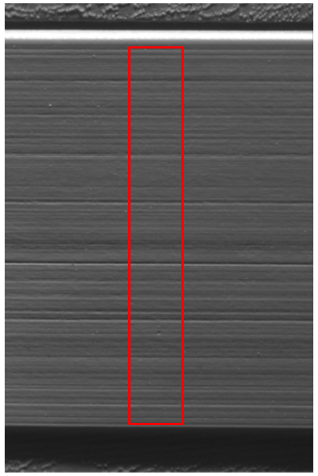
- *Shape and noise removal using Gaussian filters with cutoffs λ_{ch} = 250 μm, λ_{cl} = 5 μm (Lab C2 and G already flat)*

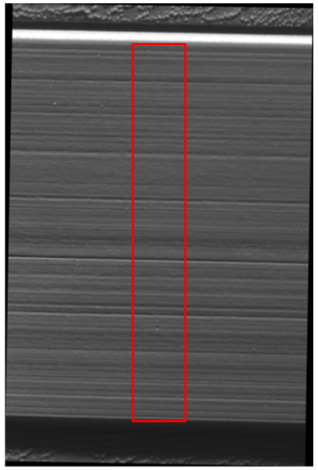


m.baiker@nfi.nl

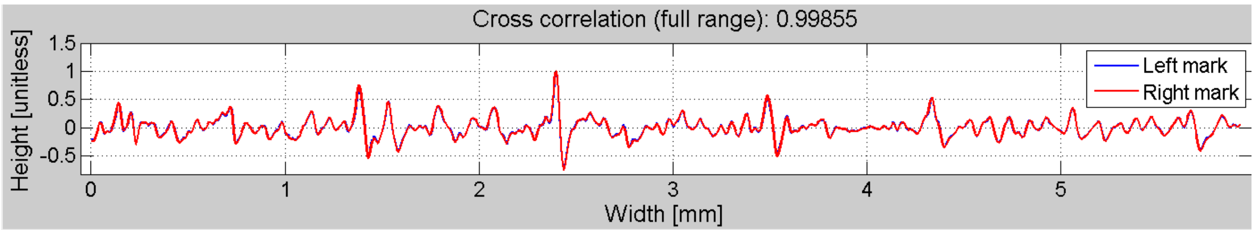
Mark rotation leads to less accurate profiles

0°

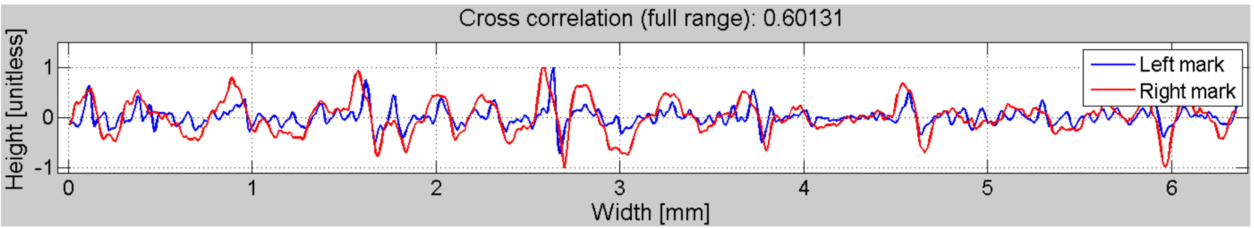




1°



With rotation correction



Without rotation correction

m.baiker@nfi.nl

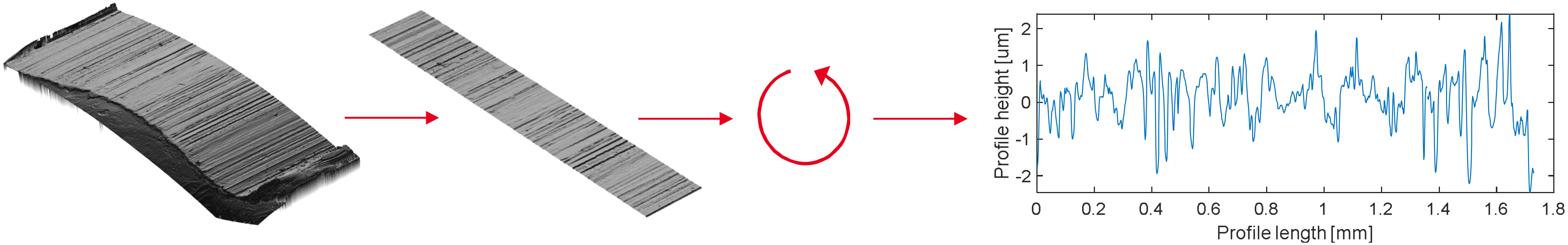
Data post processing and comparison

- *Shape and noise removal using Gaussian filters with cutoffs*

  λ_{ch} *= 250* μ*m,* λ_{cl} *= 5* μ*m*

- *Automated rotation correction*

- *Extracting profile by averaging*



m.baiker@nfi.nl

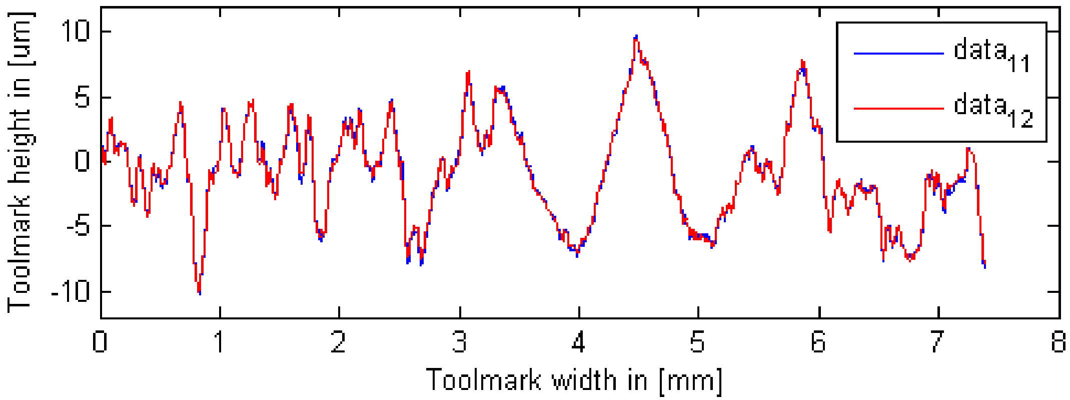
Data post processing and comparison

- *Shape and noise removal using Gaussian filters with cutoffs*
  λ_{ch} = 250 μm, λ_{cl} = 5 μm
- *Automated rotation correction*
- *Extracting profile by averaging*
- *Automated alignment using multi-scale registration with two degrees of freedom: translation and scaling (set to 1, for now…)*
- *Similarity metric: Cross-correlation*

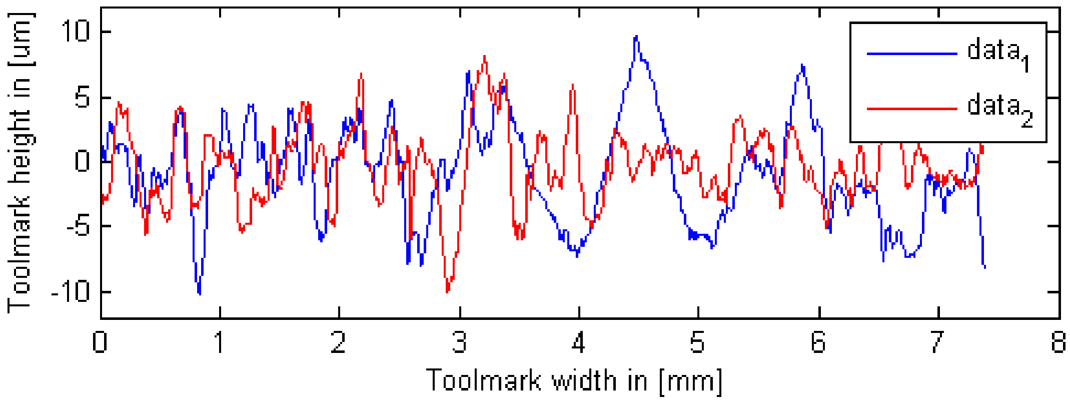
Details: [Baiker et al. 2014, Quantitative comparison of striated toolmarks]

m.baiker@nfi.nl

Similarity metric: Cross correlation



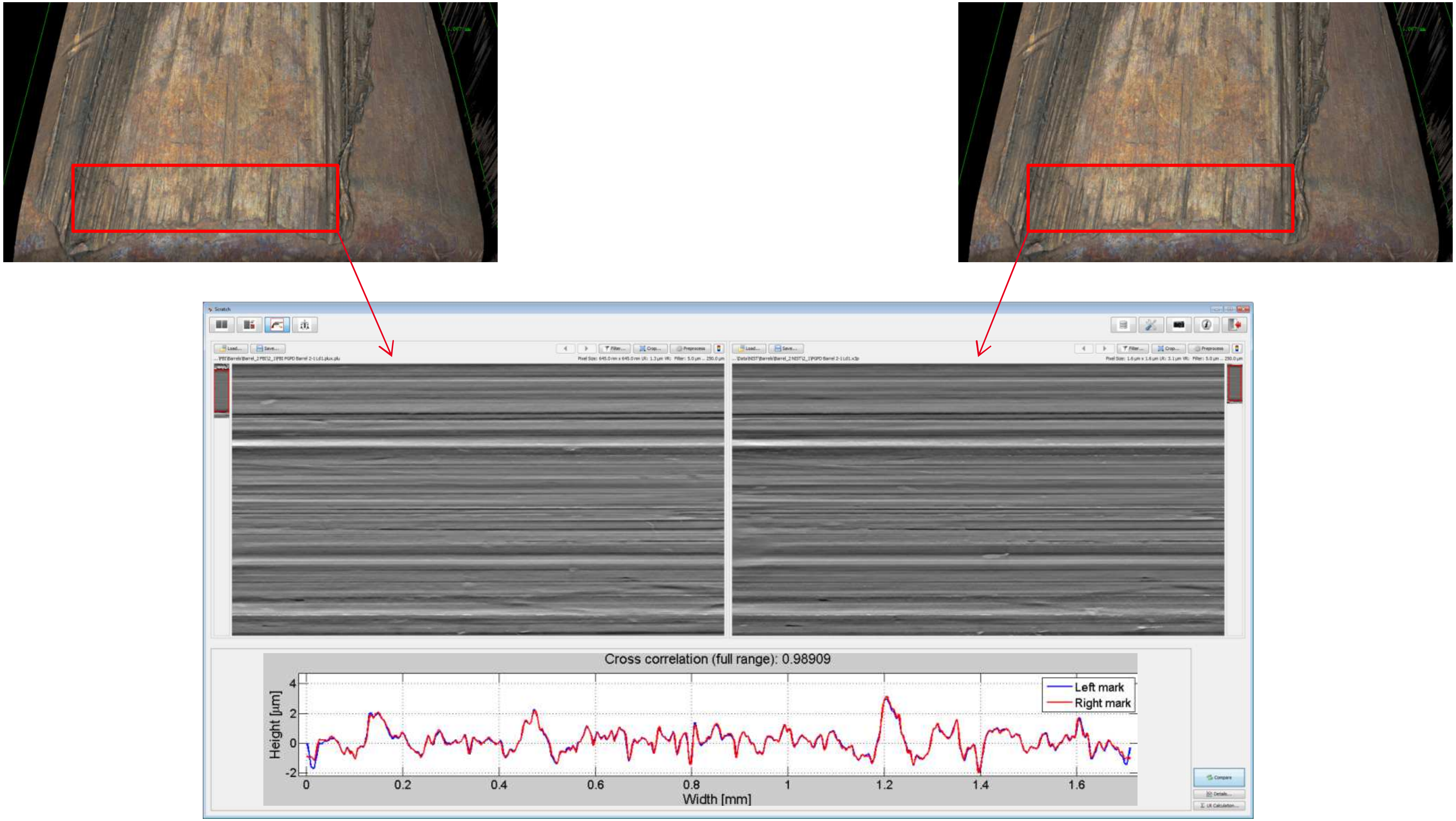
Known matches
(xcorr = 0.998)



Known non-matches
(xcorr = 0.2)

[Baiker et al., Quantitative comparison of striated marks, FSI, 2014]
[Baiker et al., Toolmark variability and quality…, FSI, 2015]

m.baiker@nfi.nl



m.baiker@nfi.nl

Qualitative and quantitative assessment of the data

- *All three repetitions were compared to each other, yielding 3 similarity scores for each barrel and each LEA*

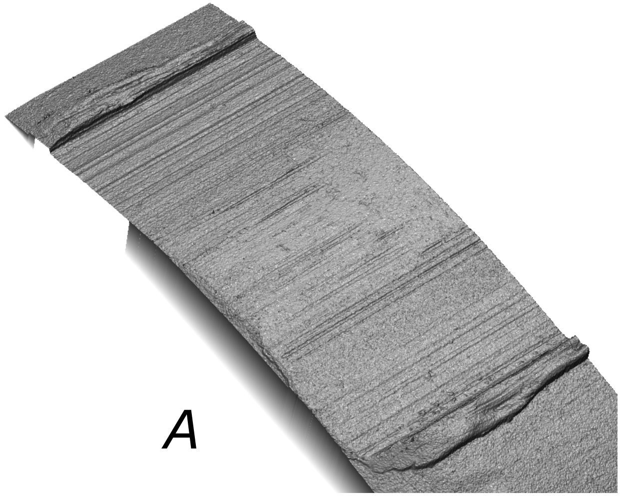
-> KM score distributions with 180 known matching scores

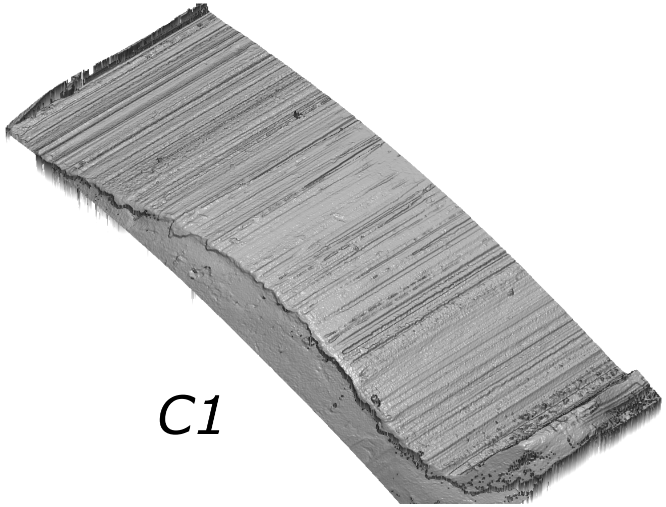
- *Each repetition of each LEA was compared to all other LEAs of the same barrel and one repetition of all LEAs of the other barrels*

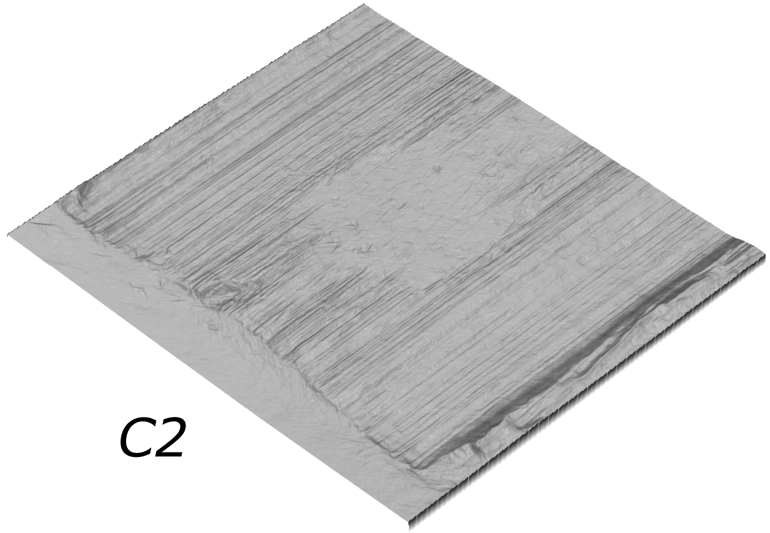
-> KNM score distributions with 10620 known non-matching scores

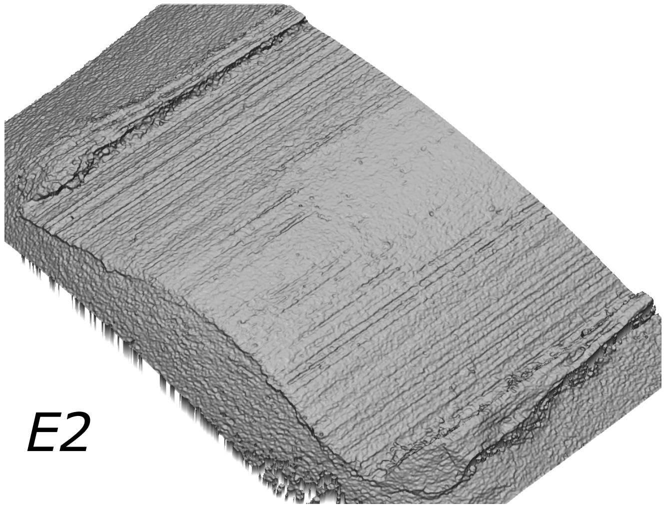
- *Some LEAs were discarded (incomplete, damaged, data missing)*
- *Testing for statistical significant differences using Kruskal-Wallis test, combined with Tukey-Kramer honest significant differences criterion*

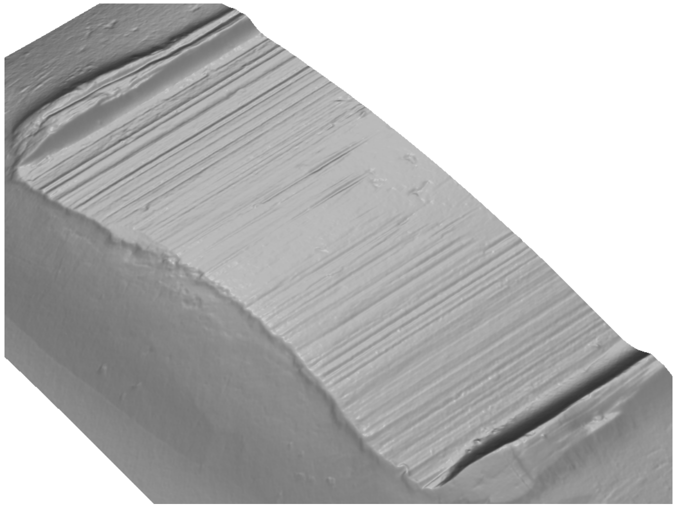
m.baiker@nfi.nl





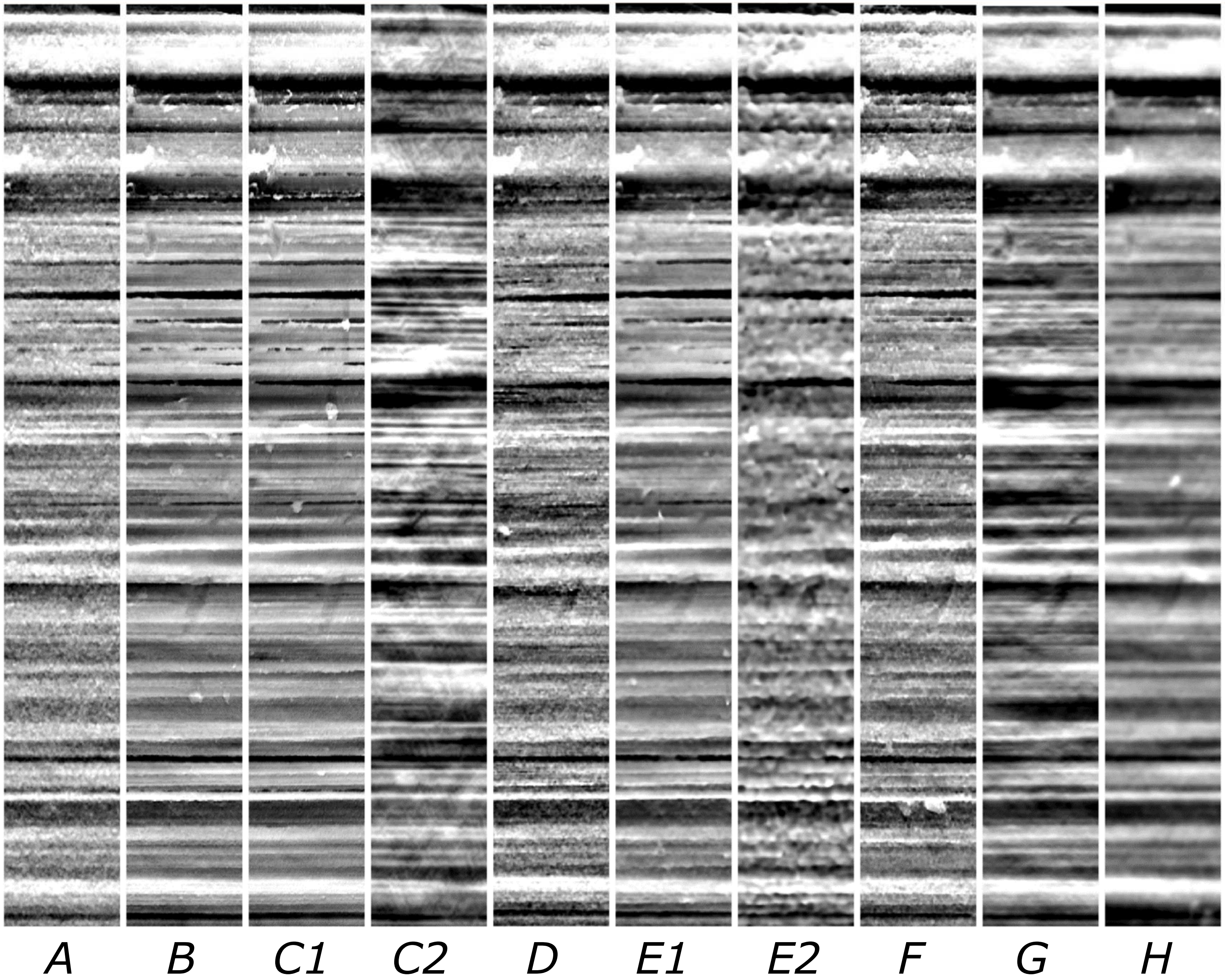






Lab	**Brand/Model** **Technology**
A	Brand 1, Model I *Focus Variation*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
E2	Brand 2, Model I *Focus Variation*
H	Brand 6, Model I *Photometric Stereo*

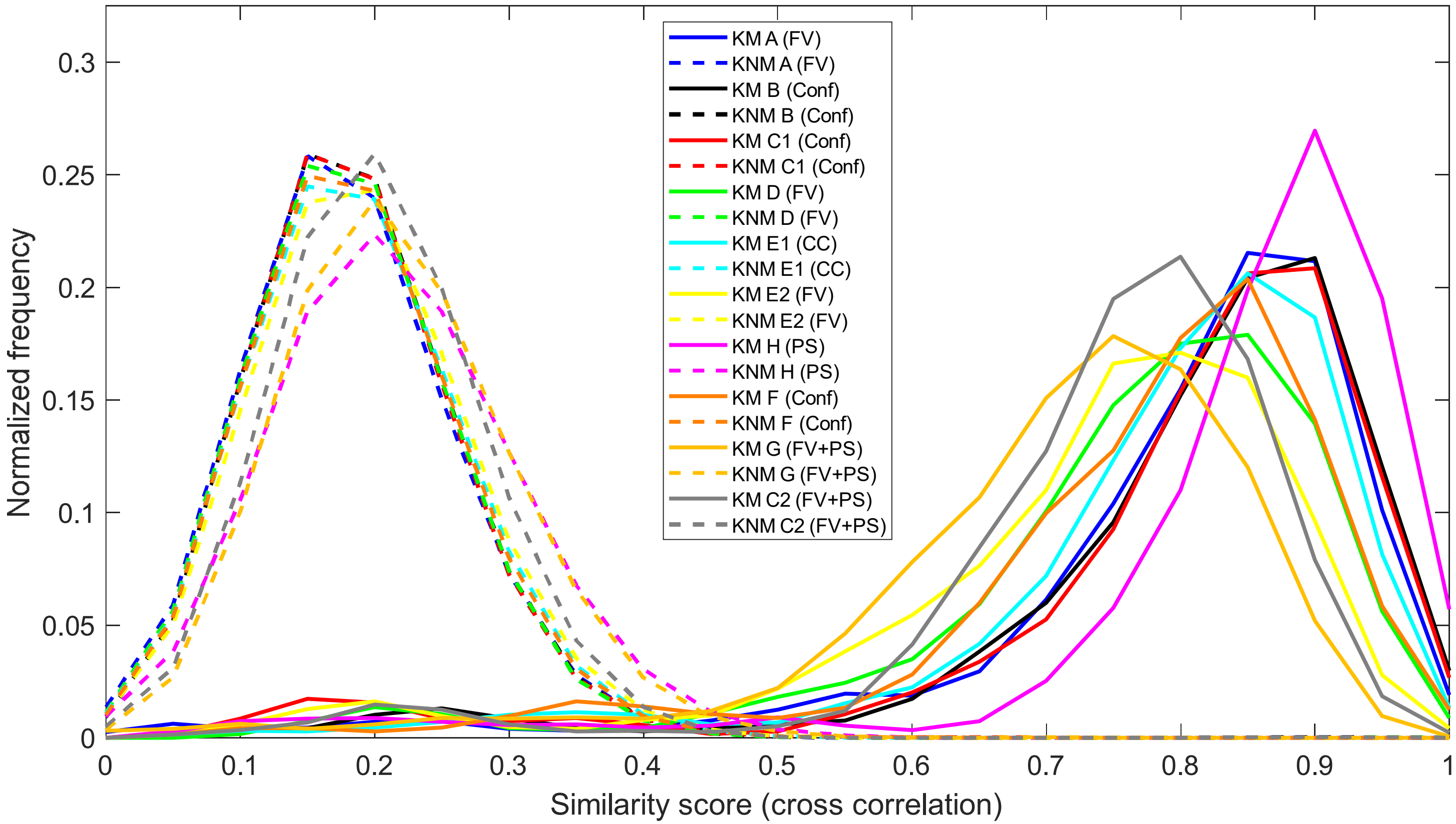
m.baiker@nfi.nl



Lab	Brand/Model Technology
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

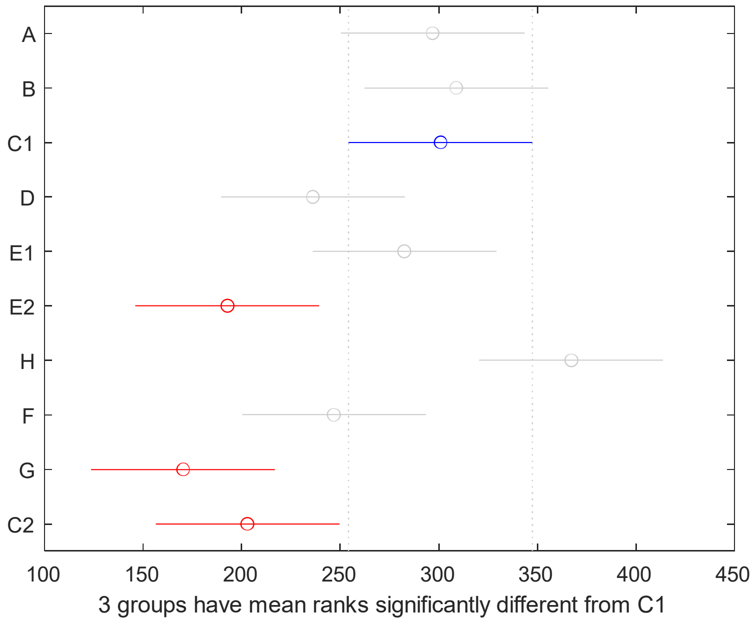
KM and KNM score distributions, all technologies



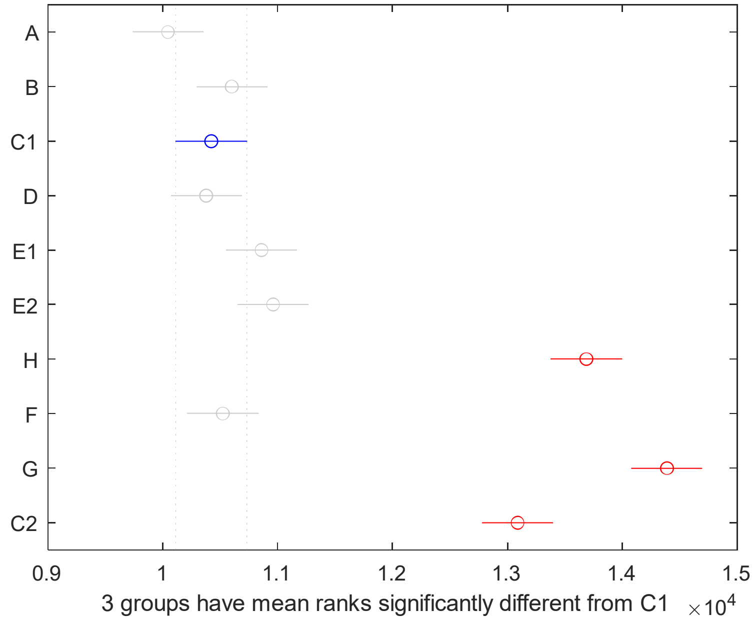
Lab	Brand/Model Technology
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

Testing for statistically significant differences



KM

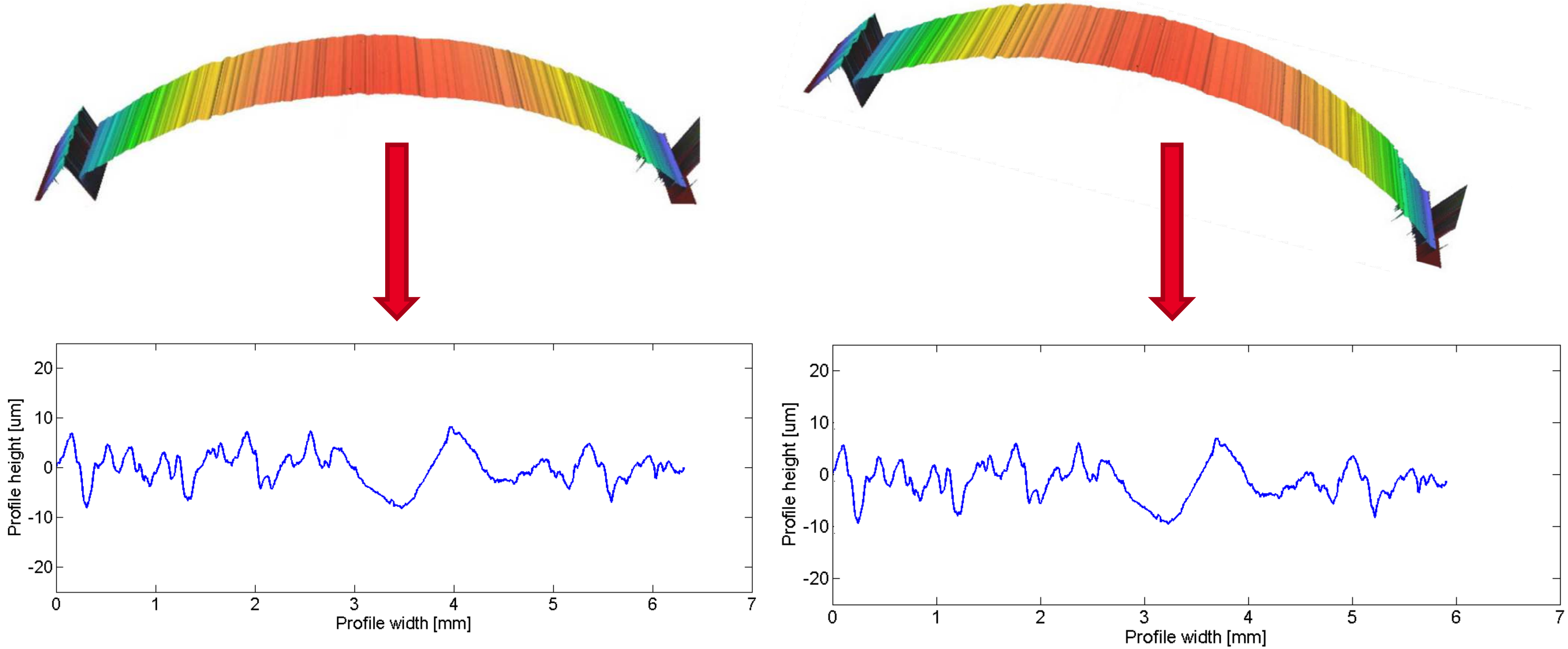


KNM

Lab	**Brand/Model Technology**
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

Axial bullet rotation causes (non-linear) compression



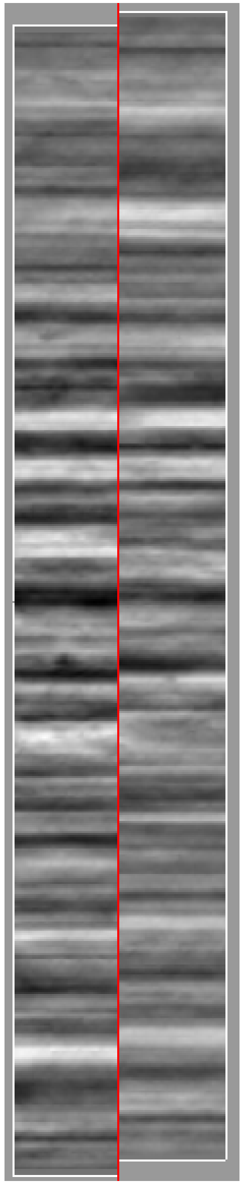
- ➢ *Correct for shape by applying 'unfolding'*

m.baiker@nfi.nl

Linear compression sometimes remains

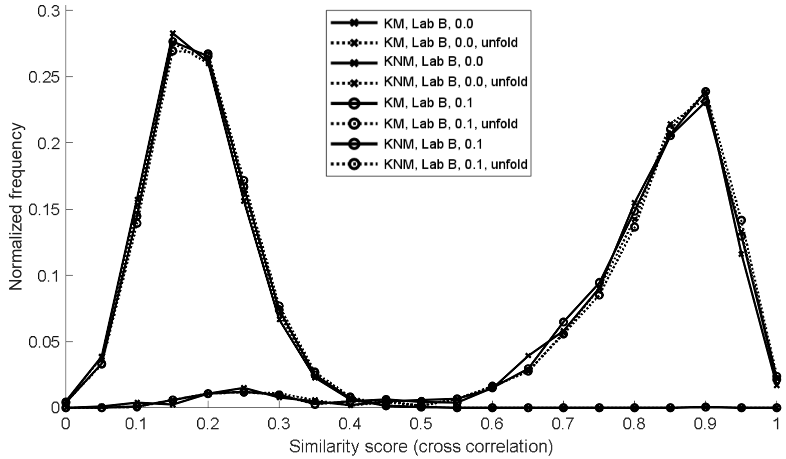
- *Still occurs sometimes, even after automated rotation*
- *Inherent to the measurement data*

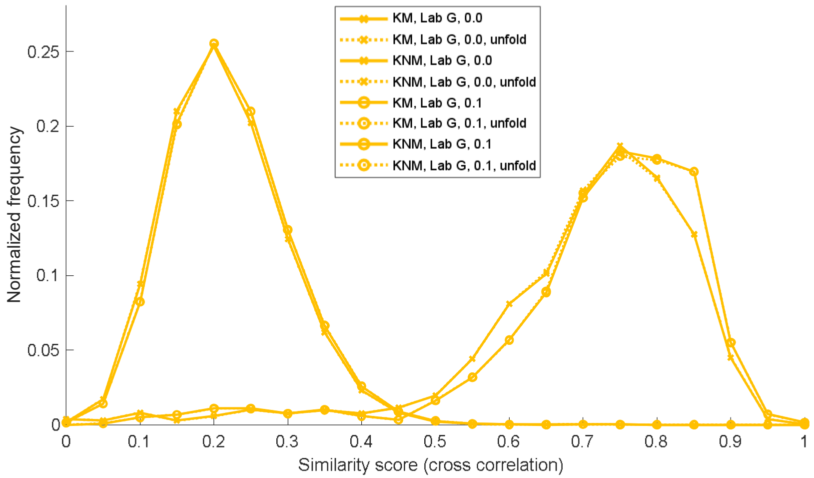
➢ *Correct for scaling as well*

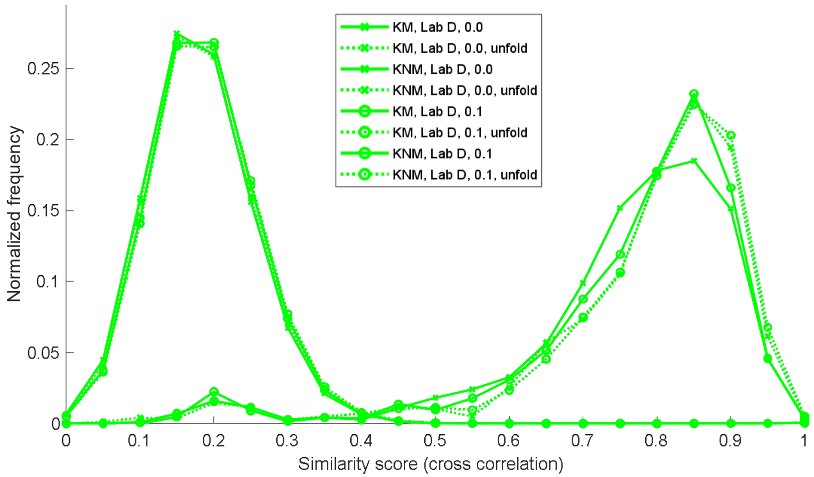


m.baiker@nfi.nl

Effect of automated scaling and 'shape unfolding'(*)



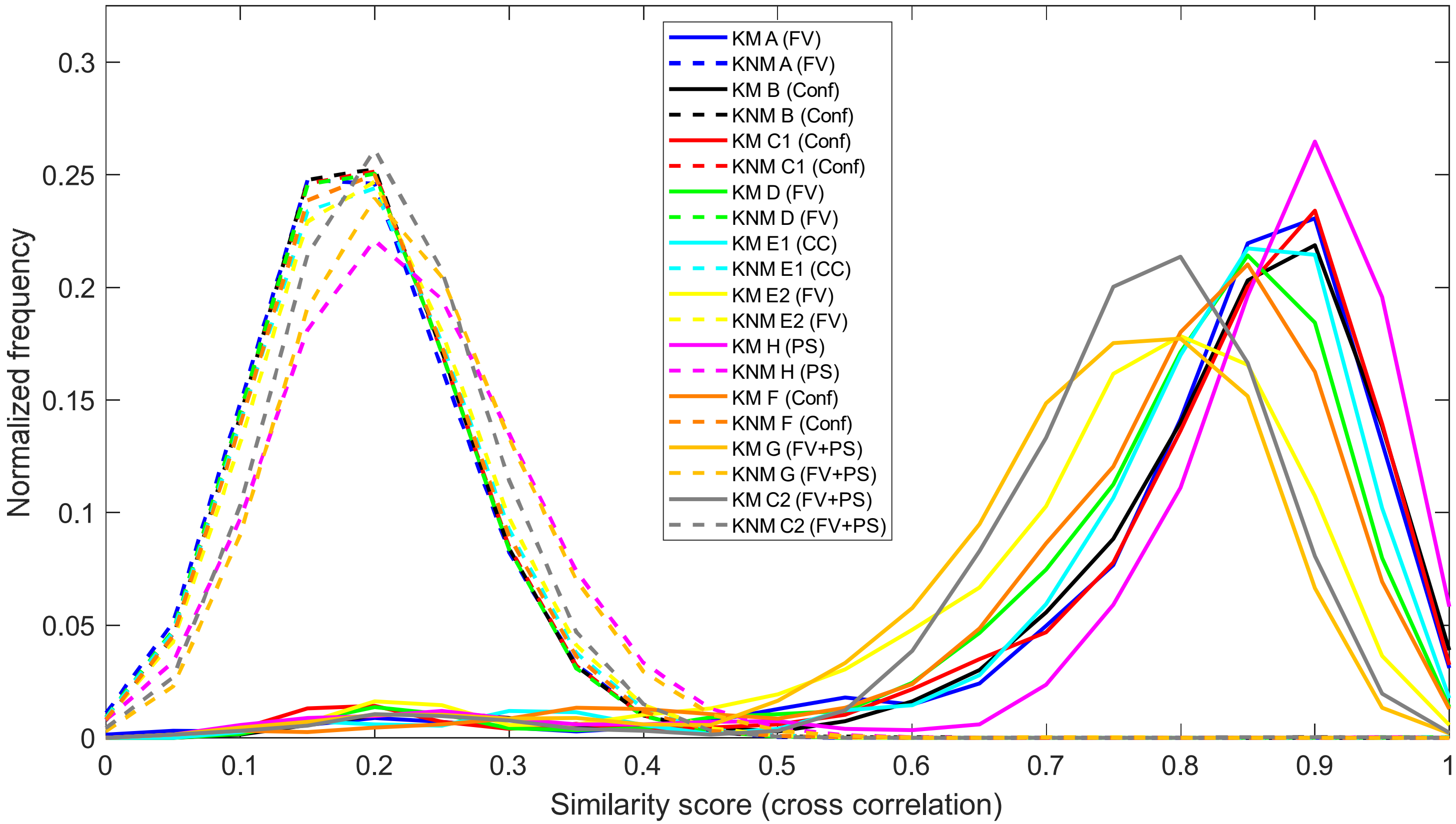




(*) Not applied on already flattened data

m.baiker@nfi.nl

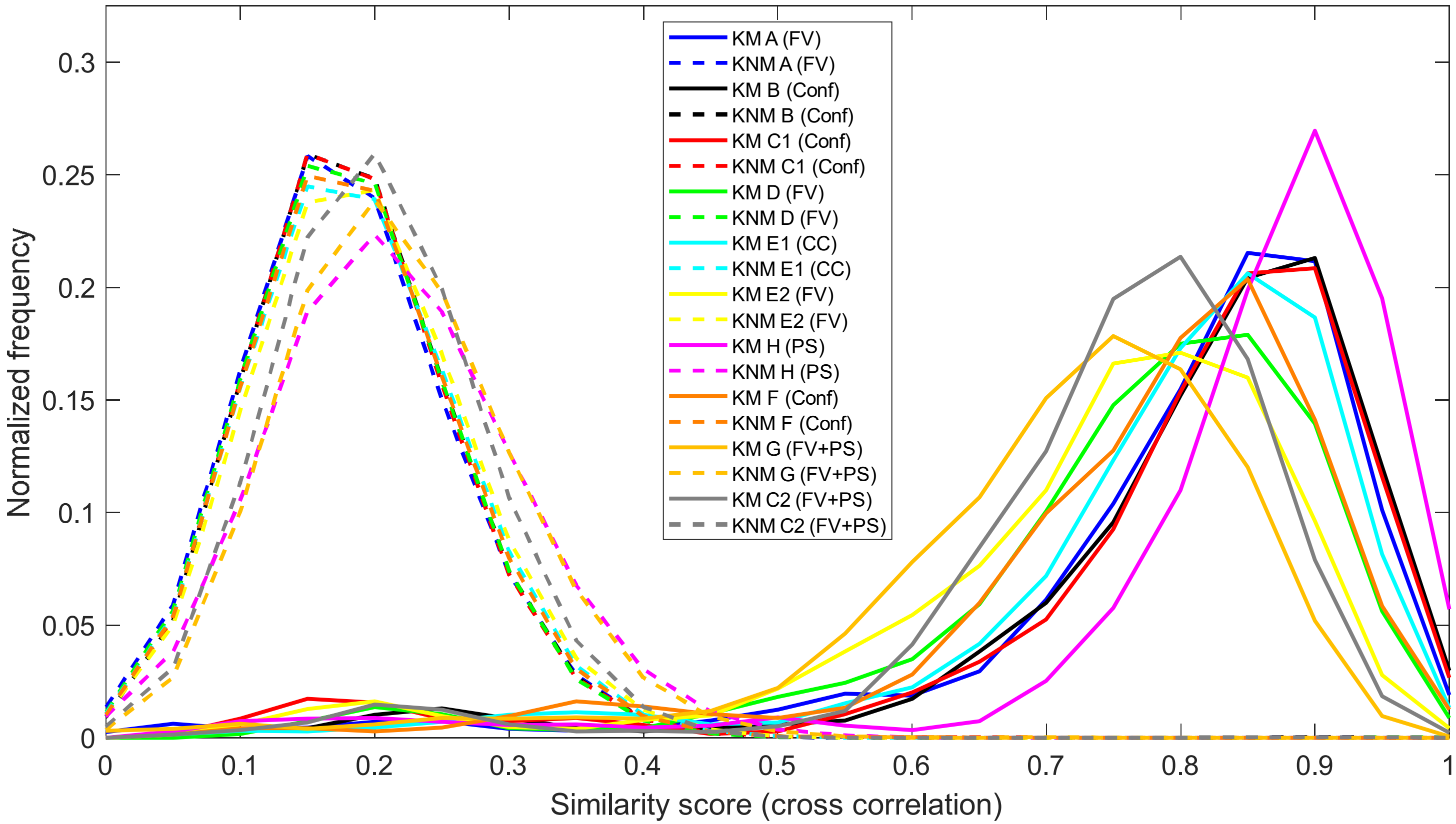
KM and KNM score distributions, with corrections



Lab	**Brand/Model Technology**
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

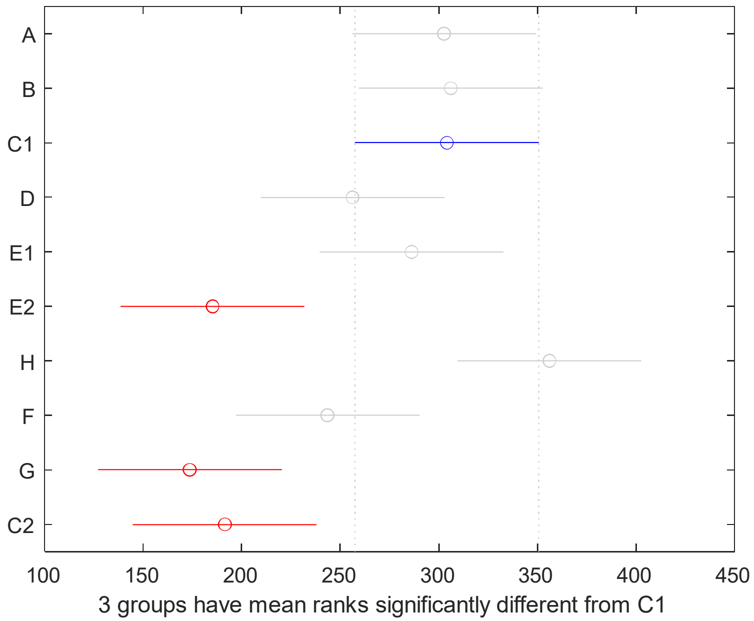
KM and KNM score distributions, no corrections



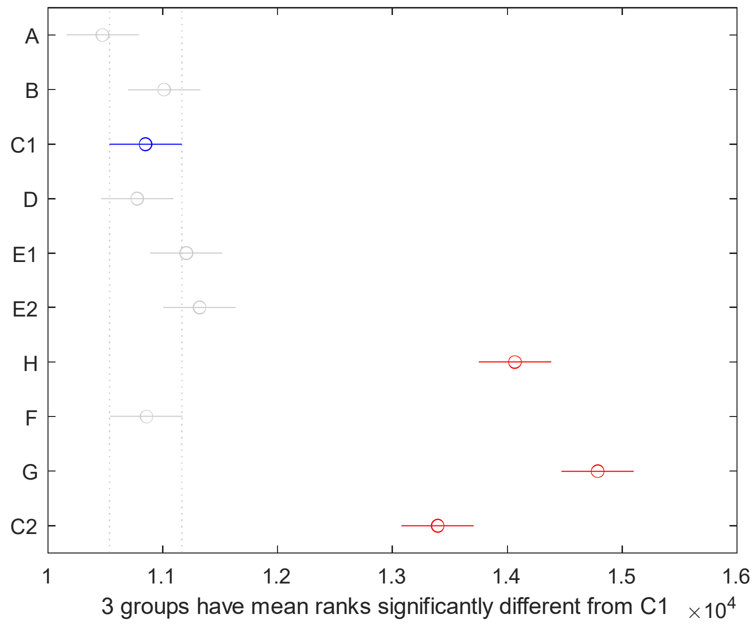
Lab	Brand/Model Technology
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

Testing for statistically significant differences



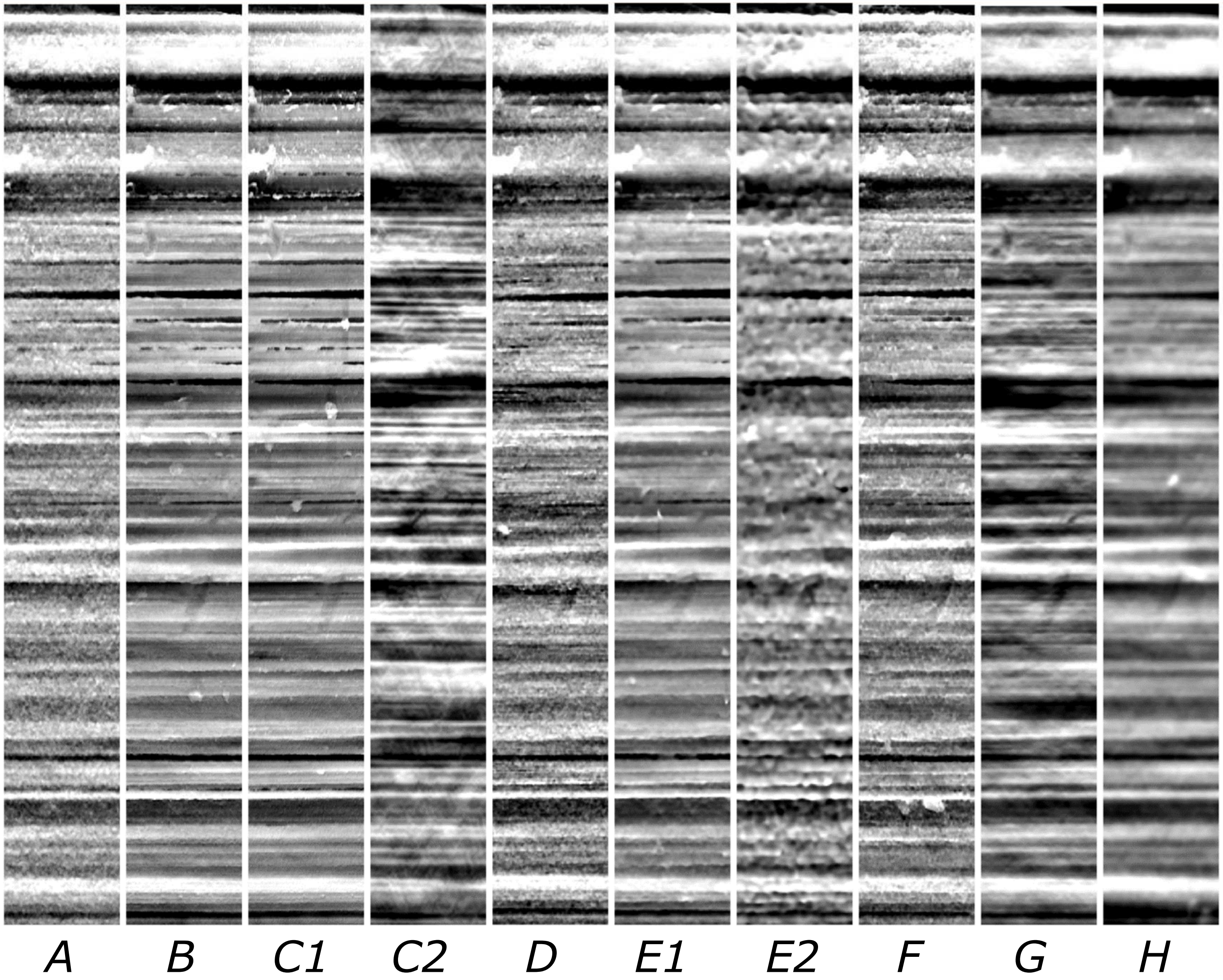
KM



KNM

Lab	**Brand/Model Technology**
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

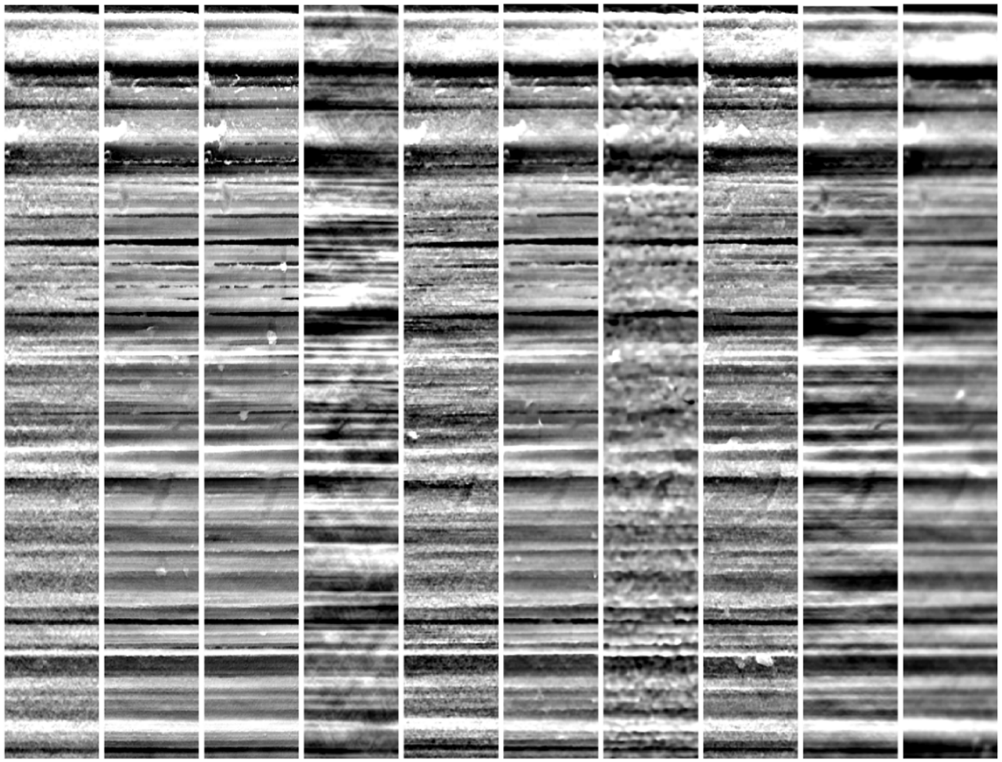
m.baiker@nfi.nl

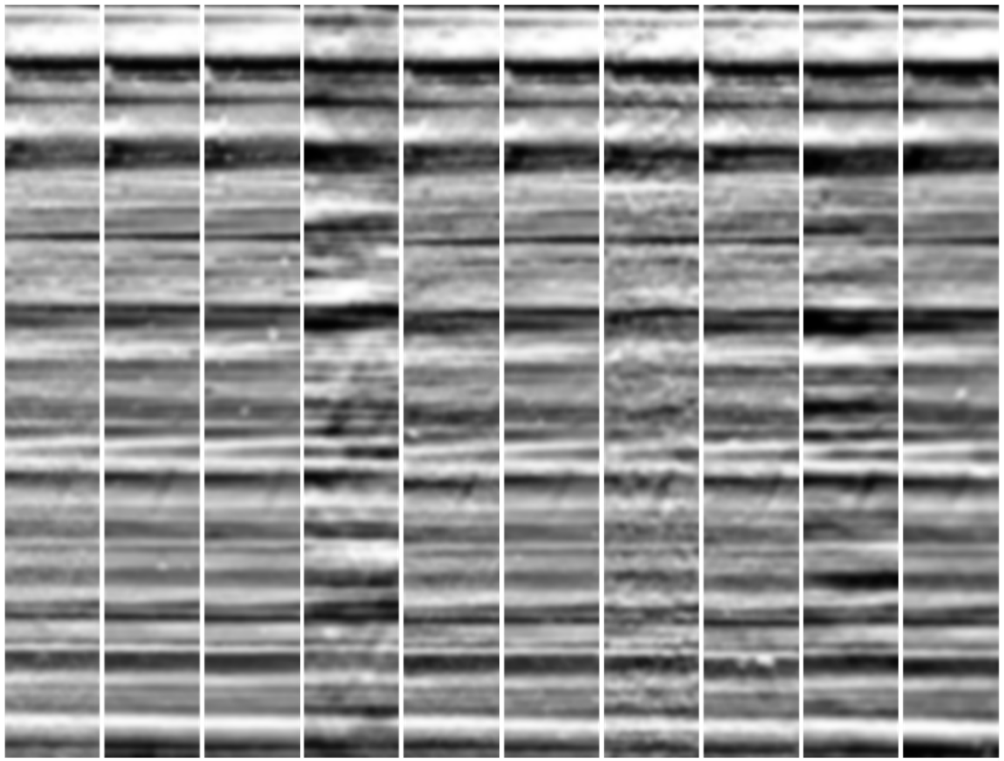


Lab	Brand/Model Technology
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

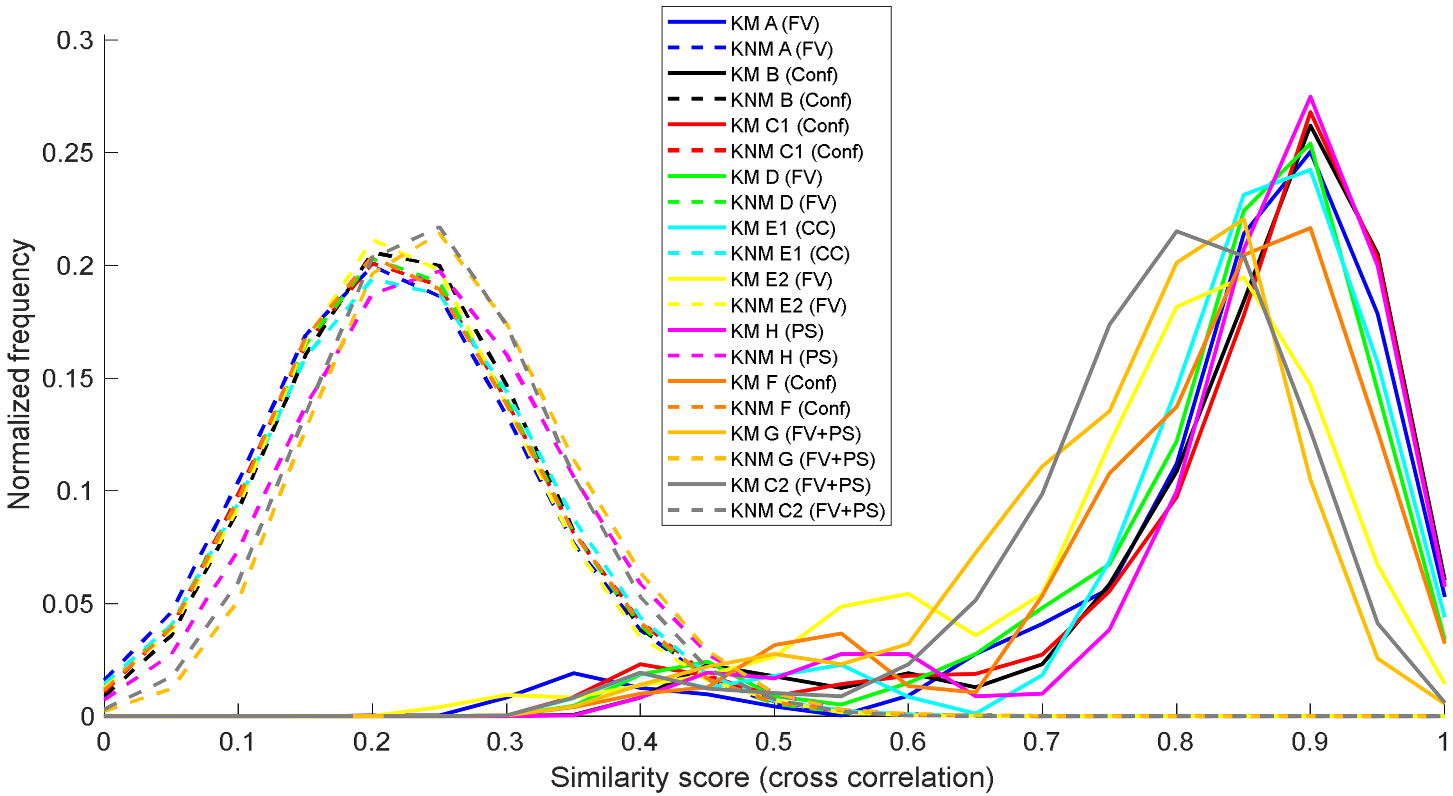
Effect of difference noise filter settings





m.baiker@nfi.nl

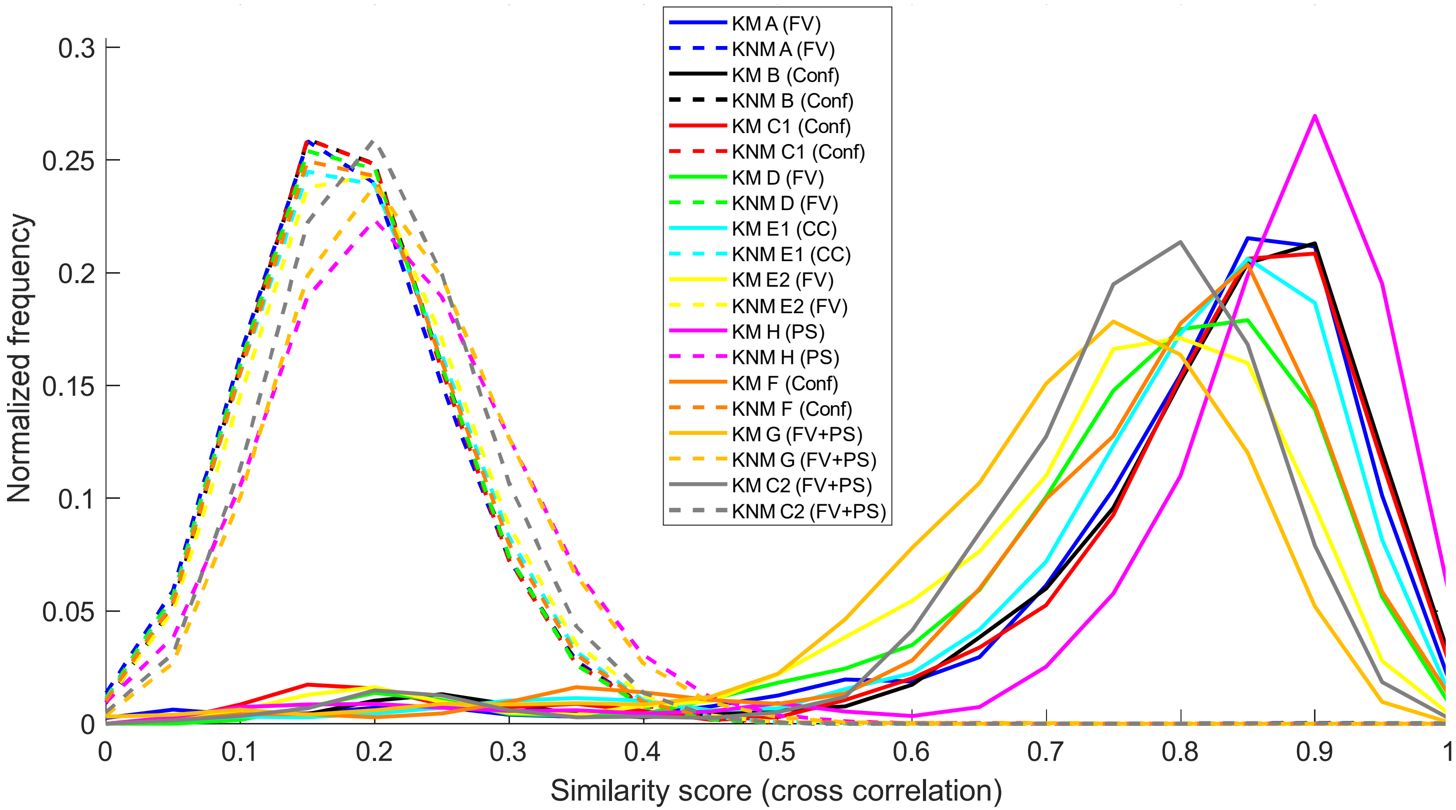
KM and KNM score distributions, noise filtered at 50 μm



Lab	**Brand/Model Technology**
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

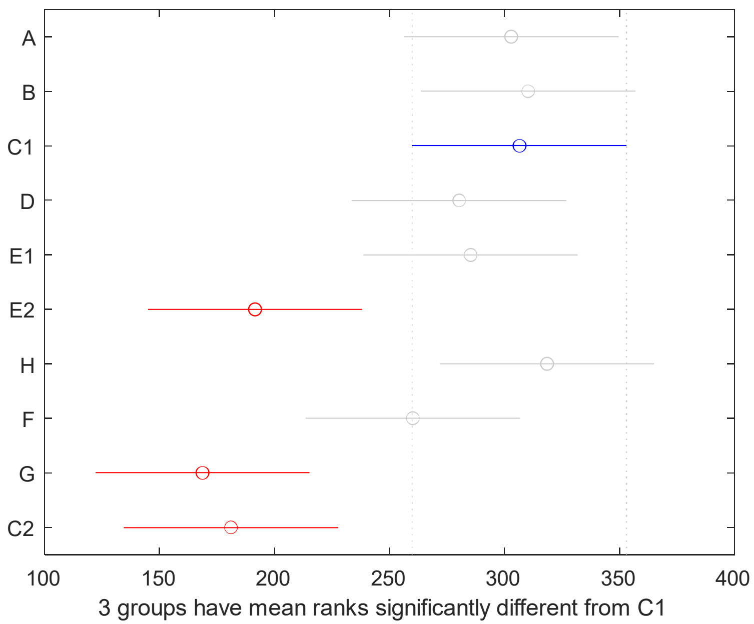
KM and KNM score distributions, all technologies



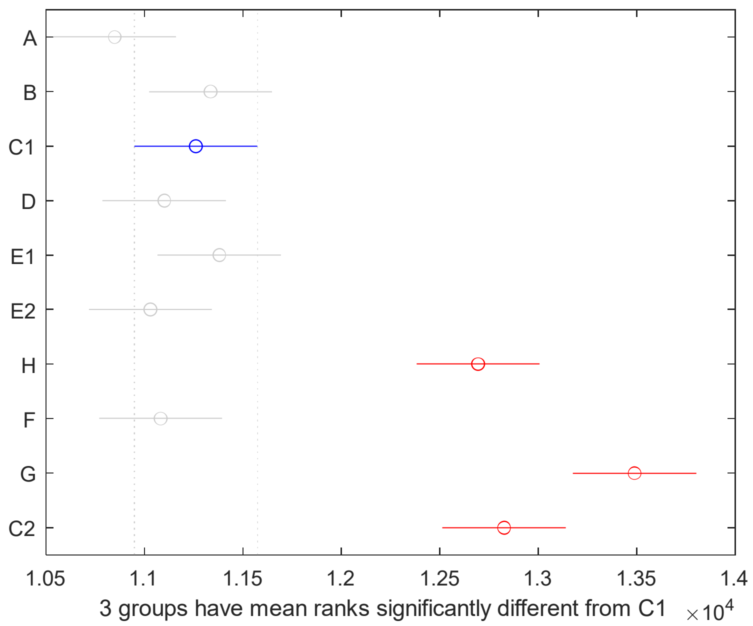
Lab	**Brand/Model Technology**
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

Testing for statistically significant differences



KM

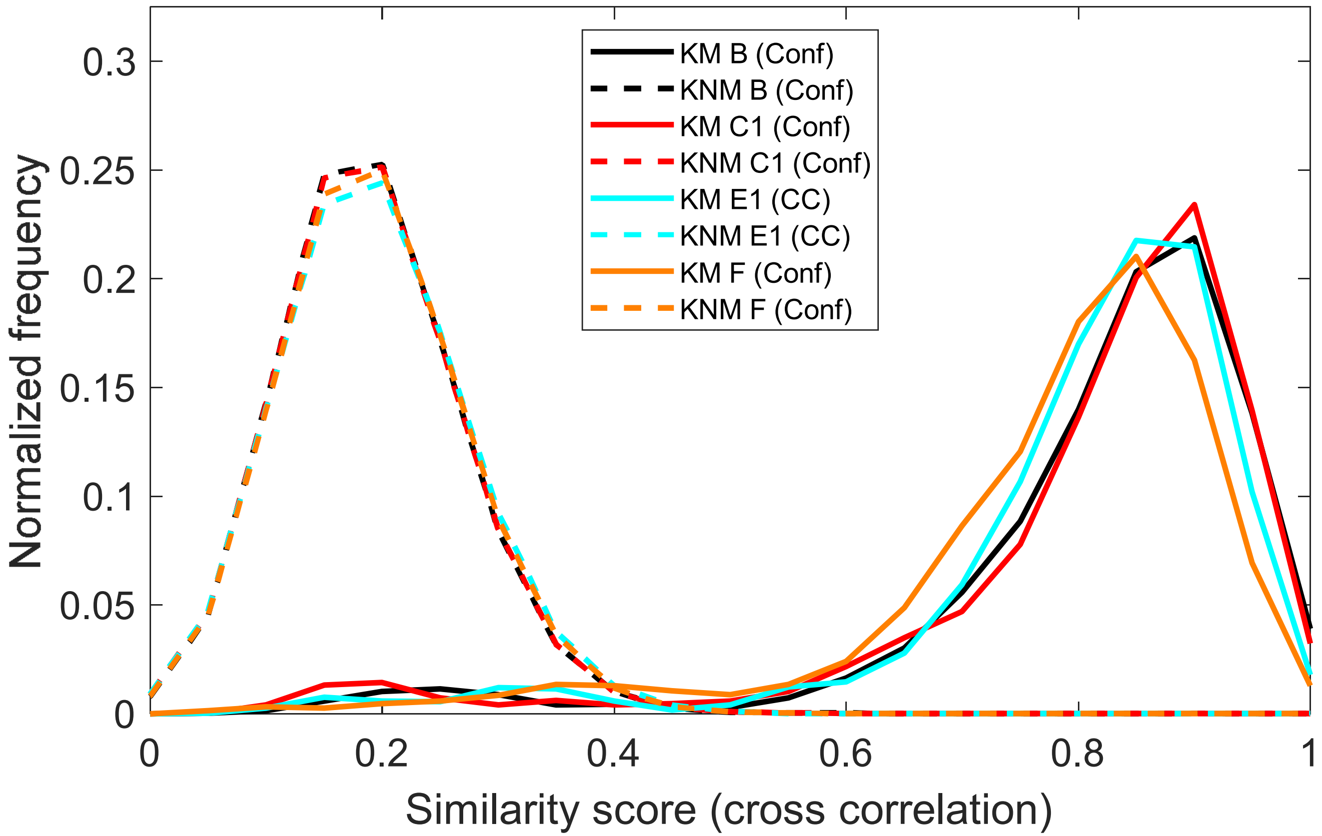


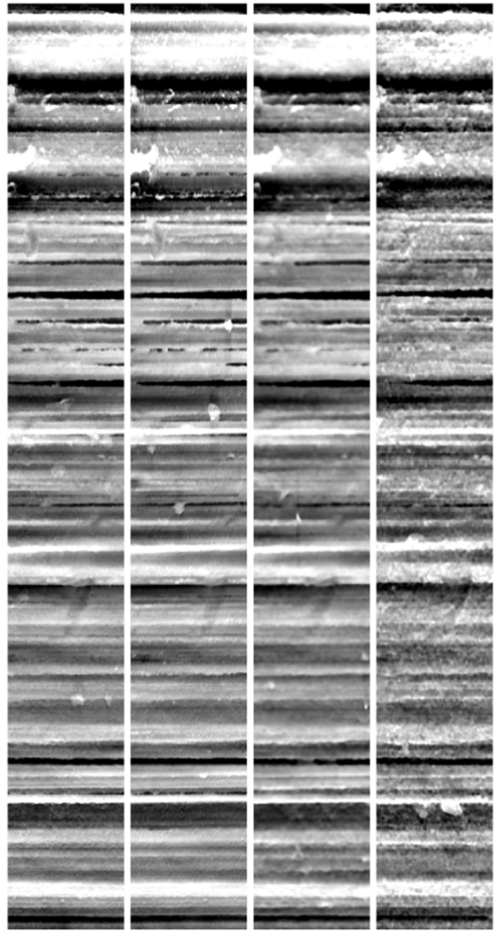
KNM

Lab	Brand/Model Technology
A	Brand 1, Model I *Focus Variation*
B	Brand 2, Model I *Confocal*
C1	Brand 3, Model I *Confocal*
C2	Brand 4, Model I *FV + Photometric Stereo*
D	Brand 1, Model II *Focus Variation*
E1	Brand 2, Model I *Continuous Confocal*
E2	Brand 2, Model I *Focus Variation*
F	Brand 2, Model II *Confocal*
G	Brand 5, Model I *FV + Photometric Stereo*
H	Brand 6, Model I *Photometric Stereo*

m.baiker@nfi.nl

Comparison Confocal microscopy



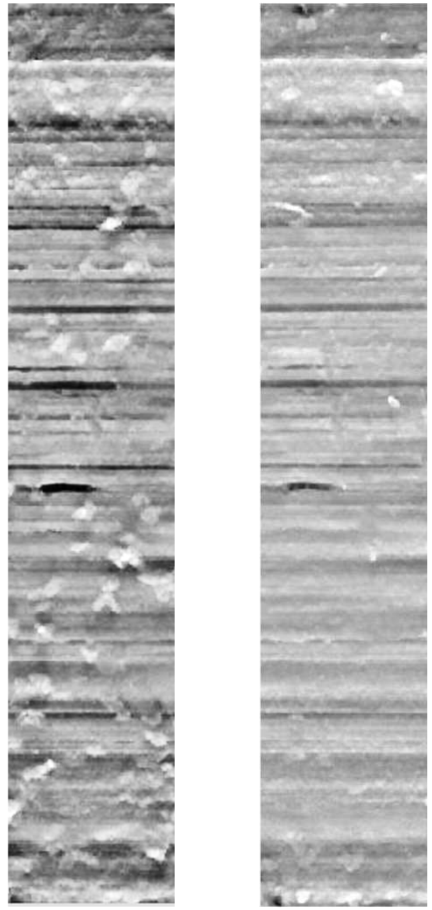


B *C1* *E1* *F*

m.baiker@nfi.nl

Comparison Confocal microscopy

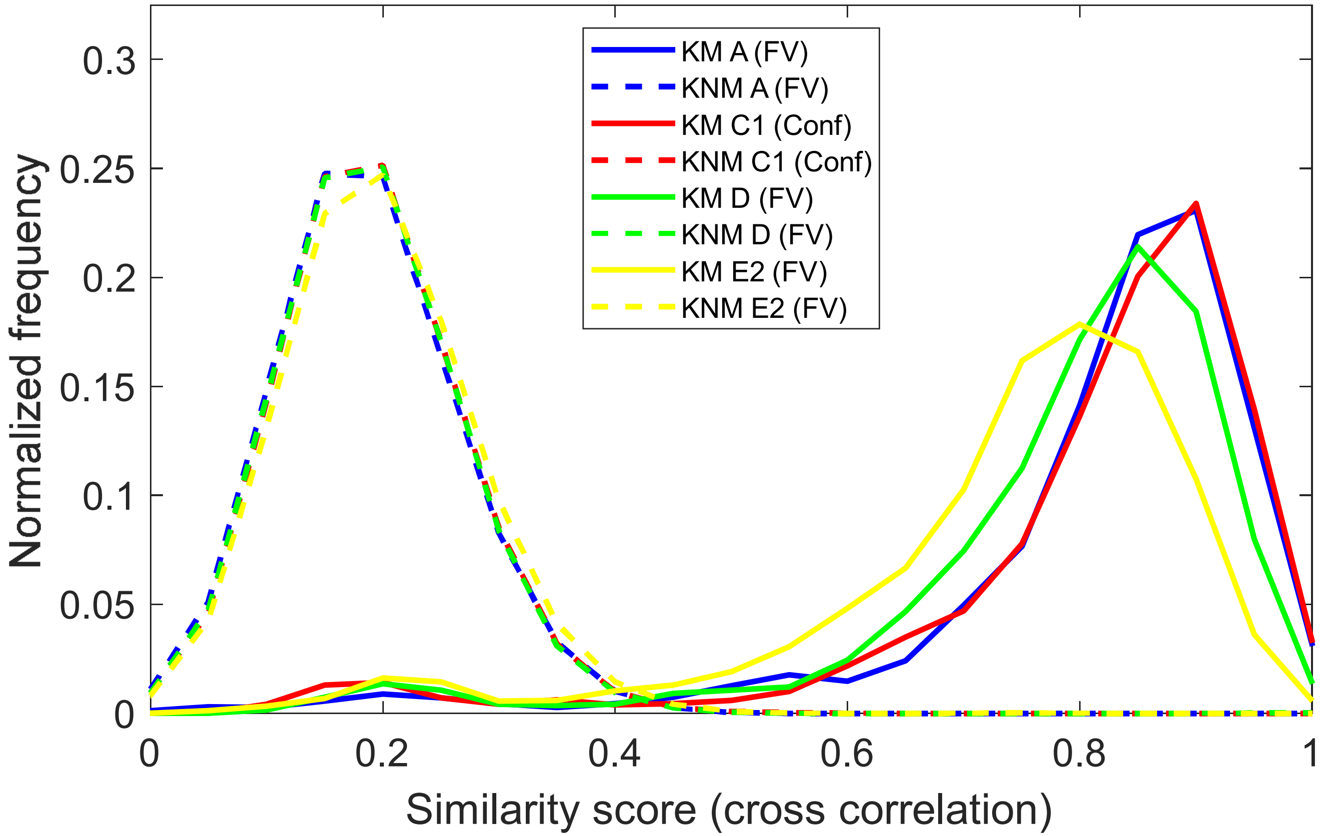
- *Lab F:*
  - *Older model*
  - *Dusty samples*

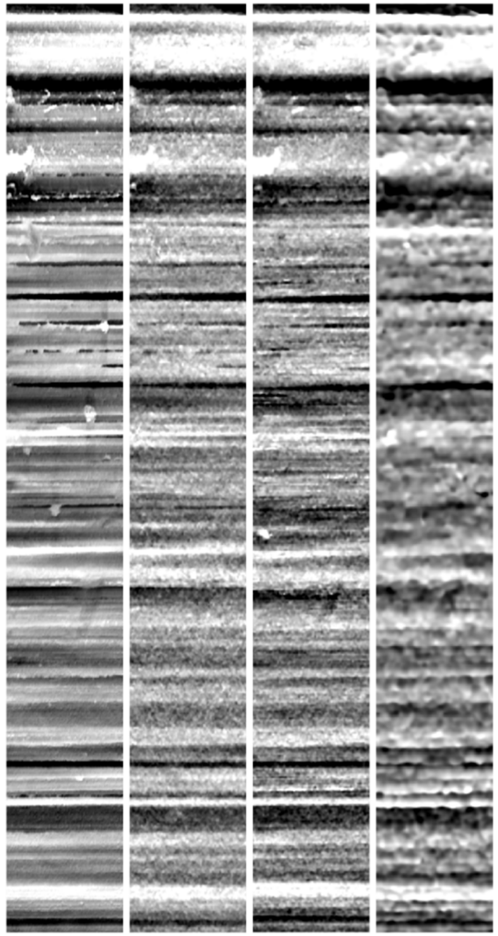


Before *After*

m.baiker@nfi.nl

Comparison Focus Variation microscopy





C1 A D E2

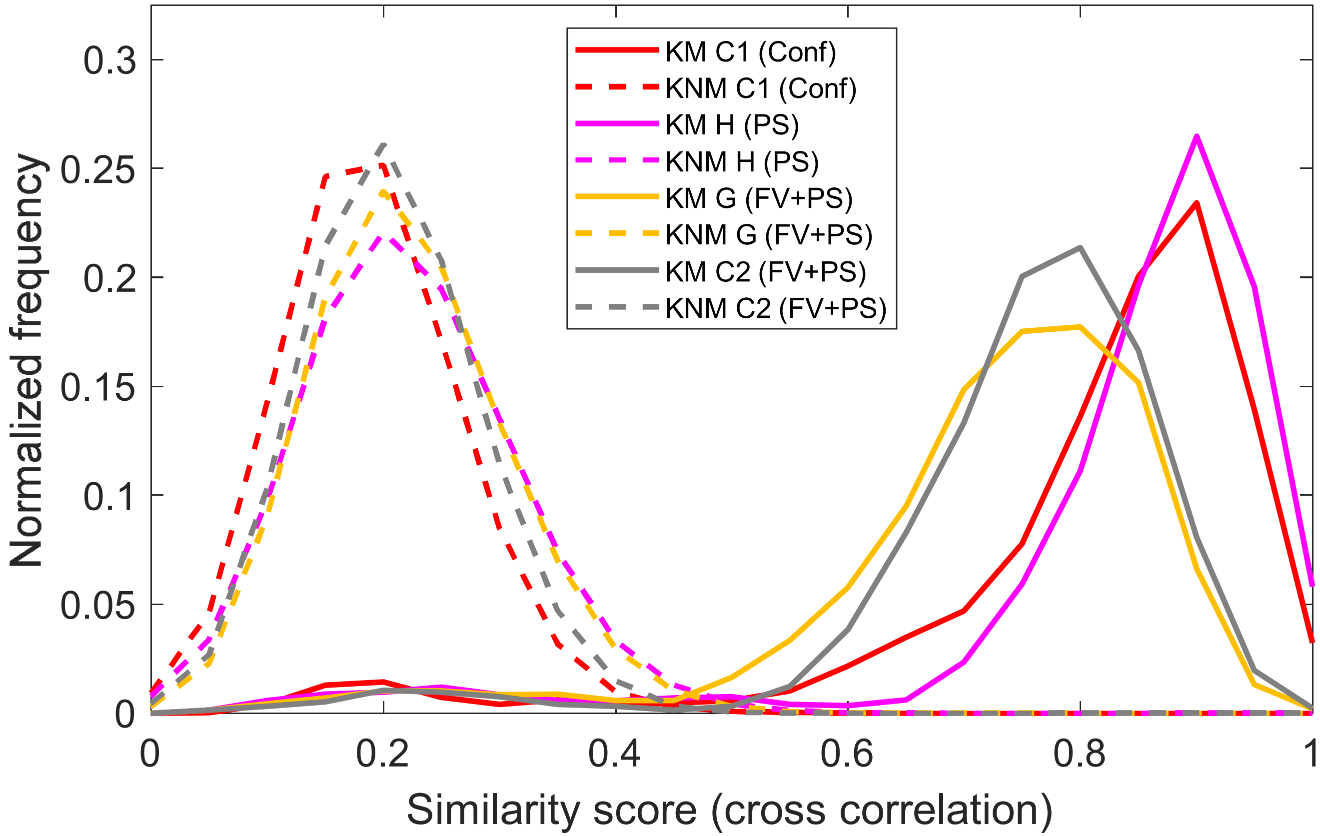
m.baiker@nfi.nl

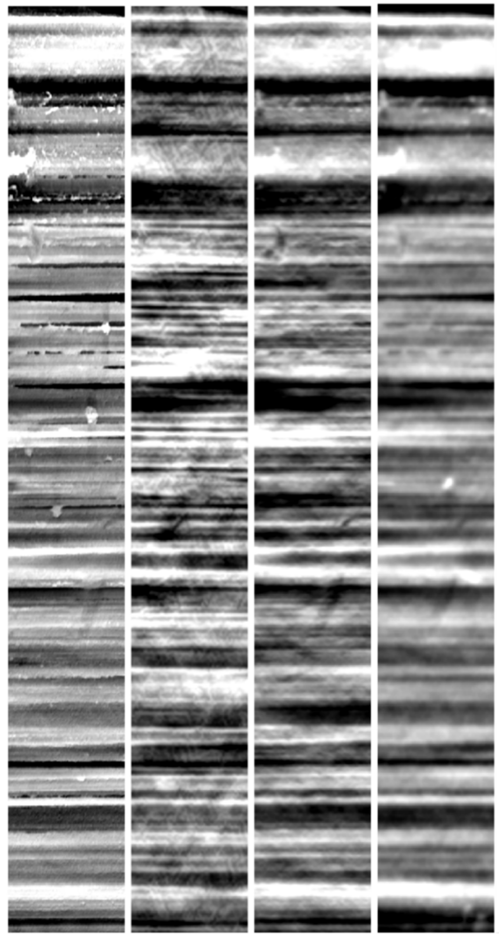
Comparison Focus Variation microscopy

- *Lab D:*
  - *Student operators*
  - *Relative large variation in positioning*
  - *Filtering with 50 μm causes KM distribution to be on top of the one from labs A and C1*
- *Lab E2:*
  - *Relatively more noise compared to other measurements*

m.baiker@nfi.nl

Comparison Photometric Stereo (+ FV) microscopy

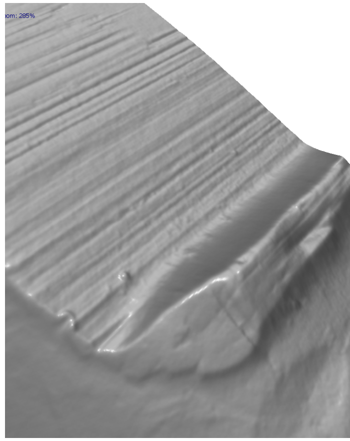




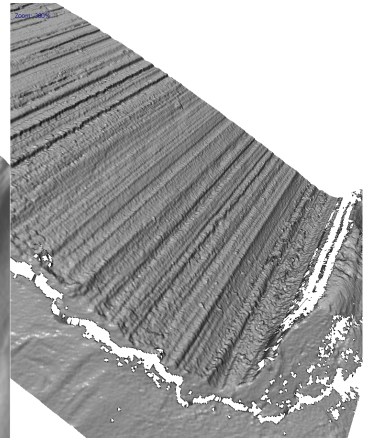
m.baiker@nfi.nl

Comparison Photometric Stereo (+ FV) microscopy

- *Lab H:*
  - *System requires gel pads for acquisition*
  - *Gel pads 'smooth' the data (remove noise and reduce strong reflections)*
  - *Gel at times doesn't contact the full LEA*



H

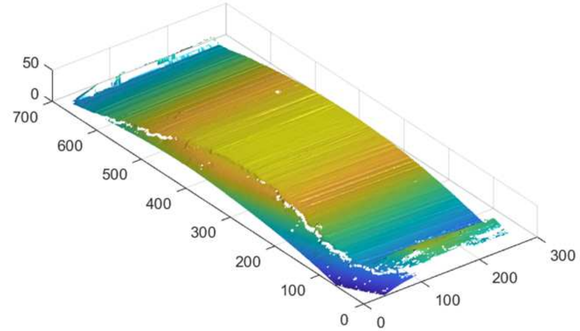


C1

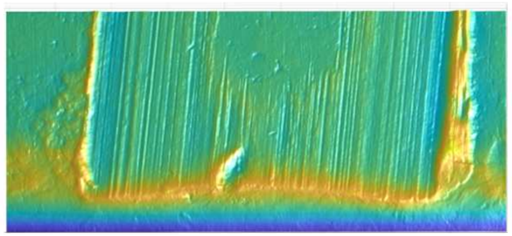
m.baiker@nfi.nl

Comparison Photometric Stereo (+ FV) microscopy

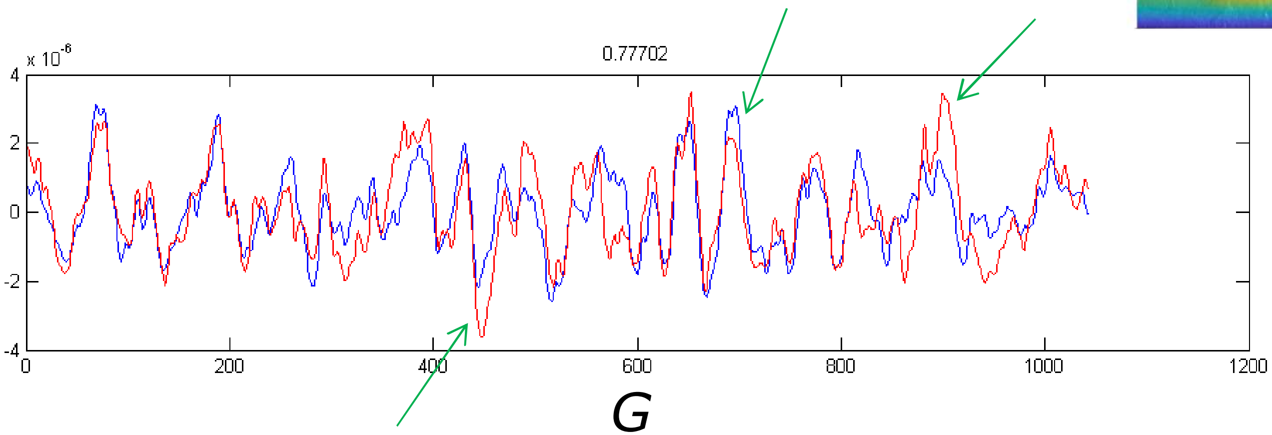
- *Lab C2 and G:*
  - *Profile amplitude accuracy might be affected by strongly reflective areas*
- *Lab G:*
  - *Filter artefacts at the onset of the mark*
  - *Small local 'feature shift'*



C1



G



G

Blue: Lab C1 (Confocal)

Red: Lab G (FV + PS)

m.baiker@nfi.nl

Comparison Photometric Stereo (+ FV) microscopy

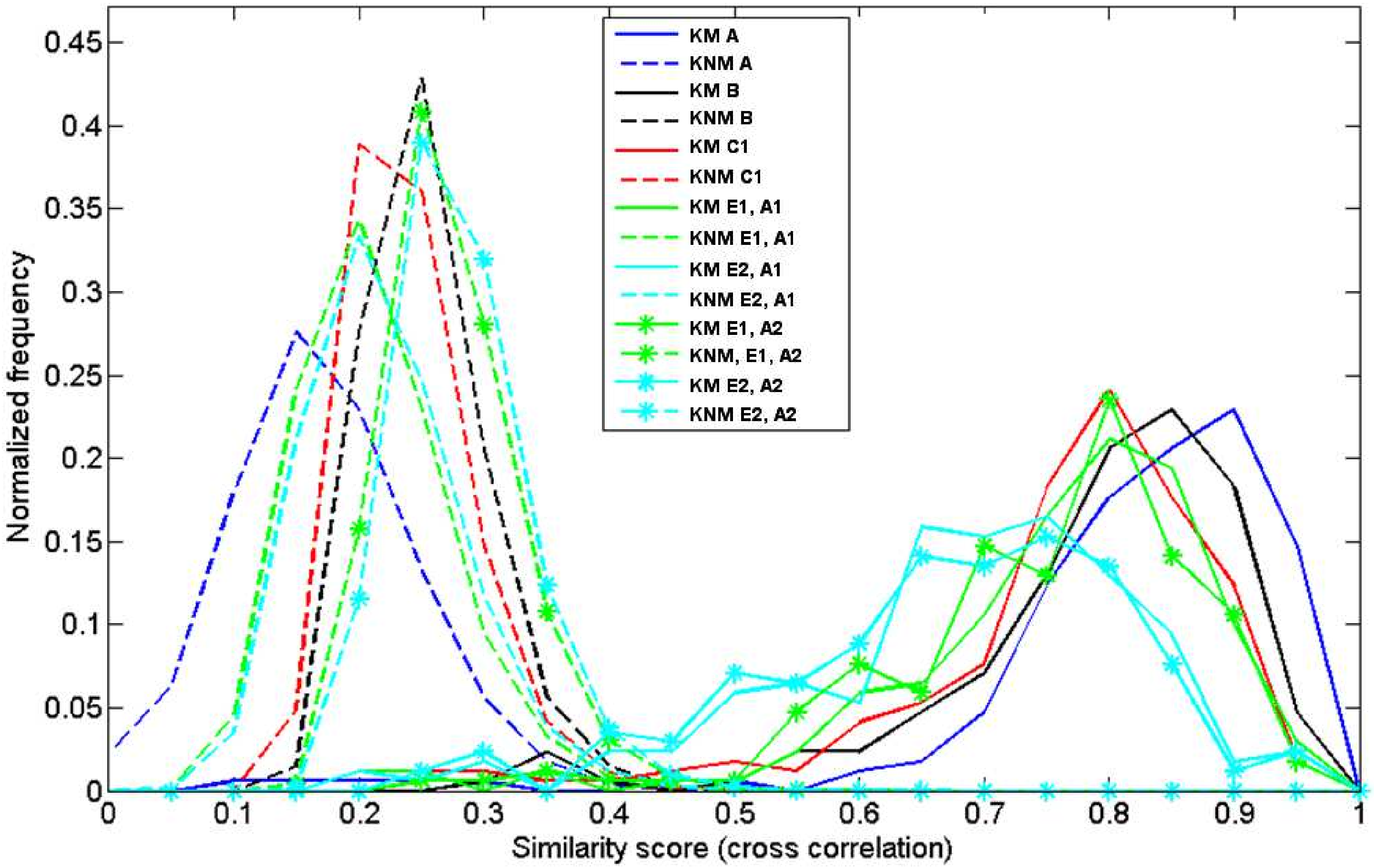
- *Systems C2 and G:*
  - *(Very) Fast and easy data acquisition*
  - *Systems targeted at database retrieval*
  - *Profile amplitude measurements might be affected by highly reflective marks surface*

Note: The shape of marks is removed by the system for C2 and G, whereas for the other systems the shape is removed by software

-> Possible differences in shape content, affecting the profiles

m.baiker@nfi.nl

Different methods yield different KM/KNM distributions



m.baiker@nfi.nl

Limitations

- *Limited data (… as always), especially in the tails of the distributions*

m.baiker@nfi.nl

Limitations

- *Limited data (… as always), especially in the tails of the distributions*
- *Only one firearm/ammunition combination so far*

m.baiker@nfi.nl

Limitations

- *Limited data (… as always), especially in the tails of the distributions*
- *Only one firearm/ammunition combination so far*
- *Study used one an application specific metric (cross correlation), not a generic surface metric*
  - ➢ *Difficult to generalize results to other mark types/scores*

m.baiker@nfi.nl

Summary

- *Results of using different acquisition methods and different algorithms is like comparing apples to pears, even using the same similarity score*

m.baiker@nfi.nl

Summary

- *Results of using different acquisition methods and different algorithms is like comparing apples to pears, even using the same similarity score*
- *Using the same comparison method yields similar score distributions, despite variation in technology, operator, settings and location*

➢ *(Automated) Interpretation results should be similar*

m.baiker@nfi.nl

Summary

- *Results of using different acquisition methods and different algorithms is like comparing apples to pears, even using the same similarity score*
- *Using the <u>same</u> comparison method yields similar score distributions, despite variation in technology, operator, settings and location*
- *Two categories of influencing factors: 1.) Technology, 2.) Protocol*

m.baiker@nfi.nl

Summary

- *Results of using different acquisition methods and different algorithms is like comparing apples to pears, even using the same similarity score*
- *Using the same comparison method yields similar score distributions, despite variation in technology, operator, settings and location*
- *Two categories of influencing factors: 1.) Technology, 2.) Protocol*
- *Possibility to (partially) address technology and protocol induced variation:*
  - *A: Experienced operator, proper sample alignment, proper cleaning*
  - *B: Automated rotation, unfolding, allowing scaling during alignment*
  - *C: Remove fine details from data by filtering and/or adjusting resolution*

m.baiker@nfi.nl

Summary

- *Results of using different acquisition methods and different algorithms is like comparing apples to pears, even using the same similarity score*
- *Using the same comparison method yields similar score distributions, despite variation in technology, operator, settings and location*
- *Two categories of influencing factors: 1.) Technology, 2.) Protocol*
- *Possibility to (partially) address technology and protocol induced variation:*
  - *A: Experienced operator, proper sample alignment, proper cleaning*
  - *B: Automated rotation, unfolding, allowing scaling during alignment*
  - *C: Remove fine details from data by filtering and/or adjusting resolution*
- *Data available in the NIST ballistics database (https://tsapps.nist.gov/NRBTD)*

m.baiker@nfi.nl

Conclusion

- *So is 3D surface data always truly objective?*
  - *Not 'out-of-the-box'*
  - *But differences between systems can be greatly reduced with proper data acquisition protocols*

- *Keep in mind:*
  - *Differences between measurements from the same system in the same lab are expected to be (very) small*

m.baiker@nfi.nl