Higher level spatial analysis of dead pixels on local grid geometry and applications to digital X-ray detector quality assessment

Julia Brettschneider

9.12.2019
A tale of dead pixels

- *Inside-out*
- The birth of dead pixels
- Dead pixels geometry
- Spatial statistics for dead pixels
- Dead pixels go shiny: *DetectorChecker*
- Dead pixels make a deal
- Dead pixels alive
Inside-out

Statistical Methods for Computed Tomography Validation of Complex Structures in Additive Layer Manufacturing

PI: Prof W Kendall
Other investigators: Prof M A Williams, Dr G J Gibbons, Dr J Brettschneider, Prof T Nichols

3 years
10/2013 - 9/2016
EP/K031066/1
X-ray computed tomography

X-ray Computed Tomography (CT) is a nondestructive technique for visualising interior features within solid objects, and for obtaining digital information on their 3-D geometries and properties.
X-ray computed tomography

[Image of X-ray computed tomography system]

https://www.researchgate.net/figure/Example-of-an-industrial-computed-tomography-CT-system_fig1_324511614
X-ray computed tomography

- X-ray shield
- X-ray generator
- X-Ray tube
- Sample
- Rotated 360°
- X-Ray detector
- Detector
- Full-angle radiographic image (2D)
- Reconstructed into CT images
- Equipment: DeskTom series
- Digital reconstruction and 3D rendering
- Volume Rendering

Projects (selection)

• Modelling the penumbra in Computed Tomography using a mixture model (Gauss + uniform) for estimating precision of radiographs from the penumbra effect in the image

• Modelling mean-variance relationship (compound Poisson for grey value, linear relationship for variance prediction)

• Detection of defects in additive manufacturing from a single x-ray projection using the empirical null filter

• Industrial uses for real-time tomography devices (e.g. airport security bag searches)

• …

• Dead pixels
X-ray detector

Perkin Elmer
XRD 1621

Readout groups (ROG):
Upper groups transferred first, starting read out from the upper row.
Lower groups starting from the last row.
Bad pixel maps

- Criteria for “underperforming” (Perkin Elmer):
  - Signal sensitivity (at different energies)
  - Noise observed in sequence of 100 bright/dark images
  - Uniformity (global, local)

- Each bad pixel map consist of a total of 10 files:
  - White images: mean, min, max, sd (.tif)
  - Grey images: mean (.tif)
  - Black images: mean, min, max, sd (.tif)
  - Bad pixel list of locations (.xml)
Modeling and analysing dead pixels

- **Spatial analysis of dead pixels:**
  - Exploratory analysis
  - Data structure for dead pixel data
  - Spatial statistics models and characteristics

- **Relationship to causes of damage:**
  - Change of perspective in the stochastic model: clusters

- **Refined analysis:**
  - Refined categories for dysfunctional pixels
  - Temporal development
Local defects: Isolated dead pixels

Singles, doubles, small clusters

A_0: Grey image [R]
A_0: bp binary image [R]
A_0: Black image [R]
Local defects: Dead lines

- Lines on bad pixel images
- From centre horizontal line outwards
- Visible on tif images of channel(s), too
Local defects: Locations of dead lines

A_0: Length of longest run

A_0: Graph of bad pixel images

A_0: Bad pixel image
Local defects: Ends of dead lines

- Most lines end in small cluster pointing to the right
- Lines are composed of dark pixels
- Lines have constant intensity, except end may differ

bmp binary image

Black image
Local defects: Corners

B_0: Binary bad pixel image [R]
Local defects: Patches

- Areas with high density area of bad pixels

E_0 Binary bad pixel image

F_0 Binary bad pixel image
Which spatial data structure?

Three common types described by Cressie (1993)

**Geostatistical data:**
Fixed study region with a random variable (observed or unobserved) in every location.
e.g. *UK with rainfall*

**Lattice data:**
Collection of fixed (nonrandom) set of points in study region with a random variable defined in each of them.
e.g. *Ising model on a lattice, crime in snap points*

**Spatial point patterns:**
Spatial locations of the observations are random, with observations itself deterministic (=1) or itself random variables.
e.g. *locations of bird nests, same with number of eggs in each nest*
Spatial model for dead pixels

Lattice or point pattern?
Detector is based on a lattice, but our interest is in locations of dead pixels. Hence, use a spatial point pattern model, but with reduced resolution (given by the detector lattice).

Point pattern X: random locations of dead pixels

Objectives:
• describe spatial distribution of dead pixels
• hypothesise causes for dead pixels

For example, look at CSR…
Complete spatial randomness (CSR)

CSR: Points are distributed independently and homogeneously, as in a homogenous Poisson process.
Exploring CSR using Ripley’s K-function

**K-function:**
expected number of extra points in circle of radius $r$ re-scaled by density

$$K(r) = \lambda^{-1} E[N_0(r)]$$

$N_0(r)$ number of points within distance $r$ from arbitrary point

$\lambda$ globally estimated density

Under CSR: $K(r) = \pi r^2$
Point pattern and K-function

Point pattern A_0

K function A_0 cropped

K(r)

r
Point pattern and K-function

\[ K(r) = \pi r^2 \]

\[ K_{\text{obs}}(r) - \pi r^2 \]

\[ K_{\text{theo}}(r) - \pi r^2 \]

\[ K_{\text{obs}}(r) - \pi r^2 \]

\[ K_{\text{theo}}(r) - \pi r^2 \]
Exploring CSR using F- and G-functions

Nearest neighbour function G: cumulative distribution function of the distance from an arbitrary point to its nearest point

Under CSR: \[ G(r) = 1 - \exp(-\lambda \pi r^2) \]

Empty space function F: cumulative distribution function of the distance from an arbitrary location to its nearest point

Under CSR: \[ F(r) = 1 - \exp(-\lambda \pi r^2) \]
Point pattern and F- and G-function

F-function, Pixels, nsim=100

G-function, Pixels, nsim=100
Are we asking the right question?

Modified question: Is it CSR after we remove all specific (known) problems?

Step 1:
Convert point process into event process by

- Reducing lines to their endpoint
- Reducing clusters to their centre point
Are we asking the right question?

Modified question: Is it CSR after we remove all specific (known) problems?

Step 1:
Convert point process into event process by

- Reducing lines to their endpoint
- Reducing clusters to their centre point

Step 2:
- Fit inhomogeneous density
- Cut out areas above threshold
Model for cause versus model for effect

Detector is based on a lattice, but damage occurs independently of the lattice structure.

The **same cause for damage** shape can hit 1, 2, 3 or 4 pixels, depending on position and orientation.
Higher level model: dead events

**Solution:** Model the damage by summarising neighbouring dead pixels into one dead event.

Convert pixel point pattern $X$ into event point pattern $Y$ by replacing each cluster of pixels $C$ in $X$ by

$$i^{\text{median}}(C) = (\rho(\text{median}_1(C)), \rho(\text{median}_2(C)))$$

Using median because of robustness:

- e.g. $i^{\text{mean}}(C) = (5, 2)$
- $i^{\text{median}}(C) = (3, 2)$
Dead pixels versus dead events

Figure 3: Pixel process and event process.

Applying the appropriate rules, configurations of connected pixels in the pixel process are reduced to one point per configuration when constructing the event process. In the example, lines provide the most striking instances of damage, but there is also damage in corners and in some other areas.

3 Quality assessment tools

There is a variety of objectives in the quality assessment of detectors which can be associated with suitable statistical measures. Our approach has several components based on global information, local configurations and spatial distributions of these. Both pixel level and event level information are used for a variety of scores we propose for usage in the context of quality assessment.

We use very simple scores for rating overall detector quality:

- Functional pixel percentage
  \[
  \text{functional pixel percentage} = \frac{\# \text{functional pixels}}{\# \text{total pixels}}
  \]

- Damage events count
  \[
  \text{damage events count} = \sum C
  \]
  where \( C \) is a cluster

Our local approach involves spatial analysis of the distribution of damage events rather than individual dysfunctional pixels. Based on the classification in Section 2.1, dysfunctional pixels belong to five categories: singletons, doubles, triplets, large clusters and lines and we summarise this using the simple scores listed below.

- Singleton count
  \[
  \text{singleton count} = \sum C
  \]
  where \( C \) is a singleton

- Line count
  \[
  \text{line count} = \sum C
  \]
  where \( C \) is a line

- Non-line cluster count
  \[
  \text{non-line cluster count} = \sum C
  \]
  where \( C \) is a singleton, double, triplet or large cluster

- Median line length
  \[
  \text{median line length} = \text{median} |C|
  \]
  where \( C \) is a line

- Median cluster size
  \[
  \text{median cluster size} = \text{median} |C|
  \]
  where \( C \) is singleton, double, triplet or large cluster

Apart from counting damage events and measuring their average size, we need to : X (dead pixels) Y (dead events)
Higher level defect model (Step 1)

Conversion of point process to event process

Defect pixels

Defect events
Density based thresholding (Step 2)

Remove areas with local density above threshold (median +1.5 IQR)
After modification: K-function

Point pattern $E_0$

K-function, Events, nsim=100
Before modification: K-function

Point pattern $E_0$

$K(r) - \pi r^2$

$K_{\text{obs}}(r) - \pi r^2$

$K_{\text{theo}}(r) - \pi r^2$

$K_{\text{hi}}(r) - \pi r^2$

$K_{\text{lo}}(r) - \pi r^2$
After modification: F-function

Point pattern E_0

F-function, Events, nsim=100
Before modification: F-function

Point pattern $E_0$

$F(r)$ function, Pixels, $n_{sim}=100$
After modification: G-function

Point pattern $E_0$

G–function, Events, nsim=100

$G(r)$
Before modification: G-function

Point pattern $E_0$

$G(r)$ function, Pixels, nsim=100
Measurement quality assessment/improvement

- Identify poor quality regions (patches with high dead pixels density) through density thresholding
- Remaining area CSR means no special causes of poor quality
- Identify causes of poor quality
- Monitor over time
- Conclusions for usage modes
Software project with the Alan Turing Institute

Objectives:
Web application “DetectorChecker”
• Feedback about state of detector through pixel damage analysis
• Detector data repository

Seed funded project:
• Working with Turing Research Software Engineer Group
• DetectorChecker R package for statistical analysis of pixel damage in CT scanners
• DetectorCheckerWebApp for useful initial graphical/analysis
• Facility to upload data in different formats (crowd sourcing)
• Hosted by Azure
Team
Dr Julia Brettschneider (University of Warwick)
Dr Oscar Giles (The Alan Turing Institute)
Dr Tomas Lazauskas (The Alan Turing Institute)
Prof Wilfrid Kendall (University of Warwick)

Contacts
julia.brettschneider@warwick.ac.uk
W.S.Kendall@warwick.ac.uk

Timeline
1.9.2018 - 29.3.2019

https://detectorchecker.azurewebsites.net
1. Select Layout
   - PerkinElmerFull

2. Visualisation
   - Layout
     - euclidean distance from centre
     - L-infinity distance from centre
     - euclidean distance to nearest corner
     - horizontal distance to nearest sub-panel edge
     - vertical distance to nearest sub-panel edge
     - L-infinity distance to nearest sub-panel edge

Display plot
Layout: PerkinElmerFull

1. Select Layout

PerkinElmerFull

2. Visualisation

- layout
- euclidean distance from centre
- L-infinity distance from centre
- euclidean distance to nearest corner
- horizontal distance to nearest sub-panel edge
- vertical distance to nearest sub-panel edge
- L-infinity distance to nearest sub-panel edge

Display plot

detectorchecker v: 0.1.9
webapp v: 0.1.7

https://detectorchecker.azurewebsites.net
coordinates of underperforming pixels (.xml)
Density
Arrows pointing at nearest neighbour
Arrows pointing at nearest neighbour
Angles distribution dominated by lines
K-function

Not CSR (completely spatially at random)
Levels: Pixels or Events?

Detector is based on a lattice, but damage occurs independently of the lattice structure.

The **same cause for damage** shape can hit 1, 2, 3 or 4 pixels, depending on position and orientation.

**Solution:** Model the damage by summarising neighbouring dead pixels into one dead event.

Convert point process into *event process* by
7. Modelling Damage Intensity

- euclidean distance from centre
- Linfinity distance from centre
- horizontal distance to nearest sub-panel edge
- vertical distance to nearest sub-panel edge

Fit model

Call:
glm(formula = as.vector(pix_matrix) ~ as.vector(dist), family = binomial(link = logit))

Deviance Residuals:
       Min          1Q       Median          3Q         Max
-0.0988  -0.0745  -0.0662  -0.0585   3.6700

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.320e+00  2.505e-02  -204.22  <2e-16 ***
 as.vector(dist) -1.002e-03  3.511e-05  -28.54  <2e-16 ***

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 133198  on 3999999 degrees of freedom
Residual deviance: 132394  on 3999998 degrees of freedom
AIC: 132398

Number of Fisher Scoring iterations: 9
Team
Dr Julia Brettschneider (University of Warwick)
Dr Oscar Giles (The Alan Turing Institute)
Dr Tomas Lazauskas (The Alan Turing Institute)
Prof Wilfrid Kendall (University of Warwick)

Contacts
julia.brettschneider@warwick.ac.uk
W.S.Kendall@warwick.ac.uk

Timeline
1.9.2018 - 29.3.2019

https://detectorchecker.azurewebsites.net
Brexit on 29.3. after all: dead pixel deal approved at 11pm in PM surprise move

Deal approved by narrow majority of Conservative and Labour MPs without DUP.

- Corbyn: British workers make dead pixels here in UK
- Rees-Mogg: I have never had an X-ray, no
- Farage: Greatest day in British history
- DUP: N. Ireland doesn’t recognise pixels of any sort
- Business: Optimistic about digital economy
- Tusk’s special place in hell better with dead pixels
- PM: Pixel means pixel
- Gove: Dead pixels key to Irish border IT solution
- Merkel and Macron seen waltzing in Brussels
Refined states (more than just just dead)

Using grey, white and black images define a variety of dysfunctional states and look at transitions.

- Healthy pixels
- Dim → No response → Dead pixel → Dead line
- Bright → Bright cluster → Supercluster → Bright line

Definitions:

- Dim: Pixels identified as dim by thresholding in grey/white images. Exclude patches identified as spots on beryllium screen.
- No response: Pixel behaves normally in black images, but response in grey/white images remains at normal black level. No response at all to presence of x-rays.
- Dead pixel: Pixelwise mean is exactly 0 in grey or white images.
- Dead line: Column of zero-valued pixels. Not yet observed in the new data set.
- Bright: Singleton pixel identified as bright by thresholding in white/grey images.
- Bright cluster: Cluster of between 2 and 9 adjacent pixels identified as right by thresholding in white/grey images.
- Supercluster: Cluster of adjacent hot and bright pixels or pixel clusters. Contains at least one hot pixel, which appears to be leaking charge into (usually) horizontally-adjacent cells - a phenomenon known as 'blooming', I think.
- Bright line: Column of cells with slightly higher values than those on either side, possibly with its root in a supercluster. May need to be identified using a separate process, since the pixels may not be very much brighter (in the two examples so far, perhaps 300 brighter than the two adjacent columns: not usually enough to be picked up by thresholding)

Still to decide...

Which images to use to define hot pixels? If px == 65535 in black image, it is definitely damaged. Perhaps classify hot pixels in white/grey images as bright instead? Then again - the user is more interested in the behaviour of the detector when in use, not in the dark images.

[How many px classified as hot in white/grey are part of a supercluster? This may suggest a solution]

1 Using grey, white and black images define a variety of dysfunctional states and look at transitions.
Model for temporal development

### Markov model with transition probabilities estimated from data:

<table>
<thead>
<tr>
<th>Initial state</th>
<th>Normal</th>
<th>No response</th>
<th>Dead</th>
<th>Hot</th>
<th>V. bright</th>
<th>Bright</th>
<th>Bright line</th>
<th>Screen spot</th>
<th>Edge</th>
<th>V. dim</th>
<th>Dim</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal</td>
<td>99.91</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>no response</td>
<td>-</td>
<td>98.83</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.17</td>
<td>-</td>
</tr>
<tr>
<td>dead</td>
<td>-</td>
<td>-</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>hot</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>96.72</td>
<td>3.28</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>v.bright</td>
<td>0.89</td>
<td>-</td>
<td>-</td>
<td>2.74</td>
<td>88.62</td>
<td>7.69</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.06</td>
<td>-</td>
</tr>
<tr>
<td>bright</td>
<td>18.07</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>8.44</td>
<td>73.44</td>
<td>0.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>line.b</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>screen spot</td>
<td>84.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>16.59</td>
<td>0.56</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>edge</td>
<td>0.06</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>0.06</td>
<td>99.89</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>v.dim</td>
<td>-</td>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>90</td>
<td>-</td>
</tr>
<tr>
<td>dim</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>94.44</td>
<td>-</td>
</tr>
</tbody>
</table>

With the exception of screen spots (which vary because the screen is replaced, not because of any problems with detector pixels), the categories identified are reasonably stable; there is some movement between bright, very bright and hot states (probably because brighter pixels are generally more variable, so perhaps should not be expected to remain in the same category). However, a more sophisticated classification approach, considering the clusters in which pixels occur, may also be useful.
Markov decision process

\[ M = (T, A, \Theta, R) \]

\[ X \] Dynamic system under partial control of DM

\[ \sigma = S_0, \ldots, S_\tau \] Subsequent states

\[ \alpha = a_0, a_1, \ldots, a_\tau \] Action sequence

\[ P_\tau^{(S,\alpha)}(\sigma) = \prod_{t=0}^{\tau-1} \theta_t(S_t, a_t, S_{t+1}) \]
Markov decision process evaluation

\[ \mathcal{M} = (T, A, \Theta, R) \]

**X** Dynamic system under partial control of DM

\[ \sigma = S_0, \ldots, S_\tau \] Subsequent states

\[ \alpha = a_0, a_1, \ldots, a_\tau \] Action sequence

\[ P_\tau^{(S, \alpha)}(\sigma) = \prod_{t=0}^{\tau-1} \theta_t(S_t, a_t, S_{t+1}) \]

\[ h = (S_0, \ldots, S_N, a_0, \ldots, a_N) \]

\[ u(h) = \sum_{t=0}^{N} \lambda^t r_t(S_t, a_t) \] Utility

\[ S_0 = S \] and \[ \delta_t(S_t) = a_t \] Usage policies

\[ P_\tau^{(S, \pi)}(h) = \prod_{t=0}^{\tau-1} \theta_t(S_t, a_t, S_{t+1}) \]

\[ \tilde{u}_\pi(S) = E_{P_N^{(S, \pi)}}(u) \]

\[ = \sum_{h \in H_N} u(h) \cdot P_N^{(S, \pi)}(h) \]
Thanks to the team!

Inside-out

Statistical Methods for Computed Tomography Validation of Complex Structures in Additive Layer Manufacturing

PI: Prof W Kendall
Other investigators: Prof M A Williams, Dr G J Gibbons, Dr J Brettschneider, Prof T Nichols

Project partners:
EOS systems, Nikon, Renishaw

PDRAs:
Audrey Kueh, Jay Warnett, David Garcia
Clair Barnes (Intern)
Sherman Ip (PhD student)
Tom Suchen (Masters student)
Goals

Bottlenecks in Additive Layer Manufacturing (ALM) (a "3D printing" technique) is quality control. Direct verification typically involves lengthy analysis of individual manufactured objects using Computed Tomography (CT) scans. Statistical methods are key to enabling engineers to do that.

- **Image quality for CT scanners**: We analysed penumbra effects and demonstrated that they can be mitigated by careful filter design (published paper, impact case).

- **Characterisation of CT noise**: We modeled the greyvalue of each pixel as a compound Possion random variance to capture the behaviour of x-ray photons and use the resulting linear relationship between the mean and variance for variance prediction (ongoing work by an EPSRC-funded research student).

- **Defect detection in ALM structures**: We have been developing a procedure for rapid assessment and location of collections of small defects in 3D-printed objects (paper in preparation).

- **Real-time tomography performance**: Industrial uses for real-time tomography devices developed in the context of airport baggage searches (published and ongoing).

- **Dead pixels and other detector damages**: We introduced a taxonomy for dysfunctional pixels based on local grid geometry. We methods from spatial statistics to establish decision rules distinguishing special causes of poor quality from common causes, which helps removing damage and avoided future problems (two technical reports online, paper in preparation).
Key outcomes

Research:


Collaborations (New academics/industry): RA Warnett now Assistant Professor at WMG, RA Kueh now Teaching Associate at Cambridge, Crevillen-Garcia now research associate at Warwick Engineering Department, Master student Jin now PhD student at OxWaSP (EPSRC funded), intern Barnes now statistics PhD student at UCL.

Impact: Formal impact case relating to publication 1 above with industrial support (Nikon).