Advanced Image Cleaning

CTA Consortium Meeting

Jérémie DECOCK

CEA Saclay - Irfu/SAp

October 25, 2016
Introduction
Subject

Try to improve image cleaning before reconstruction (*Hillas*)

Improve methods to remove:

▶ Instrumental noise
▶ Background noise

Motivations:

▶ Keep more signal (deeper into the noise)
▶ Reduce threshold
▶ Maybe eventually do cleaning and time-integration all at once
Image cleaning algorithms
The “Tailcut clean” algorithm

A very simple cleaning procedure:

- Keep pixels above a given threshold (e.g. 50% max)
- Keep some neighbors of these selected pixels: those above a second (lower) threshold (e.g. 25% max)
Remarks

- Fast and simple
- Sufficient for bright showers
- But surely we can do better for faint showers
Basic idea to go beyond

- Tailcut method: threshold in the main space
- Better idea: threshold in a different space where signal and noise can be easily separated
  - Wavelet transform
  - Cosmostat tools (iSAP/Sparse2D)
    (http://www.cosmostat.org/software/isap/)
We are considering *Wavelet Transform* method

Roughly the same idea than doing filtering with Fourier Transform
- Apply the transform on the signal
- Apply a threshold in the transformed space
- Invert the transform to go back to the original signal space

Differences with Fourier Transform
- Use functions named *wavelets* instead sin and cos functions as new bases in the transformed space
- The transformed space contains spatial information
Wavelets

Example of wavelet function (*Morlet wavelet*)

\[ \Psi(t) = e^{-x^2/2}\cos(5x) \]

“A short wave-like oscillation with a beginning and an end”
Cleaning procedure: general idea with Fourier Transform

- Input signal is converted to a **weighted** sum of sin and cos at different frequencies
- Threshold is applied on these **weights** to remove some **frequencies** in the input signal (e.g. high pass filter, low pass filter, ...)

\[
f(t) = 1.5 \cos(1t) + 0.8 \cos(2t) + 0.5 \cos(3t) + \cdots
\]

\[
+ 0.7 \sin(1t) + 1.1 \sin(2t) + 0.6 \sin(3t) + \cdots
\]
Cleaning procedure: general idea with Wavelet Transform

- Input signal is converted to a **weighted** sum of these wavelet functions at different **scales** (**dilate factor**) and **positions** (**translate factor**)
- Threshold is applied on these **weights** to remove **locally** (in space or time) some **frequencies** (or **scales**) in the input signal

![Wavelet Functions](image_url)
Find a base where signal and noise can be easily separated

In this example:

- Remove noise in direct space is difficult
- Remove noise in the transformed space is easy:
  - noise is uniformly distributed on small coefficients
  - signal is defined by few big coefficients
Example

run1001.simtel.gz (Tel. 1, Ev. 1909) 1.62E+00TeV
The same example with Tailcut

run1001.simtel.gz (Tel. 1, Ev. 1909) 1.62E+00TeV
Experimental setting
Dataset used to assess cleaning algorithms

“ASTRI mini-array” test set

- Kindly provided by the Astri team
- 33 ASTRI telescopes
- Cropped to get squared pixel arrays
Benchmark function

The error on the shape:

$$\mathcal{E}_{\text{shape}}(\hat{s}, \mathbf{s}^*) = \text{mean} \left( \text{abs} \left( \frac{\hat{s}}{\sum_i \hat{s}_i} - \frac{\mathbf{s}^*}{\sum_i \mathbf{s}^*_i} \right) \right)$$

The error on the energy:

$$\mathcal{E}_{\text{intensity}}(\hat{s}, \mathbf{s}^*) = \frac{\text{abs} \left( \sum_i \hat{s}_i - \sum_i \mathbf{s}^*_i \right)}{\sum_i \mathbf{s}^*_i}$$

Where:

- $\hat{s}$ the image "cleaned" by algorithms
- $\mathbf{s}^*$ the actual "clean" image
- $i$ is the index of a PMT (i.e. of a pixel) within an image
Preliminary results
Dataset used to assess cleaning algorithms

Realistic event set:

- Gamma photons: 4461 events, 14899 images
- Protons: 747 events, 2203 images
$E_{shape}$ (gamma photons)

Tailcut

Wavelet Transform

Total intensity in reference image (PE)

Eshape (gamma photons)
**Gammas**

\[ \mathcal{E}_{intensity} \text{ (gamma photons)} \]

**Tailcut**

- Score (the lower the better): $10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1$
- Total intensity in reference image (PE): $10^0, 10^1, 10^2, 10^3, 10^4, 10^5$

**Wavelet Transform**

- Score (the lower the better): $10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1$
- Total intensity in reference image (PE): $10^0, 10^1, 10^2, 10^3, 10^4, 10^5$

---

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
**Introduction**

Algorithms

Experiments

Results

Conclusion

References

---

**Protons**

\[ \mathcal{E}_{shape} \text{ (protons)} \]

**Tailcut**

- Score (the lower the better)
- Total intensity in reference image (PE)

**Wavelet Transform**

- Score (the lower the better)
- Total intensity in reference image (PE)

---

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
$\mathcal{E}_{\text{intensity}}$ (protons)

### Tailcut

- Score (the lower the better)
- Total intensity in reference image (PE)

### Wavelet Transform

- Score (the lower the better)
- Total intensity in reference image (PE)
Conclusion
Conclusion

This is a work is in progress...

- Optimize algorithms setting:
  - wavelet function
  - wavelet filtering methods
  - filtering thresholds
  - pre processing
  - post processing
  - ...

- Compare to optimized Tailcut
- Adapt the cleaning method to real cameras (full pixel array, hexagonal shapes, ...)
- Check ability to do real time analysis
References I

CK Bhat, Search for diffuse galactic/extra-galactic tev gamma rays.


References II


References III


______, *Novel image and non-image parameters for efficient characterisation of atmospheric cerenkov images.*
Appendix
Wavelets: why is it promising?

- Should handle more complex signal (faint signal, ...)
- May use coefficients for photon/hadron discrimination
- Data compression on site
- Require few calibration
Wavelets: mother wavelet $\Psi$

Family $\psi_{a,b}$ (where $(a, b) \in \mathbb{R}^+ \times \mathbb{R}$) is defined from the "mother wavelet" $\Psi$:

$$\forall t \in \mathbb{R}, \psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t - b}{a} \right)$$

where $a$ is the scale factor, $b$ is the translation factor.
Wavelets: general case (1D continuous case)

The original signal \( f \) defined as:

\[
f(t) = \frac{1}{C_\Psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{g(a, b)}{|a|^2} \psi_{a,b}(t) \, da \, db
\]

where \( C_\Psi \) is a constant which depends on the chosen wavelet mother \( \Psi \).

Weights are given by:

\[
g(a, b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}^*(t) \, dt
\]
Wavelets: general case (1D continuous case)
Wavelets: general case (2D hints)
Fourier transform: general case (1D continuous case)

The original signal \( f \) defined as:

\[
f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nt) + b_n \sin(nt))
\]

Weights are given by:

\[
a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) \, dt
\]

\[
b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) \, dt
\]
Fourier transform: general case (1D continuous case)

\[ f(t) = 1.5 \cos(t) + 0.8 \cos(2t) + 0.5 \cos(3t) + 0.7 \sin(t) + 1.1 \sin(2t) + 0.6 \sin(3t) + \ldots \]
Fourier transform: remarks

FFT can be applied to any $T$-periodic function $f$ verifying the *Dirichlet conditions*:

- $f$ must be continuous
- *and* monotonic
- on a finite number of sub-intervals (of $T$)

Signals defined on bounded intervals (e.g. images) can be considered as periodic functions (applying infinite repetitions)
Fourier transform: analyse

Works well:

- when the Fourier coefficients for the signal and the noise can easily be separated in the Fourier space (obviously...)
- e.g. when either the signal or the noise can be defined with few big Fourier coefficients (i.e. signal or noise have a few number of significant harmonics)
Appendix

Fourier transform

Fourier transform: a bad example

run1001.simtel.gz (Tel. 1, Ev. 1909) 1.62E+00TeV

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Different kind of “noise” in telescope images

1. Instrumental noise (Photomultiplier Tubes, ...)
   - Thermionic emission
   - Radiations
   - Electric noise

2. Background noise (Night Sky Background or NSB)
   - Parasite light (moon, stars, planes, light pollution, ...)

Appendix
Noise

Decock CEA Saclay - Irfu/SAp
Advanced Image Cleaning
MC simulations

“ASTRI mini-array” configuration

<table>
<thead>
<tr>
<th>File</th>
<th>Num. events</th>
</tr>
</thead>
<tbody>
<tr>
<td>gamma/run_1001.simtel.gz</td>
<td>4461</td>
</tr>
<tr>
<td>gamma/run_1002.simtel.gz</td>
<td>4567</td>
</tr>
<tr>
<td>gamma/run_1003.simtel.gz</td>
<td>4425</td>
</tr>
<tr>
<td>gamma/run_1004.simtel.gz</td>
<td>4401</td>
</tr>
<tr>
<td>gamma/run_1005.simtel.gz</td>
<td>4451</td>
</tr>
<tr>
<td>gamma/run_1006.simtel.gz</td>
<td>4451</td>
</tr>
<tr>
<td>gamma/run_1007.simtel.gz</td>
<td>4614</td>
</tr>
<tr>
<td>gamma/run_1008.simtel.gz</td>
<td>4423</td>
</tr>
<tr>
<td>gamma/run_1009.simtel.gz</td>
<td>4411</td>
</tr>
</tbody>
</table>
## MC simulations

### “ASTRI mini-array” configuration

<table>
<thead>
<tr>
<th>File</th>
<th>Num. events</th>
</tr>
</thead>
<tbody>
<tr>
<td>proton/run_10000.simtel.gz</td>
<td>747</td>
</tr>
<tr>
<td>proton/run_10001.simtel.gz</td>
<td>680</td>
</tr>
<tr>
<td>proton/run_10002.simtel.gz</td>
<td>763</td>
</tr>
<tr>
<td>proton/run_10003.simtel.gz</td>
<td>792</td>
</tr>
<tr>
<td>proton/run_10004.simtel.gz</td>
<td>763</td>
</tr>
<tr>
<td>proton/run_10005.simtel.gz</td>
<td>776</td>
</tr>
<tr>
<td>proton/run_10006.simtel.gz</td>
<td>738</td>
</tr>
<tr>
<td>proton/run_10007.simtel.gz</td>
<td>749</td>
</tr>
<tr>
<td>proton/run_10008.simtel.gz</td>
<td>760</td>
</tr>
<tr>
<td>proton/run_10009.simtel.gz</td>
<td>812</td>
</tr>
</tbody>
</table>
Appendix

Results (Gamma)

\[ \sum_{i} s_i^* (\gamma) \]

---

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Gamma)

$\varepsilon_1 (\gamma)$

Score

$10^{-2}$

$10^{-3}$

$10^{-4}$

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Gamma)

$E_1 (\gamma)$

<table>
<thead>
<tr>
<th>Score</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1e-3</td>
</tr>
<tr>
<td>0.2</td>
<td>1e-2</td>
</tr>
<tr>
<td>0.4</td>
<td>1e-1</td>
</tr>
<tr>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>0.8</td>
<td>10</td>
</tr>
<tr>
<td>1.0</td>
<td>100</td>
</tr>
<tr>
<td>1.2</td>
<td>1000</td>
</tr>
<tr>
<td>1.4</td>
<td>1e-3</td>
</tr>
</tbody>
</table>

WT (mr_filter)
Tailcut (JD)
Appendix

Results (Gamma)

\[ E_1 (\gamma) \]

Tailcut (JD)

WT (mr_filter)

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Gamma)

\[ \mathcal{E}_{shape} \text{ (gamma photons)} \]

- **Tailcut**
- **Wavelet Transform**

Score (the lower the better)

Total intensity in reference image (PE)

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Gamma)

\[ E_2(\gamma) \]

Score

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Gamma)

$\mathcal{E}_2 (\gamma)$

Count

Score

WT (mr_filter)
Tailcut (JD)

Decock CEA Saclay - Irfu/SAp
Advanced Image Cleaning
Tailcut (JD) scores correlation

Score 0 vs Score 1
Results (Gamma)

$E_2 (\gamma)$

Tailcut (JD)

WT (mr_filter)
Appendix

Results (Gamma)

\[ \mathcal{E}_{\text{intensity}} \text{ (gamma photons)} \]

Tailcut

Wavelet Transform

Total intensity in reference image (PE)

Score (the lower the better)

10^{-3} 10^{-2} 10^{-1} 10^{0} 10^{1}

10^{1} 10^{2} 10^{3} 10^{4} 10^{5}

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
\[ \sum_{i} s_i^* (\gamma) \]

14899 images

Count

npe

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Results (Protons)

Score

$\epsilon_1$ (protons)

10^2

10^-2

Tailcut (JD) WT (mr_transform) WT (mr_filter)

$10^{-4}$

$10^{-3}$

$10^{-2}$

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Protons)

\[ \mathcal{E}_1 \text{ (protons)} \]

<table>
<thead>
<tr>
<th>Score</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>10</td>
</tr>
<tr>
<td>0.2</td>
<td>100</td>
</tr>
<tr>
<td>0.4</td>
<td>1000</td>
</tr>
<tr>
<td>0.6</td>
<td>10000</td>
</tr>
<tr>
<td>0.8</td>
<td>10000</td>
</tr>
<tr>
<td>1.0</td>
<td>10000</td>
</tr>
<tr>
<td>1.2</td>
<td>10000</td>
</tr>
<tr>
<td>1.4</td>
<td>10</td>
</tr>
</tbody>
</table>

- WT (mr_filter)
- Tailcut (JD)

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Protons)

$\mathcal{E}_1$ (protons)

Tailcut (JD)

WT (mr_filter)

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Protons)

$\mathcal{E}_{\text{shape}}$ (protons)

Tailcut

Wavelet Transform

Score (the lower the better)

Total intensity in reference image (PE)
Results (Protons)

\[ \mathcal{E}_2 \text{ (protons)} \]

<table>
<thead>
<tr>
<th>Score</th>
<th>Tailcut (JD)</th>
<th>WT (mr_transform)</th>
<th>WT (mr_filter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^{-5}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{-4}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{-3}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{-2}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{-1}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{0}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{1}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{2}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{3}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10^{4}</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Protons)

$\mathcal{E}_{\text{intensity}}$ (protons)

Tailcut

Wavelet Transform

Score (the lower the better)

Total intensity in reference image (PE)

Decock CEA Saclay - Irfu/SAp

Advanced Image Cleaning
Appendix

Results (Protons)

WT (mr_filter) scores correlation

Decock CEA Saclay - Irfu/SAp
Advanced Image Cleaning
Papers

“Hadron suppression using Wavelet Transformations for the H.E.S.S. Telescope system” (2002, Stefan Funk)
Stefan’s Paper

Subject

- Uses Wavelets for $\gamma$-ray/hadron separation
- Mention a little bit image cleaning but no experiments (e.g. section 3.3 and conclusion)
Stefan’s Paper

Methodology

1. Add margins on the input image
2. Map the orthogonal camera coordinates into a hexagonal coordinate system
3. Apply the hexagonal wavelets to the hexagonal grid; get wavelets coefficients for each scale
4. Compute the standard deviation of wavelet coefficients for each plane
5. Give these moments to the neural network used to discriminate $\gamma$-rays to hadrons (in addition to Hillas parameters)