IPOL: A Reproducible Research Journal and Platform for Image Processing

Pascal Monasse <pascal.monasse@enpc.fr>
LIGM, École des Ponts, Univ. Gustave Eiffel, Marne-la-Vallée, France

Workshop Open Science, Bilbao, Nov. 10-12, 2022
Some definitions

**Repeatability and replicability**
Capacity to perform the same experiment as many times as needed.

→ **Repeatability**: Same team, same experimental setup
→ **Replicability**: Different team, same experimental setup

**Example:** is distilled water electrically conductive? Is salt water conductive? We can perform the experiment many times and get results (https://www.dailymotion.com/video/x21cg6a).
Repeatability and replicability

Capacity to perform the same experiment as many times as needed.
→ **Repeatability**: Same team, same experimental setup
→ **Replicability**: Different team, same experimental setup

**Example**: is distilled water electrically conductive? Is salt water conductive? We can perform the experiment many times and get results ([https://www.dailymotion.com/video/x21c6ga](https://www.dailymotion.com/video/x21c6ga)).

Reproducibility

Capacity to obtain the same results when repeating an experiment by following a detailed procedure
→ Different team, different experimental setup

**In computational sciences** (deterministic code, digital data): results obtained by following a detailed and correct pseudo-code description must coincide if the same input data is provided.
**Repeatability Examples**

<table>
<thead>
<tr>
<th>Repeatable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obtaining the classification results with a <strong>neural network</strong>. We can <strong>repeat</strong> the experiment as many times as we want. We just need the weights of the network and the input data.</td>
</tr>
</tbody>
</table>

![Diagram of a neural network](image)
Repeatability Examples

**Repeatable**

Obtaining the classification results with a neural network. We can repeat the experiment as many times as we want. We just need the weights of the network and the input data.

**Not repeatable:**

Detection of the merger of two black holes from gravitational waves. We can’t repeat the experiment as needed.
Reproducibility Examples

Reproducible:
Given:
- a detailed pseudo-code (or the source code itself),
- any associated learning or initialization data,
- the input data,
we should obtain exactly the same results each time we run the algorithm.
⇒ Exactly the same denoised image, classification results, etc.
Reproducibility Examples

Reproducible:
Given:
- a detailed pseudo-code (or the source code itself),
- any associated learning or initialization data,
- the input data,
we should obtain exactly the same results each time we run the algorithm.
⇒ Exactly the same denoised image, classification results, etc.

Not reproducible
In a paper that shows
- a pseudo-code without all the details, or its initialization,
- the source code is not available,
- neither the learning data,
other researchers can’t compare with the proposed method.
⇒ We can’t be sure about anything on the method, nor test it with our own data.
• **Non-exact sciences** (biology, medicine, ...): difficult (but *desirable*). Hard to have exactly the same *conditions* along experiments.

• Computational sciences: **no excuse**!

Why are we not all doing reproducible research?
Implementation of Reproducible Research

- **Non-exact sciences** (biology, medicine, ...): difficult (but desirable). Hard to have exactly the same conditions along experiments.
- Computational sciences: no excuse!

Why are we not all doing reproducible research?

### Several reasons in general:

- Some researchers don’t want to make public working code
  - doesn’t correspond to any version of the pseudo-codes,
  - low software quality,
  - quality software takes more time to produce: testing, documentation, objective quality metrics.
- Results of the method do not generalize
- ... (For the discussion later!)
Implementation of Reproducible Research

• Non-exact sciences (biology, medicine, ): difficult (but desirable). Hard to have exactly the same conditions along experiments.
• Computational sciences: no excuse!

Why are we not all doing reproducible research?

Several reasons in general:

• Some researchers don’t want to make public working code
  • doesn’t correspond to any version of the pseudo-codes,
  • low software quality,
  • quality software takes more time to produce: testing, documentation, objective quality metrics.

• Results of the method do not generalize

• . . . (For the discussion later!)

Not really considered for career advance

• Classic metrics: “number of high impact-factor classic publications”
• Software is considered as a 2nd class citizen
Reproducible Research Platforms

Different types of platforms

- Online execution platforms.
- Dissemination platforms.
- Peer-reviewed journals.

- **Galaxy** - https://galaxyproject.org
- **IPython** - https://ipython.org
- **Jupyter** - http://jupyter.org
- **Code Ocean** - https://codeocean.com
- **DAE** - http://dae.cse.lehigh.edu/DAE
- **IPOL** - https://www.ipol.im
- **Research Compendia** - ResearchCompendia.org
- **MLOSS** - https://mloss.org/software
- **DataHub** - https://datahub.io/
- **PaperWithCode** - https://paperswithcode.com

- **JOSS Journal** - https://joss.theoj.org
- **Insight J Journal** - https://insight-journal.org
Reproducible Research Platforms

Different types of platforms

- Online execution platforms.
- Dissemination platforms.
- Peer-reviewed journals.

- Galaxy - https://galaxyproject.org
- IPython - https://ipython.org
- Jupyter - http://jupyter.org
- Code Ocean - https://codeocean.com
- DAE - http://dae.cse.lehigh.edu/DAE
- IPOL - https://www.ipol.im
- Research Compendia - ResearchCompendia.org
- MLOSS - https://mloss.org/software
- DataHub - https://datahub.io/
- PaperWithCode - https://paperswithcode.com
- ReScience Journal - http://rescience.github.io
- JOSS Journal - https://joss.theoj.org
- Insight J Journal - https://insight-journal.org
Reproducible Research Platforms

Different types of platforms

- Online execution platforms.
- Dissemination platforms.
- Peer-reviewed journals.

- **Galaxy** - https://galaxyproject.org
- **IPython** - https://ipython.org
- **Jupyter** - http://jupyter.org
- **Code Ocean** - https://codeocean.com
- **DAE** - http://dae.cse.lehigh.edu/DAE
- **IPOL** - https://www.ipol.im
- **Research Compendia** - ResearchCompendia.org
- **MLOSS** - https://mloss.org/software
- **DataHub** - https://datahub.io/
- **PaperWithCode** - https://paperswithcode.com

- **JOSS Journal** - https://joss.theoj.org
- **Insight J Journal** - https://insight-journal.org
• Started in 2009 under the initiative of Nicolas Limare and Jean-Michel Morel (ENS Paris Saclay).

• A journal initially targeting image processing (Image Processing On Line)

• Some other data types were added: video, audio, 3D data...

• Even some articles on SARS-CoV-2 evolution!
  “A Daily Measure of the SARS-CoV-2 Effective Reproduction Number for all Countries” http://www.ipol.im/pub/art/2020/304/

• Today it it a general journal on reproducible algorithms
  → Information Processing On Line
Peer-reviewed

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.
Peer-reviewed

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

Each publication:

- A text describing the method in detail, including pseudo-codes.
- The source code, under an open-source software license.
- An online demo which allows users to test the method with their own data.
- An archive of experiments.
**Peer-reviewed**

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

<table>
<thead>
<tr>
<th>Each publication:</th>
</tr>
</thead>
<tbody>
<tr>
<td>- A text describing the method in detail, including pseudo-codes.</td>
</tr>
<tr>
<td>- The source code, under an open-source software license.</td>
</tr>
<tr>
<td>- An online demo which allows users to test the method with their own data.</td>
</tr>
<tr>
<td>- An archive of experiments.</td>
</tr>
</tbody>
</table>

- No need to be an original work. We’re interested in the math details, reproducibility, and understanding.
- ISSN, DOI, indexed by SCOPUS. Not yet an "Impact Factor".
**Peer-reviewed**

- Both the article (PDF) and the source code.
- Reproducibility: the reviewers check carefully that the source code matches the pseudo-code.

**Each publication:**

- A text describing the method in detail, including pseudo-codes.
- The source code, under an open-source software license.
- An online demo which allows users to test the method with their own data.
- An archive of experiments.

- No need to be an original work. We’re interested in the math details, reproducibility, and understanding.
- ISSN, DOI, indexed by SCOPUS. Not yet an "Impact Factor".

Multi-Scale DCT Denoising
Nicola Piazzero, Jean-Michel Morel, Gabriele Facciolo

Abstract

DCT denoising is a classic low complexity method built on the JPEG compression norm. Once made translation invariant, this algorithm was still known to be competitive at the beginning of this century. Since then, it has been outperformed by patch based methods, which are far more complex. This paper proposes a two-step multi-scale version of the algorithm that boosts its performance and reduces its artifacts. The multi-scale strategy decomposes the image in a discrete DCT pyramid, which keeps noise white at all scales. The single scale denoising is then applied to all scales, thus giving multichannel versions of the low frequency coefficients of the denoised image. A multi-scale fusion of these multiple estimates avoids the ringing artifacts resulting from the general reconstruction. The final algorithm shows a good PSNR and much improved visual image quality. It is shown to have a detail of only 1dBs with respect to state of the art algorithms, but its complexity is two orders of magnitude lower.

Download

* full text online: PDF (13.7MB) XPDF (16.2MB) F
* source code: ZIP (298MB) HEX

Preview

Leading lines are often used in the context of images, to guide the reader's eye through the document. Here, we show how to use a low resolution PDF document for a high resolution overview.

For the hard thresholding pass of the algorithm the aggregation weights are set, as in [3], by counting the number of non-zero DCT coefficients (excluding the zero frequency) in the patch after thresholding. These aggregation weights are then given by

\[(1 + N_{0})^{-1},\]

where the one is added to prevent the dividing by zero (but it is an arbitrary choice). Indeed, the number of non-zero coefficients will be small for the flat patches, compared to patches containing

LOW RESOLUTION PDF images may show compression artifacts. A full resolution PDF is available at www.ipol.im.
IPOL demo
Multi-Scale DCT Denoising

Please cite the reference article if you publish results obtained with this online demo.

2739 public experiments since 2017-01-15

This archive is not moderated. In case you uploaded images that you don’t want that appear in the archive, please contact the editor in charge. In case of copyright infringement or similar problems, please contact us to request the removal of some images. Some archived content may be deleted by the editorial board for size matters, inadequate content, user requests, or other reasons.

Experiment #507897.
2022-05-28 02:30:13 UTC

Parameters
- sigma: 5
- ps2: 8
- scales: 5
- reflector: 0.4

Files: Output text

Experiment #507898.
2022-05-28 02:50:43 UTC

Parameters
- sigma: 10
- ps2: 5
- scales: 5
- reflector: 0.4

Files: Output text
Benefits of RR

If a method is worth it, the impact is large

- Users worldwide can test the algorithm with their own data.
- Increase in the number of citations: other researchers can now compare to you.
# Benefits of RR

## If a method is worth it, the impact is large

- Users worldwide can test the algorithm with their **own data**.
- Increase in the number of citations: other researchers can now **compare** to you.

## Scientific acceleration:

- other researchers can reuse text, source code, data
Benefits of RR

If a method is worth it, the impact is large

- Users worldwide can test the algorithm with their own data.
- Increase in the number of citations: other researchers can now compare to you.

Scientific acceleration:

⇒ other researchers can reuse text, source code, data

Useful to show a landscape of our scientific activity
Main attention points:

- Consider **source code** as **part of the publication**, not supplementary material.
- Different **levels of evaluation**:
  - **Lowest**: black box (same inputs same outputs)
  - ...,
  - **Highest**: deep understanding of the method and checking that the source code matches the implementation faithfully.
Main attention points:

- Consider source code as part of the publication, not supplementary material
- Different levels of evaluation:
  - Lowest: black box (same inputs same outputs)
  - ... 
  - Highest: deep understanding of the method and checking that the source code matches the implementation faithfully.

Difficulties

- Software is not easy to review.
- Many researchers are not software engineers!

⇒ A possible solution (IPOL): use at least two reviewers, one of them being an expert reading source code.
Main attention points:

- Consider **source code** as part of the publication, not supplementary material.
- Different **levels of evaluation**:
  - Lowest: black box (same inputs same outputs)
  - ...
  - Highest: deep understanding of the method and checking that the source code matches the implementation faithfully.

Difficulties

- **Software** is not easy to review.
- Many researchers are not software engineers!

⇒ A possible solution (IPOL): use at least two reviewers, one of them being an expert reading source code.

Special case of neural networks

Focus on the **architecture**, **training**, **understanding**, and **generalization**.
• Four editors in chief: Luis Alvarez (Univ. Gran Canaria), PM, Jean-Michel Morel (ENS Paris Saclay), Gregory Randall (Univ. Montevideo)
• EiCs decide if the submission looks interesting.
• EiCs name an associate editor for the submission.
• The editor chooses reviewers and a demo editor.
• Reviewers may be asked to check different aspects: article, code, demo.
• After acceptation, an EiC checks the article and plays with the demo, testing with different input data and parameters

Published articles
Currently about 15–20 accepted submissions per year.
<table>
<thead>
<tr>
<th>TITRE</th>
<th>CITÉE PAR</th>
<th>ANNÉE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSD: a line segment detector</td>
<td>720</td>
<td>2012</td>
</tr>
<tr>
<td>RG Von Gioi, J Jakubowicz, JM Morel, G Randall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal Image Processing On Line 2, 35-55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Local Means Denoising</td>
<td>691</td>
<td>2011</td>
</tr>
<tr>
<td>A Buados, B Coll, JM Morel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal : Image Processing On Line 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TV-L1 Optical Flow Estimation</td>
<td>394</td>
<td>2013</td>
</tr>
<tr>
<td>J Sánchez, E Meinhardt-Llopis, G Facciolo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal : Image Processing On Line 3, 137-150</td>
<td></td>
<td></td>
</tr>
<tr>
<td>An Analysis of the Viola-Jones Face Detection Algorithm</td>
<td>390</td>
<td>2014</td>
</tr>
<tr>
<td>YQ Wang</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asift: An algorithm for fully affine invariant comparison</td>
<td>355</td>
<td>2011</td>
</tr>
<tr>
<td>G Yu, JM Morel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal : Image Processing On Line 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>An Analysis and Implementation of the BM3D Image Denoising Method</td>
<td>350</td>
<td>2012</td>
</tr>
<tr>
<td>M Lebrun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal Image Processing On Line 2, 175-213</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiscale Retinex</td>
<td>266</td>
<td>2014</td>
</tr>
<tr>
<td>AB Petro, C Sbert, JM Morel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal : Image Processing On Line 4, 71-88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rudin-Osher-Fatemi total variation denoising using split Bregman</td>
<td>263</td>
<td>2012</td>
</tr>
<tr>
<td>P Getreuer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal : Image Processing On Line 2, 79-95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan-vese segmentation</td>
<td>189</td>
<td>2012</td>
</tr>
<tr>
<td>P Getreuer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPOL Journal : Image Processing On Line 2, 214-224</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horn-schunck optical flow with a multi-scale strategy</td>
<td>157</td>
<td>2013</td>
</tr>
<tr>
<td>E Meinhardt-Llopis, JS Párez, D Kondermann</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Thank you for your attention