Experiences with Reproducible Research in Various Facets of Signal Processing Research

P. Vandewalle, G. Barrenetxea, I. Jovanovic, A. Ridolfi and M. Vetterli
Ecole Polytechnique Fédérale de Lausanne (EPFL)

April 18, 2007

ICASSP 2007

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Introduction

Definition:

"*Reproducibility is one of the main principles of the scientific method, and refers to the ability of a test or experiment to be accurately reproduced, or replicated, by someone else working independently*."

- Wikipedia -

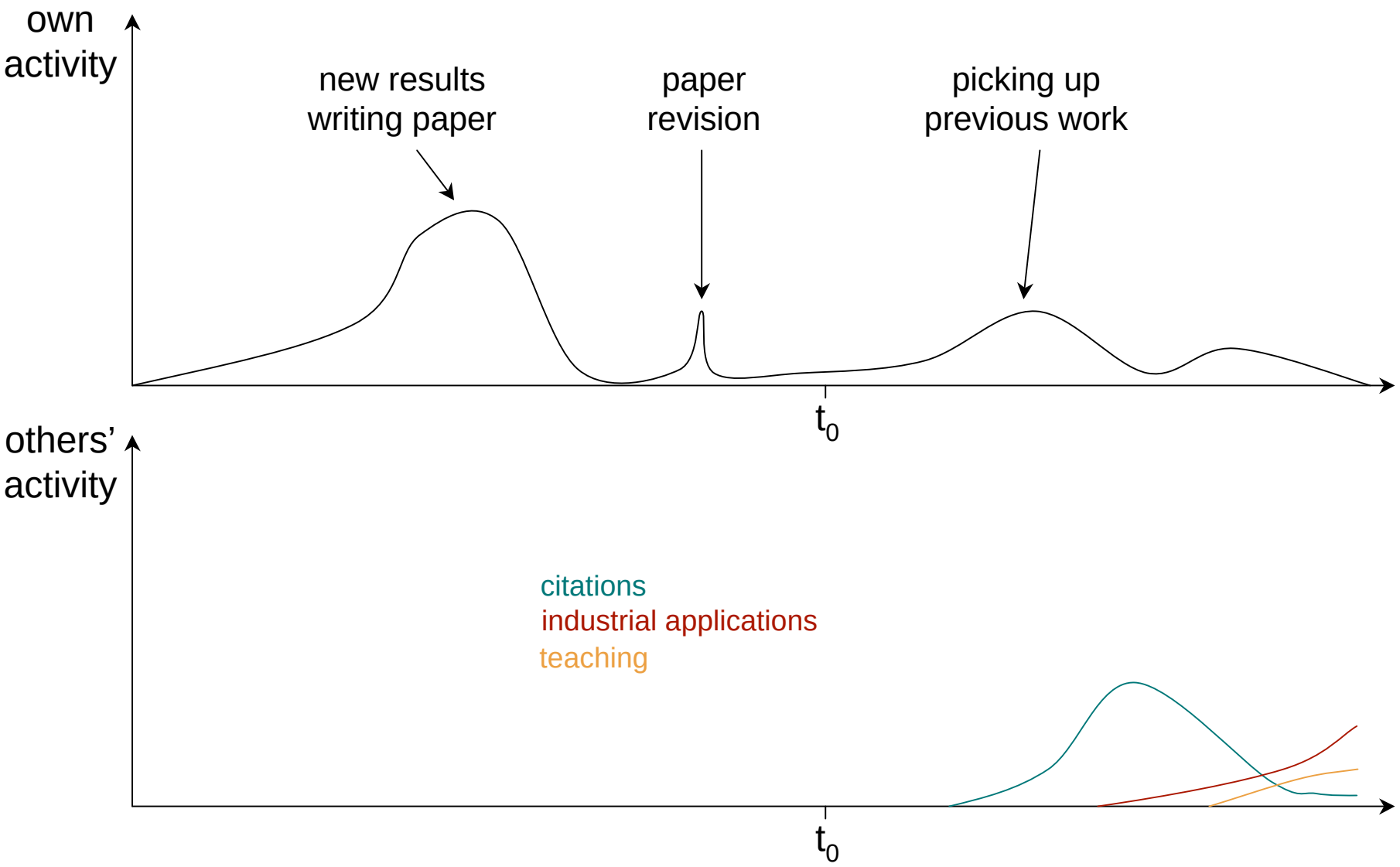
Different parts needed for reproducibility:

- Publication
- Data
- Code

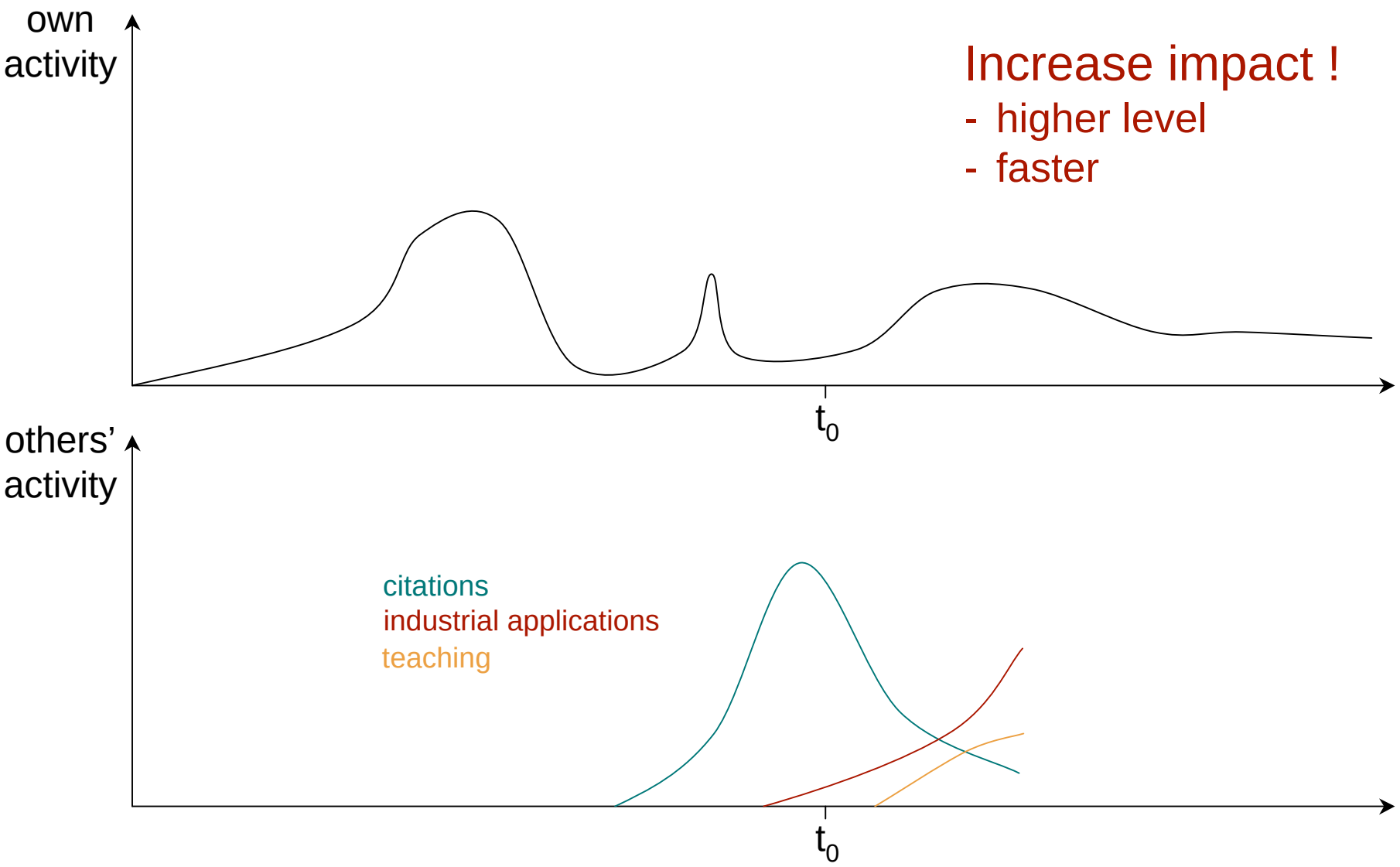
Advantages:

- Allows you to continue where you left work
- Allows others to start from the same point
- Allows others to use your work as a building block

Introduction



Introduction



Introduction

Some examples:

"I just read your paper X. It is very completely described, however I am confused by Y. Could you provide the implementation code to me for reference if possible?"

"Comment from a reviewer about a paper: Please replace Figure 3 by a higher quality version. Also, the horizontal axis should be labeled 't' instead of 'x'."

"Hi! I am also working on a project related to X. I have implemented your algorithm, but cannot get the same results as described in your paper. Which values should I use for parameters Y and Z?"

Related Work - Other Domains

- Mathematics
  - ‘Reprovable’, proofs can be verified
  - Not reproducible: Fermat’s last theorem

    “J'ai trouvé une merveilleuse démonstration de cette proposition, mais je ne peux l'écrire dans cette marge car elle est trop longue.”
- Exact Sciences (Physics, Chemistry, etc.)
  - Experimental setups
- Life Sciences
  - Experimental setups
  - Other researchers repeat experiments
  - Journal of Cell Biology checks for image manipulation

➔ We have a lot to learn!

Related Work - History of Reproducible Research

- Knuth (1984)
  - Literate Programming

    "Instead of imagining that our task is to instruct a computer *what to do, let us concentrate rather on explaining to* human beings *what we want a computer to do."*
- Claerbout @ SEP, Stanford (1990)
  - Makefiles to build and clean results
- Donoho @ Statistics Department, Stanford (1995)
  - WaveLab using Matlab
- Various other domains
  - Econometrics [Koenker96, Vinod01]
  - Neurophysiology [Pouzat05]
  - Epidemiology [Peng06]
  - Signal Processing

➔ Now is the right time to start Reproducible Research!

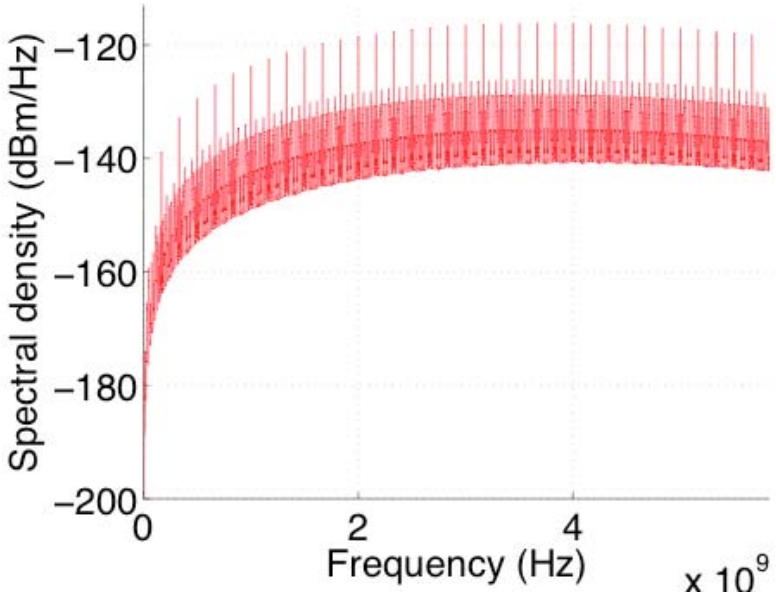
Parts of Reproducible Research - Theory

Theory

- Well described theorems and proofs
- Numerical simulations are very helpful

Example: Spectrum of Ultrawide Bandwidth Signal

$$S_X(\nu) = |\widehat{w}(\nu)|^2 \lambda \left(2\mathrm{Re} \left\{ \sum_{k \geq 0} \phi_S^k (2\pi\nu) \right\} d\nu - 1 - \lambda\delta(\nu) \right) .$$

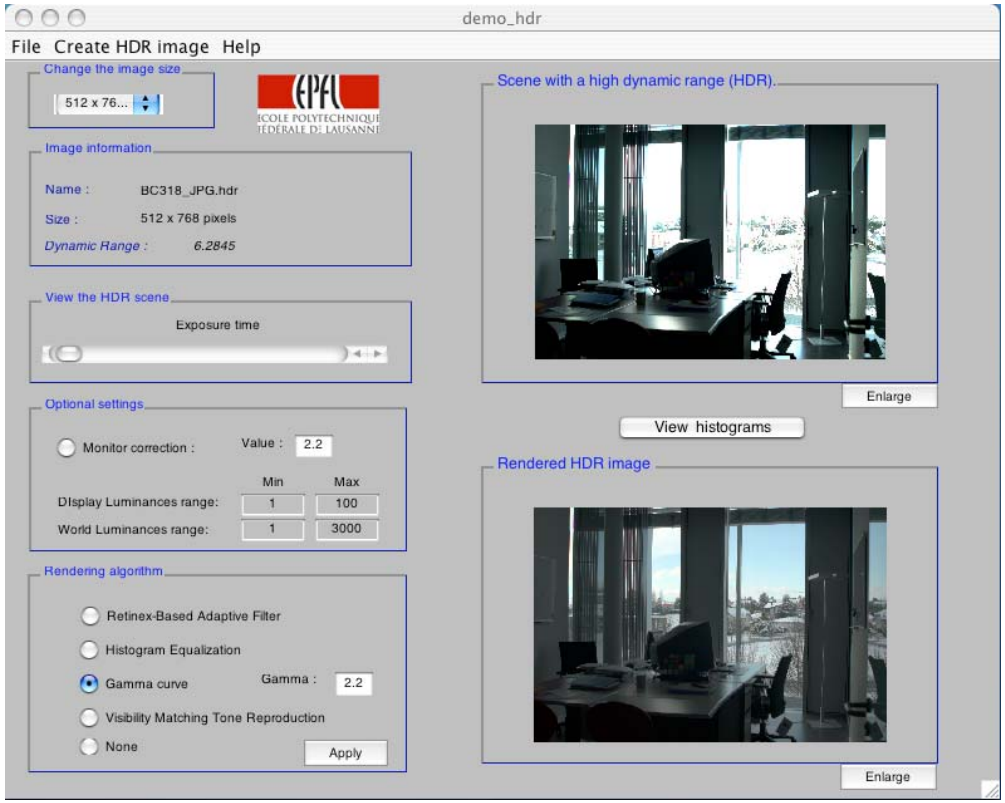


Andrea Ridolfi and Moe Z. Win, *Ultrawide Bandwidth Signals as Shot-Noise: a Unifying Approach*, IEEE Journal on Selected Areas of Communications, Vol. 24, Nr. 4, pp. 899-905, 2006.

Parts of Reproducible Research - Algorithms

Algorithms

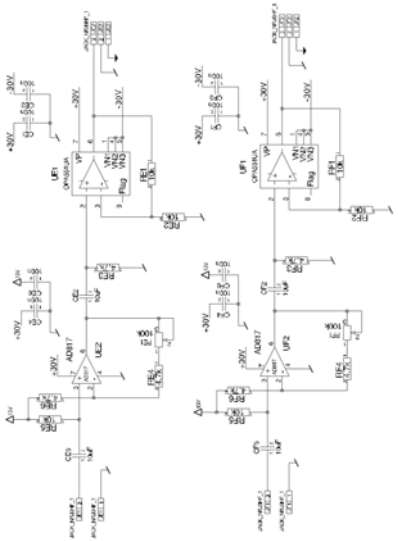
- Code
- Data
- Environment
  - Computer platform
  - Compiler and compiler flags
  - Software version
  - User interface, if available



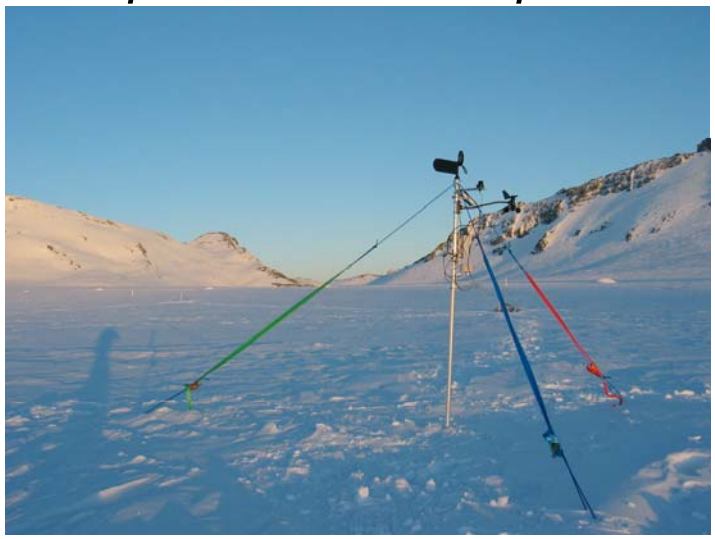
Parts of Reproducible Research - Data and Experimental Setups

Data and Experimental Setups

- Measurement setup
- Setup or calibration procedures
- Complex setups: reusable data sets



Examples: SensorScope,Acoustic Tomography

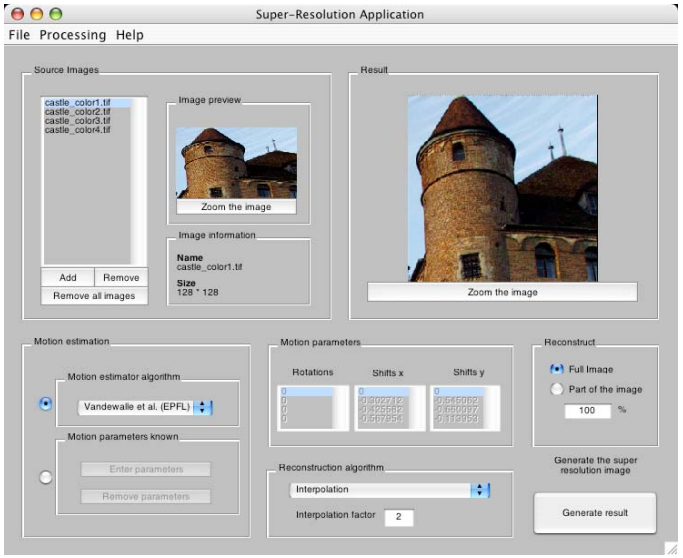


G. Barenetxea, O. Couach, M. Krichane, T. Varidel, S. Mortier, J. Mezzo, M. Bystranowski, S. Dufey, H. Dubois-Ferrière, J. Selker, M. Parlange and M. Vetterli, *SensorScope: An Environmental Monitoring Network*. In AGU Conference, 2006.

Ivana Jovanovic, Luciano Sbaiz, and Martin Vetterli, *Acoustic tomography for estimating temperature and wind flow*, 13th International Symposium for the Advancement of Boundary Layer Remote Sensing, ISARS, pp. 69-71, 2006.

Example: Super-Resolution Imaging

- First and second conference paper - 2003
  - Not reproducible
- First journal paper - April 2004
  - Let's make it reproducible…
  - 1 week work
  - Paper with Matlab code, data, figures online
- …
- Second journal paper - Dec 2005
  - Reproducible, of course!
  - 1 day work
  - Paper with Matlab code, data, and figures

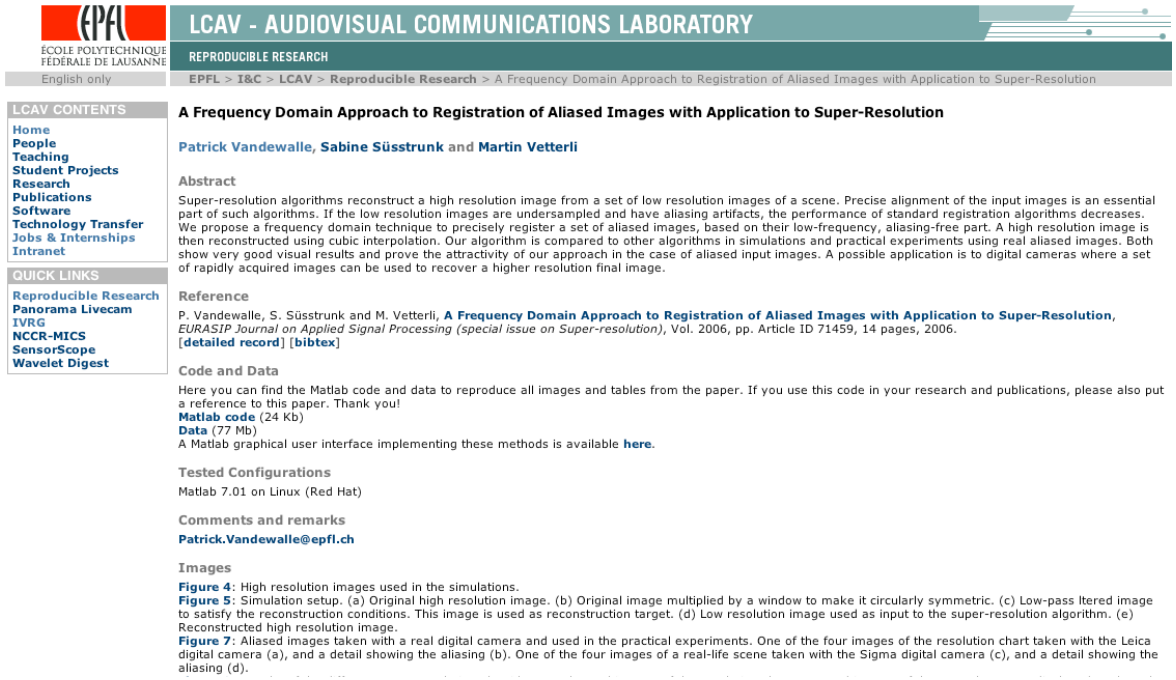


- In between (summer 2005), student wrote GUI for Matlab code
- Writing PhD thesis
  - Great, I can recycle my code!
  - Easy to create new figures by changing parameters and axis labels
  - 3 months to write thesis

Example: Super-Resolution Imaging

Web page for each paper, containing

- Full reference + PDF
- Abstract
- Code
- Data
- Tested configurations
- Contact address
- (Figures)
- (References)



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

English only

LCAV - AUDIOVISUAL COMMUNICATIONS LABORATORY

REPRODUCIBLE RESEARCH

EPFL > I&C > LCAV > Reproducible Research > A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution

LCAV CONTENTS

Home
People
Teaching
Student Projects
Research
Publications
Software
Technology Transfer
Jobs & Internships
Intranet

QUICK LINKS

Reproducible Research
Panorama Livecam
IVRG
NCCR-MICS
SensorScope
Wavelet Digest

A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution

Patrick Vandewalle, Sabine Süsstrunk and Martin Vetterli

Abstract

Super-resolution algorithms reconstruct a high resolution image from a set of low resolution images of a scene. Precise alignment of the input images is an essential part of such algorithms. If the low resolution images are undersampled and have aliasing artifacts, the performance of standard registration algorithms decreases. We propose a frequency domain technique to precisely register a set of aliased images, based on their low-frequency, aliasing-free part. A high resolution image is then reconstructed using cubic interpolation. Our algorithm is compared to other algorithms in simulations and practical experiments using real aliased images. Both show very good visual results and prove the attractivity of our approach in the case of aliased input images. A possible application is to digital cameras where a set of rapidly acquired images can be used to recover a higher resolution final image.

Reference

P. Vandewalle, S. Süsstrunk and M. Vetterli, **A Frequency Domain Approach to Registration of Aliased Images with Application to Super-Resolution**, *EURASIP Journal on Applied Signal Processing (special issue on Super-resolution)*, Vol. 2006, pp. Article ID 71459, 14 pages, 2006.
[**detailed record**] [**bibtex**]

Code and Data

Here you can find the Matlab code and data to reproduce all images and tables from the paper. If you use this code in your research and publications, please also put a reference to this paper. Thank you!
Matlab code (24 Kb)
Data (77 Mb)
A Matlab graphical user interface implementing these methods is available **here**.

Tested Configurations

Matlab 7.01 on Linux (Red Hat)

Comments and remarks

Patrick.Vandewalle@epfl.ch

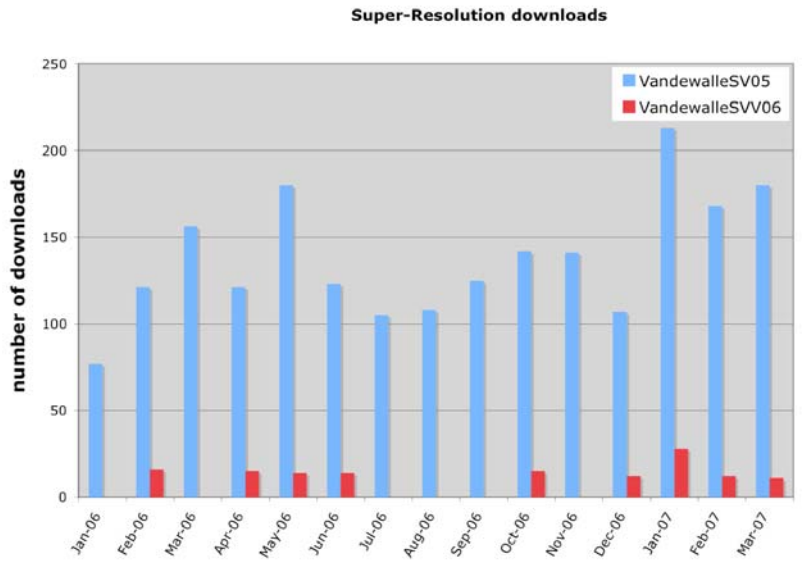
Images

Figure 4: High resolution images used in the simulations.
Figure 5: Simulation setup. (a) Original high resolution image. (b) Original image multiplied by a window to make it circularly symmetric. (c) Low-pass ltered image to satisfy the reconstruction conditions. This image is used as reconstruction target. (d) Low resolution image used as input to the super-resolution algorithm. (e) Reconstructed high resolution image.
Figure 7: Aliased images taken with a real digital camera and used in the practical experiments. One of the four images of the resolution chart taken with the Leica digital camera (a), and a detail showing the aliasing (b). One of the four images of a real-life scene taken with the Sigma digital camera (c), and a detail showing the aliasing (d).

Example: Super-Resolution Imaging

Benefits:

- Efficient reuse of my own results
- Nice demo material (using GUI)
- Many downloads

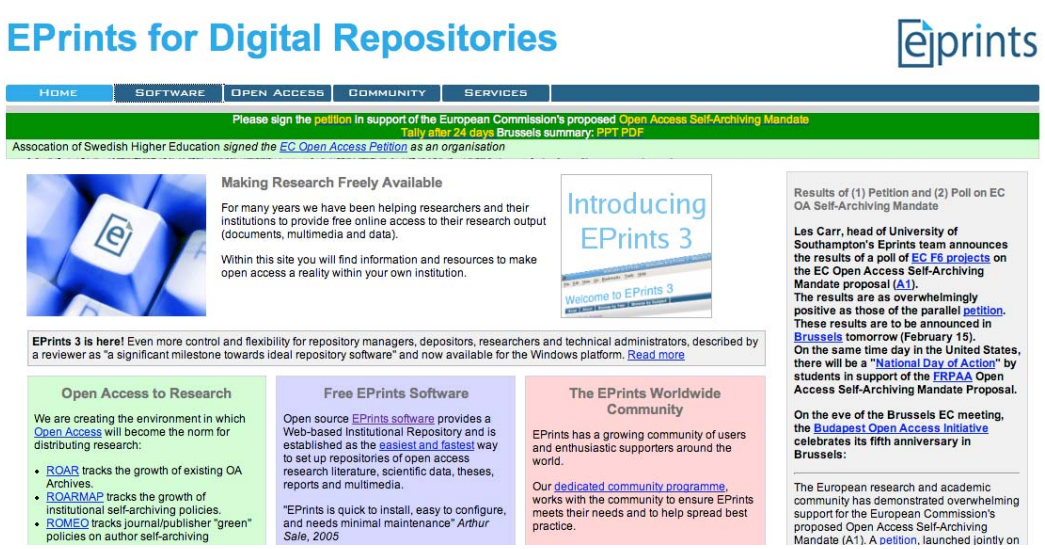


- Nice reactions
  - “When I am beginning to study image super-resolution, I read a lot of papers about it, but I found that I don't know how to realize the algorithms of these paper.Your works help me make great progress. Hope to have more super-resolution algorithms code for us to study. Thanks.”
  - “We're interested in improving the algorithm via our own methodology in registration. We're using your methods as a springboard of research for use in security applications.”
- Collaborations

Reproducible Paper Repository

How to make those data available?

- Everyone creating own web pages with more/less information
- Centralized repository with reproducible papers (cfr ArXiv)
- Distributed repository setup with standardized fields
  - Using EPrints server (www.eprints.org) and configuration files
  - Every lab/institution can build own database
  - Compatibility
  - We are currently building a setup jointly with EPrints



Licensing and Commercialization

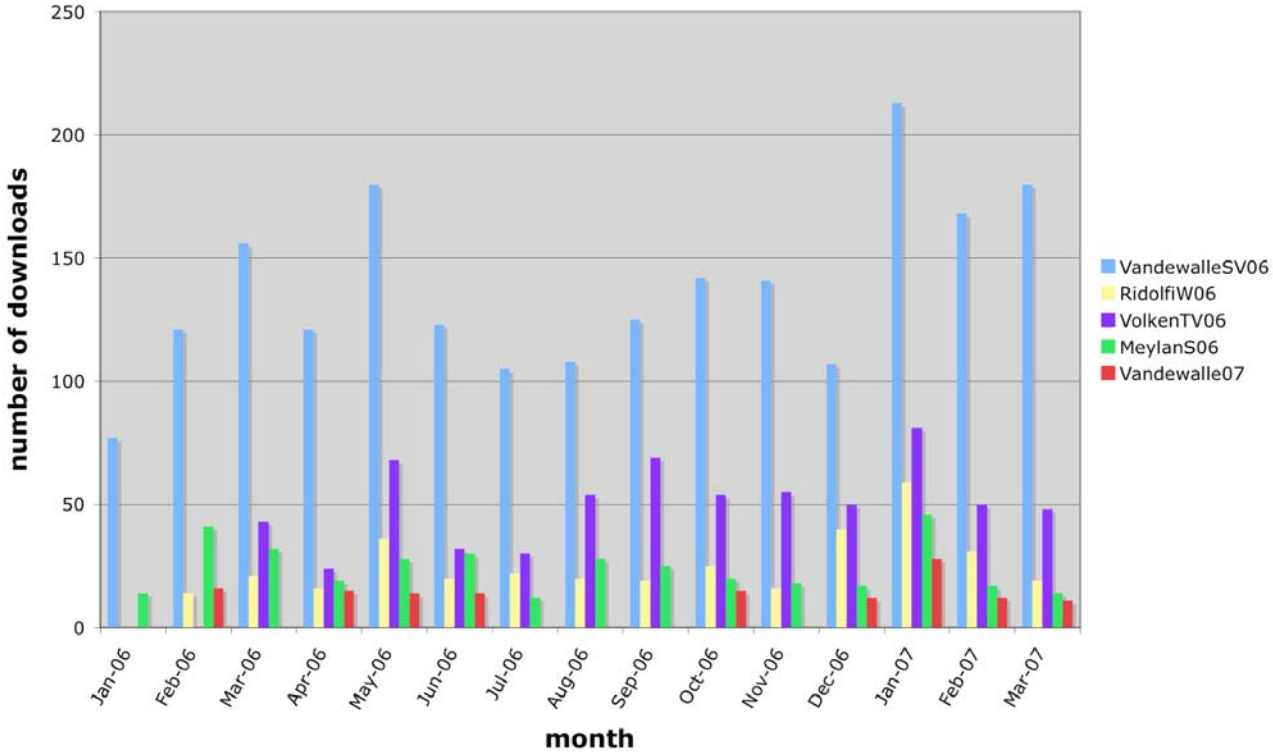
- License defines how your code can be used and distributed
- License has to be included with the code
- Many types of licenses exist
  - Open Source
    - GNU General Public License (GPL)
    - BSD License
    - MIT License
    - Common Public License (CPL)
    - … (58 listed currently on http://www.opensource.org/licenses/)
  - Proprietary Source

What if my start-up has to live from this?

- Special setup: website where people can submit data to be processed
- User can choose his own data and receives the results after processing
- Not 100% reproducible by user, but enough for comparisons and tests on other data
- Additional advantage: create large test database

Impact

- Papers available online are cited 3 x more often [Lawrence]
- Increased visibility
- More than 100 downloads/month for super-resolution code
- Red-eye removal paper - with online Java code - most popular download in EPFL database



Conclusions

- Reproducible research increases impact
- It helps both yourself and other people
- Start reproducible research now!

More information:

Have a look at

http://lcavwww.epfl.ch/reproducible_research

Or contact me:

Patrick.Vandewalle@epfl.ch