# Curriculum vitæ

## MATTHIJS DOUZE

E-mail: matthijs.douze@inria.fr
Phone: +33 (0) 6 61 91 81 01 (mobile)
Address: 14 rue Bizanet, 38000 Grenoble, France
Homepage: http://lear.inrialpes.fr/~douze
Born in 1977, Dutch
Living with my partner and my kid (born in 2013)

## Education and employment

**2005-current: Research engineer** at INRIA Grenoble, permanent position since 2010. Main topics:

- content-based image retrieval: indexing based on inverted file representations or global descriptors.
- nearest-neighbor search in Euclidean space, product quantization application to compressed-domain image classification.
- video copy detection: using bag-of-words or global video representations
- action analysis in videos, video structuring
- 3D reconstruction with a multi-camera studio (kinovis.inrialpes.fr), n-polyhedron CSG operations.

**2004:** consultant for the Fontès Réfractaires company (Revel, France). Vision-based quality control at the end of a brick production line.

**2000-2004:** IRIT (Toulouse, France www.irit.fr)
**Ph.D.** , "Estimation of inter-image homographies to generate mosaics and track in real time – applications in augmented reality". Ph.D. defended in December 2004. Real-time tracking and moasicing, inter-linked video navigation methods.

**1999-2000:** Paul Sabatier University (Toulouse)
Master degree in Imaging and Language processing. Specialisation in numerical analysis of images. Awarded a PhD grant from the CNRS (the main French research organisation).

**1997-2000:** ENSEEIHT (engineering school in Toulouse www.enseeiht.fr)
Master degree (ranked 5th among 100) in computer science and applied mathematics. Programming, scientific computing, system administration, databases.

# Main activities

As a research engineer, my activities are: implementing and evaluating research algorithms, making systems efficient at a very large scale, contributing to research papers, building demos, participating in evaluation campaigns, designing datasets, preparing code for release. I also conduct independent research, mainly on large-scale video search and geometric algorithms.

## Challenges and evaluation campaigns

To benchmark our research contributions, I was the core developer or main co-developer for the following international challenges:
- 2005: "where am I" contest (in conjunction with ICCV) – ranked 5th out of 15 participants.
- 2006: **winner** of Pascal VOC contest, ranked 1st on 4 out of 10 classes for the detection task
- 2008: **winner** of Trecvid video copy detection, out of 22, ranked 3rd in 2010 and 2011
- 2009: **winner** of the ImageCLEF Photo retrieval task with the image-only modality
- 2012: **winner** of Trecvid multimedia event detection for all 100-ex runs, out of 14 participants.
- 2013: **winner** of Trecvid multimedia event detection for all 100-ex runs, out of 17 participants. Also ranked 1st for 1 out of 4 configurations in 2014.

## Demos

I was the main developer of the following demonstrators for research results:
- Photomole: real-time image retrieval from a 500-image dataset, with a webcam (ICCV 2005).
- `bigimbaz.inrialpes.fr`: web-based image retrieval demo in 10 million images (2008)
- Bigvidbaz: video search in a set of soccer and rugby matches (Raffut demo, 2009)
- Bigimbaz mobile: 10M images indexed on a laptop (**best demo** award at RFIA 2010)
- Twinpics: similarity search on 100M images, developed together with Exalead (2010)
- Vilnius musical performance: **distributed concert** with a dancer reconstructed in 3D (ICT 2013).

## Software

This is a list of software packages I contributed to. The percentages indicate my contribution.
Transferred software (licenses sold by INRIA to companies):
- Bigimbaz (50%): inverted-file based image search engine.
- Obsidian (40%): local and global descriptor computation library.
- Pqcodes (30%): nearest-neighbor search library.
Open-source software:
- Yael (50%): efficient primitives for clustering, PCA, partial sorting, etc. Downloaded 600 times per month.
- JSGD (80%): training of linear multiclass classifiers with a stochastic gradient descent algorithm.
- MaskFishDet (20%): object class detection in images.
- FlowVR (10%): middleware for distributed interactive computations.
I also developed small demonstration and evaluation packages for research methods or datasets.

## Grants/projects

I have been involved in the following French (ANR, Oseo) and European (EU) projects:
- AceMedia (EU, 2005-2007): integrated modules for person detection and texture analysis. Developed an image organization demo (Yorg).
- Quaero (Oseo, 2008-2013): toolset for multimedia analysis. Main participant for the LEAR team: developed joint demonstrators with other partners, technology evaluations, reporting, organized meetings. Presentation of some of the project's achievements at Cap Digital 2013 in Paris.
- AXES (EU, 2012-2015): indexing audio-visual archives. Integrated image and video indexing tools in a common platform, then supervised the engineers in charge of AXES.
- Kinovis (ANR, 2013-2019): multi-camera 3D reconstruction studio (budget: 2.3 ME). **Co-author of the proposal**, set-up of the studio, supervision of the engineering team, software architect for the real-time 3D reconstruction software, organization of the inauguration and associated workshop (May 2015).

## Teaching & supervision

- 2000-2004 (PhD): I taught to master degree students in various fields (300 hours in total).

- 2008-2015: main teacher of the MSc multimedia database course at the ENSIMAG and ENSEEIHT engineering schools.
- 2011-2014: training sessions for researchers, on code wrapping and numerical programming with Python.
- 2013: presentation on research software architecture at Devexp, Lausanne.

I co-advised 15 internship students, engineers, PhDs (Mattis Paulin, Danila Potapov, Dan Oneata) and post-docs (Jérôme Revaud, Arnau Ramisa) on the technical aspects of their work. I participate in the **recruitment process** of PhDs and staff, analyzing CVs and doing the technical interviews.

## Industrial relations

- Involved in the starting phase of the Milpix company (2007-2013), that commercially exploited our Bigimbaz image search engine.
- Co-author (10%) of a **patent** "Dispositif d'aide à la reconnaissance d'image amélioré", INPI patent 08/03345, filled June 16th 2008. US version: "Assistance device for image recognition", US 2011/0164822 A1, Filled Jun 12th 2009.
- Participation in numerous evaluations of our software for industrial application cases (with Exalead, EADS, OMBLabs, etc.), some of them resulted in technology transfers.
- Presentation of our image and video indexing techniques at the MPEG standardization Workshop on Visual Search (2010, Geneva).

## Computer science skills

I installed a **50-node cluster**, its disk arrays, GPUs and batch scheduling system. I programmed a lot, both at work and through personal exploration:
- Languages: imperative languages (**C**, Fortran); object-oriented conception (**C++**, Java); scripting (**Python**, bash); assembly languages (IA-32, PowerPC); functional languages (Caml).
- Graphical components: 2D software (**GIMP**, Inkscape); 2D description languages (**PostScript**, PDF, SVG); 3D libraries (**OpenGL** + GLSL); 3D scene description languages (Povray, VRML); GUI librarires on Unix (Qt, **gtk**), MacOS (Cocoa).
- Numerical processing: BLAS, **LAPACK**, interfacing of various numerical libraries in Fortran, prototyping with Matlab/Octave and Maple, code optimization on SIMD vector units (Altivec, SSE, AVX) and with multithreading on multiprocessors or multi-core processors with pthreads, **OpenMP** and Intel TBB. **OpenCL** and Cuda on the GPU, distributed computations over several machines in a cluster with scripts, the OAR batch scheduler and the MPI communication API.
- Computer vision: OpenCV and deep neural networks with **Caffe** and cuda-convnet.
- Text/XML processing with Python, sed and awk. Documents in LaTeX and HTML. Code management with SVN or git and documentation with Doxygen. C/C++ code wrapping with **SWIG** or cython. GUI- and system-level applications for the Sharp Zaurus PDA. Dynamic web pages with AJAX and HTML5 on the client side and Python on the server side.

## Research activities

The publications where I was the first contributor are mostly on large-scale video indexing. For the others, I often designed the experiments or implemented large-scale extension of the methods. The research papers I co-authored received **3700 citations** in total. I **reviewed** for major journals (IJCV, PAMI, CVIU, etc.), and conferences (CVPR, ICCV, ECCV, ACM multimedia, etc.).

**Selected publications (main conferences and journals of the field)**

[1] **Matthijs Douze**, Jérôme Revaud, Cordelia Schmid, and Hervé Jégou. Stable hyper-pooling and query expansion for event detection. *ICCV 2013*

[2] **Matthijs Douze**, Arnau Ramisa, and Cordelia Schmid. Combining attributes and Fisher vectors for efficient image retrieval. *CVPR 2011*

[3] **Matthijs Douze**, Hervé Jégou, Cordelia Schmid, and Patrick Pérez. Compact Video Description for Copy Detection with Precise Temporal Alignment. *ECCV 2010*

[4] **Matthijs Douze**, Hervé Jégou, and Cordelia Schmid. An image-based approach to video copy detection with spatio-temporal post-filtering. *IEEE Transactions on Multimedia*, 12(4):257–266, June 2010.

[5] **Matthijs Douze**, Hervé Jégou, Sandhawalia Harsimrat, Laurent Amsaleg, and Cordelia Schmid. Evaluation of GIST descriptors for web-scale image search. In *CIVR 2009* **(oral)**

[6] **Matthijs Douze** and Hervé Jégou. The Yael library. In *ACM MM 2014*

[7] Danila Potapov, **Matthijs Douze**, Zaid Harchaoui, and Cordelia Schmid. Category-specific video summarization. In *ECCV 2014*

[8] Jérôme Revaud, **Matthijs Douze**, Cordelia Schmid, and Hervé Jégou. Event retrieval in large video collections with circulant temporal encoding. In *CVPR 2013* **(oral)**

[9] Jérôme Revaud, **Matthijs Douze**, and Cordelia Schmid. Correlation-Based Burstiness for Logo Retrieval. In *ACM MM 2012*

[10] Hervé Jégou, Florent Perronnin, **Matthijs Douze**, Jorge Sánchez, Patrick Pérez, and Cordelia Schmid. Aggregating local image descriptors into compact codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9):1704–1716, September 2012.

[11] Zaid Harchaoui, **Matthijs Douze**, Mattis Paulin, Miro Dudik, and Jérôme Malick. Large-scale image classification with trace-norm regularization. In *CVPR'12* **(oral)**

[12] Christian Wengert, **Matthijs Douze**, and Hervé Jégou. Bag-of-colors for improved image search. In *ACM MM 2011*

[13] Hervé Jégou, Romain Tavenard, **Matthijs Douze**, and Laurent Amsaleg. Searching in one billion vectors: re-rank with source coding. In *ICASSP 2011*

[14] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Product Quantization for Nearest Neighbor Search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1):117–128, January 2011.

[15] Hervé Jégou, **Matthijs Douze**, Cordelia Schmid, and Patrick Pérez. Aggregating local descriptors into a compact image representation. In *CVPR 2010* **(oral)**

[16] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Improving bag-of-features for large scale image search. *International Journal of Computer Vision*, 87(3):316–336, May 2010.

[17] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Packing bag-of-features. In *ICCV 2009*

[18] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. On the burstiness of visual elements. In *CVPR 2009*

[19] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Hamming embedding and weak geometric consistency for large scale image search. *ECCV 2008* **(most cited paper of the conference)**

**Publications related to competitions**

[20] **Matthijs Douze**, Dan Oneata, Mattis Paulin, Clément Leray, Nicolas Chesneau, Danila Potapov, Jakob Verbeek, Karteek Alahari, Zaid Harchaoui, Lori Lamel, Jean-Luc Gauvin, Christoph Andreas Schmidt, and Cordelia Schmid. The INRIA-LIM-VocR and AXES submissions to Trecvid 2014 Multimedia Event Detection, 2014. *Trecvid workshop, Orlando, November 2014*

[21] **Matthijs Douze**, Matthieu Guillaumin, Thomas Mensink, Cordelia Schmid, and Jakob Verbeek. INRIA-LEARs participation to ImageCLEF 2009, September 2009. Working Notes for the CLEF 2009 Workshop

[22] **Matthijs Douze**, Adrien Gaidon, Hervé Jégou, Marcin Marszalek, and Cordelia Schmid. INRIA-LEARs video copy detection system. In *TREC Video Retrieval Evaluation (TRECVID Workshop)*, Gaithersburg, United States, 2008.

[23] Robin Aly, Relja Arandjelovic, Ken Chatfield, **Matthijs Douze**, Basura Fernando, Zaid Harchaoui, Kevin Mcguiness, Noël O'Connor, Dan Oneata, Omkar Parkhi, Danila Potapov, Jerome Revaud, Cordelia Schmid, J.-L. Schwenninger, David Scott, Tinne Tuytelaars, Jakob Verbeek, Heng Wang, and Andrew Zisserman. The AXES submissions at TrecVid 2013, November 2013. TRECVID Workshop

[24] Dan Oneata, **Matthijs Douze**, Jérôme Revaud, Schwenninger Jochen, Danila Potapov, Heng Wang, Zaid Harchaoui, Jakob Verbeek, Cordelia Schmid, Robin Aly, Kevin Mcguiness, Shu Chen, Noel O'Connor, Ken Chatfield, Omkar Parkhi, Relja Arandjelovic, Andrew Zisserman, Fernando Basura, and Tinne Tuytelaars. AXES at TRECVid 2012: KIS, INS, and MED. TRECVID Workshop, September 2012.

[25] Hervé Jégou, **Matthijs Douze**, Guillaume Gravier, Cordelia Schmid, and Patrick Gros. INRIA LEAR-TEXMEX: Video copy detection task. In *TRECVid 2010 Workshop*, Gaithersburg, United States, 2010.

**Other publications**

[26] **Matthijs Douze**, Jean-Sébastien Franco, and Bruno Raffin. QuickCSG: Arbitrary and Faster Boolean Combinations of N Solids. Research Report RR-8687, INRIA, March 2015.

[27] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Exploiting descriptor distances for precise image search. Research Report RR-7656, INRIA, June 2011.

[28] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Recent advance in image search. In Frank Nielsen, editor, *Emerging Trends in Visual Computing*, volume 5416 of *Lecture Notes in Computer Science*, pages 305–326. Springer-Verlag, 2009.

[29] Hervé Jégou, **Matthijs Douze**, and Cordelia Schmid. Hamming Embedding and Weak Geometry Consistency for Large Scale Image Search - extended version. Research Report 6709, October 2008.

[30] **Matthijs Douze** and Vincent Charvillat. Real-time generation of augmented video sequences by background tracking. *Computer Animation and Virtual Worlds*, 17(5):537–550, 2006.

[31] **Matthijs Douze** and Vincent Charvillat. Real-time tracking of video sequences in a panoramic view for object-based video coding. In *SCIA*, pages 1022–1029. Springer Berlin Heidelberg, 2003.

[32] Christophe Dehais, **Matthijs Douze**, Geraldine Morin, and Vincent Charvillat. Augmented reality through real-time tracking of video sequences using a panoramic view. In *ICPR 2004*

[33] Romulus Grigoras, Vincent Charvillat, and **Matthijs Douze**. Optimizing hypervideo navigation using a Markov decision process approach. In *ACM Multimedia*, pages 39–48. ACM, 2002.