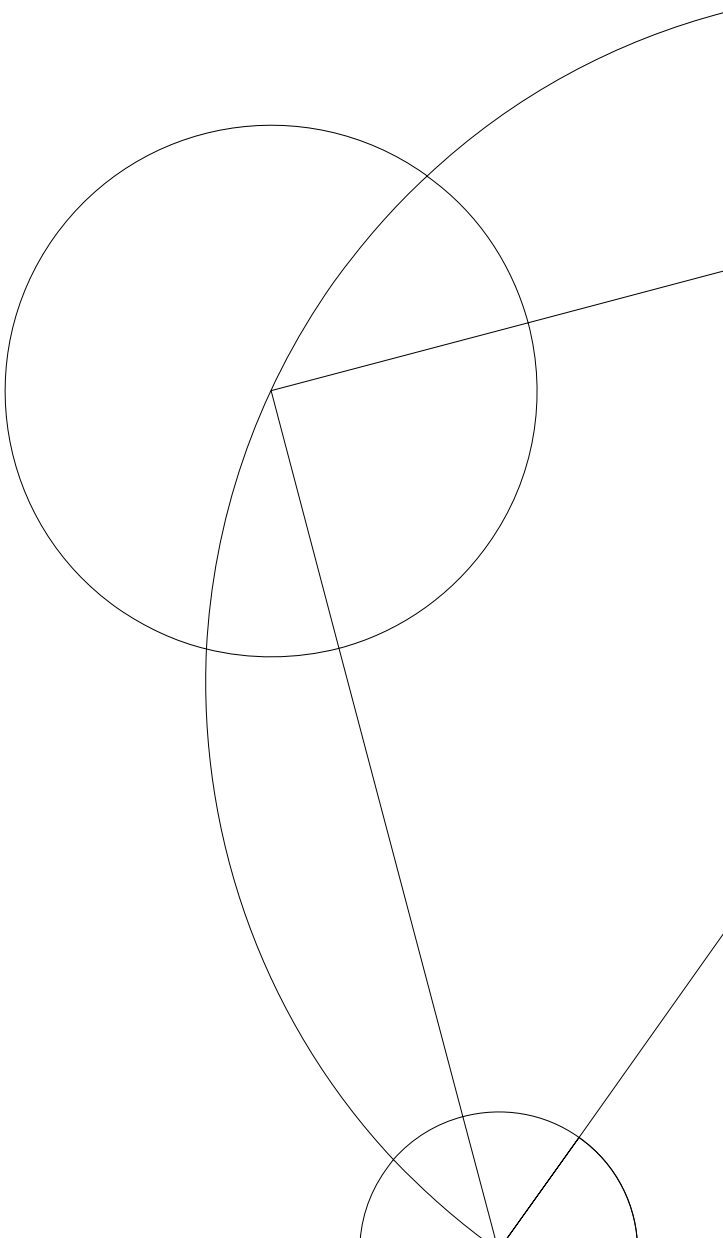
DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF COPENHAGEN

Functional Object Analysis

Toward Statistical Analysis of Functional Objects



PhD thesis

Lars Lau Raket

Thesis overview

This document constitutes my PhD thesis at Department of Computer Science, University of Copenhagen. The subject of the thesis is statistical analysis of functional data, and the main goal of the enclosed work is to go beyond the typical simple analyses of curve data, and open up the possibility of doing model-based statistical analysis of complex functional objects. The functional aspect of data naturally complicate statistical analysis; as opposed to conventional data analysis, geometric information has to be taken into account, and potentially enormous data sizes has to be handled. The contributions of this thesis are both theoretical, computational, and practical, and it offers solutions to some of the mentioned problems in various relevant cases.

During my PhD studies I have worked on a variety of topics. In particular, I have done much work on optical flow estimation and distributed video coding. Rather than writing an overview of my research output during my PhD studies, and the spurious relations between the different work, I have chosen to write a thesis on the area which I find most interesting, and on which I will continue my future research. Thus, optical flow estimation and video coding will not be considered in much detail in this thesis.

Lars Lau Raket
26 March 2014

Included papers

The main contribution of this thesis are the three papers

L.L. Rakêt and B. Markussen, "Approximate inference for spatial functional data on massively parallel processors," *Computational Statistics & Data Analysis*, vol. 72, pp. 227-240, 2014.

L.L. Rakêt, S. Sommer, and B. Markussen, "A nonlinear mixed-effects model for simultaneous smoothing and registration of functional data" *Pattern Recognition Letters*, vol. 38, pp. 1-7, 2014.

L.L. Rakêt, B. Grimme, B. Markussen, G. Schöner, and C. Igel, "Statistical analysis of

human arm movements using timing and motion separation," working paper, 2014.

Finally, Appendix ?? includes a description of an image registration algorithm, which is based on results from the works

L.L. Rakêt, *Duality based optical flow algorithms with applications,* University of Copenhagen prize thesis in Computer Science, Copenhagen University Library, 2013.

L.L. Rakêt, L. Roholm, M. Nielsen, and F. Lauze, "TV-L^1 optical flow for vector valued images," in *Energy Minimazation Methods in Computer Vision and Pattern Recognition* (Y. Boykov, F. Kahl, V. Lempitsky, and F. Schmidt, eds.), vol. 6819 of *Lecture Notes in Computer Science*, pp. 329-341, Springer, 2011.

Other scientific publications during the PhD study

While I have found that statistical analysis of functional data has been my most exciting contribution, I have spent the majority of my PhD working in other areas. I have made a number contributions to the fields of image and video processing, where I have mainly focused on variational formulations. Furthermore, I have done a considerable amount of work on integrating such variational methods in video codecs in the field of distributed video coding. These endeavors have been quite successful—for example the MORE codec (Luong et al., 2014), that I have recently co-developed, is the best performing single-view distributed video codec to date. The papers describing this work do unfortunately not fit into the topic of the present thesis, and have therefore been excluded.
The list of excluded publications are as follows:

M. Salmistraro, **L.L. Rakêt**, C. Brites, J. Ascenso, and S. Forchhammer, "Joint disparity and motion estimation using optical flow for multiview distributed video coding", *European Signal Processing Conference (EUSIPCO)* (**accepted**), 2014.

J. Petersen, M.M.W. Wille, **L.L. Rakêt**, A. Feragen, J.H. Pedersen, M. Nielsen, A. Dirksen, M. de Bruijne, "Effect of inspiration on airway dimensions measured in maximal inspiration CT images", *European Radiology* (**in press**), 2014.

H.V. Luong, **L.L. Rakêt**, and S. Forchhammer, "Re-estimation of motion and reconstruction for distributed video coding," *Image Processing, IEEE Transaction on*, vol. 23, pp. 2804-2819 2014.

M. Salmistraro, **L.L. Rakêt**, M. Zamarin, A. Ukhanova, and S. Forchhammer, "Texture side information generation for distributed coding of video-plus-depth," in *Image Processing (ICIP), 2013 20th IEEE International Conference on,* pp. 1699-1703, 2013.

M. Salmistraro, M. Zamarin, **L.L. Rakêt**, and S. Forchhammer, "Distributed multi-hypothesis coding of depth maps using texture motion information and optical flow," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on,* pp. 1685-1689, 2013.

H.V. Luong, **L.L. Rakêt**, X. Huang, and S. Forchhammer, "Side information and noise learning for distributed video coding using optical flow and clustering," *Image Processing, IEEE Transaction on*, vol. 21, pp. 4782-4796, 2012.

L.L. Rakêt and M. Nielsen, "A splitting algorithm for directional regularization and sparsification," in *Pattern Recognition (ICPR), 2012 21st International Conference on,* pp. 3094-3098, 2012.

L.L. Rakêt, "Local smoothness for global optical flow," in *Image Processing (ICIP), 2012 19th IEEE International Conference on,* pp. 1-4, 2012.

L.L. Rakêt, J. Søgaard, M. Salmistraro, H.V. Luong, and S. Forchhammer, "Exploiting the error-correcting capabilities of low density parity check codes in distributed video coding using optical flow," in *Proceedings of SPIE, the International Society for Optical Engineering*, vol. 8499 SPIE – International Society for Optical Engineering, 2012.

L.L. Rakêt, L. Roholm, A. Bruhn, and J. Weickert, "Motion compensated frame interpolation with a symmetric optical flow constraint," in *Advances in Visual Computing* (G. Bebis et al., eds.), vol. 7431 of *Lecture Notes in Computer Science*, pp. 329-341, Springer, 2012.

X. Huang, **L.L. Rakêt**, H.V. Luong, M. Nielsen, F. Lauze, and S. Forchhammer, "Multi-hypothesis transform domain Wyner-Ziv video coding including optical flow," in *Multimedia Signal Processing (MMSP), 2011 IEEE 13th International Workshop on*, pp. 1-6, 2011.

Summary

I propose a direction it the field of statistics which I denote *functional object analysis*. This subfields considers the analysis of functional objects defined on continuous domains. In this setting I will focus on model-based statistics, with a particularly emphasis on mixed-effect formulations, where the observed functional signal is assumed to consist of both fixed and random functional effects. This thesis takes the initial steps toward the development of likelihood-based methodology for functional objects. I first consider analysis of functional data defined on high-dimensional Euclidean spaces under the effect of additive spatially correlated effects, and then move on to consider how to include data alignment in the statistical model as a nonlinear effect under additive correlated noise. In both cases, I will give directions on how to generalize the methodology to more complex data setups. Finally, I consider extensions and future directions.

Contributions

The main methodological contributions of this thesis are as follows:

- An operator approximation framework for doing inference in linear functional mixed-effects models (Raket and Markussen, 2014).
- A description, and source code for doing efficient massively parallel inference in functional mixed-effects models (Raket and Markussen, 2014).
- A model for doing likelihood inference for functional data under the effect of, possibly random, alignment variation (Raket et al., 2014).

Bibliography

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997), 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic Acids Research* **25**(17), 3389–3402.

Antoniadis, A. and Sapatinas, T. (2007), 'Estimation and inference in functional mixed-effects models', *Computational statistics & data analysis* **51**(10), 4793–4813.

Bolin, D. and Lindgren, F. (2011), 'Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping', *The Annals of Applied Statistics* pp. 523–550.

Boularan, J., Ferré, L. and Vieu, P. (1994), 'Growth curves: a two-stage nonparametric approach', *Journal of Statistical Planning and Inference* **38**(3), 327–350.

Brox, T., Bruhn, A., Papenberg, N. and Weickert, J. (2004), High accuracy optical flow estimation based on a theory for warping, *in* 'ECCV', Vol. 4, pp. 25–36.

Cambanis, S., Huang, S. and Simons, G. (1981), 'On the theory of elliptically contoured distributions', *Journal of Multivariate Analysis* **11**(3), 368–385.

Chambolle, A. (2004), 'An algorithm for total variation minimization and applications', *Journal of Mathematical Imaging and Vision* **20**, 89–97.

Chambolle, A. and Pock, T. (2011), 'A first-order primal-dual algorithm for convex problems with applications to imaging', *Journal of Mathematical Imaging and Vision* **40**, 120–145.

Charon, N. (2013), Analysis of geometric and functional shapes with extensions of currents. Application to registration and atlas estimation., PhD thesis, Ecole normale supérieure de Cachan.

Chen, H. and Wang, Y. (2011), 'A penalized spline approach to functional mixed effects model analysis', *Biometrics* **67**(3), 861–870.

Cox, D. D. (1984), 'Multivariate smoothing spline functions', *SIAM Journal on Numerical Analysis* **21**(4), 789–813.

Cuevas, A., Febrero, M. and Fraiman, R. (2004), 'An anova test for functional data', *Computational statistics & data analysis* **47**(1), 111–122.

Darkner, S., Larsen, R. and Paulsen, R. (2007), Analysis of deformation of the human ear and canal caused by mandibular movement, *in* N. Ayache, S. Ourselin and A. Maeder, eds, 'Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007', Vol. 4792 of *Lecture Notes in Computer Science*, Springer, pp. 801–808.

Demetz, O., Hafner, D. and Weickert, J. (2013), The complete rank transform: A tool for accurate and morphologically invariant matching of structures, *in* 'Proc. 2013 British Machine Vision Conference, Bristol, UK'.

Dobler, G., Fassnacht, C., Treu, T., Marshall, P. J., Liao, K., Hojjati, A., Linder, E. and Rumbaugh, N. (2013), 'Strong lens time delay challenge: I. experimental design', *arXiv preprint arXiv:1310.4830* .

Dryden, I. L. and Mardia, K. V. (1998), *Statistical Shape Analysis*, John Wiley & Sons New York.

Duchamp, T. and Stuetzle, W. (2003), 'Spline smoothing on surfaces', *Journal of Computational and Graphical Statistics* **12**(2), 354–381.

Ekeland, I. and Teman, R. (1999), *Convex Analysis and Variational Problems*, SIAM.

Ferraty, F. and Vieu, P. (2006), *Nonparametric Functional Data Analysis*, Springer.

Fisher, N. I. (1993), *Statistical analysis of spherical data*, Cambridge University Press.

Glasbey, C. A. and Mardia, K. V. (1998), 'A review of image-warping methods', *Journal of applied statistics* **25**(2), 155–171.

Golub, G. and van Loan, C. (1989), *Matrix Computations*, The John Hopkins University Press, Baltimore, Maryland.

Grenander, U. and Miller, M. I. (1998), 'Computational anatomy: An emerging discipline', *Quarterly of applied mathematics* **56**(4), 617–694.

Guo, W. (2002), 'Functional mixed effects models', *Biometrics* **58**(1), 121–128.

Hastie, T., Tibshirani, R. and Friedman, J. (2009), Basis expansions and regularization, *in* 'The Elements of Statistical Learning', Springer Series in Statistics, Springer New York, pp. 139–189.

Hauberg, S., Sommer, S. and Pedersen, K. S. (2012), 'Natural metrics and least-committed priors for articulated tracking', *Image and Vision Computing* **30**(6-7), 453–461.

Hobolth, A. and Jensen, E. B. V. (2000), 'Modelling stochastic changes in curve shape, with an application to cancer diagnostics', *Advances in Applied Probability* **32**(2), 344–362.

Horn, B. K. P. and Schunck, B. G. (1981), 'Determining optical flow', *Artificial Intelligence* **17**, 185–203.

Horváth, L. and Kokoszka, P. (2012), *Inference for Functional Data with Applications*, Springer.

Joshi, S. C. and Miller, M. I. (2000), 'Landmark matching via large deformation diffeomorphisms', *Image Processing, IEEE Transactions on* **9**(8), 1357–1370.

Kiseliov, Y. (1994), 'Algorithms of projection of a point onto an ellipsoid', *Lithuanian Mathematical Journal* **34**, 141–159.

Kneip, A. and Ramsay, J. O. (2008), 'Combining registration and fitting for functional models', *Journal of the American Statistical Association* **103**(483), 1155–1165.

Kotz, S., Kozubowski, T. and Podgorski, K. (2001), *The Laplace Distribution and Generalizations: A Revisit With Applications to Communications, Exonomics, Engineering, and Finance*, number 183 *in* 'Birkhäuser Applied Probability and Statistics', Springer.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1982), *Extremes and Related Properties of Random Sequences and Processes*, Springer Berlin Heidelberg New York.

Lindeberg, T. (1993), *Scale-space theory in computer vision*, Springer.

Lindgren, F., Rue, H. and Lindström, J. (2011), 'An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**(4), 423–498.

Liu, Z. and Guo, W. (2012), 'Functional mixed effects models', *Wiley Interdisciplinary Reviews: Computational Statistics* 4(6), 527–534.

Lucas, B. D. and Kanade, T. (1981), An iterative image registration technique with an application to stereo vision, *in* 'Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)', pp. 674–679.

Luong, H. V., Raket, L. L. and Forchhammer, S. (2014), 'Re-estimation of motion and reconstruction for distributed video coding', *Image Processing, IEEE Transactions on* **23**(7), 2804–2819.

Markussen, B. (2007), 'Large deformation diffeomorphisms with application to optic flow', *Computer Vision and Image Understanding* **106**(1), 97 – 105.

Markussen, B. (2013), 'Functional data analysis in an operator-based mixed-model framework', *Bernoulli* **19**, 1–17.

Marron, J. S. and Alonso, A. M. (2014), 'Overview of object oriented data analysis', *Biometrical Journal* **56**(5), 732–753.

McCall, C., Reddy, K. K. and Shah, M. (2012), Macro-class selection for hierarchical k-NN classification of inertial sensor data., *in* 'PECCS', pp. 106–114.

Nielsen, M., Johansen, P., Jackson, A. D., Lautrup, B. and Hauberg, S. (2008), 'Brownian warps for non-rigid registration', *Journal of Mathematical Imaging and Vision* **31**(2-3), 221–231.

Olsen, O. F. and Nielsen, M. (1997), Multi-scale gradient magnitude watershed segmentation, *in* A. Bimbo, ed., 'Image Analysis and Processing', Vol. 1310 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 6–13.

Olshen, R. A., Biden, E. N., Wyatt, M. P. and Sutherland, D. H. (1989), 'Gait analysis and the bootstrap', *The Annals of Statistics* **17**(4), 1419–1440.

Paragios, N., Chen, Y. and Faugeras, O. D. (2006), *Handbook of Mathematical Models in Computer Vision*, Springer.

Petersen, I. L., Tomasi, G., Sørensen, H., Boll, E. S., Hansen, H. C. B. and Christensen, J. H. (2011), 'The use of environmental metabolomics to determine glyphosate level of exposure in rapeseed (Brassica napus L.) seedlings', *Environmental Pollution* **159**(10), 3071 – 3077.

Pizer, S. M., Fletcher, P. T., Joshi, S., Gash, A. G., Stough, J., Thall, A., Tracton, G. and Chaney, E. L. (2005), 'A method and software for segmentation of anatomic object ensembles by deformable m-reps', *Medical Physics* **32**(5), 1335–1345.

Pizer, S. M., Jung, S., Goswami, D., Vicory, J., Zhao, X., Chaudhuri, R., Damon, J. N., Huckemann, S. and Marron, J. (2013), Nested sphere statistics of skeletal models, *in* 'Innovations for Shape Analysis', Springer, pp. 93–115.

Pottmann, H. and Hofer, M. (2005), 'A variational approach to spline curves on surfaces', *Computer Aided Geometric Design* **22**(7), 693–709.

Rakêt, L. L. (2013), 'Duality based optical flow algorithms with applications', University of Copenhagen prize thesis in Computer Science, Copenhagen University Library.

Raket, L. L., Grimme, B., Schöner, G., Markussen, B. and Igel, C. (2015), 'Statistical analysis of human arm movements using timing and motion separation'. (Working paper, available on request).

Raket, L. L. and Markussen, B. (2014), 'Approximate inference for spatial functional data on massively parallel processors', *Computational Statistics & Data Analysis* **72**, 227 – 240.

Rakêt, L. L., Roholm, L., Nielsen, M. and Lauze, F. (2011), TV-L^1 optical flow for vector valued images, *in* Y. Boykov, F. Kahl, V. Lempitsky and F. Schmidt, eds, 'Energy Minimization Methods in Computer Vision and Pattern Recognition', Vol. 6819 of *Lecture Notes in Computer Science*, Springer, pp. 329–343.

Raket, L. L., Sommer, S. and Markussen, B. (2014), 'A nonlinear mixed-effects model for simultaneous smoothing and registration of functional data', *Pattern Recognition Letters* **38**, 1–7.

Ramsay, J. O. and Silverman, B. W. (2005), *Functional Data Analysis*, second edn, Springer.

Sangalli, L. M., Ramsay, J. O. and Ramsay, T. O. (2013), 'Spatial spline regression models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **75**(4), 681–703.

Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M. and Lenzen, F. (2008), *Variational Methods in Imaging*, Vol. 167, Springer.

Simpson, D., Lindgren, F. and Rue, H. (2012*a*), 'In order to make spatial statistics computationally feasible, we need to forget about the covariance function', *Environmetrics* **23**(1), 65–74.

Simpson, D., Lindgren, F. and Rue, H. (2012*b*), 'Think continuous: Markovian Gaussian models in spatial statistics', *Spatial Statistics* **1**, 16–29.

Sørensen, H., Tolver, A., Thomsen, M. H. and Andersen, P. H. (2012), 'Quantification of symmetry for functional data with application to equine lameness classification', *Journal of Applied Statistics* **39**, 337–360.

Srivastava, A., Wu, W., Kurtek, S., Klassen, E. and Marron, J. (2011), 'Registration of functional data using Fisher-Rao metric', *arXiv preprint arXiv:1103.3817* .

Steinbrücker, F., Pock, T. and Cremers, D. (2009), Large displacement optical flow computation without warping, *in* 'Computer Vision, 2009 IEEE 12th International Conference on', pp. 1609 –1614.

Sun, D., Roth, S. and Black, M. (2010), Secrets of optical flow estimation and their principles, *in* 'Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on', pp. 2432–2439.

Tucker, J. D., Wu, W. and Srivastava, A. (2013), 'Generative models for functional data using phase and amplitude separation', *Computational Statistics & Data Analysis* **61**, 50–66.

Verzelen, N., Tao, W. and Müller, H.-G. (2012), 'Inferring stochastic dynamics from functional data', *Biometrika* **99**(3), 533–550.

Vialard, F.-X. (2013), 'Extension to infinite dimensions of a stochastic second-order model associated with shape splines', *Stochastic Processes and their Applications* **123**(6), 2110–2157.

Wahba, G. (1975), 'Smoothing noisy data with spline functions', *Numerische Mathematik* **24**(5), 383–393.

Wahba, G. (1981), 'Spline interpolation and smoothing on the sphere', *SIAM Journal on Scientific and Statistical Computing* **2**(1), 5–16.

Wahba, G. (1990), *Spline Models for Observational Data*, Society for Industrial and Applied Mathematics.

Wang, H. and Marron, J. (2007), 'Object oriented data analysis: Sets of trees', *The Annals of Statistics* **35**(5), 1849–1873.

Wang, Y. (1998), 'Mixed effects smoothing spline analysis of variance', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **60**(1), 159–174.

Worsley, K. J. and Friston, K. J. (1995), 'Analysis of fMRI time-series revisited—again', *NeuroImage* **2**(3), 173–181.

Zach, C., Pock, T. and Bischof, H. (2007), A duality based approach for realtime TV-L^1 optical flow, *in* F. Hamprecht, C. Schnörr and B. Jähne, eds, 'Pattern Recognition', Vol. 4713 of *Lecture Notes in Computer Science*, Springer, pp. 214–223.