DTU

Modeling of locations in the BrainMap database: Detection of outliers

Finn Årup Nielsen, Lars Kai Hansen and Ulrik Kjems
Informatics and Mathematical Modelling, Technical University of Denmark
International Neuroimaging Consortium

INC

Introduction

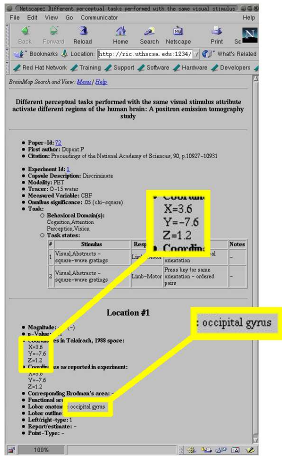
A proposal for meta-analytical modeling of functional neuroimaging locations from the BrainMap™ database [2] by modeling of the relationship between 3D Talairach coordinates and anatomical (text) labels. With novelty detection [1] based on conditional probability density modeling we are able to:

- Spot errors in the database, e.g., entry errors or errors in the original article
- Identify non-conforming and/or new terminology.

We are able to rank the novelty of each location based on the discrepancy between the 3D Talairach coordinate and the anatomical text labels. The most novel locations were investigated further by "manual" means: Comparing the BrainMap™ entry with the data in the original article and in some cases email original authors. A few of the location with high novelty were due to entry errors in BrainMap™, e.g., interchanged centimeter/millimeter. A few other errors were mistakes in the original article. Other outliers were due to variation in the terminology, e.g., "limbic lobe" is seldom used compared to "parietal lobe".
The modeling technique could also be used to model behaviroral variables to Talairach space.

Data from BrainMap™

We downloaded the BrainMap™ database via the web interface and for each location extracted the 3D Talairach coordinate (BrainMap™: "Coordinates in Talairach, 1988 space") as well as the corresponding anatomical label (BrainMap™: "Lobar Anatomy").


Netscape: Different perceptual tasks performed with the same visual stimulus
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print
Bookmarks Location: http://ric.uthscsa.edu:1234/ What's Related
Red Hat Network Training Support Software Hardware Developers

BrainMap Search and View: Menu / Help

Different perceptual tasks performed with the same visual stimulus attribute activate different regions of the human brain: A positron emission tomography study

- **Paper-Id:** 72
- **First author:** Dupont P
- **Citation:** Proceedings of the National Academy of Sciences, 90, p.10927-10931

- **Experiment Id:** 1
- **Capsule Description:** Discriminate
- **Modality:** PET
- **Tracer:** O-15 water
- **Measured Variable:** CBF
- **Omnibus significance:** .05 (chi-square)
- **Task:**
  - **Behavioral Domain(s):** Cognition,Attention Perception,Vision
  - **Task states:**

#	Stimulus	Resp		Notes
1	Visual,Abstracts - square-wave gratings	Limb-Motor	orientation	-
2	Visual,Abstracts - square-wave gratings	Limb-Motor	Press key for same orientation - ordered pairs	-

Location #1

- **Magnitude:** (-)
- **p-Value:**
- **Coordinates in Talairach, 1988 space:** X=3.6 Y=-7.6 Z=1.2
- **Coordinates as reported in experiment:** X=3.6 Y=-7.6 Z=1.2
- **Corresponding Brodman's area:** -
- **Functional area:**
- **Lobar anatomy:** occipital gyrus
- **Lobar outline:**
- **Left/right-type:** 1
- **Report/estimate:** -
- **Point-Type:** -

100%

Each label was split into *terms* constructed from all subsets of words and phrase, e.g.

```
midline occipital lobe  →
   midline
   occipital
   lobe
   midline occipital
   occipital lobe
   midline occipital lobe
```

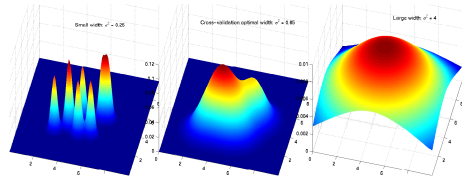
Mathematical modeling

The method estimates probability densities by a kernel method also called Parzen window or Specht kernel estimator [4]: Each location (represented by a 3-dimensional vector $\mathbf{x}$) is convoluted with an isotropic 3-dimensional Gaussian probability density function (PDF) with a width of σ^2. The combined PDF is the sum of all the Gaussian PDFs. The PDF is a *conditional on each term* w. σ^2 is the only parameter in the model and it is optimized with a Newton minimization method on the leave-one-out cross-validation negative log-likelihood costfunction

$$E(\sigma^2, w) = -\sum_{n=1}^{N_w} \log p_{-n}(\mathbf{x}_n|\sigma^2, w), \qquad (1)$$

with each location determined by

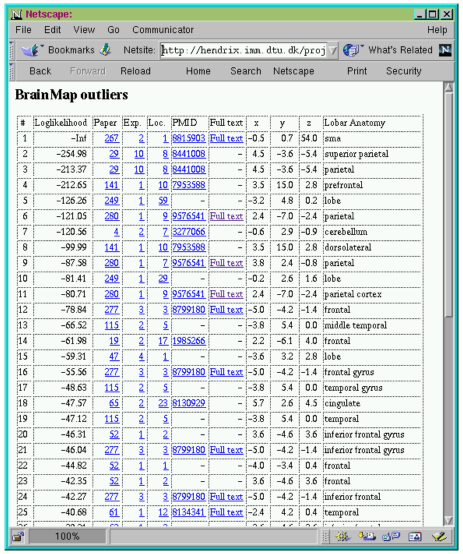
$$p_{-n}(\mathbf{x}|\sigma^2, w) = \frac{1}{N_w - 1} \sum_{n' \neq n}^{N_w} (2\pi\sigma^2)^{-3/2} \times \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{x}_{n'})^2\right). \qquad (2)$$



The figure shows three 2-dimensional kernel densities formed from 8 "locations" modeled with isotropic Gaussian kernels varying the width: small, cross-validation optimal and big.

Novelty detection

The kernel density estimation is iterated over all terms that have more than two associated locations. The procedure is two-stage with a first step including all locations containing the specific term w in the modeling, and the second stage excluding the 5% most extreme location. With the second stage density estimate the location can be ranked according to the likelihood and the table below can be listed. Sometimes it is possible to automatically obtain a link to the full text article through queries to Pubmed.

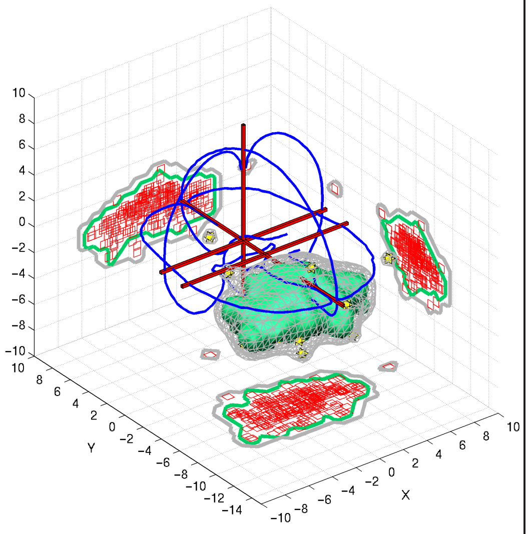

Netscape:
File Edit View Go Communicator Help
Bookmarks Netsite: http://hendrix.imm.dtu.dk/proj What's Related
Back Forward Reload Home Search Netscape Print Security

BrainMap outliers

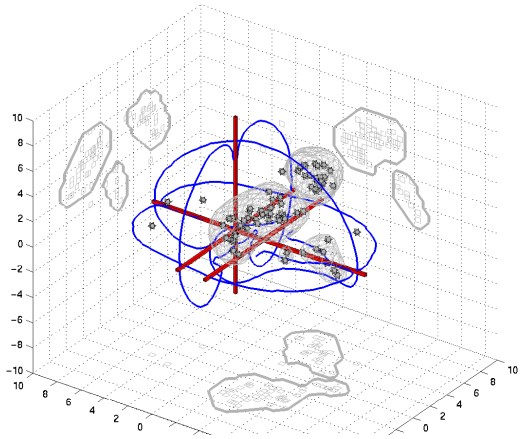
#	Loglikelihood	Paper	Exp.	Loc.	PMID	Full text	x	y	z	Lobar Anatomy
1	−Inf	267	2	1	8815903	Full text	−0.5	0.7	54.0	sma
2	−254.98	29	10	8	8441008	−	4.5	−3.6	−5.4	superior parietal
3	−213.37	29	10	8	8441008	−	4.5	−3.6	−5.4	parietal
4	−212.65	141	1	10	7953588	−	3.5	15.0	2.8	prefrontal
5	−126.26	249	1	59	−	−	−3.2	4.8	0.2	lobe
6	−121.05	280	1	9	9576541	Full text	2.4	−7.0	−2.4	parietal
7	−120.56	4	2	7	3277066	−	−0.6	2.9	−0.9	cerebellum
8	−99.99	141	1	10	7953588	−	3.5	15.0	2.8	dorsolateral
9	−87.58	280	1	7	9576541	Full text	3.8	2.4	−0.8	parietal
10	−81.41	249	1	29	−	−	−0.2	2.6	1.6	lobe
11	−80.71	280	1	9	9576541	Full text	2.4	−7.0	−2.4	parietal cortex
12	−78.84	277	3	3	8799180	Full text	−5.0	−4.2	−1.4	frontal
13	−66.52	115	2	5	−	−	−3.8	5.4	0.0	middle temporal
14	−61.98	19	2	17	1985266	−	2.2	−6.1	4.0	frontal
15	−59.31	47	4	1	−	−	−3.6	3.2	2.8	lobe
16	−55.56	277	3	3	8799180	Full text	−5.0	−4.2	−1.4	frontal gyrus
17	−48.63	115	2	5	−	−	−3.8	5.4	0.0	temporal gyrus
18	−47.57	65	2	23	8130929	−	5.7	2.6	4.5	cingulate
19	−47.12	115	2	5	−	−	−3.8	5.4	0.0	temporal
20	−46.31	52	1	2	−	−	3.6	−4.6	3.6	inferior frontal gyrus
21	−46.04	277	3	3	8799180	Full text	−5.0	−4.2	−1.4	inferior frontal gyrus
22	−44.82	52	1	1	−	−	−4.0	−3.4	0.4	frontal
23	−42.35	52	1	2	−	−	3.6	−4.6	3.6	frontal
24	−42.27	277	3	3	8799180	Full text	−5.0	−4.2	−1.4	inferior frontal
25	−40.68	61	1	12	8134341	Full text	−2.4	4.2	0.4	temporal

100%

Results



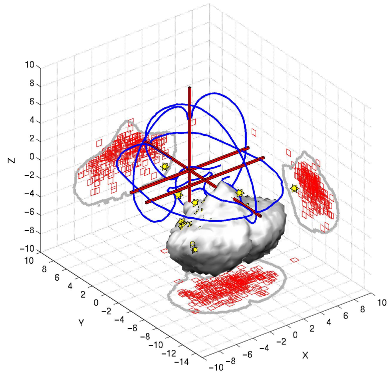
Novelty associated with "cerebellum" is shown in the figure above. The wireframe corresponds to the first stage density; the filled polygons to the second stage. The anterior BrainMap^{TM} location is associated with high novelty (#7 in the table).



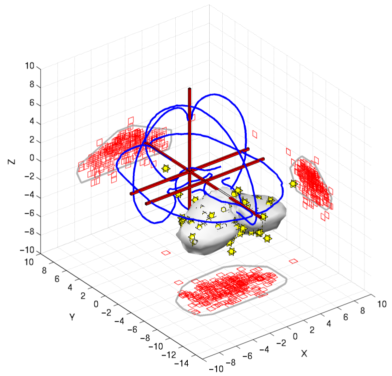
The figure above shows the probability density estimate for the term "lobe". Although "lobe" could appear over most of the brain it is more common to use it in connection with "parietal" and "occipital" rather than "frontal" or "temporal".

Atlas methods

Instead of probability density modeling the novelty detection can be based on atlases, e.g., directly on the digitized version of the Talairach atlas [5] or the probabilistic ICBM atlas.



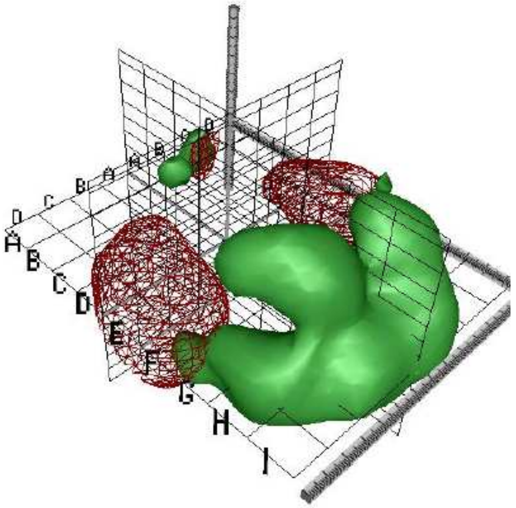
The probabilistic ICBM atlas describes a few brain structures — among others the cerebellum shown above. There are a number of locations that appear outside of the area of "ICBM cerebellum".



The discrepancy with the "cerebellum" locations from the literature to the original Talairach atlas is even larger than for the ICBM atlas: Even more locations reported outside the "cerebellum" as delineanated in the atlas.

Concluding remarks

- A system for combining analysis of spatial and text data.
- The kernel density estimator is simple, automatic and flexible for description of a set of BrainMap^{TM} locations and performs well in detecting novelty.
- Novelty detection using redundancy: A concept that can be generalized for data mining in any database with sufficient information
- "Bootstrapped Talairach Daemon". The method can be extended to make automatic anatomical labelling of locations as the Talairach Daemon [3].
- It is possible to extent the modeling to *functional* data. Below is shown probability density models on BrainMap^{TM} data when conditioning on "Vision" and "Audition" behavioral domains.



More information

- Some of the tools for this analysis is implemented in the *Brede Matlab toolbox* available at http://hendrix.imm.dtu.dk/software/brede/
- 3D visualization available on the Internet: VRML models available at http://hendrix.imm.dtu.dk/projects/brainmap/
- Technical report and submitted article to *Human Brain Mapping*.

References

[1] Chris M. Bishop. Novelty detection and neural network validation. *IEE Proceedings — Vision, Image and Signal Processing*, 141(4):217–222, August 1994. Document No. 19941330.

[2] Peter T. Fox and Jack L. Lancaster. Neuroscience on the net. *Science*, 266(5187):994–996, November 1994.

[3] Jack L. Lancaster, Marty G. Woldorff, M. Liotti, C. S. Freitas, L. Rainey, P. V. Kochunov, D. Nickerson, S. A. Mikiten, and Peter T. Fox. Automated talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10:120–131, 2000.

[4] Donald F. Specht. Probabilistic neural networks. *Neural Networks*, 3(1):109–118, 1990.

[5] Jean Talairach and P. Tournoux. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme Medical Publisher Inc, New York, 1988.

Acknowledgement

This project has been funded by The Danish Research Councils through THOR Center for Neuroinformatics and the Human Brain Project P20 MH57180 "Spatial and Temporal Patterns in Functional Neuroimaging".