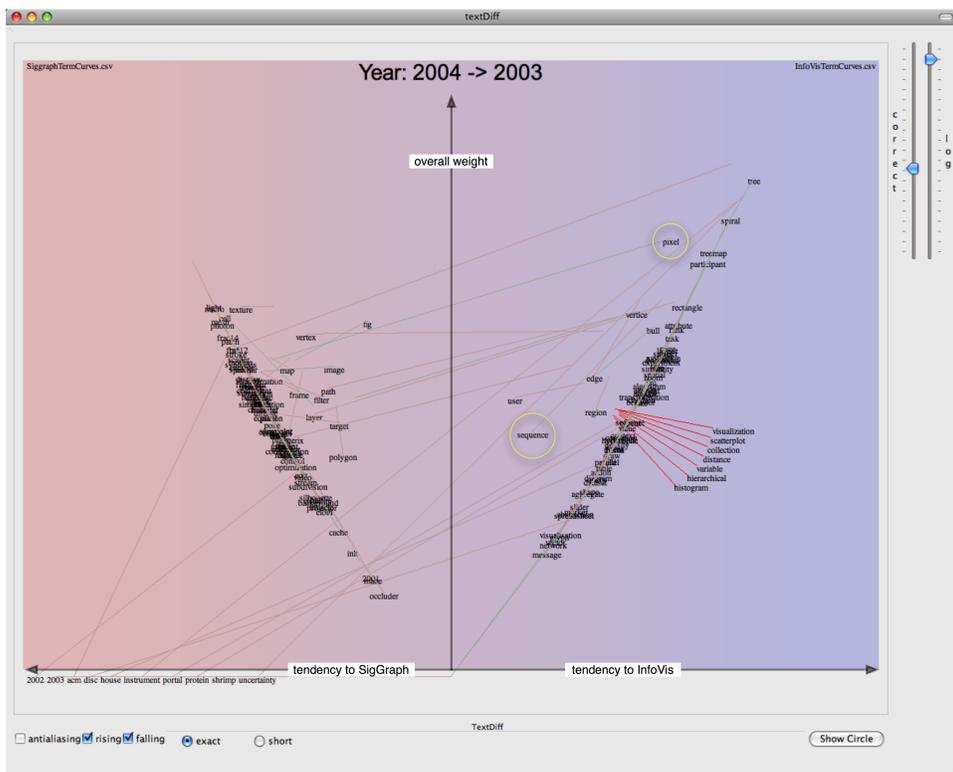


Analyzing the rise and fall of topics in conference publications and the interdependencies of the conferences with respect to the topics discussed.

In cooperation with Daniela Oelke (GK associated member, Prof. Keim's group) and Christian Rohrdantz (GK master research student, Prof. Keim's group) we evaluate methods for investigating high relevant terms of conference text corpora and especially terms belonging to more than one conference corpus. The tool allows to observe differences of a term's relevance (overall and conference related) between two years.

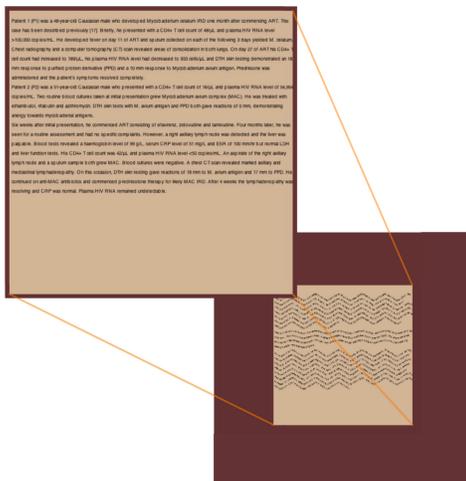
Future work: Improve the Visualization to avoid occlusions and allow selection, use short comic-style path to reduce line-crossing, use sparklines for multiple-year-preview.



The evolution of SigGraph and InfoVis terms from 2004 (line start) to 2003 (line end with term). Dark red lines mark falling tendencies (i.e. "sequence") and green lines mark rising tendencies (i.e. "pixel"). The axes are logarithmic deformed.

Case study: Read distance dependent headline deformation

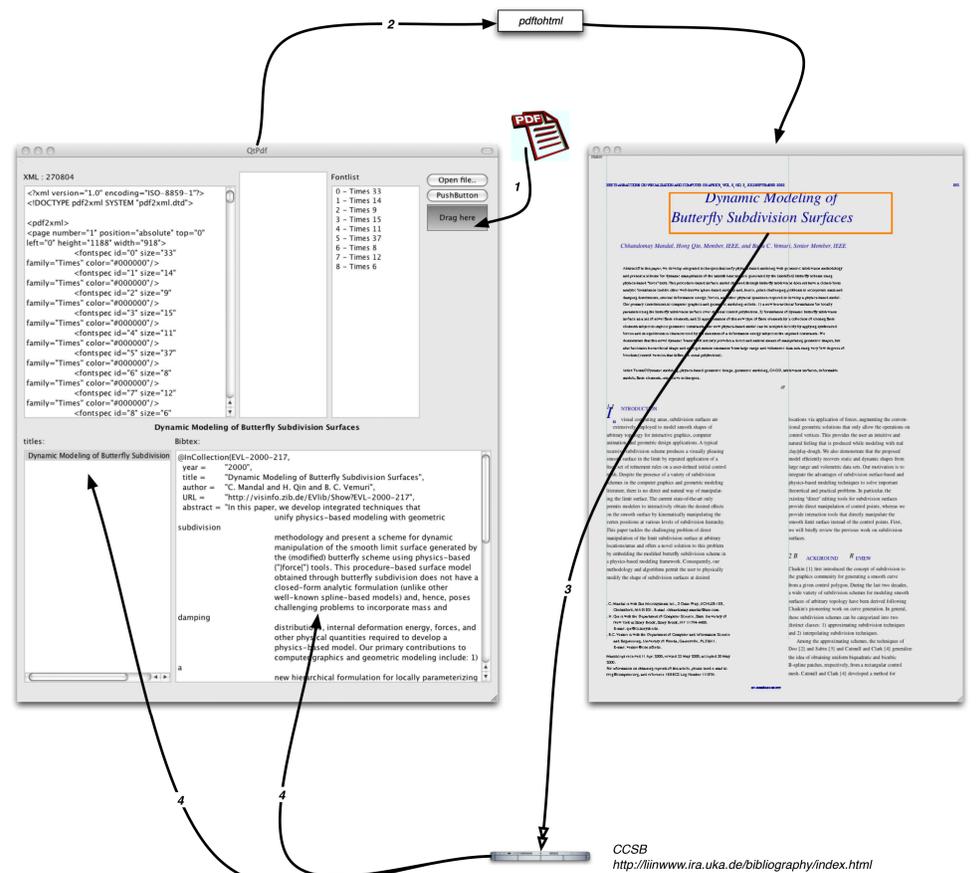
The idea was to create a tool which deforms the headlines when zooming the text away. This could be a base for a visualization of paragraphwise text features. The form-based character is less signal-like than using color highlighting, therefore its use should be with soft, subliminal text features (like for example text mood features). Future work: Find a good application :)



Bibtex Finder

The tool extracts the title of a PDF document and tries to find a bibtex entry by referring to the "The Collection of Computer Science Bibliographies" [?]. To increase the ease of use the user simply drags a PDF file on the interface (1). The PDF is transformed into a XML file which is used as input into a graphical intended heuristic (2). The extracted title is queried to the CCSB(3). The returned BibTeX entries are listed again in the tool (4).

Future work: improve stability and adapt it to other citation databases.



Future work

Conference or journal PDF corpora contain big sources of documents with mixed graphics and text. A future approach is to visualize a big portion of corpora documents on one display device (be it a monitor, a powerwall or a printer) by using the included graphics and important text as representation for one document. Finding a good zoomable visual representation (computer graphics part) and finding good text features (textmining part) will be tasks in the cooperation with Daniela Oelke and Christian Rohrdantz (Prof. Keim's group).

References

- [1] Alf-Christian Achilles and Paul Ortyl. The collection of computer science bibliographies. Website, 2005-2008. Available online at <http://liinwww.ira.uka.de/bibliography/index.html>; visited in May 2008.