

A New Visual Search Interface for Web Browsing

Songhua Xu^{1,2} Tao Jin² Francis C.M. Lau²

1 Department of Computer Science, Yale University

2 Department of Computer Science, The University of Hong Kong



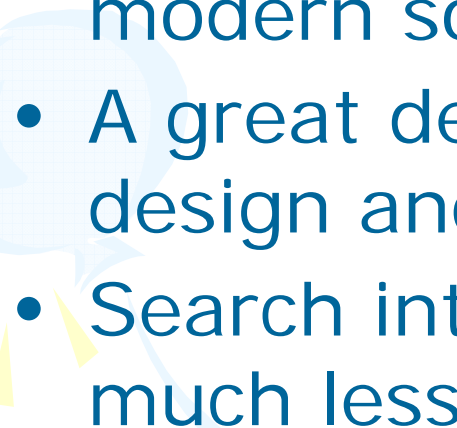
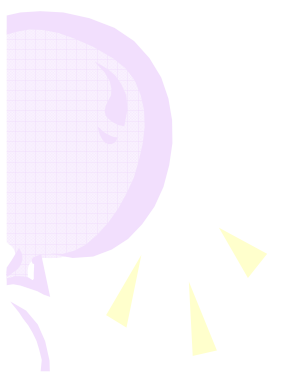
WSDM '09

Barcelona, Spain

Feb. 11, 2009

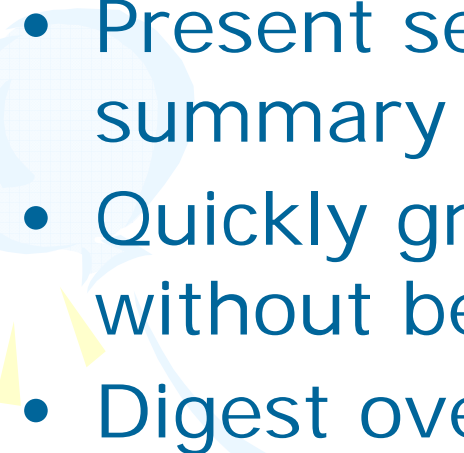
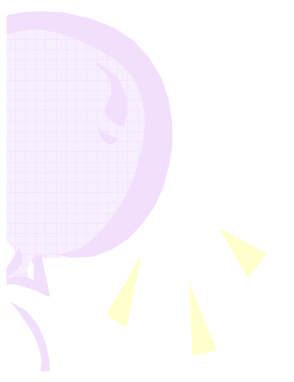


Motivation

- Search engines: one of the most useful tools computer scientists offered to the modern society.
 - A great deal of research on search engine design and backend retrieval algorithms.
 - Search interface is also important, yet much less studied.
- 
- 



A new visual search interface

- Group the search results according to their contents
 - Present selected search results in summary form
 - Quickly grasp the key topics and contents without being overwhelmed with details
 - Digest over search results
- 
- 



Organization of the talk

- Overview of our visual search interface
 - Architecture of our visual interface
 - Related work
 - Search result organization
 - Displaying region allocation for topic groups
 - Selective presentation of representative search results
 - Demo
 - Conclusion
 - Q & A
- 
- 



Novelty of our new search interface

- Traditional search interface

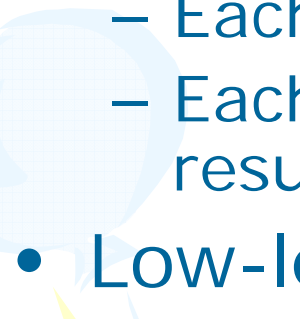

- List view
- Linearly organized

- Our new search interface

- Summarized view
- Top-down style
- Analogous to e-map
- Selected texts and images of search results



Features of our new search interface

- Multi-level view of search results
 - High-level overview of all the topics
 - Each topic mapped to a certain screen region
 - Each topic contains a few top ranked search results
 - Low-level detailed view of the topics
 - Zoom-in view of the topic
 - Sub-topics or abstracts
 - Four levels in our prototype
- 
- 

A visual search example

apple
e.g. "apple" or "computer science"

You searched for **apple**, 118 results and 23 categories

⊖ AAPL Stock (24)

- [Digg - AAPL Stock Tanking](#)
- [AAPL \(APPLE INC\): Stock analyses and research by investors on UpDown.com](#)
- [AAPL Stock - Wall Street Survivor](#)
- [AAPL Apple Inc. Stock Message Board. News. Social Network](#)
- [Quotes, Activity and Analysis for Apple Inc \(NASDAQ:AAPL\)](#)
- [Apple Inc. \(AAPL\) Stock - Seeking Alpha](#)
- [AAPL - Stock Quote for Apple Inc - MSN Money](#)
- [More...](#)

⊖ Mac News (8)

⊖ Apple Computer (7)

⊖ Apple Online Store (6)

⊖ Apple Developer Connection (5)

⊖ Apple Products (6)

⊖ More...

Apple >

AAPL Stock (24)

[Digg - AAPL Stock Tanking](#)

[digg.com/apple/AAPL_Sto](#)

[AAPL \(APPLE INC\): Stock analyses and research by investors](#)

[www.updown.com/stock.do](#)

[AAPL Stock - Wall Street Survivor](#)

[www.wallstreetsurvivor.](#)



Apple Online Store (6)

[DivX to Challenge](#)

[Apple's iTV with 'DivX Connected' - News and Analysis ...](#)

[www.pcmag.com/article2/](#)

[Apple Developer Connection](#)

[developer.apple.com/](#)



Mac News (8)

[Macworld](#)

[www.macworld.com/](#)

[Apple Investor Relations Earnings Releases](#)

[phx.corporate-ir.net/ph](#)



[Apple Developer Connection \(5\)](#)

[Top 20 Most Admired Companies - Apple \(1\) - FORTUNE](#)

[money.cnn.com/galleries](#)

[Apple - Apple](#)

[consultants.apple.com/](#)



Apple Computer (7)

[AAPL](#)

[money.cnn.com/quote/quo](#)

[Apple \(AAPL\)](#)

[www.wikinvest.com/stock](#)

[Apple Leopard Server - Full Review - Reviews by PC Magazine](#)

[www.pcmag.com/article2/](#)



Apple Products (6)

[Mac Laptops, iPods and Macintosh Computers |](#)

[Apple Online Store store.apple.com/](#)

[Apple - Appletell](#)

[www.appletell.com/apple](#)



A visual search example

apple search
e.g. "Apple" or "computer science"


You searched for **apple**. 118 results and 23 categories

- Apple Stock (24)
 - Digg - AAPL Stock Tanking
 - AAPL (APPLE INC): Stock analyses and research by investors on UpDown.com
 - AAPL Stock - Wall Street Survivor
 - AAPL Apple Inc. Stock Message Board, News, Social Network
 - Quotes, Activity and Analysis for Apple Inc (NASDAQ:AAPL)
 - Apple Inc. (AAPL) Stock - Seeking Alpha
 - AAPL - Stock Quote for Apple Inc. - MSN Money More...
- Mac News (8)
- Apple Computer (7)
- Apple Online Store (6)
- Apple Developer Connection (5)
- Apple Products (5)
- More...

Apple > [AAPL Stock \(24\) More...](#)

Digg - AAPL Stock Tanking

In the 30 days after Christmas, Apple stocks have completely tanked. ... Now isn't the time to sell AAPL, it's actually time to buy. ...
[digg.com/apple/AAPL_Ste](#)





AAPL (APPLE INC): Stock analyses and research by investors on UpDown.com

AAPL (APPLE INC): Stock analyses and research by investors on UpDown.com ... TraderNotes on Active Stocks: CELI, UVSE, GSPG, GNTA, AAPL, IMMU, EXAS, EEEK (07 ...
[www.updown.com/stock.de](#)

UPDOWN

AAPL Stock - Wall Street Survivor

AAPL Stock. Reply. Anyone know what happened? ... Re: AAPL Stock. Reply, i'm 1300 with it right now, yea i finally have something going good ...
[www.wallstreet survivor.com](#)

AAPL Apple Inc. Stock Message Board, News, Social Network


social network and stock message board discussion forums. Stop in, post a message about AAPL Apple Inc, Stock Message Board ... 181: Jason: Re: AAPL Stock up to 200 ...
[www.edutgroups.com/mg](#)

Quotes, Activity and Analysis for Apple Inc (NASDAQ:AAPL)

Detailed quote information for Apple Inc (NASDAQ:AAPL), as well as forum postings, trading activity sorted by skill, related stocks, corporate actions, etc. Worth ...
[www.marketocracy.com/mg](#)


Apple Inc. (AAPL) Stock - Seeking Alpha

Replacing P/E in Valuing Apple Stock on Jul 31, 2008 by ... e.g. MSFT, AAPL, GOOG. Enter Stock Ticker e.g. MSFT, AAPL, GOOG. See Sample " We DON'T spam! ...
[seekingalpha.com/symbol](#)




AAPL - Stock Quote for Apple Inc. - MSN Money

AAPL - Get a stock quote for Apple Inc market quotes include the latest AAPL price and fundamental investing data, price charts and news. You can also get free,
[moneycentral.msn.com/site](#)




Apple a Day: AAPL stock news - News and analysis on Apple, Macintosh ...

Concern over that incident buffeted AAPL stock for months ... AAPL stock continues rebound, boosted by Mac's growing market share ...
[weblogs.baltmoreun.co](#)




Apple St... Applied Learning Styles - Cynthia Ulrich Tobias

Seminars and materials by Cynthia Ulrich Tobias, M Ed ... Apple St. LLC, Disclaimer. Copyright A© 1998-2007 Apple St. All Rights Reserved ...
[www.apalst.com/default](#)




Apple Inc. (AAPL): Stock Quote & Company Profile - BusinessWeek

(AAPL) stock research ... Stock Quotes: Company Lookup. Recently Viewed ... to report a data error on the AAPL Snapshot page, please submit this form. ...
[investing.businessweek.com](#)




AAPL: APPLE INC Earnings Analysis

Find pertinent APPLE INC earning analysis information, including estimated and reported earnings. ... AAPL Stock price, 162.12 3.09 1.94 ...
[financa.aol.com/earnings](#)



StockTargets Inc. - Targets are your Power

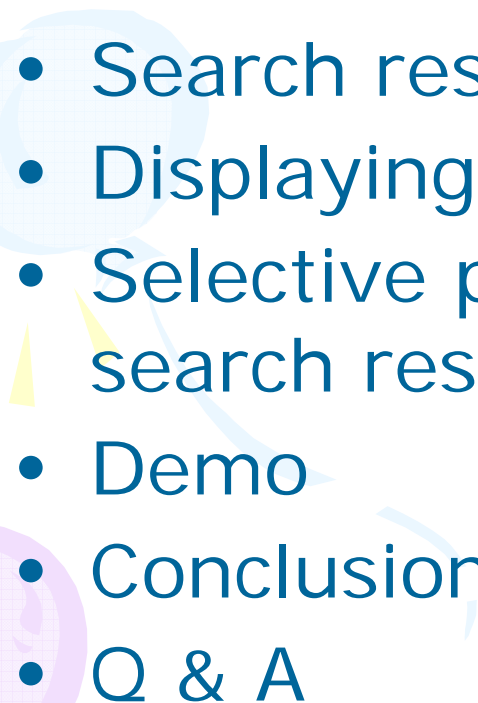
StockTargets is a leading financial destination, providing consumers with a broad range ... Stock Commentary, AAPL stock closed at a price of 163.57 yesterday ...
[www.stocktargets.com/cg](#)



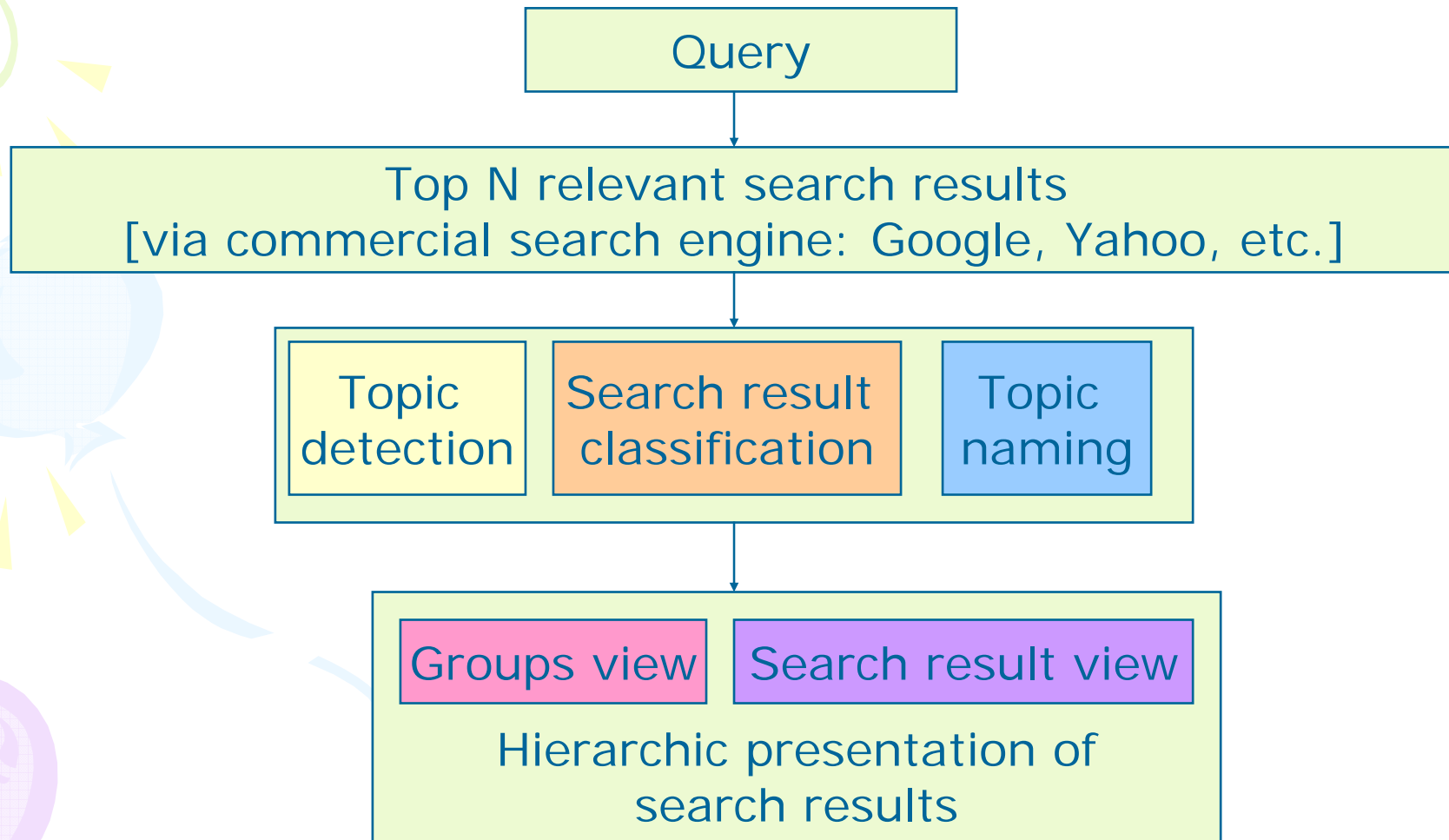
1 | 2 |



Organization of the talk

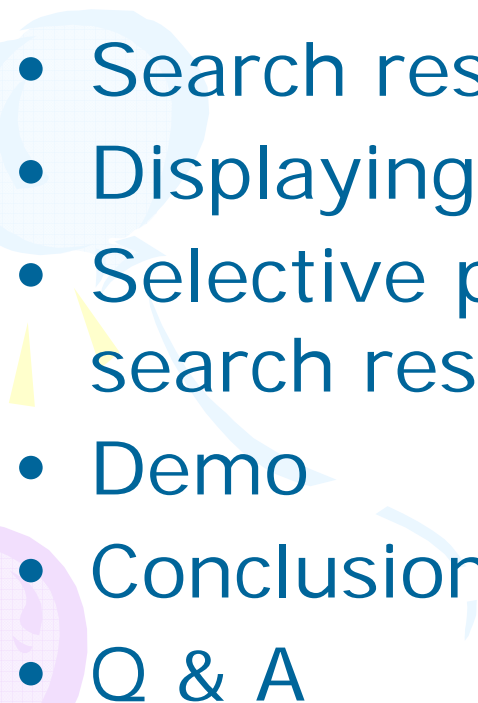
- Overview of our visual search interface
 - **Architecture of our visual interface**
 - Related work
 - Search result organization
 - Displaying region allocation for topic groups
 - Selective presentation of representative search results
 - Demo
 - Conclusion
 - Q & A
- 

Architecture of our search interface





Organization of the talk

- Overview of our visual search interface
 - Architecture of our visual interface
 - **Related work**
 - Search result organization
 - Displaying region allocation for topic groups
 - Selective presentation of representative search results
 - Demo
 - Conclusion
 - Q & A
- 

Related work: cluster-based document collection browsing

- D. R. Cutting, D. R. Karger, J. O. Pedersen, and J. W. Tukey. Scatter/gather: a cluster-based approach to browsing large document collections. SIGIR '92.
- D. R. Cutting, D. R. Karger, and J. O. Pedersen. Constant interaction-time scatter/gather browsing of very large document collections. SIGIR '93.
- P. Pirolli, P. Schank, M. Hearst, and C. Diehl. Scatter/gather browsing communicates the topic structure of a very large text collection. CHI '96.
- M. A. Hearst, D. R. Karger, and J. O. Pedersen. Scatter/gather as a tool for the navigation of retrieval results. AAAI Fall Symposium on AI Applications in Knowledge Navigation and Retrieval, 1995.
- M. A. Hearst and J. O. Pedersen. Reexamining the cluster hypothesis: scatter/gather on retrieval results. SIGIR '96.

Related work: visual search interfaces

S. McCrickard and C. Kehoe. Visualizing search results using squid. WWW, 1997.

D. A. Nation, C. Plaisant, G. Marchionini, and A. Komlodi. Visualizing web sites using a hierarchical table of contents browser: Webtoc. Conference on Human Factors and the Web, 1997.

R. Beale, R. McNab, and I. Witten. Visualising sequences of queries: a new tool for information retrieval. Information Visualization, 1997.

G. Robertson, M. Czerwinski, K. Larson, D. C. Robbins, D. Thiel, and M. van Dantzich. Data mountain: using spatial memory for document management. UIST '98.

B. Amento, W. Hill, L. Terveen, D. Hix, and P. Ju. An empirical evaluation of user interfaces for topic management of web sites. CHI '99.

Related work: visual data exploration

S. Benford, I. Taylor, D. Brailsford, B. Koleva, M. Craven, M. Fraser, G. Reynard, and C. Greenhalgh. Three dimensional visualization of the world wide web. *ACM Computing Surveys*, 31:1-16, 1999.

P. Au, M. Carey, S. Sewraz, Y. Guo, and S. M. Ruger. New paradigms in information visualization. *SIGIR '00*.

A. Becks, C. Seeling, and R. Minkenbergl. Benefits of document maps for text access in knowledge management: a comparative study. *SAC '02*.

H. Lam and P. Baudisch. Summary thumbnails: readable overviews for small screen web browsers. *CHI '05*.

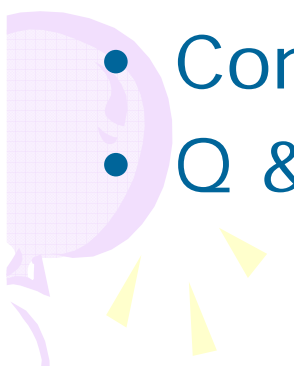
W. Kules, M. Wilson, M. Schraefel, and B. Shneiderman. From keyword search to exploration: How result visualization aids discovery on the web. Technical Report 1516920080208, School of Electronics and Computer Science, University of Southampton, 2008.

D. A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 08(1):1-8, 2002.

S. Kaasten, S. Greenberg, and C. Edwards. How people recognize previously seen web pages from titles, urls and thumbnails. In *People and Computers XVI (Proceedings of Human Computer Interaction 2002)*, 2002.



Organization of the talk

- Overview of our visual search interface
 - Architecture of our visual interface
 - Related work
 - **Search result organization**
 - Displaying region allocation for topic groups
 - Selective presentation of representative search results
 - Demo
 - Conclusion
 - Q & A
- 



Search result organization

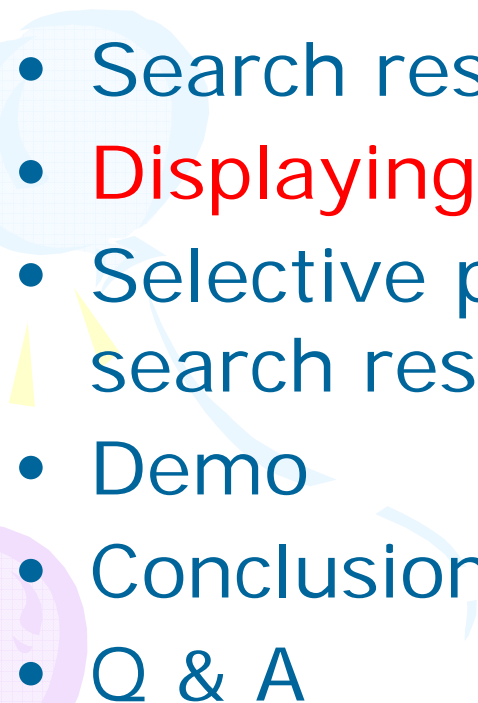
- Semantically meaningful document clustering
- Nonnegative matrix factorization analysis
 - Identifying topics present in a document corpus
 - Correspond well to semantically meaningful topics in the document content space

F. Shahnaz, M. W. Berry, V. P. Pauca, and R. J. Plemmons.
Document clustering using nonnegative matrix factorization.
International Journal of Information Processing and
Management, 42(2):373–386, 2006.

- 
- Hierarchical search result clustering



Organization of the talk

- Overview of our visual search interface
 - Architecture of our visual interface
 - Related work
 - Search result organization
 - **Displaying region allocation for topic groups**
 - Selective presentation of representative search results
 - Demo
 - Conclusion
 - Q & A
- 

Displaying region allocation: determining size of a topic group

- Utilize the webpage ranks, compute the average webpage rank for each topic group
 - dc_1, \dots, dc_n (n topic groups)
 - r_1, \dots, r_n (n average webpage ranks)
 - Assuming S is the whole screen space:

$$s_i \triangleq \frac{\frac{1}{r_i}}{\sum_{i=1}^n \frac{1}{r_i}} S$$

- Open source hierarchical document clustering software package (<http://demo.carrot2.org>)



Displaying region allocation: using rectangular regions

- Using rectangular topic regions
 - Easy to compute



Layout allocation problem



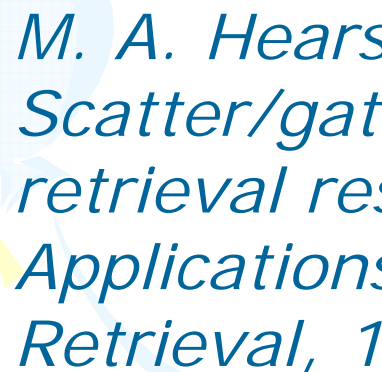
Rectangle-in-rectangle
packing problem

- Simple and visual elegance
- 

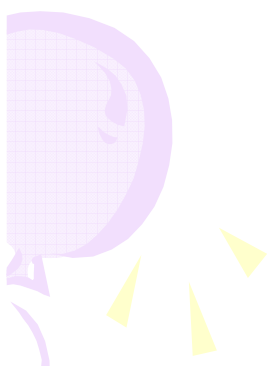


Displaying region allocation: Layout design

- Automatic yellow-page pagination and layout algorithm:



M. A. Hearst, D. R. Karger, and J. O. Pedersen. Scatter/gather as a tool for the navigation of retrieval results. In AAAI Fall Symposium on AI Applications in Knowledge Navigation and Retrieval, 1995.



Displaying region allocation: Layout design (cont.)

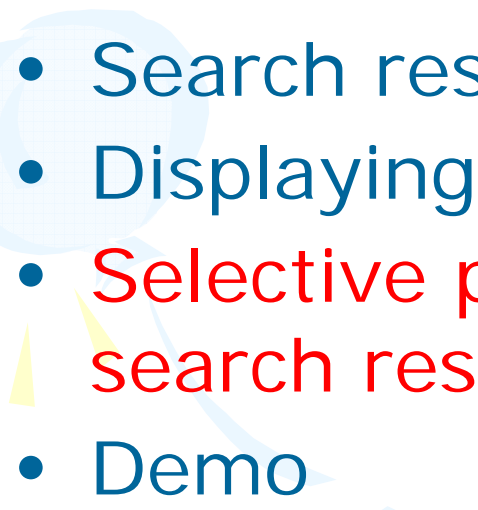
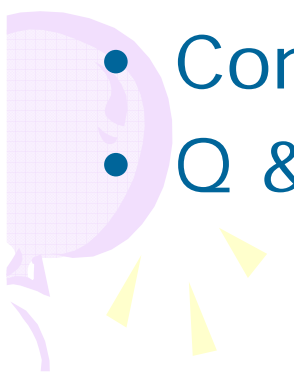
- A modification: introducing a bias term

$$\psi \triangleq \sum_{i=1}^K (GS_i(x) + GS_i(y))$$

- $GS_i(x)$ the horizontal distances of the center for the i -th topic displaying region
- $GS_i(y)$ the vertical distances of the center of the i -th topic displaying region
- topics detected



Organization of the talk

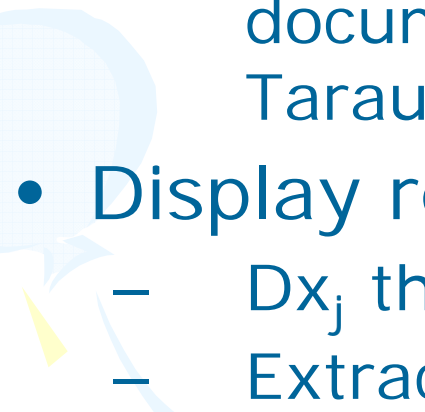
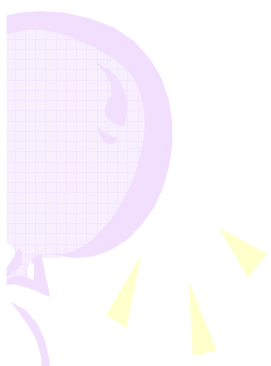
- Overview of our visual search interface
 - Architecture of our visual interface
 - Related work
 - Search result organization
 - Displaying region allocation for topic groups
 - **Selective presentation of representative search results**
 - Demo
 - Conclusion
 - Q & A
- 
- 

Selective presentation of representative search results

- Key information in the search results:
 - Text
 - Images
- A summarized version of this information
- Three modes provided in our system:
 - Text only
 - Image only
 - Mixed text and image
- Mixed text summaries and thumbnails
 - Better performance
 - Study conducted by Woodruff et al., UIST 98.



Selective presentation of search results: text only mode

- Extractive summarization
 - Iterative graph-based language independent document summarization, Mihalcea and Tarau, IJCNP 2005.
 - Display region allocate
 - Dx_j the screen area
 - Extracted key sentences length
- 
- 

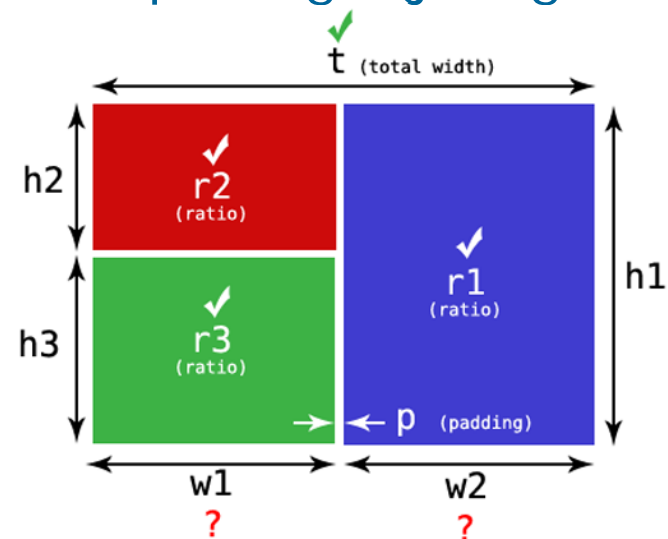


Selective presentation of search results: image only mode

- Select key image of webpage
 - No images from advertisements
 - No website navigation images
 - Image Classification techniques
 - Sophisticated methods are available
 - Keep it simple
 - Image size
 - Image position
 - Image show pattern
- 
- 

Selective presentation of search results: image only mode

- Content images
- Pick from content images
 - Less than N
 - pick the first image
 - More than N
 - Pick first a x b images to combine a rectangle image
 - Use open-source image mosaic package (jimage-mosaic.sourceforge.net)



Selective presentation of search results: image only mode

Two examples



Selective presentation of search results: mixed image & text mode

- Position of text and image
 - Left and Right
 - Upper and Bottom
 - Around

[AAPL Stock \(24\)](#)

[Digg - AAPL Stock Tanking](#)

digg.com/apple/AAPL_Sto

[AAPL \(APPLE INC\): Stock analyses and research by investors](#)

www.updown.com/stock.do

[AAPL Stock - Wall Street Survivor](#)

www.wallstreetsurvivor



Carthage Computer Science

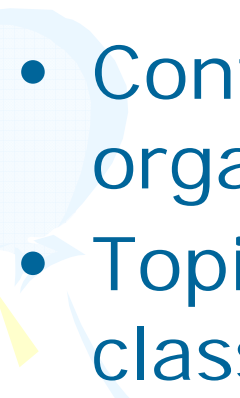
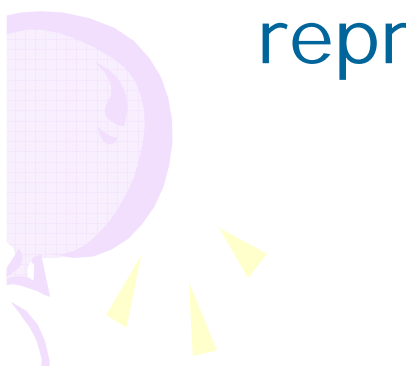
Students majoring in Computer Science at Carthage benefit from small class sizes ... The computer science program, housed in the Clausen Center for World Business, ...



www.carthage.edu/dept/c




To recap

- A user-friendly and informative graphical front-end for search engines
 - Content oriented search result organization & presentation
 - Topic detection and search result classification
 - Mixed image and text summary of representative search results
- 
- 



Organization of the talk

- Overview of our visual search interface
 - Architecture of our visual interface
 - Related work
 - Search result organization
 - Displaying region allocation for topic groups
 - Selective presentation of representative search results
 - **Demo**
 - Conclusion
 - Q & A
- 

Demo

- Our prototype visual search interface site:

<http://www.cs.hku.hk/~songhua/visualsearch/>

or

<http://www.gceel.com>

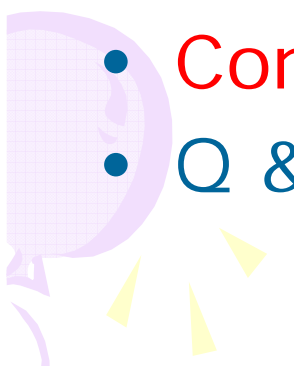
- Special thanks for:

- Google App Engine Service
- The Carrt2 Project



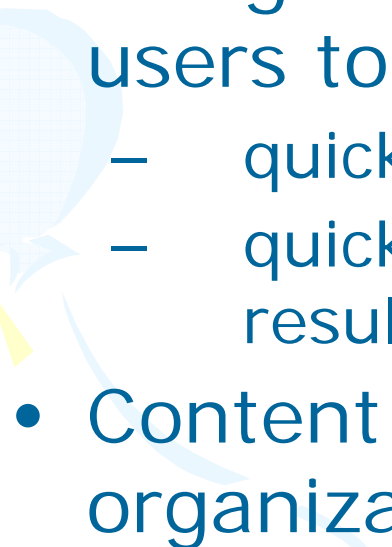
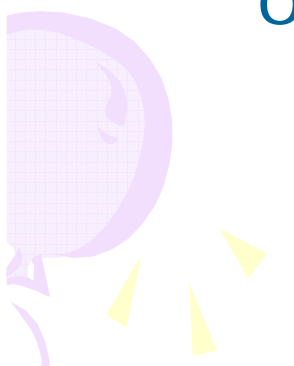


Organization of the talk

- Overview of our visual search interface
 - Architecture of our visual interface
 - Related work
 - Search result organization
 - Displaying region allocation for topic groups
 - Selective presentation of representative search results
 - Demo
 - **Conclusion**
 - Q & A
- 



Conclusion

- A visual search interface for efficient navigation of search results, facilitating users to
 - quickly grasp latent topics
 - quickly locate the interested/desired search results
 - Content oriented search result organization & presentation
- 
- 



Q & A

- Questions & comments are welcome!

songhua.xu@yale.edu



**Thank you for your time and
attention!**

