Collection Guiding: Multimedia Collection Browsing and Visualization

Stéphane Marchand-Maillet
Viper – CVML – University of Geneva
marchand@cui.unige.ch
http://viper.unige.ch

Outline

- Multimedia data context
- Searching for data
- Managing data
  - Principles
  - Framework
- Examples
  - Images
  - Text
  - … Video
- Perspectives

Context

Huge mass of digital multimedia documents

Human memory can deal online with l.t. 1K photos

- Digital camera holds few 100's of photos
- Digital camera holds few hours of video
- 1h video = 90'000 images
- 1.5h MPEG video average quality = 1 GB
- My own home video collection ~30h
  - 600+ sequences (shots), 20+GB
- In 2002, 5Mo TeraBytes (5.10^{18}) of original data created

⇒ Need automated tools:

- Query-by-example (similarity search)
- Specific document retrieval (target search)
- Collection management (collection guiding)

Searching for data

- Current systems are query-based
  - QBE, Browsing
  - Implicit goal (target) to reach (retrieve)

⇒ The user is a customer to the system

- Typical limitations:
  - "Page zero" problem
  - Semantic gap
  - Interaction protocol
  - ....

Information search (retrieval)

General assumptions:

- There exists an information need
  - Punctual: one document
  - Broad: category, ensemble
- The user can formulate a description of this need
- The information repository is finite and static

Managing data

The user is a manager of the system

- Large data collection at hand
- No specific needs
- Just wants to keep things tidy:
  - Summarizing, Filtering
  - Sorting, Organizing
  - Annotation, Description

Examples

- Visual content provider
  - Need to know well the content of the collection (overview)
  - Need to create catalogs (summaries)
- Home imagery
  - Content classification
  - Content overview
  - Content annotation
- Massive (Blind) Web Harvester
  - Filtering
  - Classification

Challenge

To create a tool (or context or framework) that would allow a (naive) user to grasp the content of a data collection as quickly as possible

Baselines:

- "Linear visit" of the collection
- Random sampling of the collection

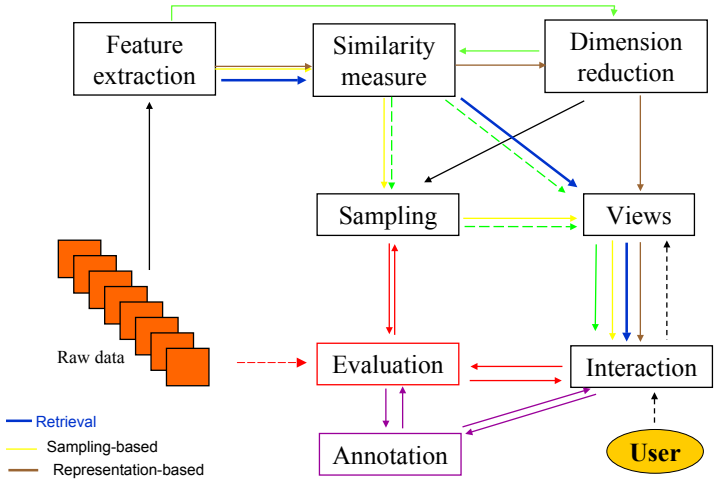
Main principles

- Intelligent sampling
  - Select a subset of the image collection that represent it well
  - Show *n* items (*n* given)
- Hierarchical visit
  - Develop when necessary only
- Organized visit
  - Follow a coherent path within the collection
  - Show *all* items

Earlier proposals (not exhaustive)

- D. Cutting *et al.* (1992)
  - Scatter/Gather for text (clustering)
- Y. Rubner (PhD, 1999)
  - EMD, MDS (2D, 3D)
- S. Craver *et al.* (1999)
  - Multi-linearization, space-filling curves
- I. Cox et al. (2000)
  - `PicHunter`, Bayesian Browsing
- K. Barnard (2001)
  - MDS, Clustering

Tools at hand



Statistical framework

- Statistical clustering (*k*-means,…)
- Hierarchical clustering (simple-link,…)
- Classification context (*x*DA, kernel-based,…)
- Information-theoretic processing
  - Coding theory
  - Sampling theory

Essentially, view the collection of items as a set of realizations of a random variable with given (estimated) pdf

- Requires a model (prior)
- Smooth out peculiarities

Statistical framework: clustering

- *k*-means clustering
- View *k* centers
- Develop interactively

+ Simple, fast
- Requires a model (prior)
- Rough approximation, smooth out peculiarities
- Not intuitive (no overlap of clusters)
- Not adaptive

Statistical framework: agglomeration

Hierarchical clustering
- Complete-/Single-link clustering

⇒ Dendogram

+ Choice of the coarseness
+ Helps exhibiting global structures of the collection
- Computationally costly
- May not behave robustly

Statistical framework: classification

- Learn *n* classes
- Sample according to classification (SV,…)

+ Trendy
+ Powerful (semantic)

- Requires (consistent/large) learning data
- Requires tuning of the technique

Statistical framework: Information

Preserve the collection entropy
- Similar to data coding

+ Theoretically sound
+ Implicit evaluation

- Smooth out outliers
- Approximation in practice (P(I)=?)

⇒ Some solutions (eg Reyni's entropy)

We keep the idea for evaluation

Discrete framework

The collection is a set of **discrete** instances in a given (possibly **non-metric**) high-dimensional feature space

⇒ Use optimal structures to characterize properties of the collection

- Minimum spanning tree:
  - Indicator of minimal proximity relationship
  - Mappable in 2D
- Set cover
  - Indicator of global item span
- Tours (eg TSP)
  - From nD to 1D
- …

OR framework

Utilizing discrete optimization techniques for optimal solutions

- Minimum Weighted Spanning Tree (MST)
- Shortest Path Spanning Tree
- Set Covering
- *k*-Median
- Traveling Salesman Problem (TSP)

Most are NP-complete ➡ Approximations
All are computationally inefficient ➡ interaction?

⇒ New Challenges…

OR-based sampling (1)

Optimal cover of image feature space

Given a "coverage power" for each image, find the minimum set of images that "covers" the complete collection

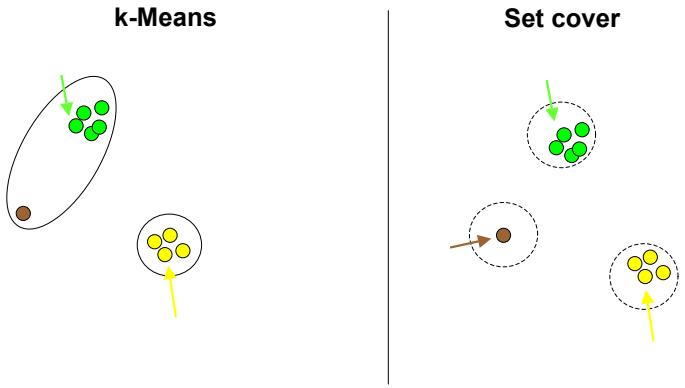
OR-based sampling (2)

+ Highlights outliers

- Coverage power (zone of influence)?
  - ⇒ Way of interaction ?
  - ⇒ Relation to k-Median?

- Computational load
  - Set cover has one nice approximation: greedy cover!

OR-based sampling (exple)



OR-based visit: formulation

To map the collection on a "near 1D" path so as to visualize it coherently
(like a guide in a museum!)

Visit a number of sites so as to minimize the sum of interdistances

⇒ Euclidean TSP (NP-complete)

⇒ Suboptimal solution from MST+DFS

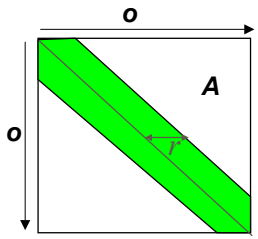
⇒ Incremental (interactive) solution from Lin-Kernigan heuristic

OR-based visit: new challenges

Alternative formulation: "Wide band" TSP

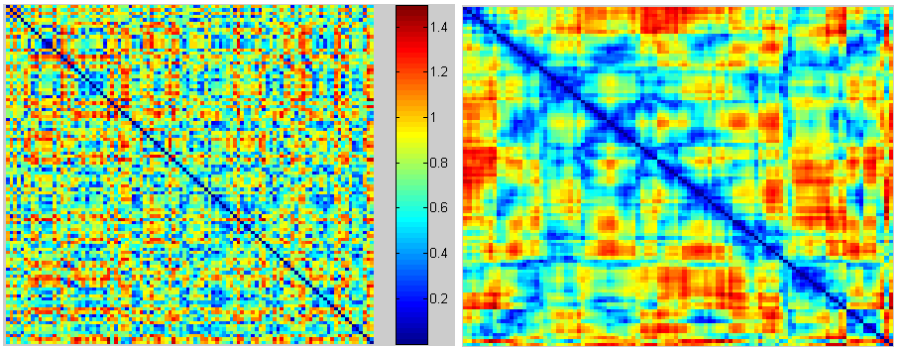
Reorganize the matrix of inter-distances so that the sum of *r* diagonals is minimum

$$o^* = \underset{o}{\operatorname{argmin}} \Sigma_{i,|k<r|} A_{ii+k}$$



- Empirically, if *r*=1 (classical TSP) we obtain a good approximation
- Solution for r>1 desirable!

Example



Dimension reduction

Idea: Preserve inter-item distances in both spaces

- Typical solutions
  - Multi-Dimensional Scaling
  - Sammon (NLM) mapping
  - CCA (from Self Org. Maps)
  - IsoMap (Geodesic distances)
  - LLE (local linear approximation)
  - FDP (physical model)
  - Relational Perspective Map (geometrical model)

Dimension reduction (tests)

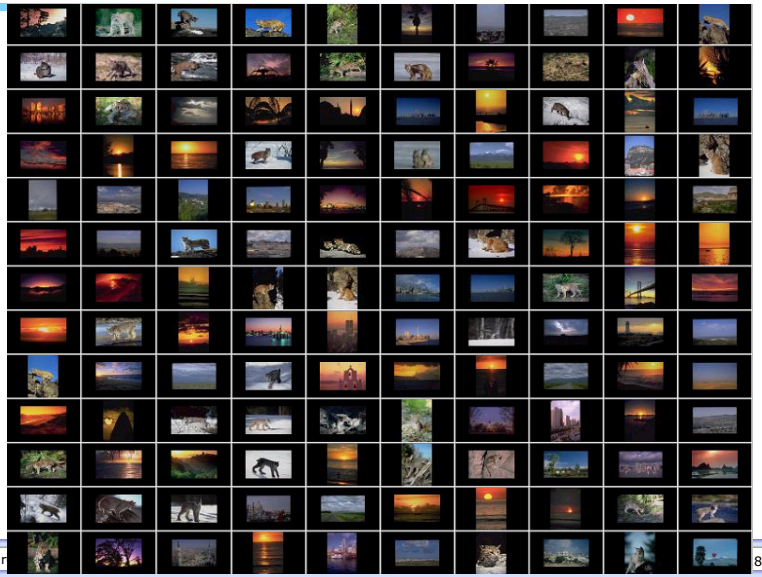
- CCA seems to work best
- However has problems with large dynamic range of distances

⇒ Hierarchical CCA
- Perform clustering
- Map every cluster onto a set of points
- At every scale, re-center point clusters

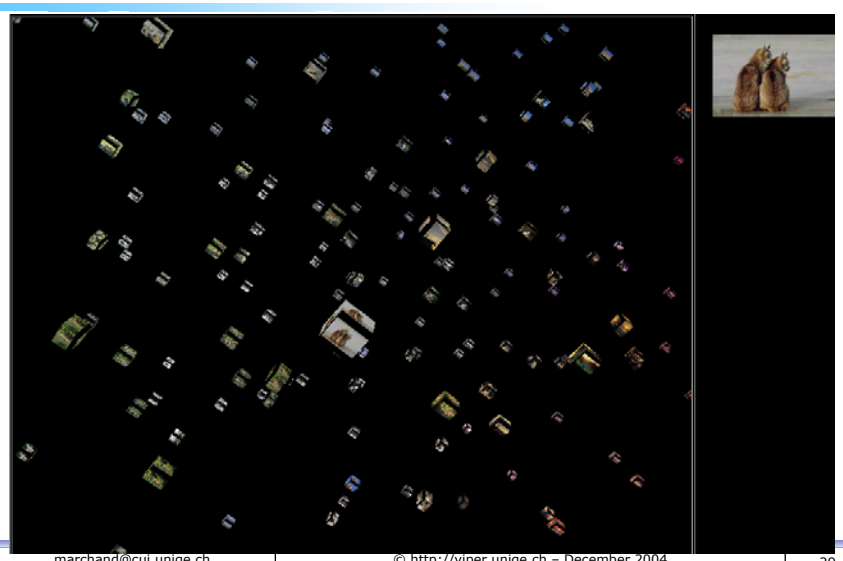
Image collection management

- Part of Corel collection
- Color and texture simple features
- « GIFT-like » similarity

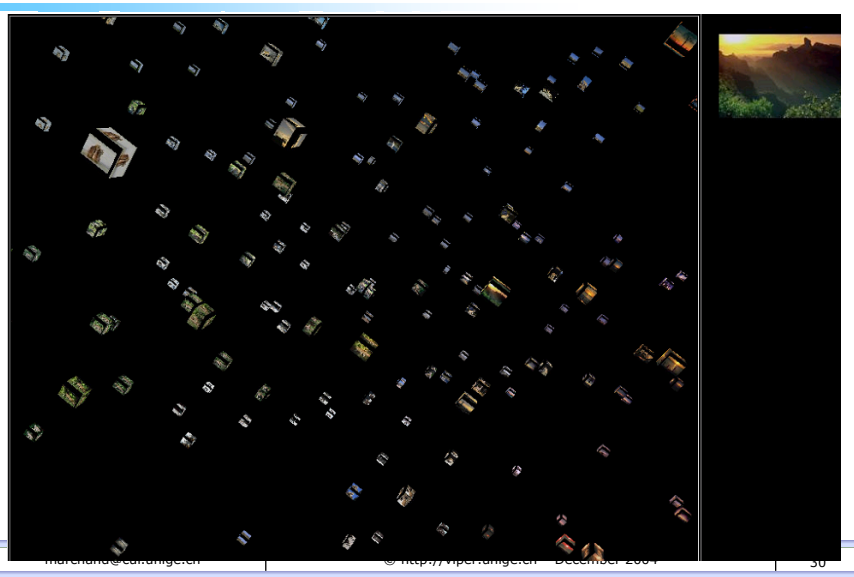
Visual Collection management (raw)



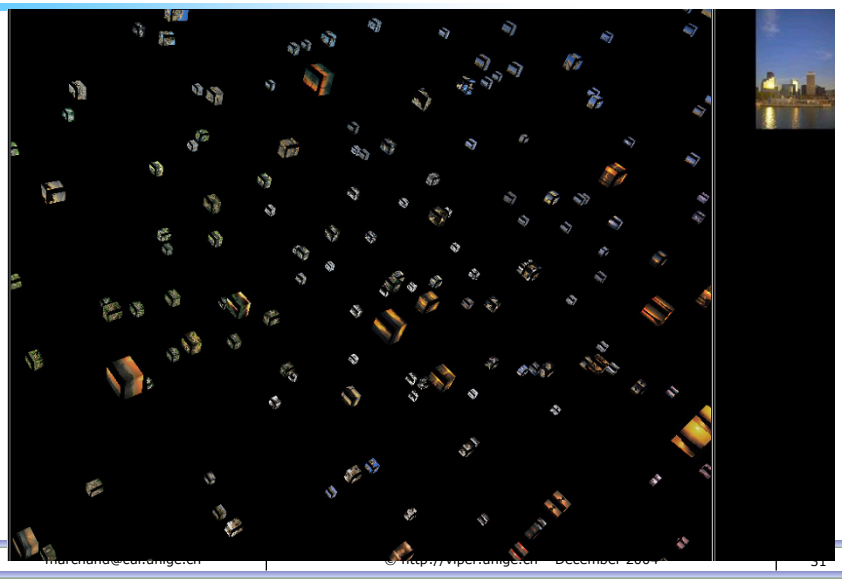
Visual Collection management (3D)



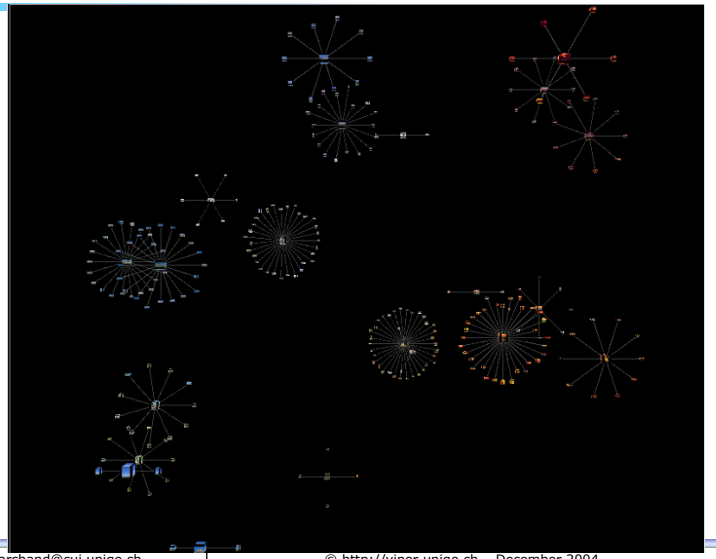
Visual Collection management (3D)



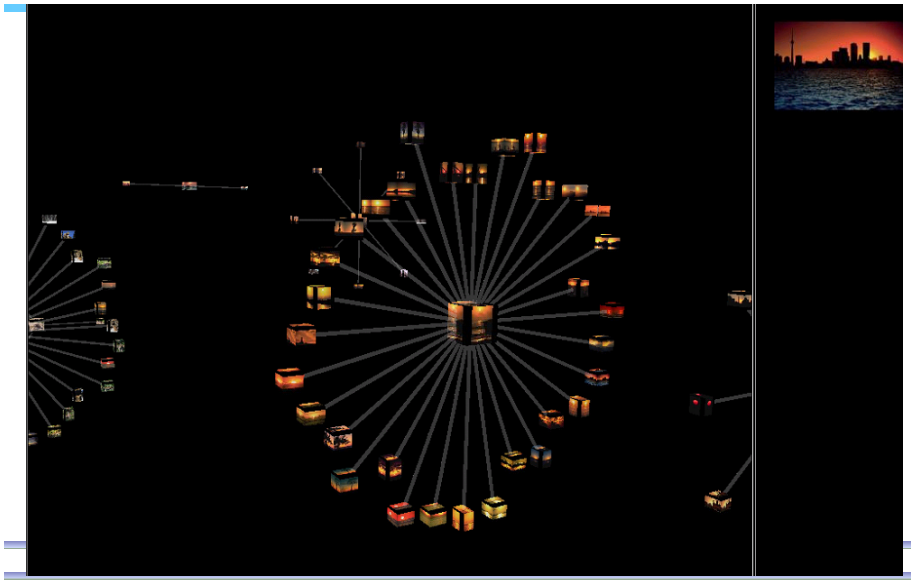
Visual Collection management (3D)



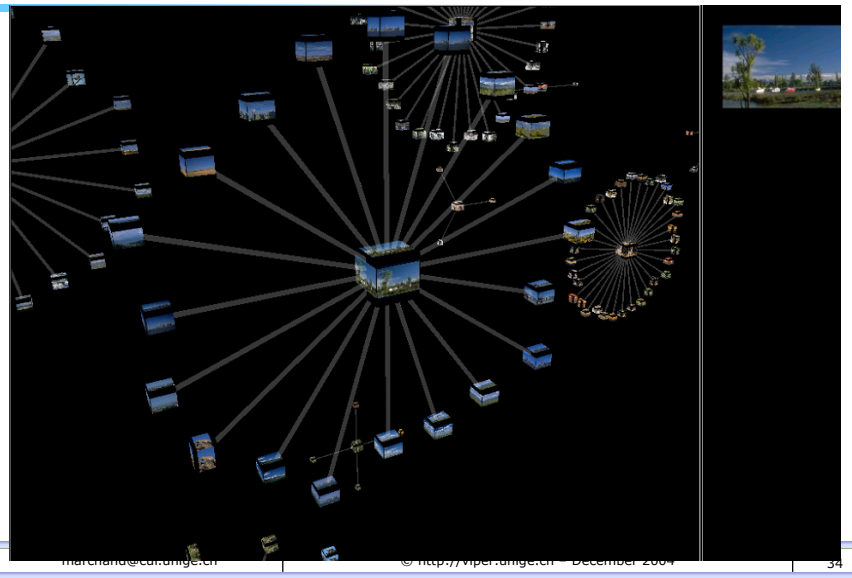
Visual Collection management (clusters)



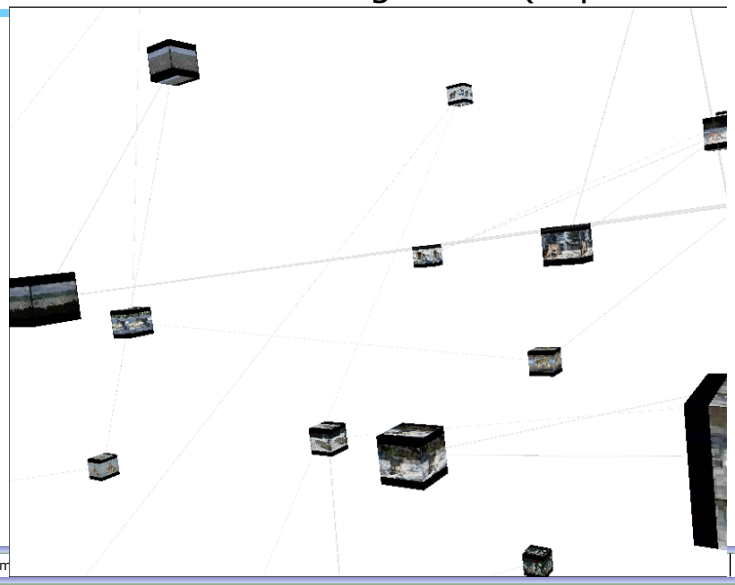
Visual Collection management (clusters)



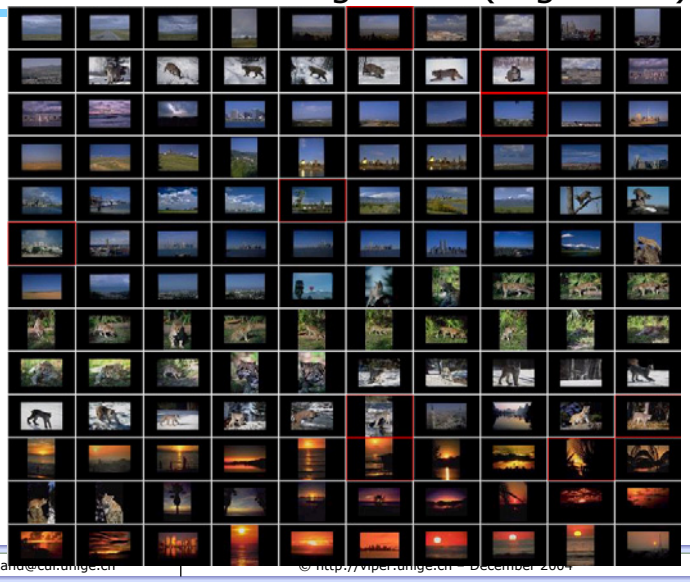
Visual Collection management (clusters)



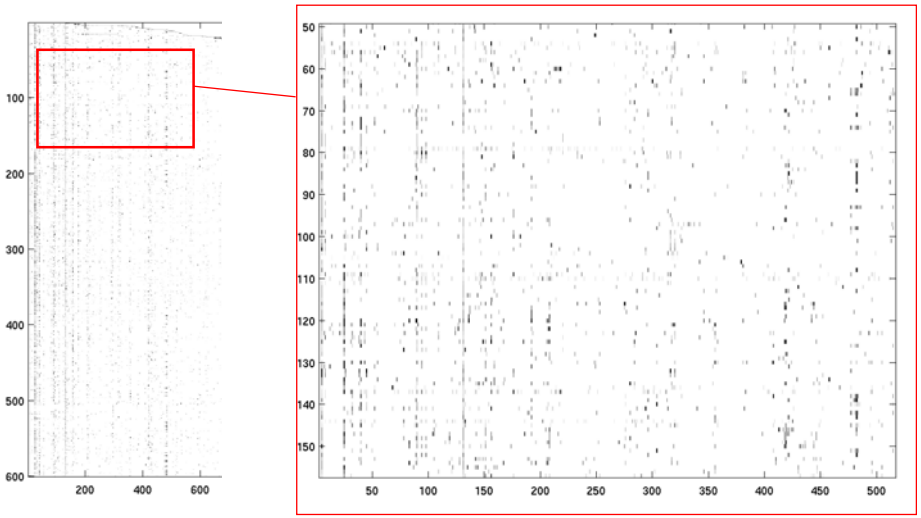
Visual Collection management (exploration)



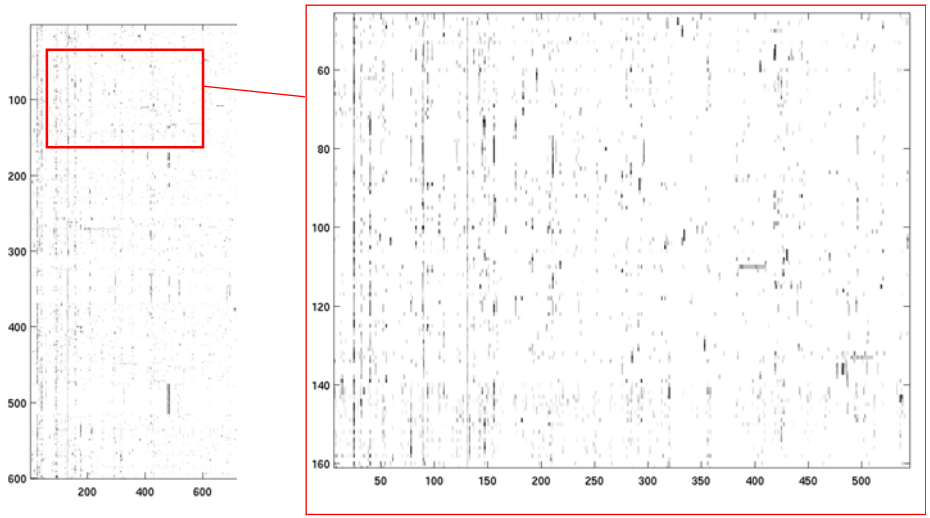
Visual Collection management (organised)



Text collection management



Text collection management



Open issues

- Estimate the **diversity** of the collection
  - Clustering
  - Entropy?
- Create intelligent (**informative**) **sampling**
  - See above
- **Organise** the collection
- Feature space **dimension reduction**
  - Similarity-based
  - Structure-based (order preserving operations)
- Create appropriate **visualization**
  - 1, 2, 3D
  - Metaphors
- Include **interaction**
- **Evaluation**
  - Tests, ground truth data

Perspectives

- Advances in previous open issues
- Efficient (approximation) algorithms for allowing interaction
- Embed this framework into an actual data management setup
  - Images
  - Texts
  - Video
  - Generic data