A continuum between serendipitous browsing and query-based search for MM information access

AMR 2009, Madrid, Spain

Julien Ah-Pine\textsuperscript{1}, Jean-Michel Renders\textsuperscript{1}, Marie-Luce Viaud\textsuperscript{2}

\textsuperscript{1}Xerox Research Centre Europe

\textsuperscript{2}Institut National de l’Audiovisuel

September 2009
1. Introduction and motivations

2. Main functionalities and architecture of our system
   - GUIs and interlinked multi-scale maps
   - Ranker/Scorer
   - Graph Layout Map Builders

3. Proof of concepts

4. Conclusions and future work
Outline

1 Introduction and motivations

2 Main functionalities and architecture of our system
   - GUIs and interlinked multi-scale maps
   - Ranker/Scorer
   - Graph Layout Map Builders

3 Proof of concepts

4 Conclusions and future work
Motivations

- Browse-based search: navigation in a dataset which is represented in a structured way (proximity relations, cluster).
- Query-based search: search of relevant objects in a dataset given a query and user feedback.
- We propose a system that attempts to couple those two search behaviors:
  - better ways for the user to express her information needs, to access information and acquire knowledge.
  - better treatment of multimodal information retrieval that deals with the problem of the semantic gap (in this talk, we deal with Image/Text data).
Outline

1. Introduction and motivations

2. Main functionalities and architecture of our system
   - GUIs and interlinked multi-scale maps
   - Ranker/Scorer
   - Graph Layout Map Builders

3. Proof of concepts

4. Conclusions and future work
Main characteristics of the system

- Interlinked multi-scale visualization and navigation: two linked maps (graph-based representations), one global (the whole dataset) the other one local (wrt to the search results).
- Multimodal views of the data: the graphs and the searches could rely on different modalities (text, image, hybrid) and can change at any time.
- Flexible multimodal relevance feedback: the labeling of the image part of a same text/image object can be independent of its text part.
- Adaptable search/development metrics: search modality, locality and forgetting factors.
Architecture of the system

GUI

Textual Query Form
Image Query Form
Global Map
Local Map
Selection and Relevance Labeling Tools
Search & Discover Control Parameters (modality, locality, recency)
Standard Text Search Engine
Standard Image Search Engine
Ranker/Scorer
Graph-layout Algorithm (Visualization)
Graphical User Interfaces

Two interlinked and multi-scale maps in order to:

- allow the user to **be aware of the boundaries of her current search direction** by visualizing on the global map the search results and by analyzing the neighbors and the clusters structure.
- allow the user to **better control the exploration/exploitation trade-off**.
Multi-scale, multimodal maps and their complementarity

**GLOBAL MAP**
- Static
- Multimodal views (text, image or cross-media similarities)

- Retrieved items are highlighted
- Search directions are tracked

- Items from global map can be added to the local map
- The multi-modal view allows a better exploration

**LOCAL MAP**
- Dynamic
- Flexible labelling of MM items
- Adaptive metrics
A relevance feedback model that integrates the following features:

- The user can label the image part of an object independently of its text part. This **flexibility** allows the user to better express her multimedia information need.

- The user can **promote any types of similarity** (text, image or hybrid) in her search at any time of a session.

- A **forgetting factor** that assumes that recently annotated items have more weight than less recently annotated ones. This aspect is similar to the Ostensive model introduced in (Campbell and Van Rijsbergen, 1996)

- A **locality factor**: the user can select a special subset of annotated objects that are given an extra weight compared to others (almost as if the user wanted to start a quite new query but without starting a new session from scratch).
Extension of Rocchio’s relevance feedback method (1/2)

\[
f^{t+1}(x) =
\gamma_T^t \left[ \sum_{y \in T^+_t} \sum_{y' \in T^+_t} \frac{\alpha_T^t(y)}{\alpha_T^t(y')} \right] \left( S_T(y, x) + \lambda_T \frac{\sum_{z \in B_T^t(y)} S_T(y, z)S_I(z, x)}{\sum_{z' \in B_T^t(y)} S_T(y, z')} \right)

- \sum_{y \in T^-_t} \sum_{y' \in T^-_t} \frac{\beta_T^t(y)}{\beta_T^t(y')} \left( S_T(y, x) + \delta_T \frac{\sum_{z \in B_T^t(y)} S_T(y, z)S_I(z, x)}{\sum_{z' \in B_T^t(y)} S_T(y, z')} \right)

+ \gamma_i^t [ \ldots ] \]
Extension of Rocchio’s relevance feedback method (2/2)

\[ \alpha^t_T(y) = \begin{cases} 1 & \text{if } y \in S^t_T \\ \frac{1}{1-loc^t} (1 - forg^t_T)^{m_T(y)} & \text{if } y \notin S^t_T \text{ and } m_T(y) \geq 0 \\ 0 & \text{if } y \notin S^t_T \text{ and } m_T(y) = -1 \end{cases} \]

where:

\[ m_T(y) = \begin{cases} t - date_T(y) & \text{if } S^t_T = \emptyset \\ \min_{z \in D^t_T(y)} (date_T(z) - date_T(y)) & \text{if } S^t_T \neq \emptyset \text{ and } D^t_T(y) \neq \emptyset \\ -1 & \text{if } S^t_T \neq \emptyset \text{ and } D^t_T(y) = \emptyset \end{cases} \]
Graph Layout Map Builders

Global maps are static and computed off-line.

1. Based upon force directed layout algorithms and the LinLog energy model (Noack, 2005)
2. And an agglomerative hierarchical clustering algorithm to identify clusters in the 2D space.

The local map layout is a dynamic process.

1. New results are appended but by slightly perturbing the previous layout (present objects keep their mutual similarity relations).
2. It is based upon the Fruchterman-Rheingold layout algorithm (Fruchterman and Rheingold, 1991).
3. Optionally, a clustering algorithm could be applied as well in the 2D local map, in order to avoid redundancy and to favor quick local exploration.
Outline

1. Introduction and motivations

2. Main functionalities and architecture of our system
   - GUIs and interlinked multi-scale maps
   - Ranker/Scorer
   - Graph Layout Map Builders

3. Proof of concepts

4. Conclusions and future work
Preliminary evaluation

Our proof of concepts is based upon:

- **Cognitive Walkthrough Inspection methodology** (achievability? efficiency?...) where an evaluator asks questions to an experimented user on how does the system help her to solve a task.

- We defined a task related to a topic, which is constituted of two subtasks: a specific search and a discovery analysis. Example *Eiffel Tower*: 1. find pictures dating from the beginning of the 20th century; 2. find all potential subtopics.

The pros and cons:

- + interlinked multi-scale maps and flexible feedback helps in solving the tasks.

- + the cross-modal similarities provides relevant results that would be difficult to reach with mono-modal similarities.

- +– the adaptable parameters, locality and forgetting factors are interesting. However, it is not easy to tune them simultaneously.
1 Introduction and motivations

2 Main functionalities and architecture of our system
   - GUIs and interlinked multi-scale maps
   - Ranker/Scorer
   - Graph Layout Map Builders

3 Proof of concepts

4 Conclusions and future work
Conclusions and future work

Summary:

- A system that couples browse-based search and query based-search by means of interlinked multi-scale maps and by letting the user the possibility to change her search direction.
- A multimedia feedback model that takes into account the different modalities in a flexible way and which supports adaptable metrics.

Future work:

- Our proof of concepts is encouraging but we need to go further in evaluating our system and particularly we should better understand to which extent each functionality improves the search process from the user viewpoint.
- We would like to design other functionalities that involve the use of multi-scale maps, in order to help the user further in her information seeking.
Many thanks for your attention!
Subtopics related to Eiffel Tower