Interaction

Seminar of multimedia retrieval (T-61.6030), TKK
Presentation 18.4.2008
Lauri Lahti

Information retrieval

- this seminar has given much attention to algorithms and techniques
- now we give focus especially to issues that are most visible to user
- emphasis on ways of interaction, their challenges and taking care of effectiveness

Overload of multimedia data

- challenges in interaction:
  - conflicts between mental and technical model
  - managing with incomplete metadata
  - helping users to express their needs
- users may learn their actual needs during the search process
- models of information seeking: “berry-picking” (clustering) and “information foraging” (strategies)

Main topics

- typical ways of user-system interaction (retrieval, dynamic query interaction, browsing and recommendation)
- different ways for a user to express needs (use of modalities)
- feedback from user about relevance of results
- personalization of information retrieval
- presenting results to the user and visualization
- adaptation of content to current conditions

Based on: Blanken et al. (2007). Multimedia retrieval. Springer.

Typical ways of user-system interaction

- a wide spectrum of level of user initiative
- from specific complex queries to automatically received suggestions
- some methods:
  - retrieval
  - dynamic query interaction
  - browsing
  - recommendation

Retrieval

- the process
  - person has an information need
  - person makes a query
  - system matches the query with the data collection
  - system returns matching items to user
- user should be able to formulate query with appropriate *syntax* for system and *semantics* for data
- limitations lead to an iterative process
- two common query types *concept-based queries* and *content-based queries*

Retrieval (cont.)

- concept-based queries
  - keywords, natural language or other semantically rich descriptions
  - matches with high-level features of data content, like Google Images search
  - easy language for user but may have problems with synonyms and incomplete metadata
- content-based queries
  - typically specified by paradigm of “query by example”, often fuzzy queries
  - matches with level-level features of data content, like similar colors, shape etc.
  - may allow to give weights to different aspects

Dynamic query interaction

- results are updated very fast by the system
- results immediately reflect the changes in user input
- iteration cycles become short and typically based on “flying through a information space”
- steering with sliders, buttons or other simple visual means of input
- posting a query and presenting results are combined
- queries less expressive than with natural language but more expressive than browsing (see next topic)

Browsing

- there is no specification of information need, just curious exploration in data collection
- aim to get impression of the search space
- useful way to try to find something without having a clear notion of its characteristics (like “beautiful scenary”)
- browsing through a collection of multimedia objects or within a multimedia object
- needs visualization, typically alternated with querying
- can be done also in a collection of keywords or for ex. in net of items linked by syntactic or semantic similarity

Recommendation

- all previous interaction types are based on initiative from the user (query or other input), belonging to so called pull category
- in recommendation the system takes the initiative, so called push category
- as soon as interesting content is available it will be offered to user
- addresses long-term interests and personal taste
- recommendation typically requires some initial training of the system and uses personalization techniques

Modalities of user input

- visual, auditory, tactile etc.
- natural interaction for humans is multimodal
- traditionally in human-computer interaction only one modality is used at time
- multimodal interaction could be beneficial (one-to-one/cross, single-multiple/multiple-single)
- use of modalities that suit best for the current content
- may allow the system to better interpret and handle the input the user provides
- problems with synchronization and fusion of the parallel inputs
- personal preferences of sequences and combinations of modalities and switching between them

Relevance feedback

- feedback from user about relevance of results
- based on feedback system can refine the original query and present better results
- for ex. clicking a result item in a list of alternative results indicates that the user finds that item promising
- specifically useful in fuzzy queries
- binary relevance feedback can offer user a list of results and the user indicates for each of them is it positive (offering an example) or negative (offering an counter-example)
- weighted relevance feedback uses more detailed scales of feedback values for initial result set

Personalization

- filtering information in a way that a person only receives relevant information addressing person's unique needs
- personalization can be successfully combined with recommendation (push category)
- also can be used in search engines (pull category), for ex. "java"
- typically requires first building an user profile
- asking profile information directly from user is problematic (user might not be able or willing to help)
- the user actions can be observed to build profile (problem of "buying a gift")
- typically asking and observing combined

Personalization (cont.)

- content filtering can be for ex. content-based filtering or social based filtering
- *content-based filtering* typically based on statistical description of preferences
- these preferences are measured on the level of the properties of the data items
- matching the user profile with multimedia items takes place in vector space

Personalization (cont.)

- both user profile and each multimedia item is expressed as a n-dimensional vector
- the closeness or similarity can be measured by the angle between the vectors
- challenges: dependent on metadata, metadata-properties mismatches, initialisation of a new user

Personalization (cont.)

- *social based filtering* uses the similar earlier experiences of a group of users
- items are recommended on the basis of user similarity rather than item similarity
- system tries to find users that have a similar user profile
- one common way of calculating the similarity is to use the Pearson correlation formula
- advantages: explicit content representation not needed, domain-independence
- disadvantages: large user group needed, problems with new items, users and unusual users

Presentation

- results of the queries need to be presented to user
- since multimedia objects are large typically just metadata of the objects is presented
- the actual object is shown after user selects a satisfactory item from the result list
- the most appropriate modality for each item should needs to be selected
- cognitive effect of simultaneous presentation of different modalities needs to be addressed

Presentation (cont.)

	visual verbal	visual nonverbal	auditory verbal	auditory non-verbal
visual verbal (text)	– – two pieces of text	+ text and image	– text and speech	+ text and music
visual non-verbal (image, animation)		– two images	++ image and speech	+ image and music
(auditory, verbal) speech, songs			– – two pieces of speech	+ speech and music
auditory non-verbal (music, environmental sound)				– two pieces of music

Combinations of modalities

Presentation (cont.)

- presentation needs attention for visualization
- visualization is more than presenting results, uses techngiques to interpret data for human needs
- visualization helps to present data in a more understandable
- basic form of visualization are treemaps (hierarchies), graphs (linked nodes), starfield display (nodes in 3D space) and browsing modes (compact presentation)

Content adaptation

- the way in which the multimedia is represented needs to be adapted to several aspects, for ex.
  - the capabilities of devices
  - the contexts the user is in
  - the capabilities of the user
- some types of content adaption are transcoding, transmoding and content summarization