Project no. 033104

MultiMatch

Technology-enhanced Learning and Access to Cultural Heritage
Instrument: Specific Targeted Research Project
FP6-2005-IST-5

D1.1.2 – State of the Art
Image Collection Overviews and Browsing

Start Date of Project: 01 May 2006
Duration: 30 Months

Organization Name of Lead Contractor for this Deliverable: ISTI-CNR

Final Version

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)

Document Information

Deliverable number:	D1.1.2
Deliverable title:	State of the Art Report
Due date of deliverable:	October 2007
Actual date of deliverable:	December 2007
Main Author(s): Participant(s):	S. Marchand-Maillet, E. Bruno, UniGE All Partners
Workpackage:	1
Workpackage title:	User Requirements & Functional Specification
Workpackage leader:	UNED
Dissemination Level:	RE (Restricted)
Version:	V1.0
Keywords:	visual information browsing, multimedia indexing and retrieval, user interaction, interface design, evaluation

History of Versions

Version	Date	Author (Partner)	Description/Approval Level
Initial	10.12.07	S. Marchand-Maillet E. Bruno (UniGE)	Initial version
Completed	12.12.07	S. Marchand-Maillet E. Bruno (UniGE)	Completed version for comments and approval of partners
V1.0	17.12.07	S. Marchand-Maillet E. Bruno (UniGE)	Amended based on comments from reviewers (Sheffield and Alinari)

Abstract

MultiMatch aims at complex, heterogeneous digital object retrieval and presentation. The development of the system implies addressing a number of significant research challenges in a multidisciplinary context. This report describes the state of the art in the development of image collection browsing and overviewing. This is motivated by the fact that such activities are complementary to search operations and may provide efficient solutions where search tools are deficient due to the lack of representative semantics within the documents.

Table of Contents

Document Information 1

Abstract 1

Executive Summary 3

1. Introduction 4

2. Image Collection Browsing 4

2.1 Browsing as extension of the query formulation mechanism 4

2.2 Browsing for the exploration of the content space 8

2.3 Browsing to aid content description 14

3. Multimedia Space Representation 14

3.1 Generic feature space representation 14

Item similarity measurement 14

Collection subsampling 14

3.2 Dimension reduction 15

4. Multimedia Collection Browsers 16

4.1 Extra image browsers 16

4.2 Related patents 17

4.3 Other media 17

5. Evaluation 21

6. Concluding remarks 22

References and relevant literature 23

Acknowledgments 25

Executive Summary

The objective of MultiMatch is to develop a multilingual search engine specifically designed for access, organization and personalised presentation of cultural heritage information. The development of the system thus implies addressing a number of significant research challenges in a multidisciplinary context. R&D expertise is required in a diverse set of system- and user-oriented research areas. On the system side, these relate to focused Internet crawling, information extraction and analysis, multilingual information access and retrieval, multimedia complex object management, and interface design. On the user side, relevant areas include user profiling, metadata and ontology studies, user/system interaction, and user-centered interface design. The technology in these areas tends to develop rapidly. For this reason, it was decided to prepare a detailed state of the art report in the initial phases of the project.

This report thus describes the state of the art in the development of image collection browsing and overviewing. This is motivated by the fact that such activities are complementary to search operations and may provide efficient solutions where search tools are deficient due to the lack of representative semantics within the documents. Further, initial evaluations of the work in MultiMatch have pinpointed the need for complements or alternatives to the Query-by-Example paradigm. Deliverable D1.1.1 has included an in-depth review of the latter. It is therefore natural to propose in this document a review of browsing technique in a context close to or departing from retrieval.

This overview is made with the view of evaluating browsing principles and technologies as useful in the context of MultiMatch.

1. Introduction

The objective of MultiMatch is to develop a multilingual search engine specifically designed for access, organization and personalized presentation of cultural heritage information. The development of the system thus implies addressing a number of significant research challenges in a multidisciplinary context. R&D expertise is required in a diverse set of system- and user-oriented research areas including, on the system side, focused Internet crawling, information extraction and analysis, multilingual information access and retrieval, multimedia complex object management, interface design, and, on the user side, user profiling, metadata and ontology studies, user/system interaction, interface design from the user perspective. The technology in these areas tends to develop rapidly. For this reason, and as part of the project activity, it was decided to prepare a detailed state of the art report in the initial phases of the project.

This report thus describes the state of the art in the development of image collection browsing and overviewing. This is motivated by the fact that such activities are complementary to search operations and may provide efficient solutions where search tools are deficient due to the lack of representative semantics within the documents. Further, initial evaluations of the work in MultiMatch have pinpointed the need for complements or alternatives to the Query-by-Example paradigm. Deliverable D1.1.1 has included an in-depth review of the latter. It is therefore natural to propose in this document a review of browsing technique in a context close to or departing from retrieval.

This overview is made with the view of evaluating browsing principles and technologies as useful in the context of MultiMatch. Discussions are made in that context.

2. Image Collection Browsing

Many current information management systems are centered on the notion of a *query*. This is true over the Web (with all classical Web Search Engines), and for Digital Libraries. In the domain of multimedia, available commercial applications propose rather simple management services whereas research prototypes are also looking at responding to queries (see section 4 for details and examples).

The notion of browsing comes as a complement or as an alternative to query-based operations in several possible contexts that we detail in the following sections.

2.1 Browsing as extension of the query formulation mechanism

In the most general case, multimedia browsing is designed to supplement search operations. This comes from the fact that the multimedia querying systems largely demonstrate their capabilities using the Query-by-Example (QBE) scenario, which hardly corresponds to a usable scenario.

Multimedia search systems are mostly based on content similarity. Hence, to fulfill an information need, the user must express it with respect to relevant (positive) and non-relevant (negative) examples (Smeulders, 2000). From there, some form of learning is performed in order to retrieve the documents that are the most similar to the combination of relevant examples and dissimilar to the combination of non-relevant examples.

The question then arises of how to find the initial examples themselves. Researchers have therefore investigated new tools and protocols for the discovery of relevant bootstrapping examples. These tools often take the form of browsing interfaces whose aim is to help the user exploring the information space in order to locate the sought items.

The initial query step of most QBE-based systems consists in showing images in random sequential order over a 2D grid (Smeulders et al, 2000). This follows the idea that a random sampling will be representative of the collection content and allow for choosing relevant examples. However, the chance for gathering sufficient relevant examples is low and much effort must be spent in guiding the system towards the relevant region of information space where the sought items may lie.

Similarity-based visualization ((Chen, 2000), (Cinque, 1998), (Leeuw, 2003), (Moghaddam, 2004), (Nazakato, 2001), (Nguyen and Worring, 2006), (Nguyen and Worring, 2008), (Rubner, 1999), (Vertan, 2002)) organizes images with respect to their perceived similarities. Similarity is mapped onto the notion of distance (Section 3.1) so that a dimension reduction technique (see section 3.2) may generate a 2D or 3D space representation where images may be organized. Figure 1 illustrates the organization of 500 images based on color information using the MDS dimension reduction (Rubner, 1999).

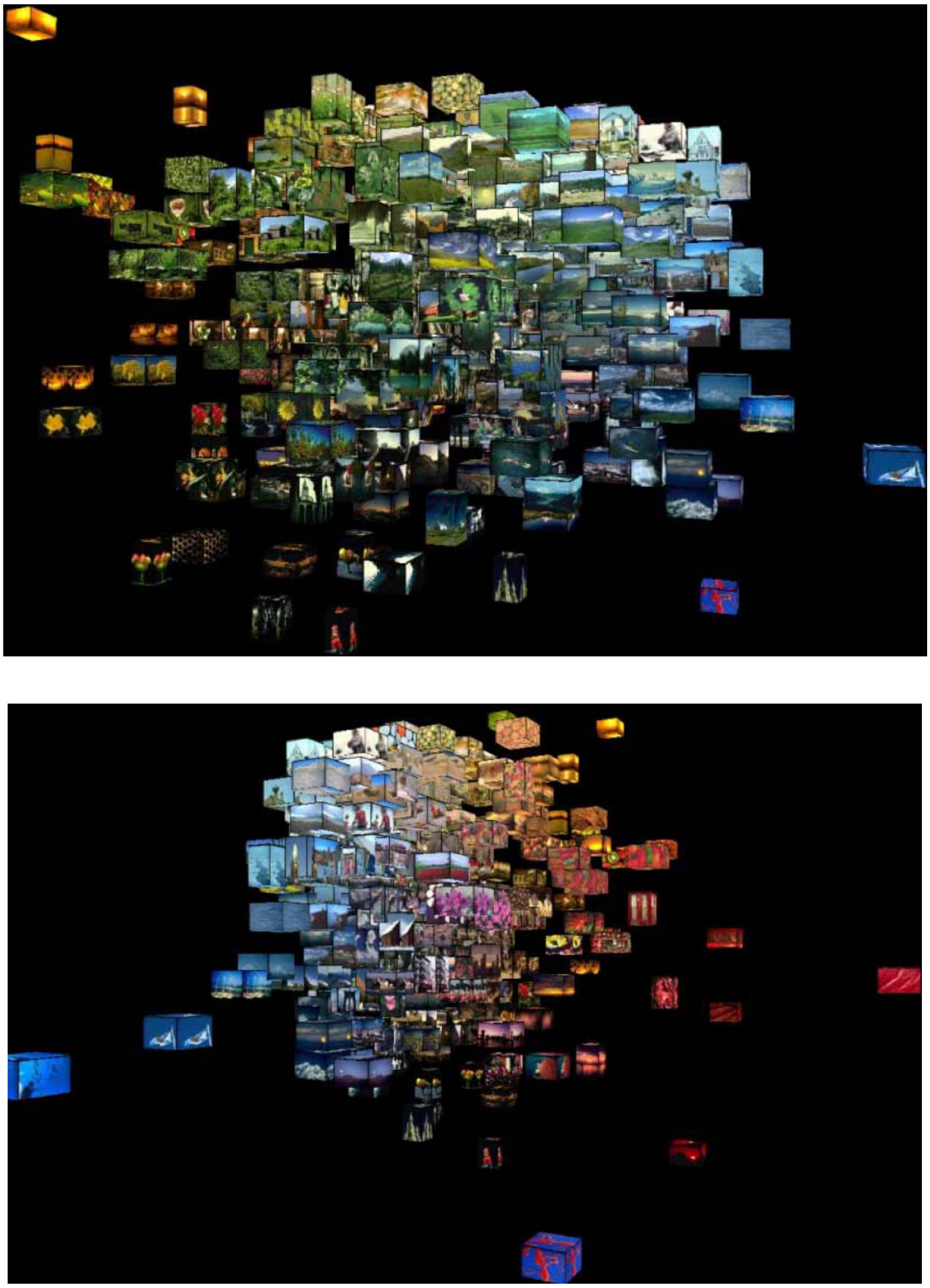


Figure 1. Two views of the MDS mapping of 500 images based on color information

This type of display may be used to capture feedback by letting the user re-organise or validate the displayed images. Figure 2 shows a screenshot of the interface of El Niño (Santini et al., 2001).

Figure 2. Interface of the El Niño system (Santini et al., 2001)
where image similarity is mapped onto planar distance

Specific devices may be used to perform such operations. Figure 3 shows operators sitting around an interactive table for handling personal photo collections (Moghaddam, 2004).

Figure 3. The PDH table and its artistic rendering (from (Moghaddam, 2004))

In Figure 5, an operator is manipulating images in front of a large multi-touch display (PerceptivePixel, 2007).

Figure 4. Manipulating images over touch.-enabled devices (from (PerceptivePixel, 2007))

Alternative item organizations are also proposed such as the Ostensive Browsers (see Figure 5 and (Urban, 2005)) and interfaces associated to the NN^k paradigm (Heesch, 2004).

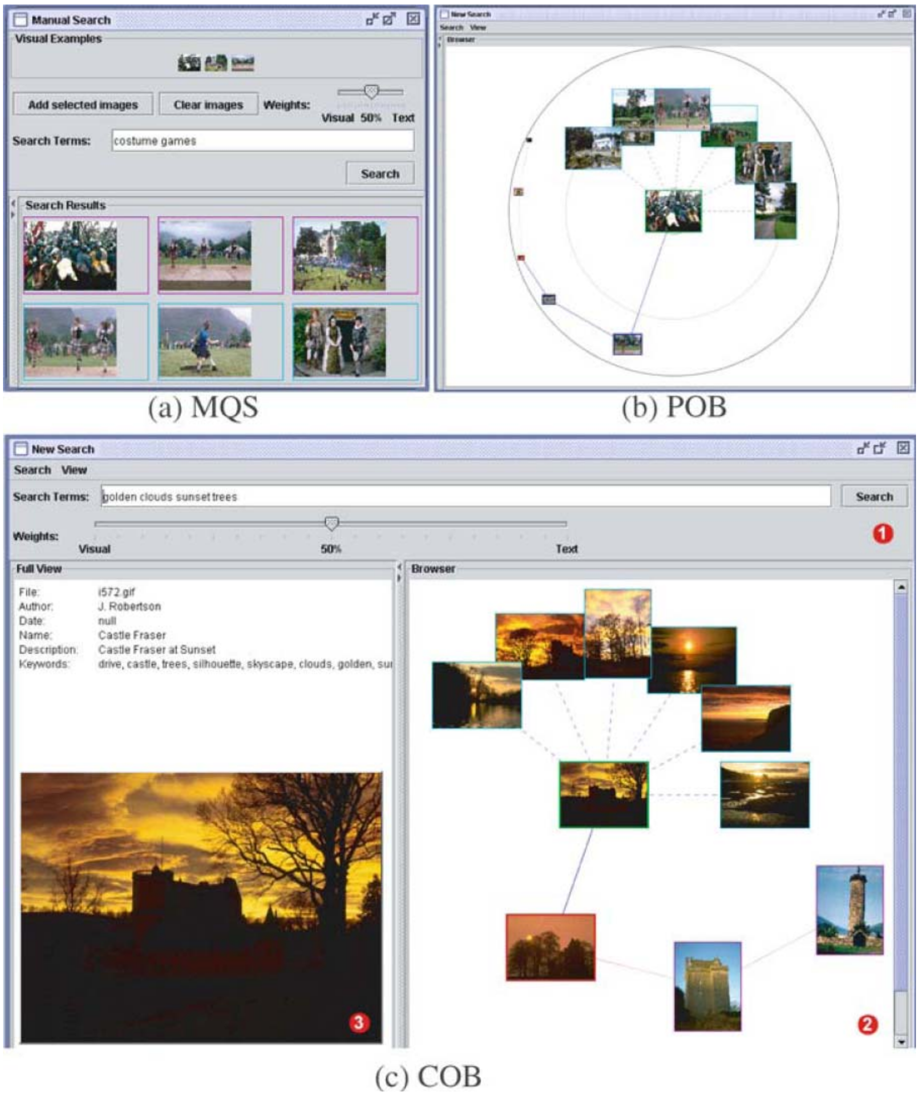


Figure 5. The Ostensive Browsers (Urban, 2005)

All these interfaces have in common the fact of placing multimedia retrieval much closer to human factors and therefore require specific evaluation procedures, as detailed in Section 5.

Although somewhat different, it is worth mentioning here the development of the Target Search browsers. Whereas using QBE-based search a user may formulate a query of the type “show me everything that is similar to this (and not similar to that)” and thus characterize a *set of images*, using Target Search, the user is looking for a *specific image* (s)he knows is in the collection. By iteratively providing *relative* feedback on whether some of the current images are closer to the target than others, the user is guided to the target image. This departs from the QBE-based search where the feedback is absolute (“this image is similar to what I look for, whatever the context”). In that sense, Bayesian search tools may be considered as focused collection browsers.

In this category, the PicHunter Bayesian browser (Cox, 2000) is one the initial developments. It has been enhanced with refocusing capabilities in (Müller, 1999) via the development of the Tracker system.

2.2 Browsing for the exploration of the content space

In the above cited works, browsing is seen as an alternative to the random picking of initial examples for the QBE paradigm. Here, we look at browsing from a different point of view.

In this setup, the user aims at overviewing the collection with no specific information need. Simply, (s)he wishes to acquire a representative view on the collection. In some respect, the above developments may be included into this category as overviews of the sub-collection representative to the query in question.

In (Kustanowitz, 2005), specific presentation layouts are proposed and evaluated (see also section 5). The interface aims at enhancing the classical grid layout by organising related image groups around a central group (see Figure 6).

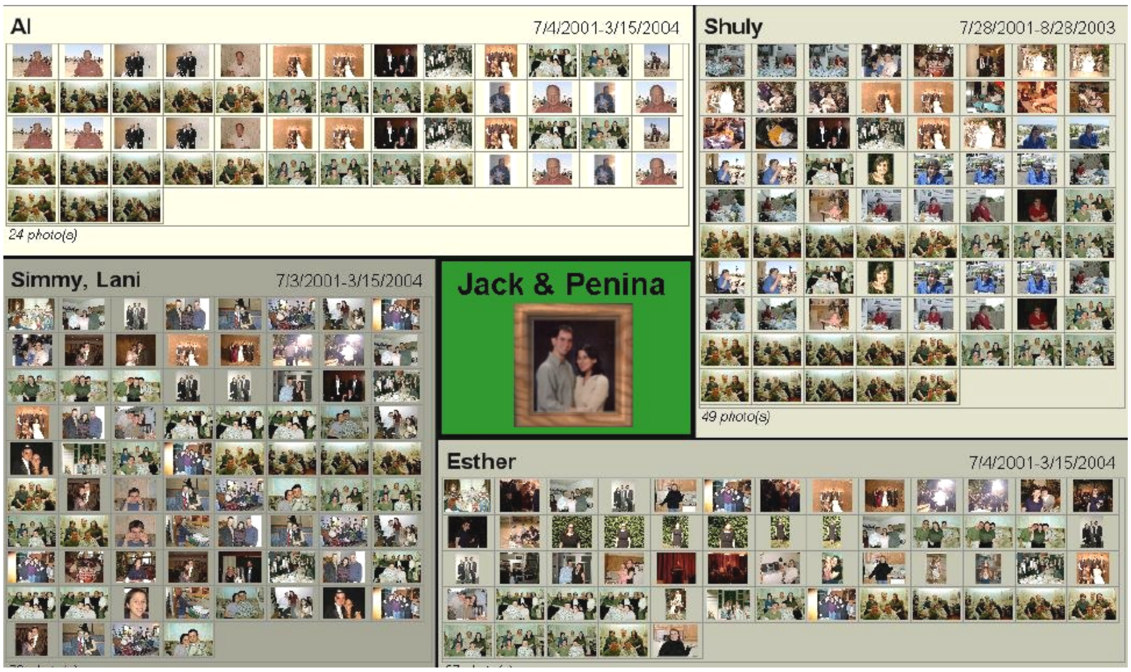


Figure 6. Bi-level radial layout (Kustanowitz, 2005).

Somewhat similar is the earlier development of PhotoMesa (Bederson, 2001) which aims at browsing image hierarchies using treemaps.

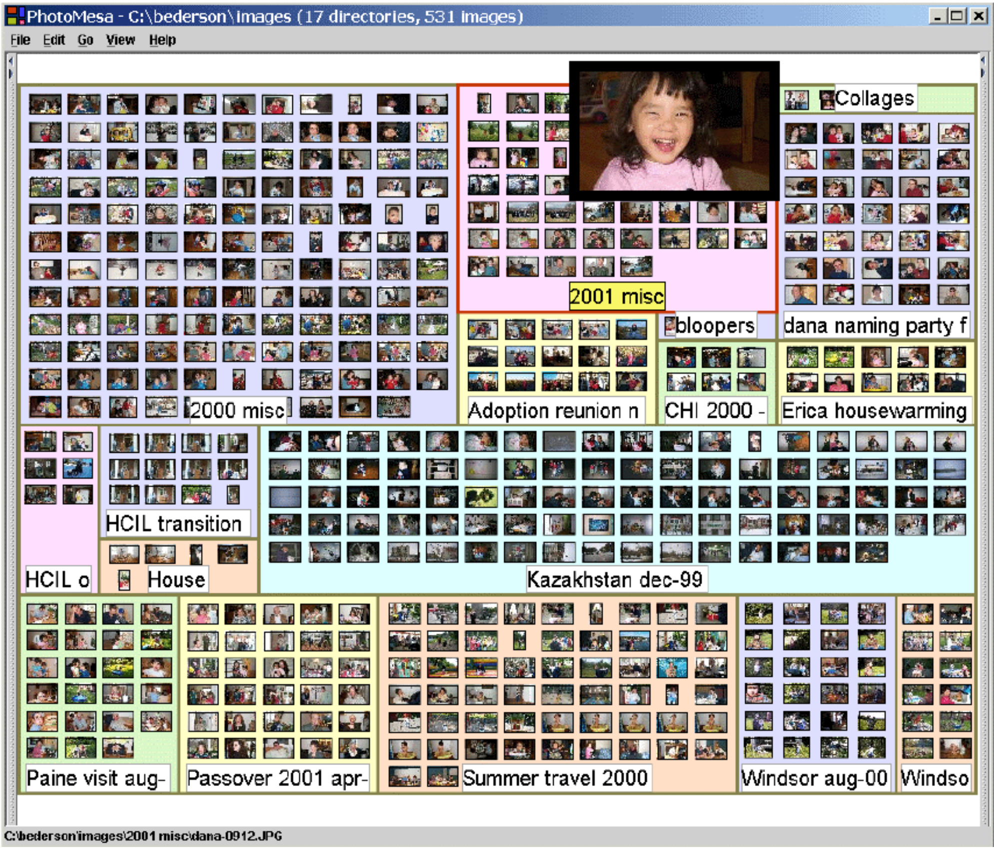


Figure 7. Screenshot of PhotoMesa, based on TreeMaps (Bederson, 2001)

Hierarchies are also studied in depth in the Muvis system, both for indexing and browsing via the Hierarchical Cluster Tree (HCT) structure (Kiranyaz, 2008). In Figure 8, an example of hierarchical browsing of a relatively small image collection (1000 images) is given.

In (Craver, 1999), the alternative idea of linearising the image collection is presented. The collection is spanned by two space-filling curves that allow for aligning the images along two intersecting 1D path. The reason for allowing two paths is that while two neighboring points on a space-filling curve are neighbors in the original, the converse is not guaranteed to be true. Hence, two neighboring points in the original space may end up far apart on the path. The use of two interweaved curves may alleviate this shortcoming.

At every image, each of the two paths may be followed in either of the two directions so that at every image, 4 directions are allowed. A browser shown in Figure 9 is proposed to materialize this visit.

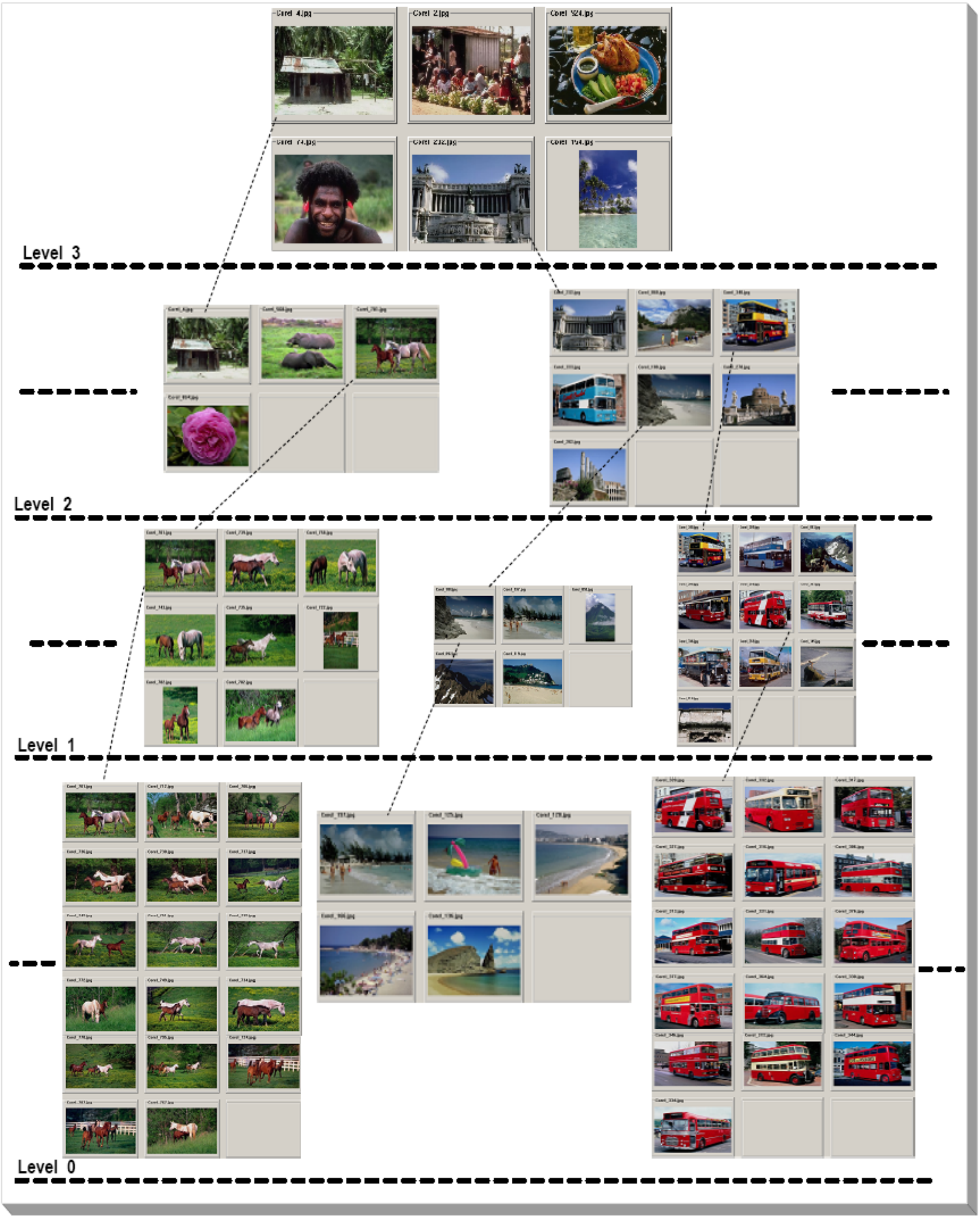


Figure 8. An example of HCT-based hierarchical navigation (Kiranyaz, 2008) on the 1k Corel image collection

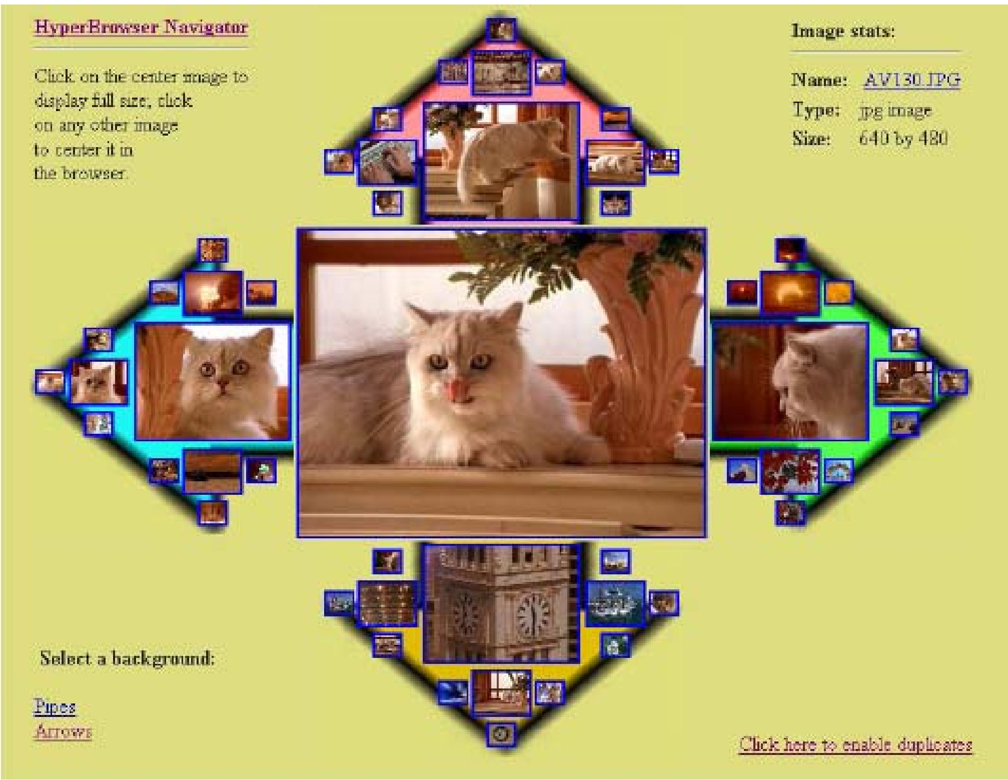


Figure 9. Multi-linearization browser (Craver, 1999)

In (Marchand-Maillet, 2005), the principle of Collection Guiding is introduced. Given the collection of images, a path is created so as to "guide" the visit of the collection. For that purpose, image inter-similarity is computed and the path is created via a Travelling Salesman tour of the collection. The aim is to provide the user with an exploration strategy based on a minimal variation of content at every step. This implicitly provides a dimension reduction method from a high-dimensional feature space to a linear ordering. In turn, this allows for emulating sort operations on the collection, as illustrated in Figure 10.

Figure 10. Image sorting via the Collection Guide (left) random order (right) sorted list

The Collection Guide provides also several multi-dimensional arrangements (see Figure 11). However, it is clear that these (as the ones presented in the above section) are conditioned to the quality of the dimension reduction strategy. In (Szekely, 2007), the underlying data cluster structure is accounted for so as to deploy valid dimension reduction operations (see section 3 below for more details).

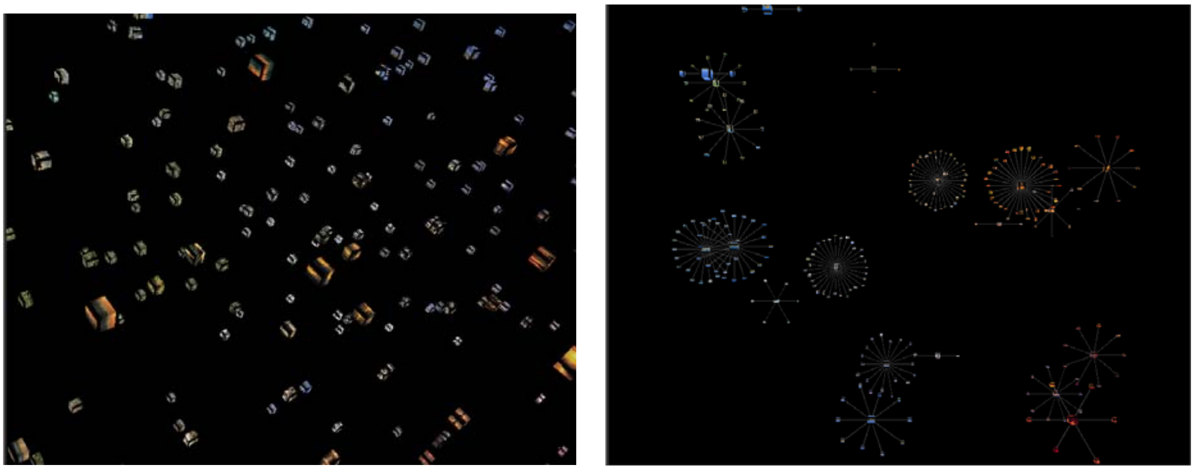


Figure 11. Examples of displays provided by the Collection Guide.
(left) generic 3D mapping (right) planet metaphor

Finally, at the border between exploration and search, *opportunistic search* is "characterised by uncertainty in user's initial information needs and subsequent modification of search queries to improve on the results" (Pu, 2003), (Janecek, 2003). In (Pu, 2003), the authors present a visual interface using semantic fisheye views to allow the interaction over a collection of annotated images. Figure 12 displays interfaces associated with this concept.

Figure 12a. Displays associated with the opportunistic search mechanism
(from (Pu, 2003) and (Janecek, 2004))

Figure 12b. Displays associated with the opportunistic search mechanism (from (Pu, 2003) and (Janecek, 2004))

Faceted browsing (Hearst, 2006), oriented towards search is also at the limit between exploration and querying as it also for filtering a collection while smoothly and interactively constructing complex queries. Figure 13 displays an example application of Faceted Search using the Flamenco toolbox for a collection of annotated images.

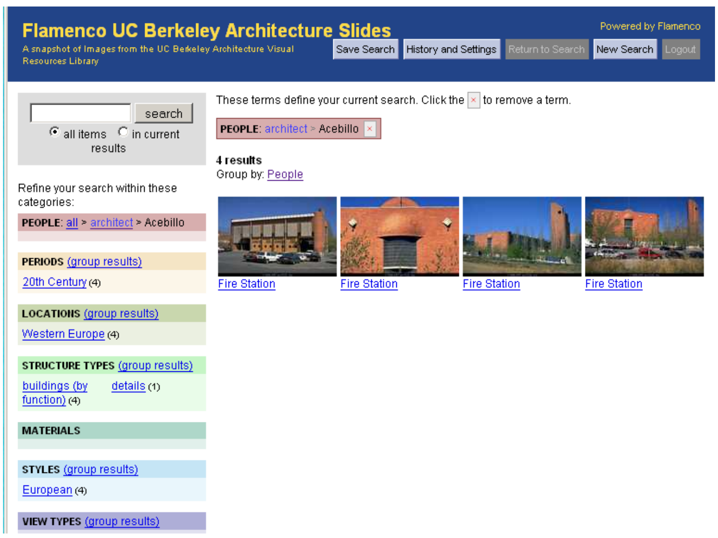


Figure 13. UC Berkeley Architecture Image Library (**Flamenco** toolbox)

2.3 Browsing to aid content description

While retrieval and browsing are in general passive to the collection (*ie* the collection stays at it is), these operations may also be used to enrich the collection content. In (Kosinov, 2003), the authors have reviewed and proposed several models that allow for the semantic augmentation of multimedia collections via interacting with them. This follows the line of the Semantic Web and associated domains of knowledge management. In this line, the work proposed in (Schreiber, 2001) relates to ontology management and image description.

3. Multimedia Space Representation

From a multimedia (image, in our case) collection, one should derive a representation that is both easy to handle via mathematical tools but which also accounts for the intrinsic meaning (semantics) of the content. From there, operations such as sampling and visualization are made possible. We overview briefly the possibilities in the next sections.

3.1 Generic feature space representation

There are well-known image representation techniques in the image compression and retrieval literature (Smeulders, 2000). Among them, features such as color, texture and shape emerge as the most global dominant cues for image content characterization.

The task of feature selection is typically associated with data mining. In our context, one may perform feature selection based on several criteria. Typical reported work is based on informative measures associated with predefined features or aimed at optimizing a given criterion by the design of abstract feature sets.

Item similarity measurement

Distance measurement depends on the space within which information is immersed. In the case of color for example, it is known that distance measurements within the RGB color cube do not correspond to any perceptual similarity. To this end, the HSV, Luv and CIELa*b* color spaces have been proposed within which simple Euclidean measurement correspond to perceptual distances.

A variety of distance functions exist and may be used for characterizing item proximity (Duda, 2000). The simplest distance functions that may be used are those derived from the Minkowsky distance (L_k norm) formula. Here, all coordinates are taken equally, meaning that we assume the fact of an isotropic space. If we assume that coordinates are realizations of a random variable with a known covariance matrix, then the Mahalanobis distance may be used. More sophisticated distance functions exist, such as the Earth Mover's Distance (Rubner, 1999).

Collection subsampling

Associated with the concepts of exploration and browsing is the concept of summarization. Summarization is an approach commonly taken for presenting large content and involves a clear understanding of the collection diversity for performing sampling.

The most common way of performing sampling is to use the underlying statistics of the collection. Typically, within the feature space, local density is analyzed. Dense regions of this space will be represented by several items whereas sparse regions will mostly be ignored within the representation. More formally, strategies such as Vector Quantization (VQ) may be used to split the space into cells and only consider cell representatives. *k*-means clustering is one of the most popular VQ techniques.

A geometrical interpretation of VQ is that of defining a Voronoi (Dirichlet) tessellation of the feature space such that each cell contains a cluster of data points and each centroid is the seed of the corresponding cell. This tessellation is optimal in terms of minimizing some given cost function, embedding the assumption over the properties of the similarity measurement function in the image representation space.

A radically different approach is to perform hierarchical clustering on the data. Initial data points form the leaves of a tree called dendrogram. The tree is built upon dependence relationships between data points. In the single-link algorithm, a point is agglomerated with its nearest neighbor, forming a new data point and a node within the tree. The algorithm stops when all points are gathered. Alternatives (complete-link and average-link) preserve the internal structure of clusters when merging.

The dendrogram obtained may then be the base for sampling the collection, as each level of the dendrogram shows a view of the collection. By defining collection samples as closest to the tree nodes at one given level, one obtains an incremental description of the collection.

3.2 Dimension reduction

So far, we have considered items as represented by vectors in the feature space. However, two aspects of this mathematical modeling should be inspected. First, we have defined distances and similarity measures irrespectively of the feature space dimensionality. However, it is known that this dimensionality has an impact on the meaningfulness of the distances defined (Aggarwal, 2001). This is known as the *curse of dimensionality* and several results can be proven that show that there is a need for avoiding high-dimensional spaces, where possible.

Further, typical visualization interfaces cannot handle more than 3 dimensions. Hence, there is a need for consistently representing items immersed in a high-dimensional space in lower dimensional spaces, while preserving neighboring properties. Dimension reduction techniques come as a solution to that problem. Methods for dimensionality reduction are employed each time high-dimensional data has to be reduced from a high to a low-dimensional space. The principle of the mapping process for methods based on distance matrices is to find the configuration of points that best preserves the original inter distances.

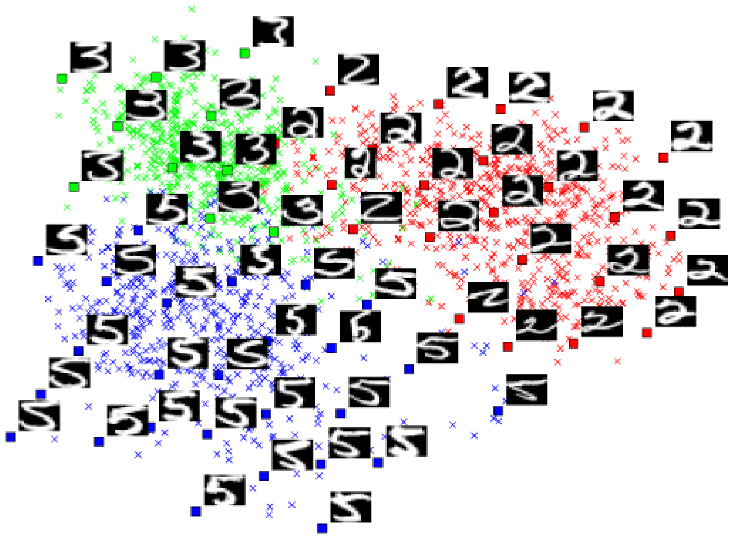


Figure 14. Dimension reduction over a database of digit images
(Illustration from http://www.merl.com/projects/dimred)

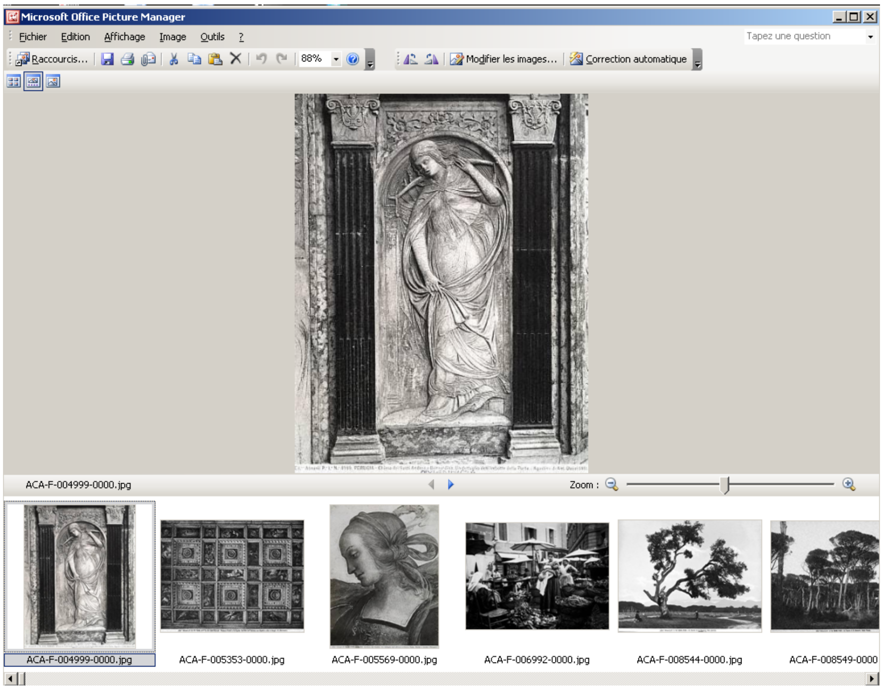
A number of methods exist. We do not detail the list and principles here but refer the reader to (Szekely, 2007), (Borg, 2005) and (Carreira-Perpiñan, 1997) for thorough reviews on the topic.

4. Multimedia Collection Browsers

4.1 Extra image browsers

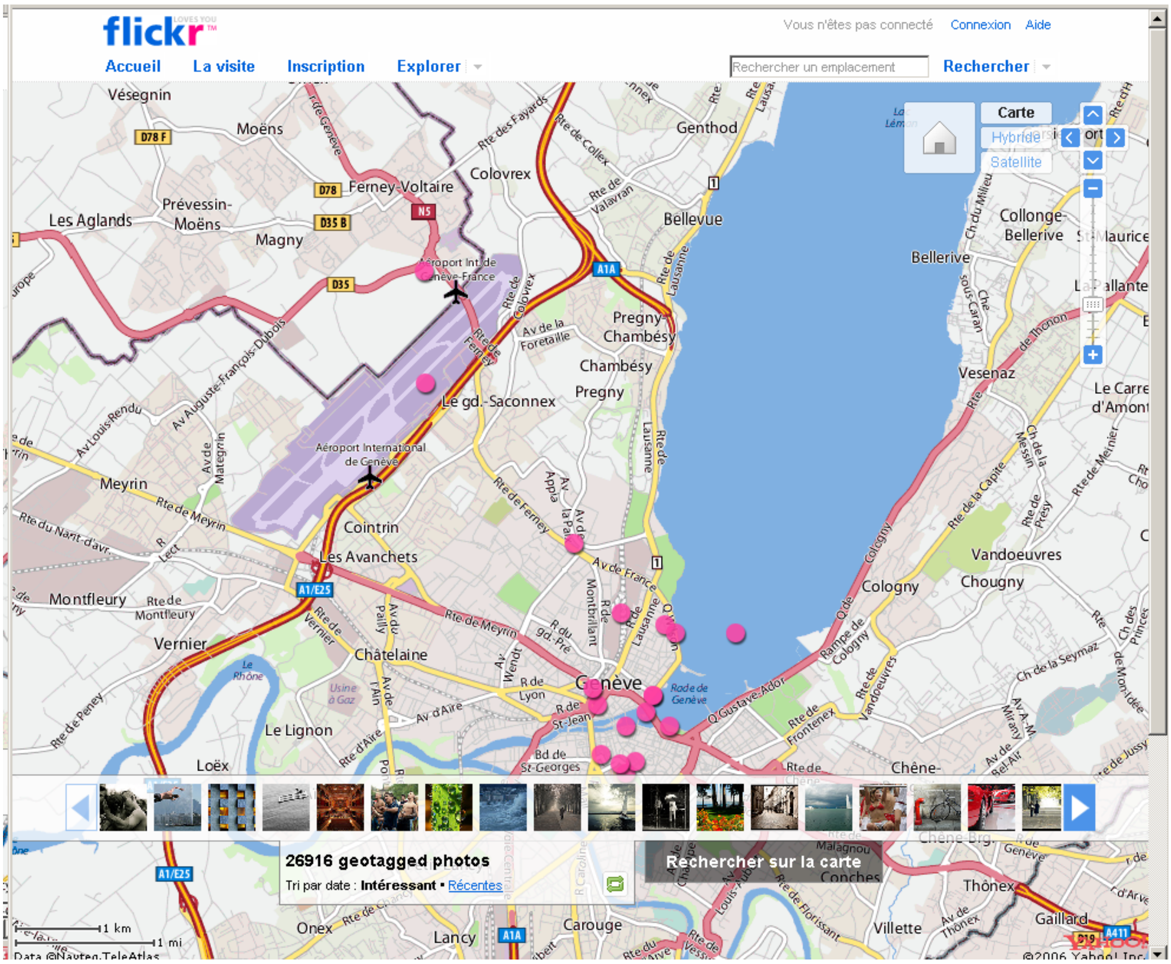
In the above pages, we have reviewed a number of strategies for image collection browsing. We list here other known browsers:

- **Microsoft**’s picture manager (filmstrip mode) is the simplest representation that can be created. It exploits a linear organization of the data. In the context of its usage, linearization is made on simple metadata, which lends itself to the ordering (eg temporal or alphabetical order)



– Google’s **Picasa** (timeline mode) also exploits the linear timeline to arrange a photo collection. An interesting feature is the near-1D organization whereby groups of pictures are arranged along the path (as opposed to aligning single pictures).

– **Flickr**'s geotagged image browser exploits the planar nature of geographic data to arrange pictures.



4.2 Related patents

Image browsing is of high commercial interest since it provides a added value over a collection of data. The following are some US patents related to image browsing, listed by US patent number.

6233367 Multi-linearization data structure for image browsing

6636847 Exhaustive search system and method using space-filling curves

6907141 Image data sorting device and image data sorting method

7003518 Multimedia searching method using histogram

7016553 Linearized data structure ordering images based on their attributes

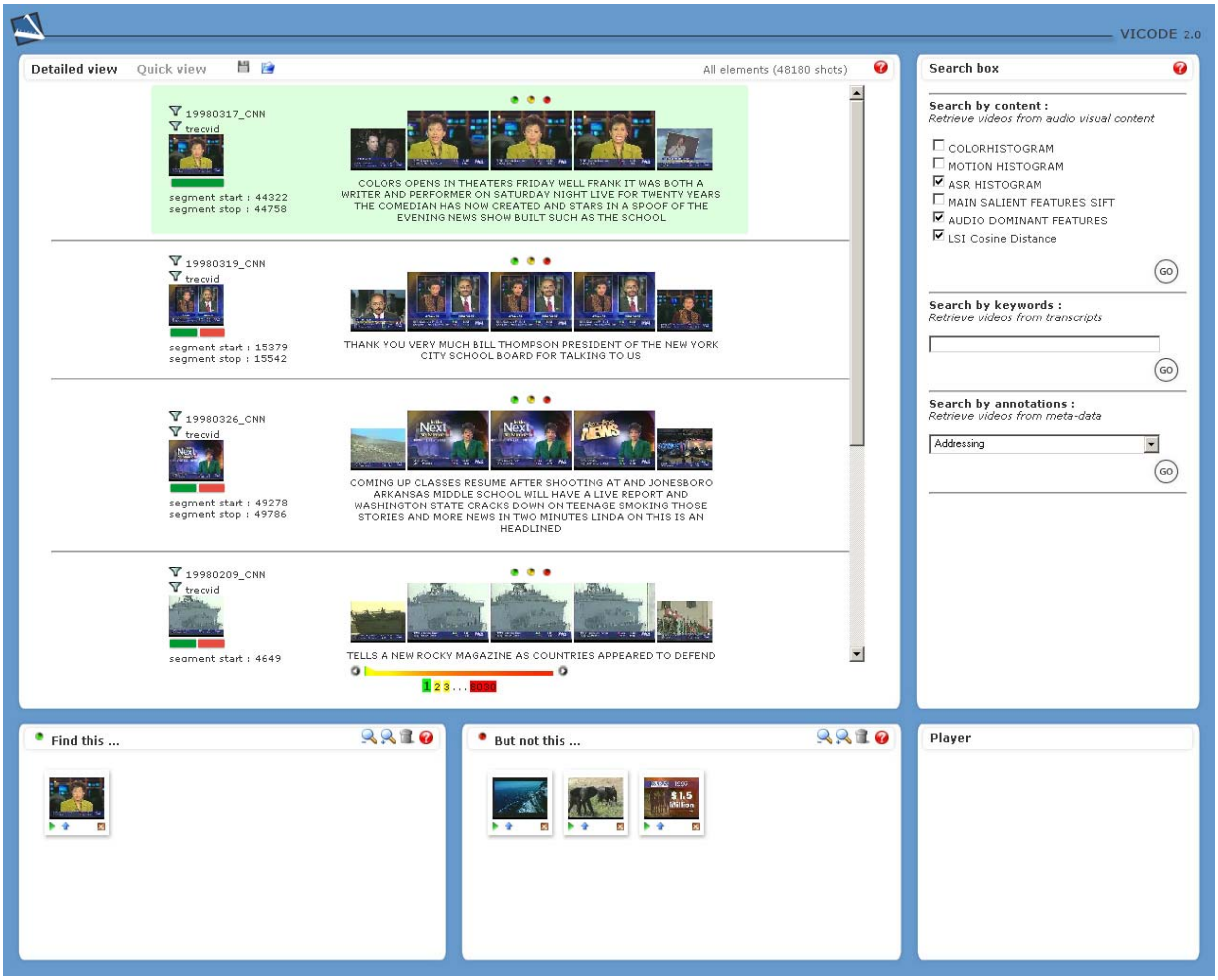
7131059 Scalably presenting a collection of media objects

7149755 Presenting a collection of media objects

4.3 Other media

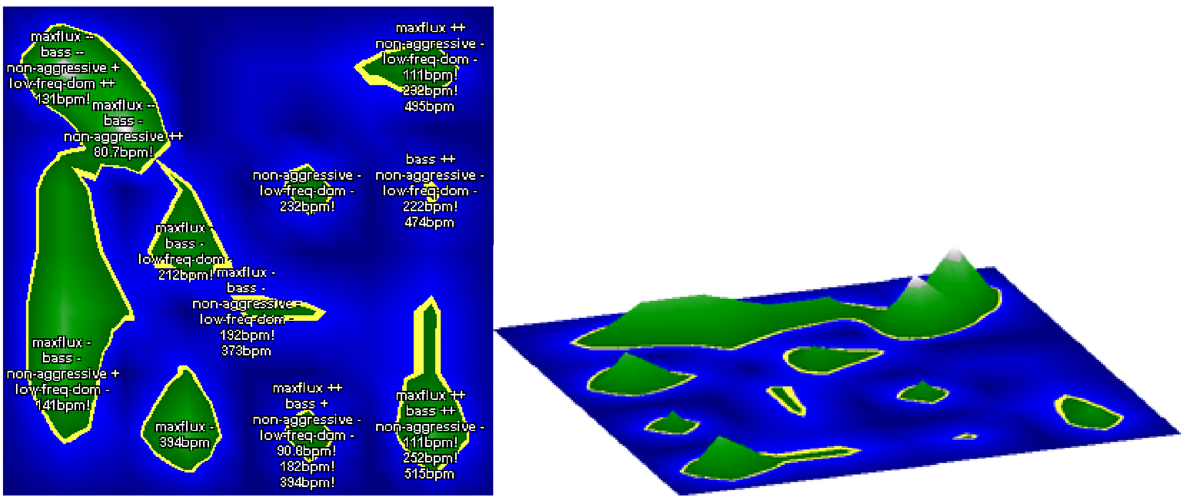
Browsing may clearly apply to media other than images. This section complements examples already provided in detail in MultiMatch deliverable 1.1 (section 7). Hence, while not detailing underlying strategies we give examples and pointers to multimedia browsers that we think provide interesting browsing functionalities.

ViCoDE (Video Collection Description and Exploration – (Bruno, 2008)) is a video search engine interface implementing the QBE paradigm and allowing some exploration functionalities.

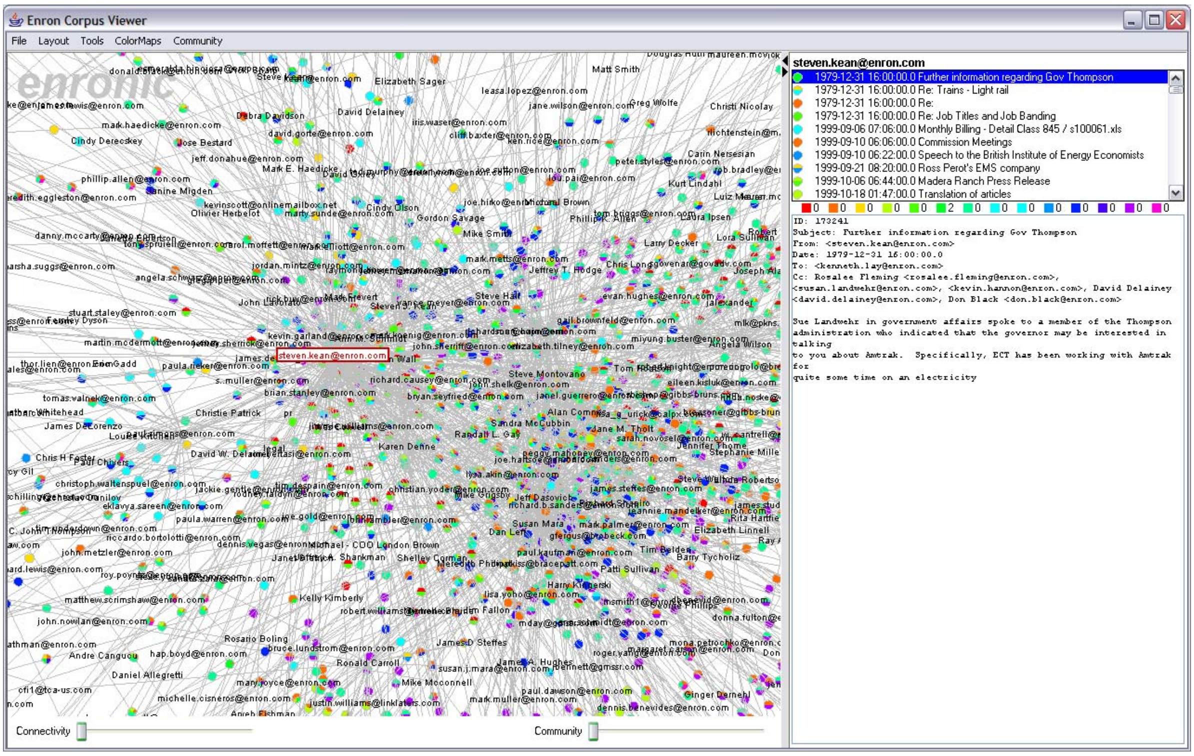


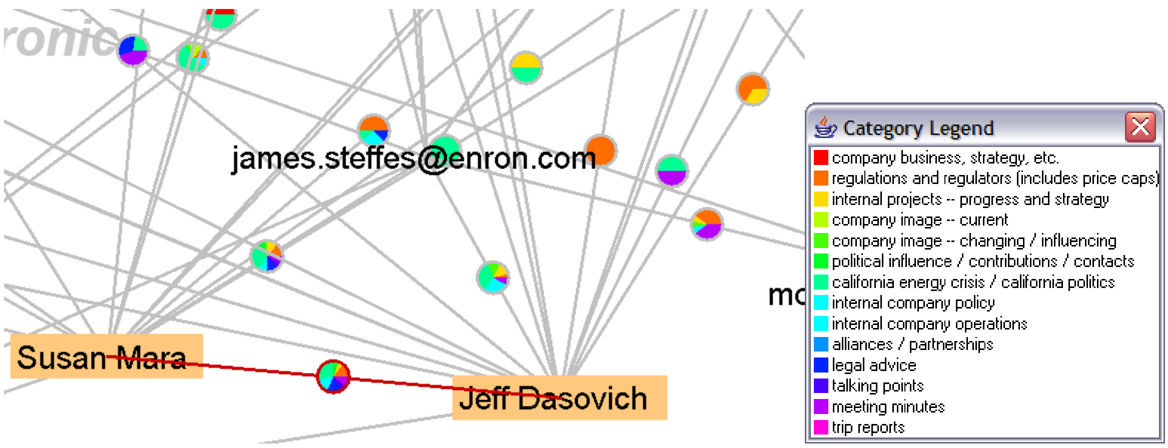
The **MediaMill** browsers (Worring, 2007) allow time and similarity based video exploration. They have been tailored to the TRECVid challenge (interactive task) and thus are relevant for news content exploration.

Islands of music (Pampalk, 2003) use Self Organising maps to arrange music pieces into a planar landscape, then used for browsing.

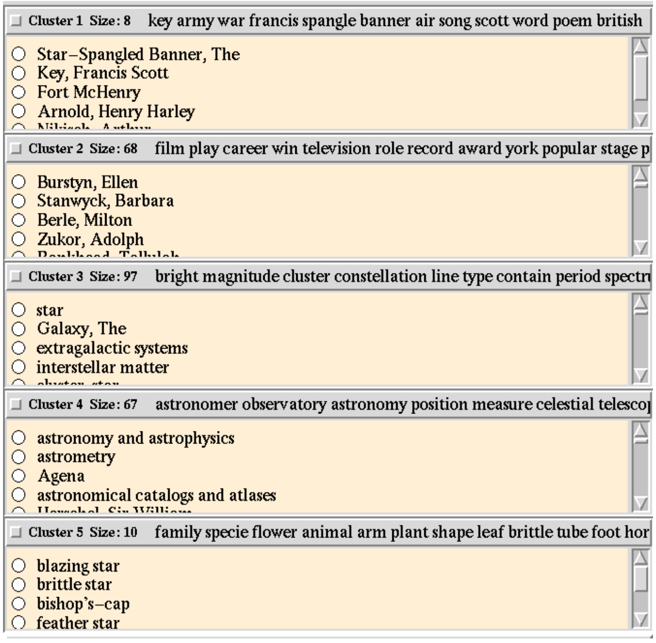


Enronic (Heer, 2004) Email collection browsing, investigation. As emails represent communications between humans, this work is related to the domain of Social Network Analysis.





Scatter/Gather (Hearst, 1996) is an early work on clustering retrieval results for their exploration by categories.



5. Evaluation

(Chen, 2002) analyzes how browsing and the more general fact of providing an efficient interface to an information system is often listed as one of "Top-ten" problems in several fields (eg, Information Retrieval (Croft, 1995), visualization and virtual reality). A new top-ten list of problems in the domain is created including benchmarking and evaluation.

Firstly, the majority of browsing tools proposed in the literature organize their content using low-level features such as color or texture. (Rorissa, 2007) demonstrates via several user studies that this is relevant and that features may indeed be used as a basis for visualization and hence browsing.

There are numerous efforts to benchmark information retrieval as a problem with a well-posed formulation. When including the quality of the interface or performance of the interaction with the information system, things are however less clear. The fact of embarking human factors in the context make the formulation less definite and prevents the automation of the performance measures (see eg (Ivory, 2001)).

Several attempts to propose evaluation protocols and frameworks have nevertheless taken place ((Black, 2002), (Rodden, 1999), (Urban 2006b)). Some particular aspects such as zooming (Combs, 1999) and presentation ((Kustanowitz, 2005), (Rodden, 2001)) have been the focus of attention for some works.

While systematic retrieval performance evaluation is possible using ground truth and measures such as Precision and Recall, having reliable performance evaluation of interfaces and interactive tools requires long-term efforts and heavy protocols. It is certainly an area where developments should be made to formally validate findings. It is often a strong asset of private companies which carefully invest in user-based testing in order to validate tools that are simpler but more robust than most research prototypes.

6. Concluding remarks

Image browsing comes as a complement to query-based search. This is valuable, due to the imperfect nature of content understanding and representation, due principally to the so-called semantic gap. Browsing is also interesting to resolve the problem of the user's uncertainty in formulating an information need. Opportunistic search and faceted browsing are example of principles and applications that bridge search and navigation.

The above analysis shows that, as a complement to classical retrieval systems, browsing and navigation should be differentiated. It is suggested here that *browsing* is directed towards an objective (information need) and thus indirectly relates to searching and acts at the *document scale*. As such, browsing is seen as assistance within similarity-based search systems, where the QBE paradigm is often deficient.

Browsing should be differentiated from *navigating* where the aim is the understanding of the collection content. Navigation-based systems thus use an absolute (global) modeling of the collection and include a global notion of similarity (ie that is driven by generic feature). This is to be opposed to browsing systems, which use a notion of similarity based on the context of the neighborhood of the sought items (ie, the interpretation of the collection is made at the light of the sought items).

Image collection browsing imposes to focus on user interaction and thus the interface design and evaluation. This refers to the work done by the Human Factors (HCI) community, which is somewhat regrettably not enough inter-weaved with the Information Retrieval and management community.

Finally, while browsing and navigation may be seen as an extension and complement to searching in image collections, it can also be applied to other media such as audio (music, eg (Pampalk, 2003)) and video (eg (Worring, 2007), (Ciocca, 2007)). These temporal media offer a temporal dimension that directly lends itself to exploration and thus makes browsing an obvious tool to use.

References and relevant literature

Aggarwal C.C. et al (2001). On the surprising behaviour of distance metrics in high dimensional space. In Proceedings of the 8th International Conference on Database Theory (ICDT'2001), 2001.

Ahlberg C. and Schneiderman B. (1994). Visual Information seeking: tight coupling of dynamic query filters with starfield displays. In CHI'94, pp 313-317, ACM Press, 1994.

Bederson D. (2001). PhotoMesa: A zoomable image browser using quantum treemaps and bubblemaps. In ACM Symposium on User Interface Software and Technology, CHI Letters, volume 3, pp 71-80. 2001.

Black JA et al. (2002) Amethod for evaluating the performance of content-based image retrieval systems based on subjectively determined similarity between images. In: Proc. Int. Conf. on Image and Video Retrieval, LNCS 2383, pp 356–366.

Borg I and P.J.F. Groenen. (2005). Modern Multidimensional Scaling: Theory and Application. Springer, 3rd Edition, 2005.

Börner K. (2000) Extracting and visualizing semantic structures in retrieval results for browsing. ACM Digital Libraries, San Antonio, Texas, June 2-7, pp. 234-235

Bruno E et al.(2008). Design of multimodal dissimilarity spaces for retrieval of multimedia documents. *To appear in IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2008.

Carey M. et al. (2003) Info navigator: A visualization tool for document searching and browsing. In Proceedings of International Conference on Distributed Multimedia Systems, September 2003.

Carpineto C. and Romano G (1996). A lattice conceptual clustering system and its application to browsing retrieval. Machine Learning, 24:95-122, 1996.

Carreira-Perpiñan M. (1997). A review of Dimension Reduction Techniques. Tech Report number CS-96-09. University of Sheffield, 1997.

Chang M.et al. (2004). Collection understanding. 4th ACM/IEEE-CS Joint Conference on Digital Libraries JCDL '04. ACM Press, New York, NY, 334-342, 2004

Chen C. et al. (2000). Content-based image visualization. In IV'00: Proceedings of the International Conference on Information Visualization. Pp 13-18. IEEE Computer Society.

Chen C. and Börner K. (2002). Top ten problems in Visual Interfaces of Digital Libraries. In Visual Interfaces to Digital Libraries, K Börner & C. Chen Eds, Springer Verlag, LNCS 2539, 2002.

Chen J.Y. et al. (2000). Hierarchical browsing and search of large image databases. IEEE Transactions on Image Processing. 9(3):442-455, 2000.

Cinque L. et al. (1998). A multidimensional image browser. Journal of Visual Language and Computing. 9(1):103-117, 1998.

Ciocca G. and Schettini R. (2007). Hierarchical Browsing of Video Key Frames. ECIR 2007: 691-694.

Combs T.T.A. and Bederson B.B. (1999). Does Zooming Improve Image Browsing? Proceedings of DigitalLibrary (DL 99) New York: ACM, 130-137.

Cox I.J. (2000). PicHunter: Theory, implementation, and psychophysical experiments. IEEE Transactions on Image Processing, 9(1):20-37, 2000.

Craver S. et al. (1999). Multi-linearisation data structure for image browsing. In SPIE Conference on Storage and Retrieval for Image and Video Databases VII, 155-166, January 1999. [US Patent 6233367]

Croft, W.B. (1995), What do people want from information retrieval? D-Lib Magazine, November 1995.

Demartines P. and Herault J. (1997). Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. IEEE Transaction on Neural Networks, 8(1):148-154, 1997.

Duda R. et al (2000). Pattern Classification. Wiley-Interscience; Second Edition.

Hearst M and Pedersen J (1996), Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results, *Proceedings of the 19th Annual International ACM/SIGIR Conference*, Zurich, August 1996

Hearst M. (2006), Design Recommendations for Hierarchical Faceted Search Interfaces. ACM SIGIR Workshop on Faceted Search, August, 2006.

Heer J (2004). Exploring Enron: Visualizing ANLP Results. In *Applied Natural Language Processing. InfoSys 290-2. University of California, Berkeley.*

Heesch D. and Ruger S. (2004). Three interfaces for content-based access to image collections. In Proceedings of the International Conference on Image and Video Retrieval (CIVR 04). LNCS volume 3115. pp 491-499. Springer Verlag, 2004.

Hinton G. and Roweis S. (2002). Stochastic Neighbor Embedding. Neural Information Processing Systems. 15:8333-840, 2002.

Ivory Y. and Hearst M. (2001). The state of the art in automating usability evaluation of user interfaces. ACM Computing Surveys, 33(4):470-516, 2001.

Janecek P. and Pu P (2003). Searching with semantics: An interactive visualization technique for exploring an annotated image collection. In Workshop on Human Computer Interfaces for Semantic Web and Web Applications (HCI-SWWA '03), November 2003.

Janecek P and Pu P. (2004). Opportunistic Search with Semantic Fisheye Views EPFL Technical Report: IC/2004/42.

Kang H. and Shneiderman B (2000). Visualization methods for personal photo collections: Browsing and searching in the PhotoFinder. *IEEE International Conference on Multimedia and Expo (III) 2000*, pages 1539–1542, 2000.

Keim D.A (2002). Information Visualization and visual data mining. IEEE Transactions on Visualisation and Computer Graphics, 7(1):100-107, 2002.

Kiranyaz M and Gabbouj M (2008). Hierarchical Cellular Tree: An Efficient Indexing Scheme for Content-based Retrieval on Multimedia Databases, IEEE Transactions on Multimedia (in Print)

Kustanowitz J. and Shneiderman B. (2005). Meaningful presentations of photo libraries: rationale and applications of bi-level radial quantum layouts. *ACM/IEEE Joint Conference on Digital Libraries*, pages 188–196, 2005

Kosinov S. and Marchand-Maillet S (2003). Overview of approaches to semantic augmentation of multimedia databases for e_cient access and content retrieval. In 1st International Workshop on Adaptive Multimedia Retrieval (AMR2003), 2003.

Laaksonen J. et al (2000). PicSOM- Content-based image retrieval with self-organising maps. Pattern recognition Letters, 21(13-14):1199-1207, 2000.

Leeuw W. and Liere R. (2003). Visualization of multidimensional data using structure preserving projection methods. In Data Visualization: The State of the Art. FH Post and GM Nielson and GP Bonneau eds, pp 213-224. Kluwer 2003.

Le Saux B. and Boujemaa N. (2002) Unsupervised Robust Clustering for Image Database Categorization, IEEE-IAPR International Conference on Pattern Recognition (ICPR'2002), Quebec, Canada, August 2002.

Li J.X. (2004). Visualization of high-dimensional data with relational perspective map. Information Visualization, 3:49-59, 2004.

Marchand-Maillet S. and Bruno E (2005). Collection Guiding: A new framework for handling large multimedia collections. Seventh International Workshop on Audio-Visual Content and Information Visualization in Digital Libraries (AVIVDiLib'05), Cortona, Italy.

Moghaddam B et al. (2004). Visualization and user-modeling for browsing personal photo libraries. International Journal of Computer Vision, 56(1-2):109-130, 2004.

Müller W. et al (1999). Hunting moving targets: an extension to Bayesian methods in multimedia databases, In *Multimedia Storage and Archiving Systems IV (VV02)*, Vol. 3846 of SPIE Proceedings, Boston, Massachusetts, USA, 20-22 September 1999. (SPIE Symposium on Voice, Video and Data Communications)

Nakazto M. and Huang T.S. (2001). 3D MARS: Immersive virtual reality for content-based image retrieval. In IEEE ICME'01. IEEE Computer Society, 2001.

Nakazato, M. et al. (2003) ImageGrouper: A group-oriented user interface for content-based image retrieval and digital image arrangement. Journal of Visual Languages and Computing 14, 363–386 (2003)

Nguyen GP and Worring M. (2005). Scenario optimization for interactive category search. In ACM-MIR'05. ACM Press, 2005.

Nguyen GP and Worring M. (2006). Interactive access to large image collections using similarity-based visualization. Journal of Visual Languages and Computing (to appear).

Nguyen GP. and Worring M. (2008), *Optimization of interactive visual similarity based search*, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP). Vol. 4, Issue 1, February 2008.

Pampalk. E (2003). Islands of Music - Analysis, Organization, and Visualization of Music Archives. Journal of the Austrian Society for Artificial Intelligence, Vol. 22, No. 4, pp 20-23.

PerceptivePixel, (2007). Introduction video at http://www.perceptivepixel.com

Pu P. and Janecek P. (2003). Visual interfaces for opportunistic information seeking. In C. Stephanidis and J. Jacko, editors, 10th International Conference on Human -Computer Interaction (HCII '03), pages 1131{1135, Crete, Greece, June 2003.

Rodden K. et al. (1999). Evaluating a visualization of image similarity as a tool for image browsing. In INFOVIS'99. IEEE Computer Society, 1999.

Rodden K. et al. (2001). Does organization by similarity assist image browsing? In ACM-CHI'01. pp 190-197, ACM Press, 2001.

Rorissa A et al (2007). Exploring the relationship between feature and perceptual visual spaces. Journal of American Society for Information Science and Technology (JASIST), 58(10): 1401-1418 (2007).

Rubner Y (1999). Perceptual Metrics for Image Database Navigation. PhD thesis, Stanford University, 1999.

Rubner Y. et al. The Earth Mover's distance as a metric for image retrieval. International Journal of Computer Vision, 40(2):99-121, 2000.

Santini, S et al. (2001). Emergent semantics through interaction in image databases. IEEE Transactions on Knowledge and Data Engineering, 13(3):337-351, 2001.

Schreiber A.T. et al. (2001). Ontology-based photo annotation. IEEE Intelligent Systems, 16(3):66-74, 2001.

Székely E. et al. (2007). Clustered Multidimensional Scaling for Exploration in IR. In 1st International Conference on the Theory of Information Retrieval, Budapest, Hungary, 2007.

Smeulders AWM et al. (2000). Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12): 1349-1380, 2000.

Stan D. and Sethi I. K (2003). *e*ID: A system for exploration of image databases. Information Processing and Management Journal, 39(3):335-361, 2003.

Tenebaum J.B. et al. (2000) A global geometric framework for nonlinear dimensionality reduction. Science, 290(5500):2319-2322, 2000.

Urban, J et al. (2006a) An adaptive technique for content-based image retrieval. Multimedia Tools and Aplications, 31(1):1-28, 2006.

Urban, J. and Jose, J. M. (2006b) Evaluating a workspace's usefulness for image retrieval. *Multimedia Systems* 12(4-5):pp. 355-373.

Vendrig J. et al (2001). Filter image browsing: interactive image retrieval by using database overviews. Multimedia Tools and Applications, 15(1):83-103, 2001.

Vertan C. et al. (2002). Browsing image databases with 2D image similarity scatter plots: update of the IRIS system. International Symposium on Communications, 397-402, 2002.

Wang J.Z. et al. SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9):947-963, 2001.

Worring, M. (2007) The MediaMill semantic video search engine. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages -. Honolulu, Hawaii, USA, April 2007

Yang J. et al. (2006). Semantic Image Browser: Bridging Information Visualization with Automated Intelligent Image Analysis. IEEE Symposium On Visual Analytics Science And Technology, pp 191-198, 2006

Acknowledgments

The authors would like to thank all their colleagues in the MultiMatch project for many useful discussions and much input.