Preference Based Feedback for Collaborative Image Retrieval

Martin Halvey
Glasgow Interactive Systems Group
School of Computing Science
University of Glasgow
Glasgow, UK, G12 8QQ
+44-141-330-8430
martin.halvey@glasgow.ac.uk

ABSTRACT

There are a many of image related search tasks that can be collaborative in nature and require input from more than one person e.g. organisation of photographs or videos from multiple views, students working on a class project etc. In this work in progress paper we describe a new image search paradigm, namely preference based image search. We describe initial implementations of this approach and some evaluations of this paradigm for image search. We also describe how this approach can be adapted for collaborative image retrieval, the potential benefits of a preference based paradigm for image retrieval and some limitations. Finally we describe how future work will realise a collaborative preference based image retrieval system.

Categories and Subject Descriptors

H.5.1 Multimedia Information Systems, H.5.3 Group and Organization Interfaces

General Terms

Design, Experimentation, Human Factors

Keywords

Synchronous, collaboration, image, search, retrieval, visual.

1. INTRODUCTION

The majority of multimedia search and retrieval services e.g. YouTube, Flickr etc. rely on users searching using textual queries. However, this places a burden on the users of such systems to be able to create effective text based queries. Annotations and tags also do not present a complete solution to problems associated with retrieving images. It is usual for users to have different perceptions about the same image and as such will annotate those images differently. This can result in synonyms, polysemy and homonymy, which make it difficult for other users to retrieve the same media objects [1]. Indeed how retrievable a piece of media is can depend on the number of annotations that piece of media has [2]. To overcome the problems of a lack of text, content-based image retrieval (CBIR) systems [3-4] can be used. CBIR systems index image data using the visual features of the images e.g. colour, shape etc. CBIR systems can be difficult to use, however, as a query must be expressed visually. For example, sketch interfaces [5] force the user to draw the target images required, which can be dependent on the users artistic talents and thus difficult for some users. Alternatively the user could possess an exemplar image which they submit as a query. In an effort to overcome some of the shortcomings highlighted above we have developed an image retrieval system that can use a number of different preference feedback mechanisms for preference based image retrieval, these feedback mechanisms are inspired by conversational recommendation systems. Conversational recommender systems attempt to engage the user in an extended, interactive recommendation dialog during which the system attempts to elicit additional query information in order to refine recommendations. In our system users are presented with a small number of image search results and simply have to express whether any of these results match their information need or not. We have examined the suitability of a number of feedback approaches based on conversational recommendation approaches for semi automatic and interactive image retrieval. In particular for interactive retrieval we have compared our preference based approaches with text based search (where we consider text to be an upper bound).

One of the main benefits of these preference based approaches, is that they are easily adaptable to allow more than one person to give feedback. As such in this paper we also discuss the application of preference based image retrieval for collaborative image retrieval. In particular we highlight the potential benefits and some limitations of using preference based feedback for collaborative image retrieval. We also discuss our future work and how we will examine the application of a preference based feedback paradigm for collaborative image retrieval. The remainder of this paper is outlined as follows: in the following section we outline related work on collaborative multimedia retrieval and preference based recommendation. Next we discuss the application of preference based recommendation for image retrieval. Following this we discuss the potential application of preference based retrieval for collaborative image retrieval, including benefits, limitations and ongoing work. Finally we present some conclusions.

CIR'11, October 28, 2011, Glasgow, Scotland, UK.

2. RELATED WORK

2.1 Collaborative Multimedia Retrieval

In the area of interactive video retrieval, a great deal of research has taken place on the subject of co-located synchronous search. Smeaton et al. [6-7] have developed the Fischlar DiamondTouch system. This system makes use of a large table top touch screen display surface, allowing users to interact with each other and with search results simultaneously. Smeaton et al. [6] explore a number of user interface design elements targeted specifically at increasing awareness between pairs of users. Awareness is the ability of users to understand and interpret the activity of others engaged in a cooperative effort without causing interruptions through explicit communication, such as asking questions [8]. The interface elements were evaluated in a user evaluation that compared two user interfaces: the first interface was designed to maximise search efficiency at the expense of awareness, the second interface was designed to maximise awareness of the other user's actions at the expense of efficiency. The results of this evaluation show that the interface designed for awareness outperforms the interface designed for efficiency in relation to a number of measures. In a related approach, Adcock et al. [9] have also designed a system for collocated, synchronous collaboration between searchers for video search. However, when using their system, users adopt a pre-defined role for the collaboration. This approach was compared with an individual user searching. The results of a user evaluation illustrate that in some situations, collaborative search can outperform individual search performance. Villa et al. [8, 10] have developed an interface that allow users to see each other's search results and queries while performing the same video search task. Using this interface they investigated the role of awareness and its effect on search behaviour in collaborative multimedia retrieval. In particular they look at varying search behaviour in multiple awareness scenarios. The results suggest that balanced awareness scenarios provide the best retrieval results in comparison with scenarios where unbalanced awareness exists between users. Halvey et al. present the ViGOR system [11], a video retrieval system that allows users to group videos in order to facilitate video retrieval tasks. ViGOR also provides mechanisms for groups of users to collaborate both asynchronously and remotely on video search tasks.

2.2 Preference Based Recommendation

Conversational recommender systems attempt to engage the user in an extended, interactive dialog during which the recommender system endeavours to extract additional query information in order to refine recommendations. There are two forms of conversational recommender systems; navigation *by asking* or navigation *by proposing* approaches. Our work focuses on navigation by proposing. The main feature of navigation by proposing is that the user is presented with one of more recommendation alternatives during each recommendation cycle. There are 3 important types of feedback that can be used for navigation by proposing, two of which we use in this work, preference based feedback and critique based feedback.

2.2.1 Preference Based Feedback

Perhaps the simplest form of feedback is *preference based feedback;* during interaction with a preference based feedback system the user simply expresses a preference for one alternative over the others. It is extremely suitable for domains where users have very little domain knowledge, but where they can easily express a preference. In this way this approach to conversational navigation is well suited for image or video retrieval. Unfortunately, while this approach carries a small feedback cost for the user, it is can be limited in its ability to direct the interaction process, i.e. it will not always be clear why a user has selected one object over another, both objects may have many features in common and many features that differentiate the objects. In an attempt to address this issue the *comparison-based recommendation* work of McGinty and Smyth [12] propose a number of query revision strategies that are designed to revise the current query/recommendation as a result of preference-based feedback. The most clear-cut of these strategies (*more like this*) simply adopts the preferred case as a new query and proceeds to retrieve the *k* most similar cases to it for the next interaction cycle. However, this approach has drawbacks and indeed might not be very efficient as it does not really attempt to elicit the user's true preferences at a feature level. An alternative approach (*partial more like this*) transfers features from the chosen case, if these features are absent from all of the rejected cases, thus allowing the system to focus on those aspects of the preferred case that are unique in the current cycle. A third strategy attempts to weight features in the updated query according to how confident the recommender system can be that these features are responsible for the user's preference (*weighted more like this*). A final strategy was to allow users to give return a case as negative feedback to retrieve cases less like the returned case (*less like this*).

2.2.2 Critique Based Feedback

As with the preference based approach the critique based feedback mechanism is one that requires limited domain knowledge on the part of the user. Critiquing-based recommenders allow users to provide feedback in the form of a directional feature constraint. The FindMe systems [13-14] were amongst the first recommender systems to use critiquing as an approach for a conversational recommender system. The Entree recommender [13] suggests restaurants in Chicago and each recommendation allows the user to select from seven different critiques. When a user selects a critique such as *cheaper*, Entree eliminates cases (restaurants) that do not satisfy the critique from consideration in the next cycle, and selects that case which is most similar to the current recommendation from those remaining; thus each critique acts as a filter over the cases. The FindMe systems evaluated this form of conversational recommendation and feedback in a number of contexts including movie, car, and accommodation recommendation [14]. In the Car Navigator recommender [14] system, individual critiques were also designed to cover multiple features, so that, for instance, a user might request a *economic* car than the current recommendation, simultaneously constraining features such as *engine size* and *emissions*. These *compound critiques* obviously allow the recommender to take larger steps through the information-space, eliminating many more cases than would be possible with a single-feature, *unit critique*, in a single recommendation cycle. More recently the work of McCarthy et al. [15-16] has investigated the possibility of automatically generating dynamic compound critiques based on the remaining cases and the user's progress so far. This *dynamic critiquing* approach uses data mining techniques to identify groups of unit critiques that reflect common difference patterns between the remaining cases.

3. PREFERENCE BASED IMAGE SEARCH

3.1 Application for Image Retrieval

In this section we describe how various preference based methods can be adapted to use low level features e.g. colour histogram etc. for image retrieval.

3.1.1 Compound Critiquing

The main problem with applying Compound Critiquing to image retrieval was finding a convenient and efficient way of creating unit critiques that would start off the algorithm. For images it is much harder to define features of low complexity which could still hold some meaningful information. Instead, it is normal for some kind of multidimensional vectors to be used to describe visual information e.g. colour histograms, edge histograms etc.

In order to overcome this problem the following solution was adopted, which can be summarised in the following steps:

1. Separately for each feature, distance to the current prototype is computed and a sorted list of images is created using these results.
2. For each feature the corresponding list is split into two sub lists, where the point of divide is determined by multiplying the amplitude, i.e. the biggest distance minus the smallest distance, by a fraction which is called the *split factor* for the purposes of this work.
3. The sub lists for each feature become two unit critiques. One sub list can be described as more similar to the prototype or as simply being smaller than the split point. The other sub list is considered as less similar or being bigger than the split point.

The second part of the Compound Critiquing algorithm involves creating association rules from the unit critiques which are later turned into compound critiques. In general, association rules are used to mine `interesting combinations of features' or popular patterns and they focus on examining the frequency of similar values for some features appearing together.

The algorithm implemented in the project uses the Apriori algorithm [17] to generate association rules. It is one the most popular and relatively simple techniques which is based on calculating the support (frequency) for each rule and creating complex rules from lower-order ones. Eventually, the algorithm finishes returning a list of critiques which all have support levels exceeding the threshold. The Apriori algorithm in the worst case can have an exponential complexity.

3.1.2 xMLT Approaches

Three of the approaches presented by McGinty and Smyth [12], namely More-Like-This (MLT), partial More-Like-This (pMLT) and weighted More-Like-This (wMLT), were implemented as proposed query processing methods. These approaches were chosen as they were described as "promising" variants in the original paper. The basic operation of the algorithms involves the following of steps with some variations as outlined by McGinty and Smyth [12]:

1. For each image in the available set, normalised distances to the given prototype are computed with respect to each feature. These are then summed to give the overall distance across all features.
2. A list is created containing all images sorted by their distances to the prototype.
3. A *reduction factor* can optionally be applied to the list of images. This *reduction factor* removes a certain percentage of most dissimilar images, thus preserving some information from the current feedback in future cycles.

3.2 Analysis for Image Retrieval

A simulated study was conducted using the ImageCLEF 2007 collection and tasks [18]. The results of the simulated study had important consequences for the development of an interactive system. The most important design decision was eliminating the pMLT and wMLT approaches from further consideration, for now, as they essentially worked just like the basic MLT version but was slightly less effective. The reduction factor in MLT was also set at 30% on the basis that reducing the set of results by too much proved to negatively affect the overall performance. Compound Critiquing approach will be assigned specific values for the support threshold and split factor.

Following the simulated study, a user evaluation was conducted. 12 users carried out 4 image search tasks from the ImageCLEF collection using 4 different search methods. The search tasks varied on difficulty and whether they were visual or semantic search tasks. The four search systems were (1) a text based search system, (2) a system using MLT feedback, (3) a system using compound critiquing and (4) a system using a combination of MLT and compound critiquing. In terms of task performance we measured precision and recall. While the text based system was considered to be an upper bound, it was found that for some task that the preference based image retrieval methods outlined outperformed the text based system. The participants indicated in their post task questionnaires that the MLT approach had the best interface and the best search process. They also felt that the text based interface was best in terms of results; however, this may be that they found the results and interaction more familiar. In the exit questionnaire it is also shown that the users have a definite preference for the preference based image retrieval systems, in particular the MLT system. In conclusion, the results of the evaluation highlighted the promise of this approach to alleviate the major problems that users have while searching for multimedia, thus presenting a potential work around to the semantic gap and some other problems associated with image search.

4. APPLICATION FOR COLLABORATIVE IMAGE RETRIEVAL

In the previous section the application of preference based feedback to image retrieval has been discussed. In this section we discuss the possible application of the search paradigm to collaborative image retrieval, including potential benefits and limitations.

4.1 Benefits

One of the main reasons for selecting preference feedback is that research has already been conducted into using preference based feedback and in particular compound critiquing for collaborative tasks [19]. As has been demonstrated as these approached can be applied to image retrieval, then these collaborative methods [19] could also be applied to image retrieval.

Another benefit to using preference based feedback is that it is relatively lightweight and easy way to provide feedback. This has some good and bad points, it means that a sharing of expertise is not necessary for collaboration, but it does mean that it may be difficult for an expert to demonstrate or take advantage of their expertise. However, it may be possible to account for this when combining feedback from multiple users; this is a potential area for future work. In addition this simple feedback mechanisms means that this search paradigm can easily be used on different devices i.e. feedback from a mobile phone, a desktop computer or a tablet PC would be the same, meaning that users are not limited by the hardware available to them. This also means that the approach can easily be adapted for either remote or co-located search; however this search would always probably be synchronous. Finally it would be easy to gamify this approach. It could easily be imagined that image collections could be classified visually using some sort of crowd sourcing approach.

4.2 Limitations

While there are many benefits to this approach there are obvious limitations. It is only suitable for certain types of search in particular search where users can easily give quick feedback and do not require expertise. This search paradigm is more suited to synchronous search and not really suitable for asynchronous search, although in theory it could be applied for asynchronous search. Also this preference based approach does not support division of labour, or it does not support discussion or in depth feedback.

4.3 Ongoing Research and Future Work

The next main focus is to integrate some of the collaborative methods described by McCarthy et al. [19]. In addition thus far we have created a system for preference based image search that works in a desktop environment. Our next stage of research is to develop and Android application that allows the same type of interaction as the desktop application. This will allow us to investigate the application of preference based feedback for collaborative image search. In addition by using different platforms e.g. desktop, tablet, phone, we will be able to examine if these approach will truly be platform independent for collaborative image retrieval. Finally as part of the move to mobile Android devices we plan to present the interaction as a game to see if we can create some sort of visual classification of an image collection.

5. CONCLUSIONS

In this work in progress paper we have presented a new image search paradigm, namely preference based image search. We have described and initial implementations of this approach and some evaluations. We have also outlined how this approach could be beneficial for collaborative information retrieval and the steps that we plan on taking to realise this. It is hoped that this simple interaction paradigm could be adapted to provide effective, lightweight and easy collaborative image search.

6. REFERENCES

[1] Guy, M. and Tonkin, E. Folksonomies: Tidying Up Tags? D-Lib Magazine, 12, 1 2006).

[2] Halvey, M. and Keane, M. T. Analysis of Online Video Search and Sharing. In Proceedings of the ACM Conference on Hypertext and Hypermedia (Manchester, UK, 2007)

[3] Urban, J. and Jose, J. M. EGO: A personalized multimedia management and retrieval tool. International Journal of Intelligent Systems, 21, 7 2006), 725-745.

[4] Nakazato, M., Manola, L. and Huang, T.-S. ImageGrouper: A group-oriented user interface for content-based image retrieval and digital image arrangement. Journal of Visual Language Computing, 142003), 363-386.

[5] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P. Query by image and video content: the QBIC System. City, 1997.

[6] Smeaton, A. F., Foley, C., Gurrin, C., Lee, H. and McGivney, S. Collaborative Searching for Video Using the Fischlar System and a Diamond Touch Table. In Proceedings of the ACm Tabletop (2006)

[7] Smeaton, A. F., Lee, H., Foley, C. and McGivney, S. Collaborative video searching on a tabletop. Multimedia Systems, 12, 4 2007), 375-391.

[8] Villa, R., Gildea, N. and Jose, J. M. A study of awareness in multimedia search. In Proceedings of the ACM/IEEE Joint Conference on DIgital Libraries (2008)

[9] Adcock, J. and Pickens, J. FXPAL collaborative exploratory video search system. In Proceedings of the ACM Conference on Image and Video Retrieval (2008)

[10] Villa, R., Gildea, N. and Jose, J. M. Collaborative awareness in multimedia search. In Proceedings of the ACM Multimedia (2008)

[11] Halvey, M., Vallet, D., Hannah, D., Feng, Y. and Jose, J. M. An asynchronous collaborative search system for online video search. Information Processing and Management, 46, 6 2010), 733-748.

[12] McGinty, L. and Smyth, B. Comparison Based Recommendation. In Proceedings of the European Conference on Case-Based Reasoning (2002)

[13] Burke, R., Hammond, K. and Young, B. Knowledge based navigation of complex information spaces. In Proceedings of the AAAI National Conference on Artifical Intelligence (Portland, OR, 1996)

[14] Shimazu, H. ExpertClerk: Navigating Shoppers' Buying Process with the Combination of Asking and Proposing. In Proceedings of the International Joint Conference on Artifical Intelligence (2001)

[15] McCarthy, K., Reilly, J., McGinty, L. and Smyth, B. On the Dynamic Generation of Compound Critiques in Conversational Recommender Systems. In Proceedings of (2004)

[16] McCarthy, K., Reilly, J., McGinty, L. and Smyth, B. Experiments in Dynamic Critiquing. In Proceedings of the International Conference on Intelligent User Interfaces (2005)

[17] Agrawal, R. and Srikant, R. Fast algorithms for mining association rules in large databases. In Proceedings of the 20th International Conference on VLDB (Santiago, Chile, 1994)

[18] Grubinger, M. and Clough, P. On the creation of query topics for ImageCLEFPhoto. In Proceedings of the Third Workshop on Image and Video Retrieval Evaluation (Budapest, Hungary, 2007)

[19] McCarthy, K., Salamo, M., Coyle, L., McGinty, L., Smyth, B. and Nixon, P. Group recommender systems: A critiquing based approach. In Proceedings of the IUI (2006)