Using Color and Texture Indexing to improve Collaborative Filtering of Art Paintings

Arnd Kohrs and Bernard Merialdo

Institut EURECOM – Department of Multimedia Communications
BP 193 – 06904 Sophia-Antipolis – France
{kohrs,merialdo}@eurecom.fr

Abstract. Information filtering is a key technology for the creation of Web sites, which are adapted to the user's needs. In this paper we identify collaborative filtering and content-based filtering as independent technologies for information filtering. We apply both technologies in our prototype user-adapting Web site, the Active WebMuseum, a recommender system for art paintings. Our new approach extends existing user profiles with content-based information gained through automatic image indexing. These extensions lead to a better performing collaborative filtering system. We validate our approach in off-line experiments.

Keywords: Collaborative filtering, content-based filtering, image indexing, Web museum

1 Introduction

More and more services are available on the Internet through Web sites. In general, these Web sites are focused on providing information or to sell products. Often both is done at the same time, e.g., E-commerce sites which provide detailed information about their products. The user (potential customer) has been recognized as a very valuable asset to the Web sites. Therefore, the Web sites try to tie the users to their service by letting the users more efficiently (e.g., less time-consuming) access the pieces information and products, which they prefer. Often, users may define preferences through user profiles, which are then used to personalize the visits to that Web site, so that the Web site presents a customized view adapted to the users interests. This trend of user-adapting Web sites, in contrary to the static collection of hypertext documents, necessitates new technologies and tools to adapt to users. One key technology is information filtering, so that important objects can be automatically identified and presented to the user. We evaluate the use of information filtering for user-adapting Web sites in our prototype the Active WebMuseum.

Techniques, which have been proposed for information filtering fall in two independent categories: Content-based filtering and collaborative filtering. In content-based filtering the filtering is based on content analysis of the considered objects, e.g. term frequency for text documents, and its relation to the user's preferences. For content-based filtering it is therefore necessary, that the results of content analysis and user preferences can reliably and automatically be determined. While recent research shows good results for the content-based filtering of text documents, filtering of other media, as audio and video, is hard due to the limitations of content analysis technology available.

Collaborative filtering is another approach to identify objects, which are relevant to a user. In collaborative filtering objects are selected for a particular user, which are relevant to similar users. Generally, in collaborative filtering the content of the objects is ignored and only other users opinions on the considered objects are relevant. Therefore, collaborative filtering is especially interesting for objects, for which content analysis is weak or impossible. However, the performance of collaborative filtering relies on the amount of available opinions on the considered objects and it therefore fails when few or no opinions are known.

In order to build better performing filtering systems both techniques can be combined. In recent research several approaches of the combining both techniques have been studied [2, 8]. However, these approaches are limited to text documents. For objects, such as images, these approaches are not appropriate. In order to explore the combination of content-based and collaborative filtering we implemented the Active WebMuseum. The museum is a Web based service, that gives personalized tours through a virtual collection of art paintings. Personalization is achieved through collaborative filtering. The opinions on paintings collected through the Active WebMuseum serve now for further studies of combining content-based and collaborative filtering.

First, we describe the Active WebMuseum in more detail in section 2. Then, the content-based filtering, which is based on color and texture indexing techniques for images, is described in section 3. Collaborative filtering is explored in more detail in section 4. We then present our combination approach together with an evaluation in section 5. Finally, we conclude and describe future steps in section 6.

2 The Active WebMuseum

The Active WebMuseum [1]is a project that aims at studying how the combination of content-based and collaborative filtering techniques can be used to build user-adapting Web sites, sites where the navigation through information is dynamically adapted to each user. The chosen application is a Web site that allows users to browse through a collection of art paintings.

Numerous WebMuseum sites already exist on the Internet. They use hyper-media technologies to display pictures of objects, and hypertext navigation to simulate walking along the museum corridors. Existing sites are static, which means that the hypertext structure linking the objects has been defined once for all, and is the same for all users, in the same way that the topology of buildings does not change.

In the Active WebMuseum project, we use filtering techniques to create an user-adapting Web site, in which the hypertext structure is created for each specific user, based on predictions of what this user should prefer. The basic idea is that the user is asked to provide ratings for the paintings that he views during his visit. The system then selects paintings similar to paintings with high ratings (this is content-based filtering). Collaborative filtering is also used to make predictions based on the ratings that other users have provided during previous visits. These predictions are then used to present paintings to the user accordingly, so that more relevant paintings are seen first.

Once logged in, a user may browse the collection of art paintings in several directions: Depending on a reference painting the user may choose to see further paintings according to several criteria, for which personalized virtual corridors are created using filtering techniques which are described later in this paper (see Figure 1 and Figure 2 for examples).

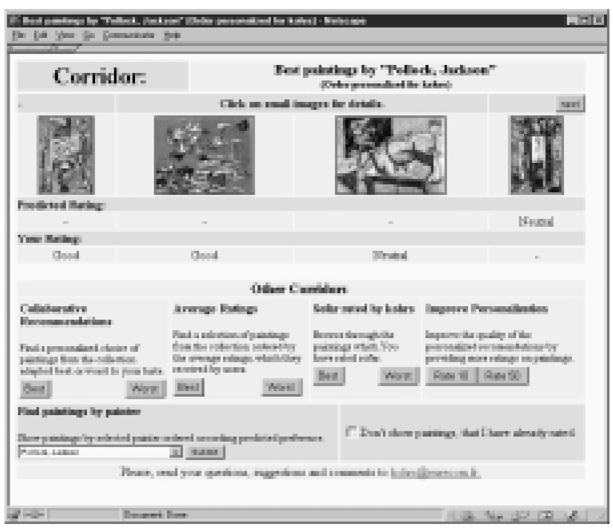


Fig. 1. Browsing dynamic corridors: If the user has chosen a dynamic corridor (in the example a corridor containing paintings by Jackson Pollock), he is presented iconized paintings ordered according to his preference. From here the visitor may choose to get closer to a painting by clicking on it.

In the Active WebMuseum, visitors can express preferences by giving symbolic ratings to paintings (*excellent*, *good*, *neutral*, *bad*, *terrible*). For historic reasons, the symbolic ratings are then mapped on numerical ratings in the interval [0..10]. For paintings which have not been rated by the visitor, the ratings are predicted using other users ratings and collaborative filtering technology.

[1] The Active WebMuseum (accessed through http://www.eurecom.fr/ kohrs/museum.html) uses the collection of paintings from the *WebMuseum, Paris*(accessed through http://metalab.unc.edu/wm/), which has been created by Nicolas Pioch and contains roughly 1200 paintings by about 170 painters.

Fig. 2. A single painting in detail close-up: When the user choses an iconized painting from a corridor it is presented in more detail (artist, title, creation date).

3 Content-Based Filtering

It is reasonable to expect that images with similar content will be almost equally interesting to users. The problem is that defining image content and image similarity is still an open problem. Ongoing research in multimedia indexing is focusing on two directions:

- either each image is described by a textual caption, and captions are compared using techniques derived from document retrieval,
- or analysis and recognition techniques are applied to the image pixels to extract automatically features which are compared using some distance measure in the feature space.

We focus on the second approach, because it can be entirely automated. In our prototype, we have currently implemented two feature extraction components, derived from the work described in [10, 11]: Color histograms and texture coefficients.

3.1 Color Histograms

The original paintings are available in RGB format, where each pixel is defined by the values (0-255) of the three components red, blue and green. We project these values in the HSV space (Hue, Saturation, Value) which models more accurately the human perception of colors. The HSV coefficients are quantized to yield 166 different colors. For each image, the histogram of these 166 colors is computed (proportion of pixels with a given quantized color).

To compare two images, we compute the L_1 distance (equation 1) between their color histograms:

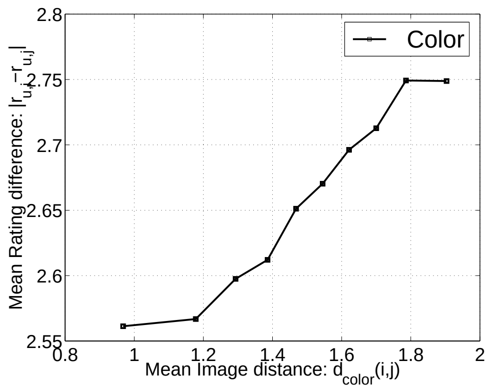
$$
\begin{aligned}
h_i(j) &: \text{percentage of number of pixels} \\
&\quad\ \text{of painting } i \text{ with the color } j. \\
L_1(h_k, h_l) &= \sum_j |h_k(j) - h_l(j)| \\
d^{color}(p, p') &= L_1(h_p, h_{p'}) \qquad (1) \\
d^{color} &\in [0..2]
\end{aligned}
$$

3.2 Texture Coefficients

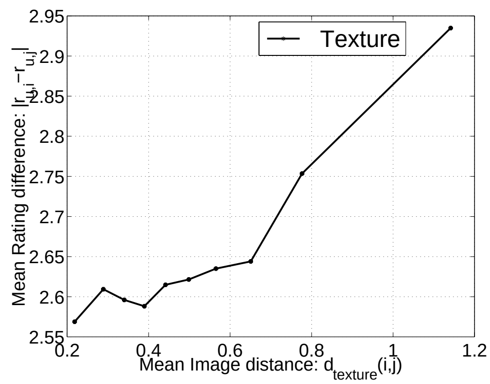
While color histograms do not take into account the arrangement of pixels, texture coefficients can be computed to characterize local properties of the image. We are using a wavelet decomposition using the Haar transform, by which a number of sub-images corresponding to a frequency decomposition are generated . These sub-images are quantized to binary values, so that each pixel of the original image is associated with a binary vector of length 9. The histogram of these vectors (it has length 512) is the feature vector associated to the texture analysis of the image. As previously for color distance, the L_1 distance (see equation 1) is used to measure the distance between images.

3.3 Correlation between Content-based Distances and Ratings assigned by Users

Using the individual ratings, that users assigned to paintings, and the previously described content-based distances between paintings, we measured a correlation between image distance and the difference of ratings, which the same user assigned to the paintings. The results for color histograms and texture coefficients are plotted in figure 3.



(a) Color-histogram distances



(b) Texture-coefficients distances

Fig. 3. Correlation of distances between color-histograms(texture-coefficients) of paintings and rating differences. For each user and for each painting, that a user rated, all occurring distances between the paintings were collected together with the according absolute difference of ratings. The distances are then sorted and grouped. The mean distance of each group determines the values for the x-axis. For each group the mean absolute rating difference determines the y-coordinate.

These measurements suggest, that paintings which are close in color or in texture receive in general similar ratings by the same users. Later, we describe how we derive artificial ratings considering this relationship, which can then be used to improve collaborative filtering.

4 Collaborative Filtering

Collaborative filtering systems select items for a user based on the opinions of other users. Generally, collaborative filtering systems do not rely on content-based information about the items, considering only human judgments on the value of items. Collaborative filtering systems consider every user as an expert for his taste, so that personalized recommendations can be provided based on the expertises of taste-related users.

Collaborative filtering has been applied to several domains of information: News articles, GroupLens [6–8]. Music, Ringo [9]. Movies, MovieCritic[2].

Most collaborative filtering systems collect the users opinions as ratings on a numerical scale, leading to a sparse matrix $rating(user, item)$ (in short $r_{u,i}$). Collaborative filtering systems then use this rating

[2] http://www.moviecritic.com

matrix in order to derive predictions. Several algorithms have been proposed on how to use the rating matrix to predict ratings [4, 9, 3]. For the Active WebMuseum we derived a collaborative filtering algorithm from a commonly used technique, also used in the GroupLens project and in Ringo, which is based on vector correlation. In the following we will describe the underlying formulas in more detail to make the general idea of automatically using other users as expert recommenders understandable.

Usually the task of a collaborative filtering system is to predict the rating of a particular user u for an item i. The system compares the user u's ratings with the ratings of all other users, who have rated the considered item i. Then a weighed average of the other users ratings is used as a prediction.

If I_u is set of items that a user u has rated then we can define the mean rating of user u as:

$$\overline{r_u} = \frac{1}{|I_u|} \sum_{i \in I_u} r_{u,i} \tag{2}$$

Collaborative filtering algorithms predict the ratings based on the ratings of similar users. When Pearson correlation is used, similarity is determined from the correlation of the rating vectors of user u and the other users u':

$$\rho(u, u') = \frac{\sum_{i \in I_u \cap I_{u'}} (r_{u,i} - \overline{r_u})(r_{u',i} - \overline{r_{u'}})}{\sqrt{\left(\sum_{i \in I_u \cap I_{u'}} (r_{u,i} - \overline{r_u})^2\right) \left(\sum_{i \in I_u \cap I_{u'}} (r_{u',i} - \overline{r'_u})^2\right)}}$$

It can be noted that$\rho \in [-1, +1]$.

The value of ρ measures the similarity between the two users' rating vectors. A high absolute value signifies high similarity and a low absolute value dissimilarity.

The general prediction formula is based on the assumption that the prediction is a weighed average of the other users ratings. The weights refer to the amount of similarity between the user u and the other users.

$$\begin{aligned} U_i &: \text{Users, who rated item } i. \\ p^{collab}(u, i) &= \overline{r_u} + k \sum_{u' \in U_i} \rho(u, u')(r_{u',i} - \overline{r_{u'}}) \\ k &= \frac{1}{\sum_{u' \in U_i} \rho(u, u')} \end{aligned}$$

The factor k normalizes the weights.

Sometimes the correlation coefficient between two users is undefined because they have not rated common objects, i.e. $I_u \cap I_{u'} = \varnothing$. In such cases the correlation coefficient is estimated by a default value ($\rho_{default} = 0.2$), which is the measured mean of typically occurring correlation coefficients.

5 Extending Collaborative Filtering with Content-based Information

In earlier work we discovered typical problematic cases for collaborative filtering systems [4], cases when not enough ratings are available, due to an insufficient amount of users or to few ratings per user. In the contrary, content-based schemes are less sensible to sparsity of ratings, since the performance for one user relies exclusively on his user-profile and not on the number of users in the system. However, comparative studies have shown, that collaborative filtering can outperform content-based filtering [1]. Collaborative filtering should therefore be favored over content-based filtering. In cases, where collaborative filtering limited by an insufficient amount of users and ratings, an integration of content-based filtering into collaborative, might lead to better filtering performance. In the following we present briefly recent research, which pursues the combination of content-based and collaborative filtering.

5.1 Existing Approaches

Fab[2] is an agent-based document filtering system. An agent society adapts through genetic algorithms and machine learning to topics and users. The agents are grouped into two different groups: The selection agents adapt to the preferences of a specific user and the collection agents adapt to topics. The collaborative aspect is achieved by the use of the same specialized collection agents for a group of related users and by forwarding highly rated documents to similar users. Similarity between users is determined from the keywords of preferred documents, by the use document retrieval techniques.

Sarwar [8] suggested filter-bots, specialized agents which detect features (spelling accuracy and message length) in news articles for the GroupLens Usenet filtering system. According to the feature detection the agents insert artificial ratings into the system.

The previously described projects combine collaborative and content-based filtering technology. However, a coherent method for combining collaborative filtering with content-based filtering has not been described yet. Further, both approaches were designed to operate on textual documents. For text documents the IR community has produced powerful analysis models, e.g. the Vector Space Model. For other media, such as images, it is not obvious how collaborative filtering can gain from existing content-indexing techniques.

In previous research we began to explore the combination of content-based filtering and collaborative filtering [5]. We used content-based criteria to design content-based predictors. These predictors were then linearly combined with the collaborative filtering predictor. In the following we present a different approach which integrates content-based filtering into a collaborative filtering system by the use of information gathered using automatic indexing of color and texture. Our approach extends the idea of Sarwar[8], in such using content-based criteria to create artificial ratings, but it differs as to we create many artificial user profiles, whose ratings are derived by content-based filtering and corresponding original users' profiles, which are similar to the selection agents of Fab[2], which specialize to the preferences of specific users.

5.2 Deriving Artificial Users from Image Metrics

Following the findings of section 3.3 we pursue to extend the database used for collaborative filtering, so that artificial ratings are inserted, which are coherent with the content-based distances. For each described distance metric of section 3 and for each real user u a corresponding artificial user u^{color} and $u^{texture}$ is derived. The artificial users are assigned the same ratings as the original user u, so that if $r_{u,i}$ is defined, then $r_{u^{color},i} = r_{u^{texture},i} = r_{u,i}$.

Additionally, artificial ratings are derived for some images, which the original user u had not rated. The artificial ratings are content-based predictions for that particular user. That means that some unrated items are assigned a predicted rating , based on similarity between the rated items and the item whose score should be predicted. In order to perform a content-based prediction, we define a restricted neighborhood for of a painting i within user-profile of a user u, which contains paintings rated by u, which distance is below a threshold T:

$$N_{u,i}^{color} = \{j \in I_u | d^{color}(i,j) \leq T^{color}\}$$

These neighboring images are then used to predict a score for the artificial ratings. The prediction formula for color is described below:

$$p^{color}(u,i) = mean_{j \in N_{u,i}^{color}}(r_{u,j}).$$

In summary the database is extended for color as follows(for texture analogous).

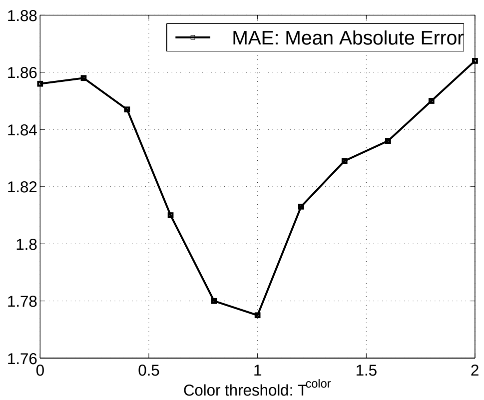
$$r_{u^{color},i} = \begin{cases} r_{u,i} & \text{if } r_{u,i} \text{ is defined.} \\ p^{color}(u,i) & \text{if } N_{u,i}^{color} \text{ is not empty.} \\ undefined & \text{else.} \end{cases}$$

The extended collaborative filtering database is then used with collaborative filtering algorithm, which has been described earlier. By extending existing users the possibility of correlation with the artificial users is increased. In fact, a user u correlates perfectly with his counterparts u^{color} and $u^{texture}$, which causes the content-based prediction to be strong part of the collaborative prediction of u and transitively also of all other users according to their similarity to u.

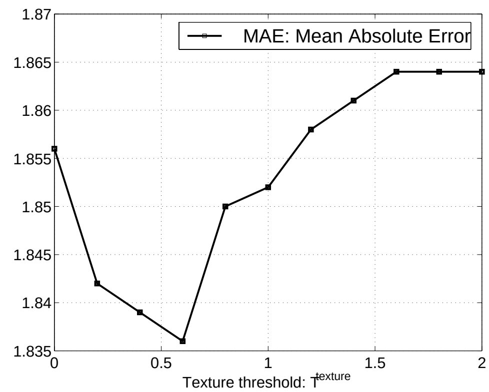
5.3 Evaluation

After the Active WebMuseum had been online for several months, roughly 4000 ratings by 140 users were collected. In order to verify that the previously described extension approach is valid, we measured the performance of the prediction algorithm in off-line evaluations on the collected data.

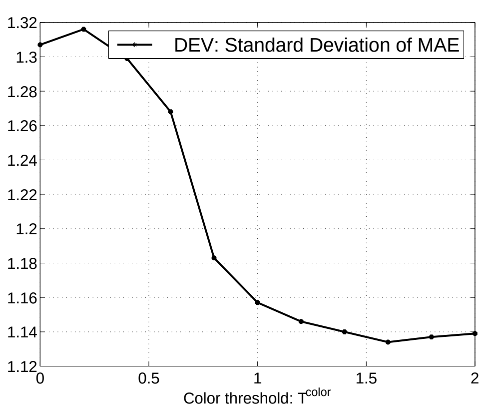
For the measurements 10 ratings for each user of a subset of 24 users were randomly chosen and separated in a test set of 240 ratings. The remaining ratings were used as a training set for the collaborative filtering algorithm and as a basis to create content-based artificial ratings. Then the extended training database was used by the collaborative filtering algorithm to predict ratings in the test set. The quality of the collaborative prediction was then derived by comparing the predictions with the actual ratings in the test-set. We used the mean absolute prediction error (MAE) as a metric for the precision, which is commonly used to evaluate collaborative filtering systems. The separation into training and test set was repeated 20 times (for the same parameters of $T^{color/texture}$) and the measurements were averaged.



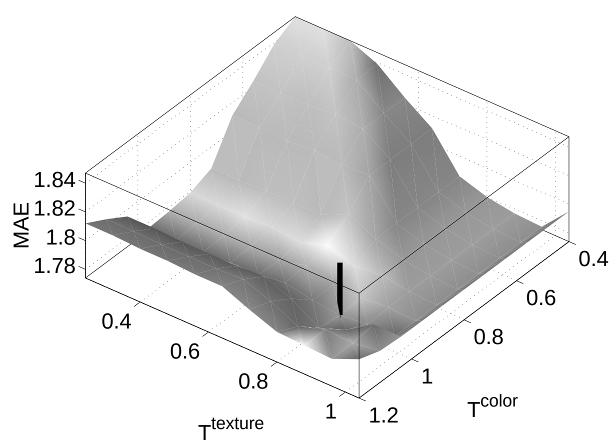
(a) Mean absolute error (MAE).



(b) Mean absolute error (MAE).



(c) Standard deviation of MAE (DEV).



(d) Using color and texture at the same time.

Fig. 4. Variation of the color and texture thresholds: $T^{texture}$ and $T^{texture}$

The graphs, figure 4(a) and figure 4(b), illustrate how the prediction precision in terms of MAE changes dependent on the choice of a parameter T^{color} and $T^{texture}$. A lower MAE can be observed for both parameters T^{color} and $T^{texture}$ greater than zero (a threshold equal to zero corresponds to no artificial ratings), indicating that each criteria can lead to extensions of the rating database which improve the results of the collaborative filtering algorithm. However, the impact of color seems to be more significant than from texture. Figure 4(c) depicts the standard deviation of the MAE of figure 4(a). It is notable that with increasing T^{color} the standard deviation of the MAE decreases, indicating that the predictions become more robust, i.e. higher absolute errors are more likely avoided (for texture similar measurements were made).

Figure 4(d) depicts the measurement of the MAE when color and texture extension was applied at the same time. The little black bar indicates the global minimum. The minimum is located within the plane, which indicates that through the extension with color and texture at the same time the overall performance can even be improved.

6 Conclusion and Future Work

In this paper we identify filtering as a key technology for user-adapting Web sites, sites which allow users to access information more efficiently by adapting to them. However, filtering is a hard problem, and cannot be addressed by one filtering technology alone. Due to limitations of both collaborative and content-based filtering, it is useful to combine these independent approaches to achieve better filtering results and therefore better user-adapting Web sites. We validate the combined use of collaborative and content-based filtering in several steps: First, we established a prototype user-adapting Web site, the Active WebMuseum, in order to apply our findings and to collect data about users. Second, we describe for the example of art paintings how image indexing can be used to extend collaborative filtering systems. And third, we validate the approach in off-line performance measurements.

In the future we plan evolve the extension algorithm to achieve better performance, e.g. by only extending users who are especially sensible to a certain image metric. Also, in addition to texture and color, other automatic indexing techniques shall be used to integrate more content-based schemes, e.g. fractal distances between images. Another important issue is a reliable measure of the systems performance. Monitoring the users' ratings during the use of a user-adapting Web site, will lead to a measure, which comes closer to user satisfaction.

Acknowledgments: This research was supported by a regional research grant, and by Eurecom's industrial members: Ascom, Cegetel, France Telecom, Hitachi, IBM France, Motorola, Swisscom, Texas Instruments, and Thomson CSF. The Active WebMuseum is based the collection of paintings of the *WebMuseum, Paris*, which has been created by Nicolas Pioch.

References

1. Joshua Alspector, Alexander Kolcz, and Nachimuthu Karunanithi. Feature-based and clique-based user models for movie selection: A comparative study. *User Modeling and User Adapted Interaction*, (7):279–304, 1997.
2. Marko Balabanovic and Yoav Shoham. Fab: Content-based, collaborative recommendation. *Communications of the ACM*, 40(3):66–72, March 1997.
3. Jack Breese, David Heckerman, and Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Madison, WI, July 1998. Morgan Kaufmann Publisher.
4. Arnd Kohrs and Bernard Merialdo. Clustering for collaborative filtering applications. In *Computational Intelligence for Modelling, Control & Automation*. IOS Press, 1999.
5. Arnd Kohrs and Bernard Merialdo. Improving collaborative filtering with multimedia indexing techniques to create user-adapting web sites. In *Proceedings of the 7th ACM Multimedia Conference*. ACM, 1999.
6. Joseph A. Konstan, Bradley N. Miller, David Maltz, Jonathan L. Herlocker, Lee R. Gordan, and John Riedl. Grouplens: Applying collaborative filtering to usenet news. *Communications of the ACM*, 40(3):77–87, March 1997.
7. Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of ACM CSCW'94 Conference on Computer-Supported Cooperative Work*, Sharing Information and Creating Meaning, pages 175–186, 1994.
8. Badrul M. Sarwar, Joseph A. Konstan, Al Bochers, Jon Herlocker, Brad Miller, and John Riedl. Using filtering agents to improve prediction quality in the grouplens research collaborative filtering system. In *Proceedings of ACM CSCW'98 Conference on Computer-Supported Cooperative Work*, 1998.
9. Uprendra Shardanand. Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of Human Factors in Computing Systems, CHI '95*, 1995.
10. J. R. Smith and S.-F. Chang. Local color and texture extraction and spatial query. In *IEEE Proceedings International Conference on Image Processing*, September 1996. Lausane, Switzerland.
11. J. R. Smith and S.-F. Chang. Tools and techniques for color image retrieval. In *Symposium on Electronic Imaging: Science and Technology - Storage and Retrieval for Image and Video Databases IV*, volume volume 2670, San Jose, CA, February 1996.