Creating User-Adapted Web Sites by the use of Collaborative Filtering

Arnd Kohrs – Bernard Merialdo

Institut Eurecom – Department of Multimedia Communications
BP 193 – 06904 Sophia-Antipolis – France
{Arnd.Kohrs,Bernard.Merialdo}@eurecom.fr

Abstract: The information globalization induced by the rapid development of the Internet and the accompanying adoption of the Web throughout the society leads to Web sites which reach large audiences. The diversity of the audiences and the need of customer retention require active Web sites which expose themselves in a customized or personalized way: We call those sites User-adapted Web sites. New technologies are necessary to personalize and customize content. Information filtering can be used for the discovery of important content and is therefore a key-technology for the creation of user-adapted Web sites.

In this article we focus on the application of collaborative filtering for user-adapted Web sites. We studied techniques for combining and integrating content-based filtering with collaborative filtering in order to address typical problems in collaborative filtering systems and to improve the performance. Other issues which are mentioned but only lightly covered include user interface challenges. To validate our approaches we developed a prototype user-adapted Web site, the Active WebMuseum, a museum Web site, which exposes its collection in a personalized way by the use of collaborative filtering.

Keywords: Collaborative filtering, personalization, content-based filtering, user-adapted Web sites, Web museum

More and more information is available on the Internet through Web sites. Generally, these Web sites focus on providing information on services and products. Sometimes, they also allow to perform transactions. The site owner compiles a vast amount of information and organizes it in a hypertext structure of Web pages in a way that is most desirable for expected users. Users have been recognized as a very valuable asset to Web sites, e.g. as advertisement audience or potential customers. Therefore, Web sites try to tie their users to their services by providing more efficient (e.g., less time-consuming) access to preferred content. A single hypertext structure for all users fails to achieve broad satisfaction because the users have different backgrounds and focuses. Therefore, Web sites which adapt to individual users are likely to be more attractive and successful. One approach for building such sites consists in letting the users consciously define preferences through user profiles, which are then used to personalize their visits to that Web site. Then, the Web site presents a customized view adapted to the user's interests. However, such an approach does require the user's effort to define his user-profile, the designer's effort to select the possible parameters for the profile, the developer's effort to implement the necessary procedures, and the content provider's effort to categorize every item of information. It seems to be more desirable to use machine learning techniques for automatically adapting the Web-site with little or no effort left for the user and then expose the Web site with a user-adapted structure. We call a Web site which adapts its structure to the individual user *user-adapted Web site*.

This trend of user-adapted 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, which can identify relevant pages for each user.

Information filtering techniques fall in two independent categories: Content-based filtering and collaborative filtering. Content-based 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.

Another important issue for adaptation deals with the impact on the user interface. First, the information that is displayed should be adapted to the user without violating editorial constraints. Second, the user feedback that is required to adapt the filtering system should be smoothly integrated in the interface.

Our work focuses on the use of information filtering techniques to construct user-adapted Web sites. In particular, we address the issues of combining content-based and collaborative filtering, as well as user feedback and dynamic presentation. A prototype Web site, the Active WebMuseum illustrates our approach.

In this article, we first present our prototype user-adapted Web site, the Active WebMuseum in section 2. In particular we focus on typical elements which are of interest for a broader application in other domains. In section 3 after describing the underlying collaborative filtering algorithm we present content-based techniques to address some weaknesses of collaborative filtering and to improve its performance. Further, we provide results from experiments. Finally, we discuss a more general application of our findings, conclude and describe further steps of this research in section 4.

2 The Active WebMuseum

Here, we first motivate the need for user user-adapted Web sites and at the same time give a broad definition. Then we present our prototype the Active WebMuseum.

2.1 A user-adapted Web Site

The design and creation of Web sites has become an industry by itself. A few years ago the creation of a hypertext web for presenting products or information was the task of the site's owner. It has been recognized that with growing competitiveness and demand for originality, besides other pressing constraints, the creation of a Web site needs to be performed by professionals in order to present the content in the most attractive way.

While a few years ago the Web was mostly used by the academic world, it is nowadays open to a broad heterogeneous audience spread over the entire society. The differences in the audience demand often different structures of the presented information in a Web site. Some Web sites anticipate several different groups of users and prepare predefined structures for the corresponding groups, e.g. most company Web sites let users access as customers, as investors or as other specific groups. For example, businesses provide several structures to their site's visitors in the roles of clients or investors. This approach of pre-clustering follows the prevailing assumption that users can be categorized in a small number of categories according to their interests or information needs. However, it is impossible to anticipate the wishes of every individual user through a set of predefined, usually manually prepared, structures of Web pages. Other Web sites allow users to define their personal profile, e.g. most Web portals* let users create personalized information pages, based on the user's selections from the portal's information resources and therefore providing a user-customized view to the site's content. However, this technique requires some effort by the user, defining his user profile, before he can benefit by receiving a personalized selection of information.

In order to avoid the effort of Web site creators to anticipate possible kinds of visitor-clusters, and to avoid the effort by visitors to customize their view to the site by defining their user-profile, we evaluate in this article another approach: a user-adapted Web site, which adapts by the use of collaborative filtering.

In general, people explore a Web site differently according to their personal interests or information needs. Doing so they continuously make decisions, e.g. setting bookmarks for content they like and ignoring content they dislike. Mostly, these decisions are related to personal good and bad experiences in finding desired information they are not further exploited to enhance the structure of the Web site in order to improve future visits by the same or other users. Collaborative filtering is a promising technology which predicts users' preferences based on other users experiences. Collaborative filtering seems to be an appropriate technology to automatically adapt Web-sites by presenting an personalized structure to each user.

Collaborative filtering systems are generally used for recommender services, which provide personal recommendations based on ratings provided by the user. So two important tasks can be identified when collaborative filtering systems are used:

- capturing feedback from the user.

* The *My Yahoo* service of the Web portal *Yahoo!* is a good example for a customized user-customized Web-site (`http://www.yahoo.com`).

- presenting recommended information to the user.

When meeting these requirements other constrains might interfere. Providing explicit feedback might be perceived as inconvenient for the user. Also, the provisions for providing feedback could disturb the structure of a Web site, if done uncarefully. Therefore, it is important to allow user feedback in a convenient and non-disturbing way. On the other side, the Web site needs to create an adapted presentation based on predicted preferences for a particular user. It is necessary to integrate the predictions in presentation elements without disturbing the editorial conception of the Web site.

In the next section we describe our prototype user-adapted Web site, the Active WebMuseum. We describe how collaborative filtering is integrated in a Web museum site, in order to provide a personal experience to the user, while addressing challenging issues presented above. The particular type of information that is contained in a Web-based museum, namely paintings, makes it simple to define an adaptive presentation of this information.

2.2 The Application: The Active WebMuseum

In an ideal world a visitor of a museum would enter a museum and then find in the first corridor exactly those items, which he would find most interesting. Given that real museums serve many people at the same time, it is not feasible to rearrange the collection for individual visitors. Often, real museums offer guided tours, which might be covering a particular topic or addressing a particular interest group, but having personal guides who show exactly the items of high interest seems to be impractical. When a museum's art collection is presented through the Web, it is at least feasible to rearrange the collection for each individual visitor.

Numerous museum sites already exist on the Web. They present images of arts contained in a hypertext structure, so that the navigation within a Web-based museum emulates strolling through the corridors of a real museum. 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 contrast, our Active WebMuseum has a dynamic topology which is adapting to the museum visitor's taste and choices.[**]

2.3 The Content Model

While filtering techniques allow automated recommendation of what users might prefer, using these techniques introduces new constraints for the design and presentation of the Web site. The content of the site should be rearrangeable according to filtering results while still offering attractiveness for users by providing typical features of the Web, interlinked content accessed through self-directed browsing. In our prototype we use art paintings as entities which can be rearranged for the user. At a higher level we use a *Multi-Corridor-Access-Paradigm*. Corridors are ordered containers of paintings with some common characteristics. The user may choose from several corridors. The results of the filtering are then reflected in a personalized order of the paintings within the corridor.

The user is assumed to behave as follows:

- choose a corridor
- enjoy the paintings in a corridor
- at any time leave the corridor and choose another one.

The corridors are interlinked. E.g., when a user visits a corridor with painting by Van Gogh, he can at any time switch to a related

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

corridor with paintings of the same time-period and from there to painting of a painter who lived at the same time.

By reordering and inter-connecting existing corridors, which contain references to pages showing paintings, it is possible to dynamically restructure the museum site in a way adapted to the user. See figure 1 and figure 2 for examples.

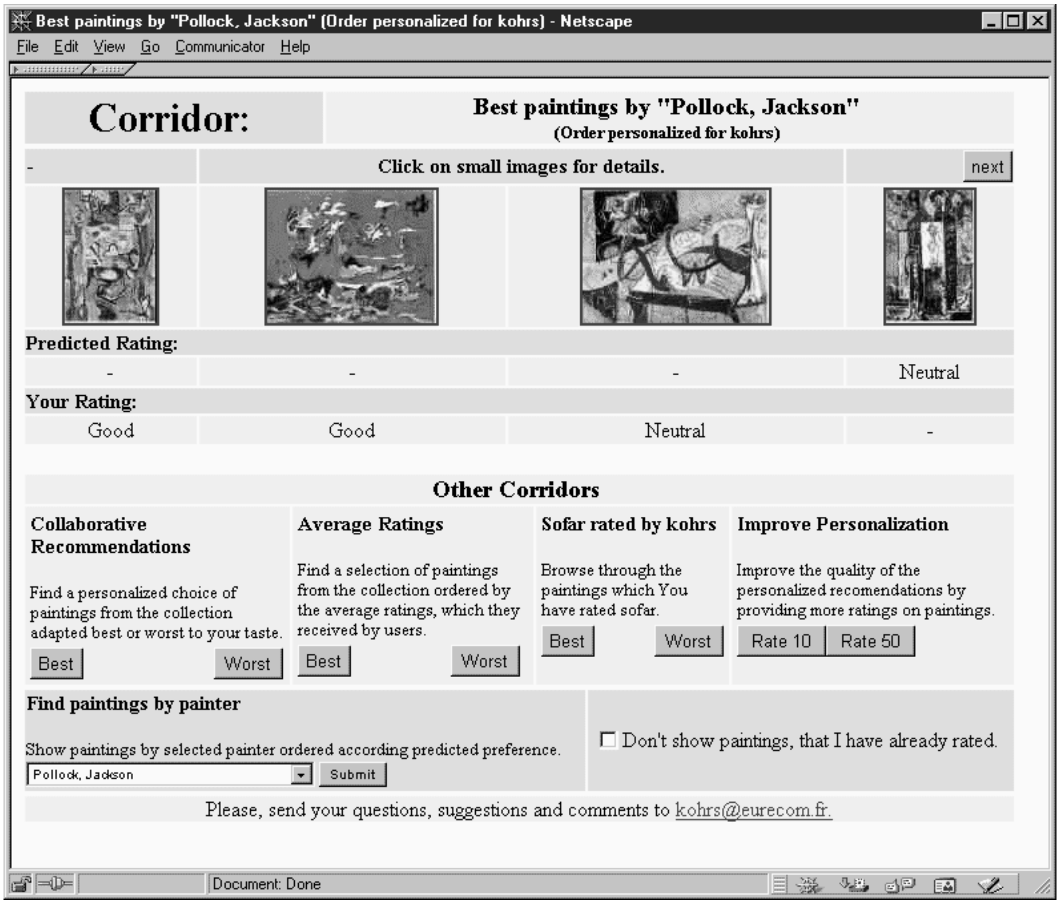


Fig. 1. Browsing dynamic corridors: When a user has entered a dynamic corridor (in this example a corridor containing paintings by Jackson Pollock), he is presented iconized paintings ordered according to his preference. From here the visitor continues in the same corridor until he looses interest, or may choose to see more details of one painting, or enter a new corridor.

2.4 User Profiling: Acquiring Preferences

In the previous section we explained how preferences are used in order to transform corridors into personalized corridors. This

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). From here the user may return to current corridor or enter new corridors related to this painting.

approach is based on the existence of the visitor's preferences. In the Active WebMuseum visitors can express preference by giving ratings to paintings (5 values: *excellent*, *good*, *neutral*, *bad*, *terrible*). For paintings which have not been rated by the visitor, the ratings are predicted using other users' ratings and collaborative filtering technology.

Rating paintings should not be the main occupation of a user. Therefore, the ratings can be conveniently provided while wandering within the museum. We noticed in initial trials, that users are hesitant in giving ratings, because giving a rating demands the inconvenience of having to make a decision. Therefore, we provided in the user-interface that

- the ratings can be provided with very little effort (one mouse click) without disrupting the users chosen tour, and

- if a painting is viewed in detail a rating is mandatory so that the visitor has to provide a rating or otherwise he cannot continue his tour.

In the next section we describe the used collaborative filtering algorithm and our content-based combination techniques for its performance.

3 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. This allows these techniques to be used on any type of objects and thus build a large variety of services. 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 [13–15]. Music, Ringo [16]. Movies, MovieCritic[⋆⋆⋆]. Books at Amazon[⋆⋆⋆⋆]. Restaurants [7].

We first describe collaborative filtering as it is commonly used. And then we address some improvements of collaborative filtering by introducing our content-based extension based on multi-media indexing.

⋆⋆⋆ http://www.moviecritic.com
⋆⋆⋆⋆ http://www.amazon.com

Most collaborative filtering systems collect the users' opinions as ratings on a numerical scale, which leads to a very sparse matrix $rating(user, item)$ (in short $r_{u,i}$). Collaborative filtering systems then use this rating matrix in order to derive predictions. Several algorithms have been proposed on how to use the rating matrix to predict ratings [9,16,6]. 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 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 weighted 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:

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

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':

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

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.

$$p^{collab}(u,i) = \overline{r_u} + k \sum_{u' \in U_i} \rho(u,u')(r_{u',i} - \overline{r_{u'}})$$

$$\text{with} \begin{cases} U_i : \text{Users, who rated object } i. \\ k = \frac{1}{\sum_{u' \in U_i} \rho(u,u')} \end{cases}$$

(The factor k normalizes the weights.)

Sometimes the correlation coefficient between two users is undefined because the users have not rated common objects, i.e. $I_u \cap I_{u'} = \varnothing$. We found in our experiments that assuming a default value for the correlation between the user rating vectors is helpful when the data-set is very small. For example, we measured in experiments $\rho_{default} = 0.2$ as the mean of typically occuring correlation coefficients in our data-set. The application of a default correlation biases the prediction toward the mean, which might stabilize the prediction results for very small data-sets. To avoid this bias, we did not use default correlation in our later described experiments, instead, if correlation between a target and a peer user can not be measured, the peer user is ignored for the prediction for the target users.

3.2 Extending Collaborative Filtering with Content-based Information

In earlier work we identified typical problematic cases for collaborative filtering systems [9], cases when not enough ratings are available, due to an insufficient amount of users or to few ratings per user. In these cases, the above formulas are bad estimations (in the statistical sense) of the real correlation and rating values.

In summary, shortages of rating data appear in the following situations:

New user case: A new user with no or very few ratings can not be reliably matched against other users.

New object case: If a new object is entered to the collection, it has not been rated by any user and therefore it might neglected.

Bootstrap case: This is a combination of the above cases and is typical for new recommender sites with no or very few ratings. It is hard to attract users because they get very poor recommendations in the beginning and therefore the recommender system is likely to stay in a bootstrap dilemma.

In the contrary, content-based schemes are less sensible to sparsity of ratings. While the new user case (and consequently partially the bootstrap case) remains a problem for content-based recommender systems the bootstrap and the new object case can be handled easier since new objects can be compared to objects one user has recently rated based on a content-based criteria.

However, pure content-based schemes are only appropriate when objects can be reliably automatically compared, i.e. the content can automatically be indexed or other meta data is available. While it is possible to obtain meta information about movies (actors, director, genre etc), it is hard to video-index a movie with current technology in order to derive the content. Once the content of a object is indexed it is crucial for content-based filtering, to decide as to how much two objects are similar. E.g., in order to recommend the second object to a user if he likes the first and the first is similar to the second object.

Comparative studies [2,1] indicate that content-based filtering nearly performs as well as collaborative filtering in the domain of movie recommendations, where the content of the objects (movies) is described in terms of textual attributes obtained from movie databases e.g. director, genre etc. In other studies content-based filtering and collaborative filtering have been combined in the context of document recommendation in order to overcome inherent disadvantages of both technologies [5,8]: Advanced techniques for indexing and comparing textual documents, as the vector space model, were borrowed from the "Information Retrieval" community. In the newsgroup article recommendation track of the project GroupLens, weak content-based criteria, such as spelling and article length, is used to fabricate ratings. Our research has a similar focus: In order to improve collaborative filtering we address key problematic cases which result from lack of ratings. We chose a domain, art paintings, where content-based filtering and indexing techniques are more promising. Combination techniques for content-based and collaborative filtering for art paintings should therefore have similar properties as for other domains where indexing techniques are less advanced.

In the following we motivate the use of content-based filtering in combination with collaborative filtering by first describing the studied criteria and observations made from the users ratings. Then, we present our approaches of integrating content-based criteria into collaborative filtering in order to improve the prediction performance. The proposed techniques are evaluated in off-line experiments.

3.2.1 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 [17,18]: Color histograms and texture coefficients.

3.2.2 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:

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

3.2.3 Texture Coefficients

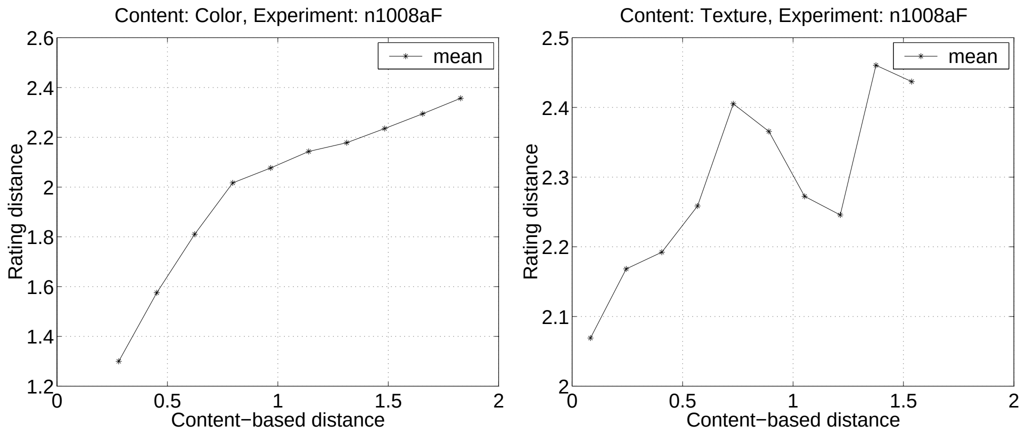
While color histograms do not take into account the arrangement of pixels, texture coefficients can be computed to characterize local properties of the image. We are using a wavelet decomposition using the Haar transform, by which a number of sub-images corresponding to a frequency decomposition are 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, the L_1(see equation 1) distance is used to measure the distance between images.

3.2.4 Relationship between Ratings and Color and Texture

After advertisement in several mailing lists, which are related to recommender systems, we attracted initial users to the *Active WebMuseum*. The (ongoing) trial lead to a database of approximately 11500 ratings of 468 users. We use this data-set now to analyze the relationship between the users rating behavior and the chosen content-based criteria color and texture.

From the rating database we sampled instances when a user rates two different paintings. For each instance we measured the absolute difference between the rating (a score between 0 and 10) and the absolute content-based distance (according to the distance measures, described in sections 3.2.2 and section 3.2.3). In a second step these instances were clustered into ten equal-length intervals of content-based distances. For each interval the mean absolute rating distance was measured, so that a statistic is approximated as to what absolute rating distance to expect for one user for two paintings with a given content-based distance.

The results of these measurements for the color histogram distance and the texture coefficient distance measure are plotted in figure 3. It can be noted that there is a clear positive correlation between color histogram distance and the rating difference, suggesting that paintings which are close in color are more likely to get similar ratings by one user as paintings which are dissimilar in color. The plot for texture is less indicative: while there seems to be a trend towards a positive correlation there is some fluctuation. Further, it can be noted that the correlation of color is stronger than the correlation for texture in terms of magnitude of rating difference.



(a) Color-histogram distances (b) Texture-coefficients distances

Fig. 3. Correlation of distances between color-histograms(texture-coefficients) of paintings and differences of ratings assigned by users.

It should be noted that the observations above averages all the users from the dataset, while it is probable that users vary strongly respective to the correlation between color, texture and ratings, e.g., for some users the correlation might apply weaker while for others it applies stronger.

In the next sections we suggest techniques which combine collaborative filtering with content-based measures in order to take advantage of the correlation between color, texture and ratings.

3.2.5 Combining Linearly with Content-based Prediction

Based on the findings of the previous measurements concerning the relationship between content-based painting distance and rating difference we derived a content-based prediction model. We use a linear estimator for the content-based prediction which we model as follows:

$$
\begin{aligned}
r_{u,i} &: \text{user } u\text{'s rating for painting } i \\
I_u &: \text{Paintings, rated by user } u.
\end{aligned}
$$

Distance intervals:

$$j = 1..n_\lambda$$
$$interval_1 = [0, 1); interval_2 = [1, 1.5)...$$

Distance classes:

$$C_j(i) = \{i' \in I_u : d^{color}(i, i') \in interval_j\}$$

Prediction for painting i for user u:

$$p^{color}(u, i) = \sum_{j \in 1..n_\lambda} \lambda_j \cdot \frac{\sum_{i' \in C_j(i)} r_{u,i'}}{|C_j(i)|}$$

Expressed in words, the prediction works as follows: If a prediction is to be made for a user u and a target painting i, all the paintings previously rated by user u are grouped into distance classes ($C_j(i)$) according to color-based distance to target painting i (see figure 4 for an illustration). Each class is associated with

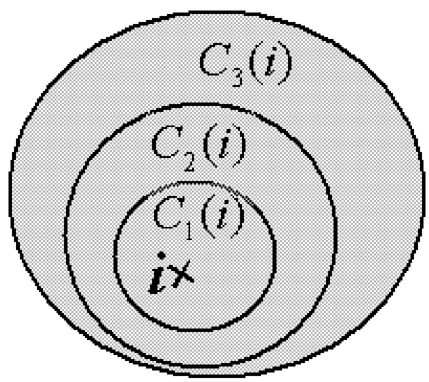


Fig. 4. All paintings are classified in distance classes according to painting i, so that paintings that have a similar distance to painting i fall in the same class.

a weight λ_j, the degree of confidence. The prediction is then the weighted sum of the mean ratings of each class. The weights λ_j are estimated through linear regression by using a priorly separated subset of the ratings. For each content-based criteria, color histogram and texture coefficients, a predictor was created: p^{color} and $p^{texture}$.

In a second step these content-based predictors are combined with the collaborative filtering predictor p^{collab}, as described in section 3, linearly using the following formula:

$$\begin{aligned} p^{comb}(u,i) = & \mu^{collab} \cdot p^{collab}(u,i) + \\ & \mu^{color} \cdot p^{color}(u,i) + \\ & \mu^{texture} \cdot p^{texture}(u,i) \end{aligned}$$
$$\text{with } \sum_k \mu^k = 1$$

The weights $\mu^{\{collab,color,texture\}}$ are estimated by the use of linear regression with a set-aside subset of the ratings, so that the weights are adapted to the relevance of a predictor, e.g., higher for color based than for texture based prediction. The estimation of the weights should be repeated as the rating database grows, in order to take into account a precision-gaining collaborative predictor.

As research leads to additional content-analysis tools for paintings, this combination approach allows unlimited inclusion of new content-based predictors. Later in this article, we present experimental results for this approach, after presenting a second technique in the following section.

3.2.6 Deriving Artificial Users from Image Metrics

While the previous approach combined the collaborative filtering and content-based linearly, assuming that these are independent, we present now another approach which does not change the collaborative prediction algorithm (see section 3.1) but instead alters the rating database based on content-based criteria. The rating database is extended with artificial users whose ratings are based on content-based criteria and ratings of real users. This approach was inspired by Sarwar's *rating-bots* approach [15] for the GroupLens news filtering project: For that project software agents which used simple content-based criteria (spelling and article length) to rate news articles automatically and to increase the amount of ratings in the database.

For each described distance metric of section 3.2.1 and for each real user u a corresponding artificial user u^{color} and $u^{texture}$ is derived. The artificial users inherit the ratings from 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. This means that some unrated items are assigned a predicted rating, based on similarity between the rated items and the item for which the rating is missing. In order to make a content-based prediction, we define a restricted neighborhood $N_{u,i}$ around a painting i containing paintings j whose rating is defined by the user u and whose content-based distance to painting i is below a threshold θ:

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

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

$$p^{color}(u,i) = \sum_{j \in N_{u,i}^{color}} \frac{r_{u,j}}{|N_{u,i}^{color}|}$$

In summary the database (or rating matrix) is extended for color as follows:

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

A similar process is conducted with the texture distance. Figure 5 illustrates the extension technique.

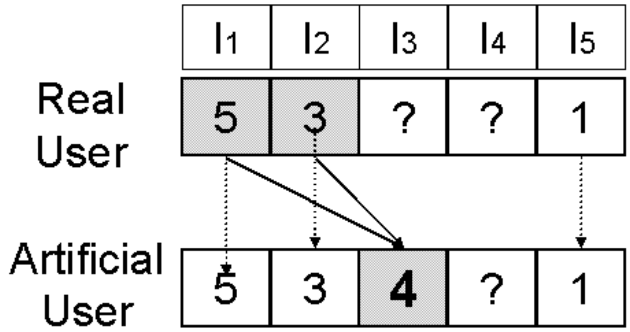


Fig. 5. Illustration of the extension technique: The artificial user inherits ratings from the real user for the paintings I_1, I_2 and I_5. The rating for I_3 is estimated based on ratings for I_1 and I_2 which are related to I_3 according to a content measure (not shown in the illustration).

The extended rating database is then used with the collaborative filtering algorithm, which has been described earlier (see section 3.1). By extending existing users the possibility of correlation with the artificial users is increased, because the number of commonly rated items is maximal. 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 who are similar to u.

3.2.7 Evaluation

We evaluate our approaches for integrating content-based information in the collaborative filtering task in off-line experiments. The experiments are based on rating data which we collected from users who participated in our ongoing public online trial of the Active WebMuseum. The Active WebMuseum had been advertised in related mailing lists to attract users.

3.2.7.1 The Dataset While the dataset is still growing, we used a snapshot of the dataset, which contains 11500 ratings by 468 users for 1082 paintings.

Figure 6 depicts a histogram of the collected user ratings.

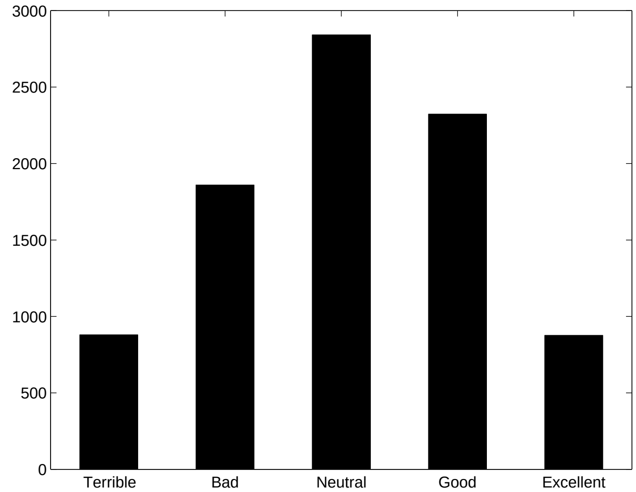


Fig. 6. Distribution of the user ratings for the paintings in the Active WebMuseum.

3.2.7.2 Measurements: In order to evaluate various approaches of collaborative filtering we divided the rating dataset in test-set (r^{test}) and training-set ($r^{training}$). The training-set is used to predict ratings in the test-set. The predictions are then compared to the ratings that users provided in the test-set using a commonly used error measure:

Mean Absolute Error (MAE): The mean absolute error is calculated as follows:

$$MAE(u) = \frac{\sum_{i \in I_u/r^{test}} |r_{u,i} - p(u,i)|}{|I_u/r^{test}|}$$
$$I_u/r^{test} = \{i : r_{u,i} \in r^{test}\}$$

The MAE for several users is then accumulated as follows:

$$MAE = \sum_{u \in U} \frac{MAE(u) \cdot |I_u/r^{test}|}{|U|}$$

The test-set contains 50 users for which more than 40 ratings are known. From each user in the test-set 10 ratings are sampled and put aside into the test-set the remaining ratings are used in the training set.

3.2.7.3 Parameter Estimation Both presented combination approaches use parameters to control the mix of collaborative filtering and content-based filtering. For the linear combination approach the parameters ($\mu^{collab,color,texture}$) determine the weight in the sum of each predictor. The algorithm is designed so that the parameters are adjusted automatically. We found through our experiments the following values for these parameters:

$$\mu^{collab} = 0.53, \quad \mu^{color} = 0.41, \quad \mu^{texture} = 0.06 \tag{2}$$

This shows that the most important component in the linear combination is collaborative filtering closely followed by color-based prediction. Texture-based prediction has only negligible importance.

For the approach with artificial users, thresholds values (θ^{color} and $\theta^{texture}$) are needed in order to determine the neighborhood paintings of a target painting which should determine the artificial rating. The threshold values need to be assigned for the extension algorithm to reasonable values. In order to find reasonable values we searched in the possible range: The previously described dataset was divided twenty times into test-set and training-set by sampling each time a different subset as the test-set. The extension algorithm was used to predict each test-set.

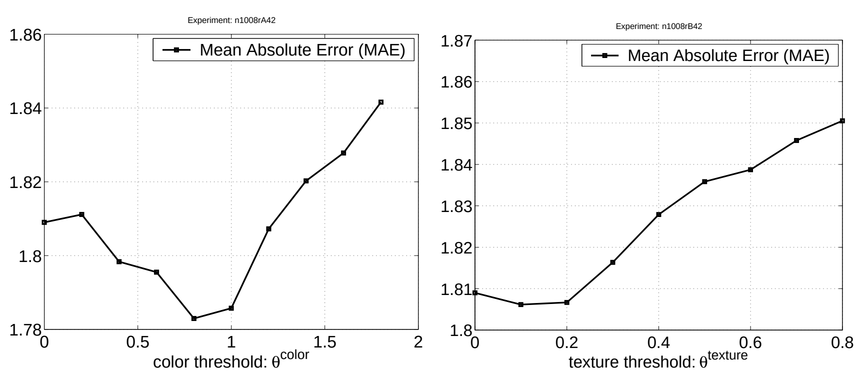


Fig. 7. Variation of the parameters θ^{color} and $\theta^{texture}$

The graphs of figure 7 plot the MAE as a function of θ^{color} and $\theta^{texture}$. The experiments suggest that good settings for these parameters are as follows:

$$\theta^{color} = 0.8, \qquad \theta^{texture} = 0.1$$

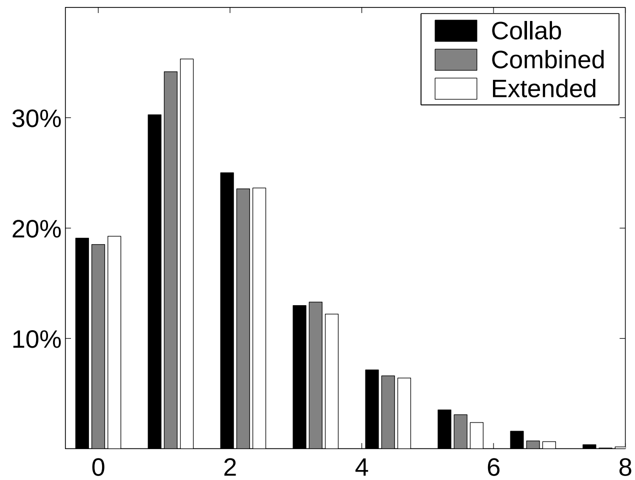


Fig. 8. Histogram of the absolute prediction errors.

3.2.7.4 Comparing Extension and Linear Combination: Figure 8 shows the histogram of the absolute prediction errors created by using pure collaborative filtering (collab) , the linear combination method (combined) and the extension with artificial users technique (extended). It can be noted that while the collaborative predictor shows more frequent smaller errors, the combined predictor avoids large errors. However, it is hard to judge which one should be better.

Table 1 lists the measured precision for the previously discussed predictors. When the combined or the extended technique is used lower errors are obtained than by the use of the pure collaborative predictor.

Prediction Method	**MAE**
Collab	1.787
Combined	1.772
Extended	1.771

Table 1
Prediction precision of collaborative, combined and extended predictors: The mean absolute prediction error(MAE) for ratings of the test-set was measured.

The observed improvements in prediction precision through the combination of collaborative filtering and content-based filtering, based on color and texture, suggest that content-based information should be exploited if available. However, the presented combination models do not distinguish between individual users, i.e. for each user the same mix of predictors is used. For some users the content-based measures might be less appropriate as for others. In further developments of the combination models the differences in sensibility of different users to content features should be considered.

The goal of the prediction is to present users only relevant paintings (paintings, which they would rate highly), so that the users get the most satisfaction from visiting the site of the Active WebMuseum. Therefore, a good metric for the comparison of different prediction strategies should be focused on user satisfaction. We believe that the previously used metric which has been commonly used in the literature, is related to user satisfaction but does not focus on the goals of the users. User-satisfaction can be derived from how well the museum's collection had been restructured to meet the user's preferences. Current research is directed at developing metrices to measure the recommendation performance with respect to the *multi-corridor-paradigm*.

4 Discussion

Now, we step back from our prototype application of a user-adapted Web site and provide more general insight, then we conclude this article.

4.1 General Application

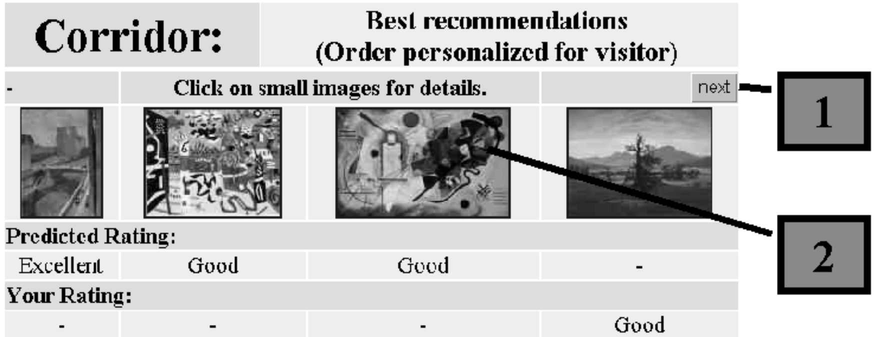
In this section we summarize some general observations made while designing the Active WebMuseum. In order to create a more attractive Web site, we added adaptability-to-the-user to it. This should lead to a win-win situation for

- the user, because the site is more convenient to use (more interesting and less time-consuming) and
- for the Web-site provider, because the site becomes more attractive. An attractive Web site leads to better leverage of underlying resources and therefore a more valuable Web site for the owner, e.g., the information can be spread to a larger audience leading to more advertising revenues.

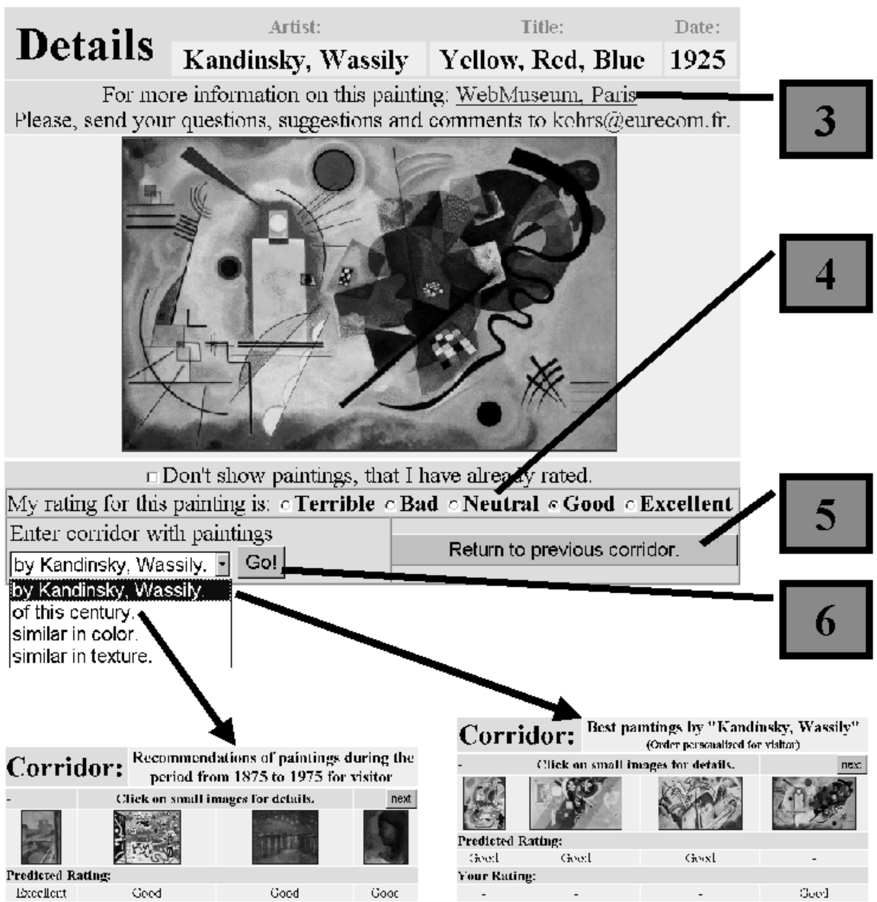
Here, we point out important issues when adding adaptability to a web-site by the use of collaborative filtering. The general scenario can be outlined by the following steps:

- The user visits a site containing Web pages of some sort of information.
- The server offers predefined structures to access the pages.
- The user navigates through the server while providing feedback.
- The server adapts subsequent pages reflecting a structure adapted to the user.

The most important issue is how to integrate (predicted) preferences about items into the structure of pages. In the case of the Active WebMuseum we use the multi-corridor-access-paradigm [12]: Interconnected corridors containing objects which are ordered according to the filtering algorithm. Figure 9 depicts how the multi-corridor-access-paradigm is applied and implemented in the Active WebMuseum. This paradigm can be applied in a similar way to other domains and is not exclusively suitable for paintings only. For example in the general case, when the corridor is in fact a regular Web page, the links contained in a page could be highlighted according to importance, e.g. the color intensity would depend on importance (to reflect the ordering). Another possibility is to always accompany a page with a box of recommended links.



(a) As a starting point the visitor is presented the beginning of a corridor containing all paintings in the order best suited (predicted) for the visitor. From here the visitor may move within the corridor(1). When the visitor finds an interesting painting he may select it in order to see more details(2).



(b) When viewing a painting in detail the visitor is provided with information about the painting and links to more information (3). Now the visitor expresses his opinion about the painting by selecting the corresponding score(4). From here the visitor may return to the previous corridor(5) or he might want to follow another direction by selecting a corridor related to the current painting(6).

Fig. 9. Multi-Corridor-Access.

Another important issue is the acquisition of user feedback (ratings) so that the underlying recommender system can learn about the user's taste, upon which recommendations are based. In the case of the Active WebMuseum we use a little bit of force to motivate the user to provide an explicit rating, by denying to proceed to another corridor if a viewed painting does not receive an evaluation.

Two general approaches are possible to attain user feedback:

- Implicit feedback: The user's actions are observed and the feedback is derived based on assumptions on the user's behavior. In the GroupLens[13] news article filtering project implicit feedback was used in terms of measuring the time a user spend reading an article, assuming that users would spend more time with articles they enjoy. Balabanovic[4] proposed a system where the users interactions with an application are used to derive implicit information, e.g., the user bookmarks, files or deletes an object.
- Explicit Feedback: Explicit feedback in collaborative filtering systems is usually provided as single numerical score.

The advantage of implicit feedback lies in the ease for the user. The price for ease is paid in inaccuracy, since the interpretation of the user's actions might be wrong, e.g. when a user views a painting for several minutes it is uncertain that he was not interrupted by a phone call, so that the conclusion from this observation is flawed with an estimation error.

While explicit feedback is more precise it might appear unnatural in the context, e.g. some visitors of museums might find it annoying to decide on an appropriate ratings for a painting. Further, people might not see the direct benefit for providing feedback. It is the nature of collaborative filtering that user feedback is compensated by a better performance of the collaborative filtering system for other users but also especially for the rating-providing users themselves. However, users might not always understand the technology and therefore incentives to provide feedback are more promising. Consequently, in the Active WebMuseum the user is reminded of the importance to provide ratings and may not proceed each time he neglects to provide a score for a painting. Avery[3] argues voluntary provision of evaluations leads to a suboptimal supply. Therefore, models which compensate the users directly for providing ratings are more promising.

A hybrid approach would contain the good elements of both approaches: reliable feedback and little or no disruption of the users activity. If the user interface allowed more expressiveness, then the valuable information could be derived, for example by providing alternative *back* button in a browser:

- back, because I don't like this page,
- back, because I like the previous page,
- back, but I like this page, or
- back, for no particular reason.

Even when the Web-site is designed in a way that user feedback can be captured, typical situations may arise when the collaborative filtering performs poorly due to a lack of ratings, especially in the beginning (see section 3.2). Then a combination of two independent technologies collaborative-filtering and content-based filtering leads to better prediction performance. While our approaches focus on paintings and related content features, i.e. color-histograms and texture-coefficients, the techniques can be easily adapted to other domains by replacing the feature comparison part. For example, if the algorithms should be used for adapting a web-based message board, word frequency vectors of messages can be used as content-based features, and content-based distance metrices can be derived from these.

In this article, we have identified filtering, especially collaborative filtering, as a key technology for the creation of user-adapted Web sites, sites which allow users to access information more efficiently by exposing a personalized structure of the site. We identified problematic issues when using filtering technology in order to adapt a Web-site to individual users: How can user feedback be achieved, and how can the filtering recommendations be used to augment the site's structure. We proposed alternative solutions. We implemented a prototype user-adapted Web site, the Active WebMuseum, which has been tested by a number of users.

However, our key interest lies in the enabling technology. 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-adapted Web sites. We described how content-based techniques (in this case based on image indexing) can be used to add content-based capabilities to the collaborative filtering and then evaluated our approach in off-line experiments.

This work opens a number of directions for further research:

- improve the combination with content-based filtering with respect to individual sensibility to content features,
- propose evaluation measures more related to user satisfaction in particular based on the *multi-corridor access paradigm*,

We believe that collaborative filtering will appear as a powerful tool in the future for the construction of user-adapted Web sites.

4.2.0.5 Acknowledgments: This research was supported by a regional research grant, and by Eurecom's industrial members: Ascom, Cegetel, France Telecom, Hitachi, ST Microelectronics, Mo-

torola, 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*, pages 279–304, 1997.

[2] Joshua Alspector, Alexander Kolcz, and Nachimuthu Karunanithi. Comparing feature-based and clique-based user models for movie selection. In *DL '98. Proceedings of the third ACM Conference on Digital libraries*, pages 11–18, 1998.

[3] Christopher Avery, Paul Resnick, and Richard Zeckhauser. The market for evaluations. *American Economic Review*, (89(3)):564–584, 1999.

[4] Marko Balabanovic. An interface for learning multi-topic user profiles from impicit feedback. In *Recommender Systems: AAAI 1998 Workshop*, 1998.

[5] Marko Balabanovic and Yoav Shoham. Fab: Content-based, collaborative recommendation. *Communications of the ACM*, 40(3):66–72, March 1997.

[6] 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.

[7] Robin Burke. Semantic ratings and heuristic similarity for collaborative filtering. In *AAAI Workshop on Knowledge-based Electronic Markets. Austin, TX. July, 2000*, 2000.

[8] Joaquin Delgado, Naohiro Ishii, and Tomoki Ura. Content-based collaborative information filtering. In *Cooperative Information Agents II*, Lecture Notes in Artificial Intelligence 1435, pages 206–215. Springer Verlag Berlin Heidelberg New York, 1998.

[9] Arnd Kohrs and Bernard Merialdo. Clustering for collaborative filtering applications. In *Computational Intelligence for Modelling, Control & Automation (CIMCA'99), Vienna*, pages 199–204. IOS Press, February 1999.

[10] 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, Orlando*. ACM, 1999.

[11] Arnd Kohrs and Bernard Merialdo. Using color and texture indexing to improve collaborative filtering of art paintings. In *Proceedings of the European Workshop on Content-Based Multimedia Indexing (CBMI'99)*, 1999.

[12] Arnd Kohrs and Bernard Merialdo. Using category-based collaborative filtering in the active webmuseum. In *IEEE International Conference on Multimedia and Expo, New York*, 2000.

[13] 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.

[14] 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.

[15] 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.

[16] Uprendra Shardanand and Pattie Maes. Social information filtering: Algorithms for automating "word of mouth". In *Proceedings of Human Factors in Computing Systems, CHI '95*, 1995.

[17] 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.

[18] 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.