PAPER

# Image Recommendation Algorithm Using Feature-Based Collaborative Filtering

**Deok-Hwan KIM**[†a)], ***Member***

**SUMMARY** As the multimedia contents market continues its rapid expansion, the amount of image contents used in mobile phone services, digital libraries, and catalog service is increasing remarkably. In spite of this rapid growth, users experience high levels of frustration when searching for the desired image. Even though new images are profitable to the service providers, traditional collaborative filtering methods cannot recommend them. To solve this problem, in this paper, we propose feature-based collaborative filtering (FBCF) method to reflect the user's most recent preference by representing his purchase sequence in the visual feature space. The proposed approach represents the images that have been purchased in the past as the feature clusters in the multi-dimensional feature space and then selects neighbors by using an inter-cluster distance function between their feature clusters. Various experiments using real image data demonstrate that the proposed approach provides a higher quality recommendation and better performance than do typical collaborative filtering and content-based filtering techniques.

***key words:*** *feature clustering, image segmentation, recommendation, collaborative filtering, region-based image retrieval*

## 1. Introduction

As the multimedia contents market grows more and more rapidly, the amount of image contents used in mobile phone services, digital libraries, and online catalog services has increased remarkably and these services have taken a large portion of the market.

Although the popularity of the image contents has increased rapidly, many users often fail to search for the image that they really want because a user's preference with respect to images is ambiguous and more changeable over time than that with respect to the usual items. It is hard to explain images' features such as color, texture and shape, etc. Therefore, many users search for the desired image by scanning the offered image list one by one or for a keyword directly. To reduce the users' searching efforts and time, service providers adopt a recommender system. The recommender system is a software system developed to identify particular items that are likely to match each user's tastes or preferences and the system then recommends the items to the users.

One of the most successful recommendation techniques to date is collaborative filtering (CF) [5], [12], [14], [16], which identifies users (i.e. neighbors) whose tastes are similar to those of a given user and recommends items those users have liked in the past. A general CF technique uses a user-item binary matrix as its profile.

However, CF techniques have three major shortcomings: (1) most similarity measures used in CF work properly only when there is a sufficient number of rating on common images from the target user's neighbors; (2) it cannot recommend new images since new image have not been purchased by anyone so they cannot have any ratings; (3) it is hard to find the real neighbors since CF uses only common images to form the target user's neighbor with respect to a user-item binary matrix. The former addresses a sparsity problem [2], [9], [14]. The more the number of users and images is increased, the worse this problem becomes because the likelihood that different users will rate common images decreases. Such sparsity in ratings results in poor recommendations because it makes neighborhood formation inaccurate [7]. The latter addresses a new item problem [2], [9]. In a mobile environment, digital library, or online catalog site, new images are very frequently supplied and their purchasing ratio is high. However, new images that are added to a site recently cannot be recommended. Because CF recommends an image on the basis of previous user's ratings with respect to that image, it cannot recommend a newly added image until sufficient ratings of that image are available. Beyond these problems the CF techniques have a radical problem related with input data representation. When a user has purchased an item, a general CF technique gives a rating to the item and its purchase sequence is represented by the $m \times n$ user-item binary matrix with $m$ user's rating about $n$ items. It may be possible that someone whose preference is similar to the target user's preference did not purchase common images. That is, the CF technique cannot find such hidden neighbors.

As a solution to these problems, we propose a new feature-based collaborative filtering (FBCF) method. In this method, images purchased by the user in the past are represented as the feature clusters in the multi-dimensional feature space. Neighbors are selected by using an inter-cluster distance function between their feature clusters and the target user's feature clusters. To reduce the semantic gap between a high level user's concept and the low level feature representation of an image, we segment an image into distinct regions and extract the image's various visual features like the color, texture, and shape of individual regions. Then each region of an image is represented as a point in the multi-dimensional feature space. The main objective of using local features is to enhance the ability of captur-

Manuscript received November 26, 2007.

Manuscript revised July 14, 2008.

[†]The author is with the School of Electrical Engineering, Inha University at Incheon, Korea.

a) E-mail: deokhwan@inha.ac.kr

DOI: 10.1587/transinf.E92.D.413

ing user's perception with respect to an image. To reduce the user's search effort, the proposed method can reflect the user's most recent preference by representing his purchase sequence and preference in the visual feature space. The contribution of the proposed method is as follows:

- It can identify the neighbors whose preferences are really close to the target user even though the ratings on common images from the target user's neighbors do not exist.
- It can recommend new images by giving them virtual ratings in the feature space.

The rest of the paper is organized as follows: Section 2 reviews the related work addressing issues of CF and region-based image retrieval. In Sect. 3, we give a proposed methodology. In this section, we describe the model building process, recommendation process, new image recommendation, and the proposed image recommendation algorithm, respectively. Section 4 presents experimental results demonstrating the efficacy of our technique using real image data. Finally, Sect. 5 concludes this paper by pointing out extensions to the current work and future research directions.

## 2. Related Work

We review the past studies by two subjects: collaborative filtering and region-based image retrieval.

### 2.1 Collaborative Filtering

CF is an information filtering method that depends on evaluations of human beings. It is defined as one that makes recommendations by finding correlations among users of a recommender system [11].

The typical CF recommendation methods consist of the following three steps [5], [12], [13], [15]: (1) A user provides a system with preference ratings of products that may be used to build a user profile of his or her likes and dislikes. (2) The system applies statistical or machine learning techniques to find a set of users, known as neighbors, who in the past had exhibited behaviors similar to the target user who needs recommendations based on past transactions and product feedback information (i.e., either they had similarly rated or they had purchased similar set of products). A neighborhood is formed based on the degree of similarity between users. (3) Once a neighborhood is formed for a target user, the system predicts whether the target user would like a particular product by calculating a weighted sum of the neighbors' ratings on the product, or it generates a set of products that the target user is most likely to purchase by analyzing the products that neighbors have purchased.

Although CF has proven its success in various areas of its application, it yields lower quality recommendations for images than it does for ordinary items because of visual content's distinct characteristics [7].

### 2.2 Region-Based Image Retrieval

Most early image retrieval systems represent images by a set of global features such as color, texture, and shape, and they perform retrieval based on similarity in the feature space. However, the retrieval accuracy is still far from the user's expectation because of the large gap between the high-level concept and the low-level visual features of images. Furthermore, the similar images that are ranked in terms of a user's high-level concept may not be clearly clustered in the global feature space. To narrow down this gap, region-based local features are widely used. Region-based image retrieval (RBIR) extracts local features from segmented regions of an image, and then images are retrieved according to local similarity among regions. The main objective of RBIR is to perform a more meaningful search that is closer to a user's perception with respect to an image's content. Instead of looking at an image as a whole, we look at the objects in the image and their relationships.

Most of the existing region-based image retrieval approaches can be classified based on the definition of the image similarity measure. The former uses region-to-region similarity and users are required to select one or several regions from the query image under the query-by-example scheme. In the case of the Blobworld, for example, each region of an image is a blob associated with color and texture features [3]. Users are forced to specify the attributes of some specific regions as the query, rather than specifying a description of the entire image. Then, the system responds with images having regions that are similar to the query regions. Most early systems adopting region-to-region similarity tend to partition one object into several regions, none of which is representative for the semantic object because it is difficult to achieve automatic and precise object-level segmentation. Consequently, it is often difficult for users to determine which regions should be used for retrieval. As a result, such systems require the users to perform a substantial amount of work to achieve good performance.

To reduce the influence of inaccurate segmentation and to relieve the users from puzzling decisions, the latter systems adopting an image-to-image similarity measure provide a more simpler querying interface and use information from the whole image for the query. Such systems only require the users to assign a query image without having to specify the attributes of some regions as the query. For example, A. Natsev et al. proposed the WALRUS system using image-to-image similarity, which is defined in terms of the fraction of the area covered by matching regions of the two images [10]. In this measure, one region of an image can match only one region of another image. J.Z Wang proposed integrated region matching (IRM) as the image similarity measure of the SIMPLIcity system [18]. IRM can reduce the influence of inaccurate segmentation since it allows many-to-many matching of the regions.

## 3. Proposed Methodology

### 3.1 Overall Procedure of Feature-Based Collaborative Filtering

In this paper, we propose a new image recommendation method based on the visual characteristics of images when the image purchase sequence and buying history of a target user during the past $T$ periods are given. For solving the above problem, our recommendation procedure is divided into two components: a model building phase and a recommendation phase. Figure 1 presents the overall procedure. The model building phase is performed once by the periodic time unit to create a reliable model from the user transaction database, while the recommendation phase is used to recommend images that the target user is most likely to purchase.

The model building phase is divided into the following three steps: image segmentation, visual feature extraction, and clustering of the purchased images in the feature space using the user transaction database. First, image segmentation is conducted on all images in the database. An image is segmented into several meaningful regions. In this paper, we find regions corresponding to objects in an image by using the modified version of normalized cut segmentation algorithm [1]. Second, we extract various visual features such as color, texture, and shape from regions of images [8]. Since localized features based on regions can represent objects well, they catch a user's high level concept better than do global features extracted from whole pictures of images. Finally, by analyzing the user transaction database, we can group images that are purchased by the same user into the same cluster in the feature space and construct a user profile using the clustered data. Since an image is decomposed into several regions and each region is represented as a point in the feature space, an image purchased by a certain user is represented as multiple points in the user's cluster in the feature space. We construct feature clusters by grouping them with respect to each user. A user profile is represented as a set of feature clusters reflecting the preference between users.

A typical CF approach uses a user-item binary matrix as a user profile, while the proposed FBCF approach uses feature clusters in the feature space (which represent images purchased by users) as a user profile. The preference between users can be measured by using an inter-cluster distance function in the feature space.

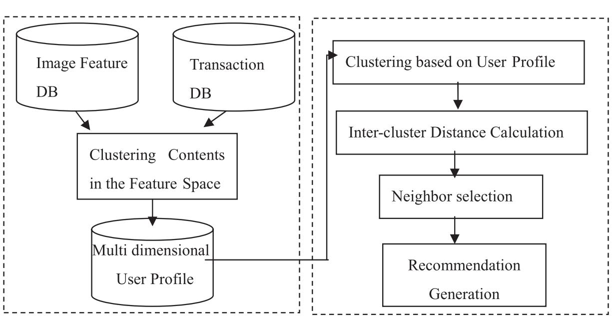


**Fig. 1** Overall procedure of the proposed method.

The recommendation phase searches neighbors of the target user using a set of feature clusters created from the model building phase. A $k$-nearest neighbor search using the user profile is conducted to find neighbors whose feature clusters are close to the cluster of the target user. Finally, the system recommends images having the top-$l$th purchasing ratio among images in clusters of the neighborhood and new images covered by the cluster radius of the target user. That is, the proposed method has an advantage of recommending similar images using visual characteristics.

### 3.2 Model Building Phase

In this section, we create a new user profile based on the feature space by representing visual characteristics of images purchased by users as feature vectors. A user profile, which is a key component of the FBCF algorithm, represents information of the user's preferences about images as a set of clusters in the multi-dimensional feature space. Creating the user profile consists of (a) a process of segmenting an image and extracting features from its regions and (b) a process of representing images purchased by users as a set of feature clusters.

#### 3.2.1 Image Segmentation Process

The normalized cuts method is a grouping criterion that aims to partition a set of points into coherent subsets, originally developed by the Berkeley segmentation research group [16]. It follows a graph theoretic approach to partition a point set. Given a similarity measure between each pair of points in the set, it tries to group together points that have a higher affinity between each other. This criterion has been applied to the domain of image segmentation. A possible approach is to treat each pixel in an image as a point in some arbitrary feature space and group together those pixels that are very similar to each other in terms of the chosen contour, texture, and color features. We used a modified version of a normalized cuts segmentation algorithm to achieve better grouping of regions in natural images [1].

#### 3.2.2 Image Feature Extraction Process

In the image segmentation process, an image is partitioned into a variable number of regions. We extract color, texture, and shape features from each individual region to characterize the objects implied by those regions, respectively. The color feature of a region is encoded by computing the mean and standard deviation of the $rgS$ color space. The $rgS$ color values are obtained from $RGB$ values as $S = R + G + B$ and $g = G/S$. Hence, the color of a region is represented by a set of 6 numbers. The shape of a region is instituted by region size, region location, second moment, and compactness. The second moment is a standard deviation of region

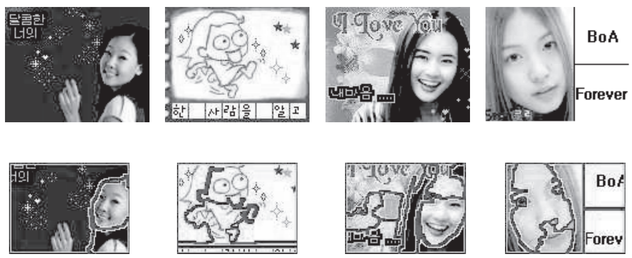


**Fig. 2** Wallpaper images and their regions using image segmentation.

pixels from the region center of mass. The compactness of a region is given by the ratio of its area to the square of its outer boundary length. Therefore, $Compactness = A/P^2$ where $A$ is the area of the region and $P$ is its perimeter. The texture of a region represents the variation of intensity patterns within a region. To extract texture information of a region, the responses to a filter bank at different scales and orientations at all the pixels within the region are averaged separately.

Figure 2 shows the original images provided from a wallpaper image download service at SKT and their segmented regions using the normalized cuts segmentation method.

### 3.2.3 Dynamic User Profile Creation

The objective of constructing a user profile is to find the neighborhood more efficiently. In contrast to CF, clustering user transactions based on the multi-dimensional feature space does not require explicit rating or interaction with users. A user profile $P_a^T$ for user $a$ is generated using feature cluster(s) of images purchased by the user $a$ during $T$ periods. It can be represented as one or several clusters according to user $a$'s preference. The more user $a$'s preference diverges, the more the number of clusters of user $a$ is increased. The user profile includes the centers and the variances of clusters and a buyer's information.

**Definition 1:** If image $x$ was purchased during $T$ periods, it could be given a rating $R_x^T$.

$$R_x^T = \begin{cases} n, \text{ the number of users purchasing} \\ \quad \text{image } x \text{ during } T \text{ period} \\ 0, \text{ otherwise} \end{cases} \tag{1}$$

A rearrangement of feature clusters by means of the user and the time interval is necessary for identifying the dynamic behavior of each user. A feature cluster can be represented as a hyper-ellipsoid in the feature space and its mean vector determines the location of the hyper-ellipsoid, while its covariance matrix characterizes its shape and orientation. The region weights of an image are determined by region area and the mean vector of the region and the covariance of each cluster is calculated using the region weights and region feature vectors. Therefore, the characteristics of a feature cluster depend on the regions whose importances are high in the cluster.

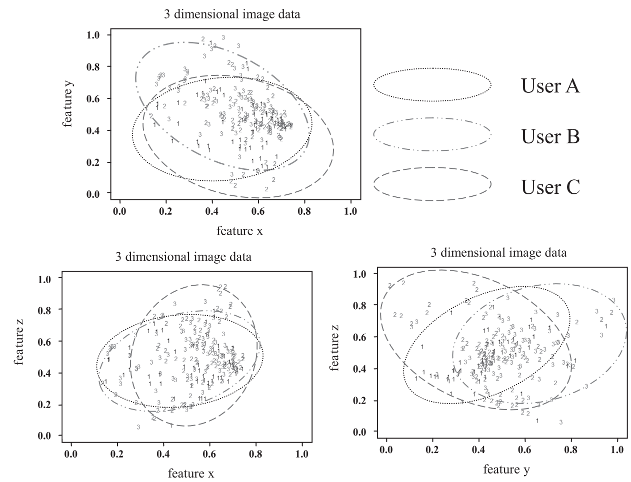


**Fig. 3** A set of images preferred by three users.

## 3.3 Image Recommendation Phase

In this phase, the proposed user profile is used to find neighbors and top-$l$ images are recommended according to the purchase likeness score values.

### 3.3.1 Neighborhood Formation

Existing CF methods calculate the correlation between users using a cosine function or a Person Coefficient. However, when using these methods, it is difficult to find neighbors whose preferences are similar to that of the target user. Existing CF methods recommend items in accordance with a correlation that is calculated by using the buyer information and the web log history while the proposed method can recommend images having similar visual characteristics since it can represent images as points in the feature space and their distance becomes small if they have similar characteristics.

Figure 3 represents a set of images preferred by user $A$, user $B$, and user $C$, respectively, in a three dimensional feature space. Regions of all images in the user transaction database can be represented as points in the feature space. The set of images purchased by each user forms one or several feature clusters. As shown in Fig. 3, a set of images purchased by user $A$ includes eight images and consists of fifty four regions. Among them, the number of common images purchased by both user $A$ and user $B$ is three, that purchased by both user $A$ and user $C$ is two, and that purchased by user $A$, user $B$, and user $C$ is two. As an image can be represented as points in the multi-dimensional feature space, we can find real neighbors by using the distance between the target user and the other user. Since the set of images purchased by the same user can be represented as clusters in the feature space, we use the inter-cluster distance to calculate the similarity between the target user and the other user.

We assume that a user has a preference that is similar to that of the target user if the cluster of the user is close

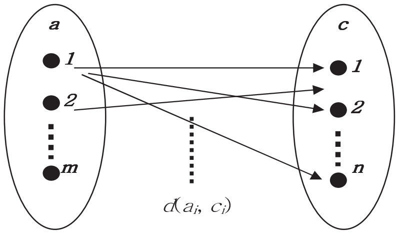


**Fig. 4** Earth Mover's distance.

to that of the target user. There are many inter-cluster distance functions. Among them, the favorite function is the Euclidean distance function. This function is simple and easy to calculate, but it works well in a condition where each cluster is uniformly distributed and its form is a circle. However, the individual user's preference is not the same and their distributions are different from each other as shown in Fig. 3.

In this section, we propose two types of inter-cluster distance functions for neighborhood formation: the former is used when the number of clusters for each user is one; the latter is used when the number of clusters for any user is not less than two.

(Case 1) A profile of target user $c$ forms a feature cluster and those of other users also forms different clusters, respectively. To find a neighbor, a $k$-nearest neighbor search using a set of feature clusters is performed. Basically, the similarity between target user $c$ and other user $a$ is calculated using the inter-cluster distance function $T^2_{ca}$ between cluster $C_c$ and cluster $C_a$. We find $k$ neighbors $H = \{h_1, h_2, \ldots, h_k\}$, $c \in H$ for target user $c$ according to the ascending order of the $T^2_{ca}$ value.

**Definition 2:** The inter-cluster distance function between two clusters $h_i$ and $h_j$ is defined as:

$$T^2_{ij} = \frac{n_i n_j (n-2)}{(n_i + n_j)^2} (\bar{x}_i - \bar{x}_j)^T S^{-1}_{p_{ij}} (\bar{x}_i - \bar{x}_j) \qquad (2)$$

where $n_i$ is the number of regions in the $i$th cluster, $n_j$ is the number of regions in the $j$th cluster, $\bar{x}_i$ is the centroid of the $i$th cluster, $\bar{x}_j$ is the centroid of the $j$th cluster, respectively, and $S_{p_{ij}}$ is the pooled covariance matrix.

(Case 2) A profile of target user $c$ forms several feature clusters and those of other users also forms more than two clusters, respectively.

In order to measure the distance between two users, the Earth Mover's distance function [4] is used to measure feature clusters of target user $c$ and those of other user $a$ as shown in Fig. 4. The weights $w_{c_i}$ and $w_{a_j}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ are calculated using the number of images in each cluster. $T^2_{c_i a_j}$ is used as the ground distance between cluster $c_i$ and $a_j$. See the detailed description of the EMD measure in the Appendix.

#### 3.3.2 Recommendation List Generation

As a final step for image recommendation, we generate a list of $l$ images, $R = \{r_1, r_2, \ldots, r_l\}$ such that $r_j \notin$ {the images that target user $c$ has already purchased} and $PLS(c, r_1)$ is the highest $PLS$, $PLS(c, r_2)$ is the next highest and so on. $PLS(c, x)$ denotes the purchase likeliness score of the target user $c$ for image $x$ and is computed as follows:

$$PLS(c, x) = \frac{\sum_{a \in H} R_x \times sim(c, a)}{\sum_{a \in H} sim(c, a)} \qquad (3)$$

User $a$ is included in the neighborhood set $H$ and $R_x$ is the rating of purchased image $x$. The $sim(c, a)$ function denotes the similarity between target user $c$ and user $a$, and we convert the distance values into similarity values and normalize them as follows:

$$sim(c, a) = \frac{Max_{u,w \in H}[d(u, w)] - d(c, a)}{Max_{u,w \in H}[d(u, w)] - Min_{u,w \in H}[d(u, w)]} \qquad (4)$$

where $u, w$ are users in the target user's neighbor $H$ and $d()$ is a reciprocal function of $T^2_{ij}()$ .

### 3.4 New Image Recommendation

As the multimedia contents market continues its rapid expansion, new images are frequently provided on the Internet and/or the mobile Web and their purchasing ratio becomes very high. A shortcoming of the traditional CF methods is that they cannot recommend a new item even though the purchasing ratio of the new item is very high. The traditional CF methods use common items to find neighbors of the target user. So, they cannot recommend new items that have not been purchased yet because they can not be included in a traditional user profile and they are excluded from the neighborhood formation process.

If a new item has a rating, however, it can be recommended. To solve the new item problem, we propose to grant a virtual rating to a new image. Our basic assumption is that two images $x$ and $y$ have very similar preference rating with respect to a given user if they have been purchased by the same user and they are very close in the feature space.

Let $g$ clusters $C_1, \ldots, C_g$ be the set of images purchased by $g$ users, respectively. We determine the candidate cluster, which includes new image $x_{new}$, using the Bayesian classifier function.

**Definition 3:** The Bayesian classifier for cluster $C_i$ is defined as:

$$\hat{d}_i(x_{new}) = -\frac{1}{2}(x_{new} - \bar{x}_i)' S^{-1}_{pooled}(x_{new} - \bar{x}_i) + ln(w_i) \qquad (5)$$

where $\bar{x}_i$ is the centroid of the $i$th cluster, $w_i$ is a normalized weight of the $i$th cluster, and the weight is calculated using the sum of the user preference rating.

We select cluster $C_k$ which has the largest value among $\hat{d}_1(x_{new}), \hat{d}_2(x_{new}), \ldots, \hat{d}_g(x_{new})$ and verify whether $x_{new}$ is located within the effective radius of the cluster $C_k$ using Eq. (6).

$$(x_{new} - \bar{x}_k)' \left(\frac{1}{n} S\right)^{-1} (x_{new} - \bar{x}_k) < \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha) \qquad (6)$$

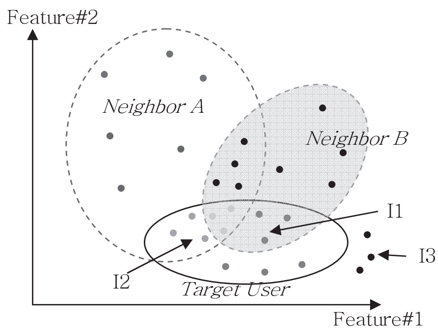


**Fig. 5** A representation of new images $I_1, I_2, I_3$ in the feature space.

Statistically, the effective radius of the cluster follows an $F$ distribution with degree of freedom $p, n - p$ and significance level $\alpha$. That is, if Eq. (6) is satisfied, we can determine cluster $C_k$ as the candidate cluster of the new image $x_{new}$.

The purchasing ratio of the new image will become very high if the candidate cluster is chosen among neighbor clusters or the target user's cluster since each cluster represents each user's preference. If it is, we grant a virtual rating to the new image and calculate the *PLS* value of the new image. Then, we can generate a recommendation list with a set of images purchased by neighbors and the new image. It is reasonable in that the recommendation of the new image follows the same process of image recommendation that was determined by what the neighbor wants to purchase.

For example, we assume that $I_1, I_2, I_3$ are new images. In Fig. 5, they are represented as points in the feature space.

When three feature clusters for neighbors A, B, and the target user are given, if new image $I_2$ is included in the feature cluster preferred by the target user, it can be recommended. As shown in Fig. 5, $I_1$ and $I_2$ can be recommended, respectively, and $I_2$ has a virtual rating of 1.0 since an image close to $I_2$ has the target user's preference rating 1.0.

### 3.5 Proposed Algorithm

To summarize our method, the algorithm is as follows:

**Algorithm** Image Recommendation using FBCF
**input:** Region feature DB $FDB$, transaction database, user profile $P$
**output:** recommended image list $R$
**begin**
**step:** Model Building
**(1)** Carry out region segmentation for all the images.
**(2)** Extract local features from each region of each image.
**(3)** Store them into feature database $FDB$.
**(4)** Construct a set of feature clusters by grouping images purchased by each user during certain periods.
**(4)** Store the center, variance, effective radius, user information, and feature vectors of the images into a user profile. That is, the user profile is represented as the multi-dimensional feature clusters.
**step:** Neighborhood formation
**(5)** Compute the inter-cluster distance function $T_{ca}^2$ between the cluster of target user $c$ and the cluster of any other user $a$.
**(6)** Rank the users in ascending order of $T_{ca}^2$ values.
**(7)** Select $k$ neighbors $H = \{h_1, h_2, \ldots, h_k\}$, $c \notin H$ for target customer $c$.
**step:** Recommendation list generation
**(8)** Generate a list of top $l$ contents, $R = \{r'_1, r'_2, \ldots, r'_l\}$, such that $r'_j \notin$ {the contents that c has already purchased} and $PLS(c, r'_1)$ is the highest PLS, $PLS(c, r'_2)$ is the next highest, and so on.
**(9)** If target user $c$ purchases one of the recommended images, update the transaction database.
**(10)** Repeat steps (5) to (9) for several times
**step**: New item recommendation
**(11)** When a neighboorhood $H = \{h_1, h_2, \ldots, h_k\}$ of target user $c$ is given, determine the target cluster including new item $x_{new}$.
**(12)** Add the new item into recommendation list if the new item is located in the effective radius of the target cluster.
**end**

## 4. Experimental Results

We conducted experiments to answer the following questions:

- How does the proposed method perform, compared to the traditional collaborative filtering scheme?
- Can the proposed method outperform the traditional content-based filtering scheme?
- How effective is the proposed method in terms of new image recommendation?

### 4.1 Experimental Environment

For the experiments, we used real-world data to examine the performance of the proposed approach. The data used in the experiment were 25,680 transaction records and the 5,300 wallpaper images that SK Telecom(SKT), a leading Korean CDMA (code division multiple access) carrier, offered at the time of our experiment. The transaction records include purchase sequences for the wallpaper images sold by a mobile Web site during the three-month period from June to August 2004 in order to establish the behavioral characteristics of the users over time. The input data from the SKT purchase database consist of 25,680 transactions and 520 users. Each user have purchased an average of 55.6 images. The experiment involved only 476 users who had purchased more than 40 wallpaper images from SKT. The participants selected as suitable to receive recommendations were restricted active users who had purchased images frequently because it is

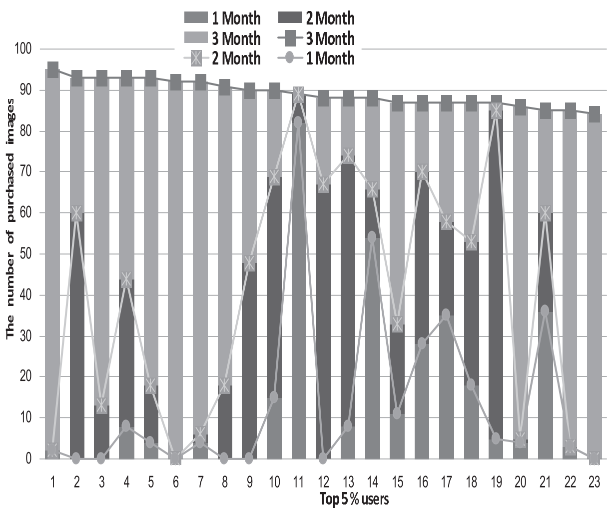


**Fig. 6** Distribution of images purchased by the top 5% of users.

difficult to identify the dynamic purchase behavior of users who purchase images rarely. Figure 6 shows the distribution of images purchased by the top 5 % of users during 1 month, 2 month, and 3 month periods, respectively. It illustrates that the behavior of users changed dynamically. The number of recommendations was fixed as eighteen. To characterize images, we used various kinds of visual features extracted from individual regions of the image. The feature data consist of a six dimensional color feature, a six dimensional shape feature, and an eight dimensional texture feature as mentioned in Sect. 3.2.2.

To evaluate the quality of the recommendation method, we devised the hit ratio metric $h_i$, which is defined as the ratio of the number of purchased and recommended images to the number of recommended images at iteration $i$. Then, the average hit ratio is defined as follows:

$$Average\_Hit\_Ratio = \frac{\sum_{i=1}^{n} h_i}{n} \tag{7}$$

where $n$ is the iteration number of the recommendation and $h_i$ is the hit ratio at iteration $i$. Basically, the hit ratio measures the user's effort for a successful search.

We used the hit ratio to compare the proposed FBCF method with a traditional CF-based recommender system (pure-CF) and a typical content-based filtering system (CB). The strategy of the pure-CF system is identical to that of FBCF, except that pure-CF uses the correlation between the target user and his or her neighbors to evaluate how many of the ratings on common images are from the target user's neighbors, whereas the strategy of CB system is different from that of FBCF since CB selects items based on the correlation between the content of the items and the user's preferences.

To evaluate the performance of the proposed method in terms of new image recommendations, we devised the new item ratio metric $hnew_i$, which is defined as the ratio of the number of recommended new images to the number of recommended images at iteration $i$. Then, the average new

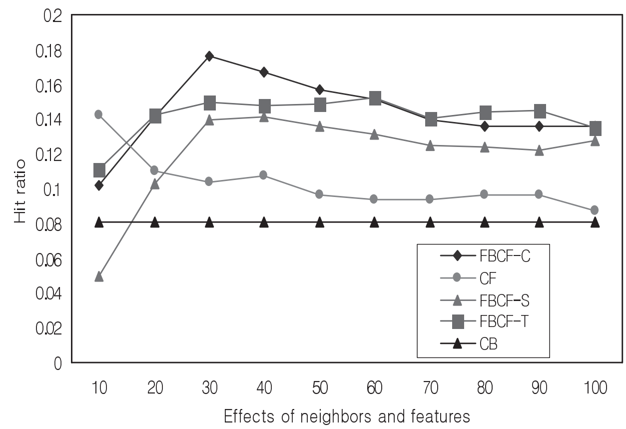


**Fig. 7** Effects of neighbors and features.

item ratio is defined as follows:

$$Average\_NewItem\_ratio = \frac{\sum_{i=1}^{n} hnew_i}{n} \tag{8}$$

where $n$ is the iteration number of recommendation for new images and $hnew_i$ is the new item ratio at iteration $i$.

## 4.2 Results and Discussion

In the case of FBCF, participants performed the experiment for color, shape, and texture features to see how changes in features affect the overall performance. We also performed an experiment to determine the optimal neighborhood's size because the quality of CF recommendations varies according to the neighborhood's size.

Figure 7 shows the performance of FBCF and pure-CF with respect to various neighborhood's size when color, shape, and texture features are used. FBCF-C is a feature based CF recommendation procedure using a color feature, FBCF-S is that using a shape feature and FBCF-T is that using a texture feature. The experimental results show that FBCF-C yields better performance than the other three methods and it yields the best performance when the neighborhood size is 30. We averaged the hit ratio whose values are obtained according to ten different neighborhood sizes. The average hit ratio of FBCF-C is approximately 40.3% higher than that of pure-CF while the average hit ratio of FBCF-T is 37.8% higher than that of pure-CF and the average hit ratio of FBCF-S is 16.9% higher than that of pure-CF whereas that of CB is 21% lower than that of pure-CF.

To evaluate the system's performance change with respect to the training period, six training periods were used: June 01 - June 14, June 01 - June 29, June 01 - July 14, June 01 - July 30, June 01 - Aug 14, June 01 - Aug 30. We measured the average hit ratio per each period, respectively. This was to see how the system's performance changes over time.

According to the ANOVA results, the variation in hit ratio over the six periods (that is, the period effect) is significant. ($F = 17.0, p < 0.05$). The $F$ statistic provides a test for the statistical significance of the observed FBCF

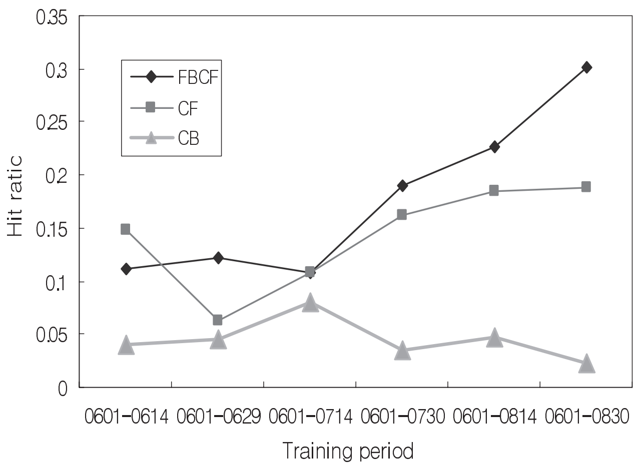


**Fig. 8** Effects of training periods.

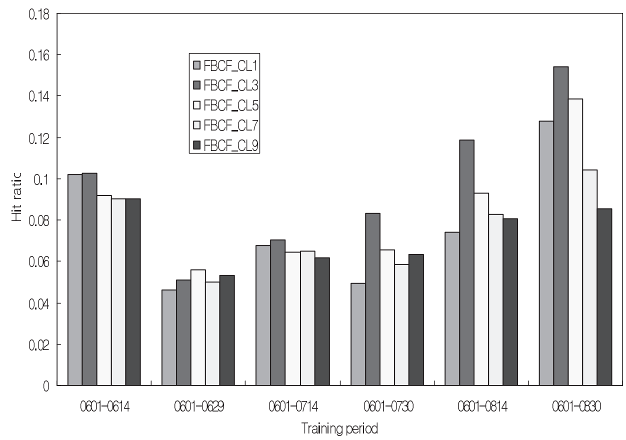


**Fig. 9** Effects of the number of clusters.

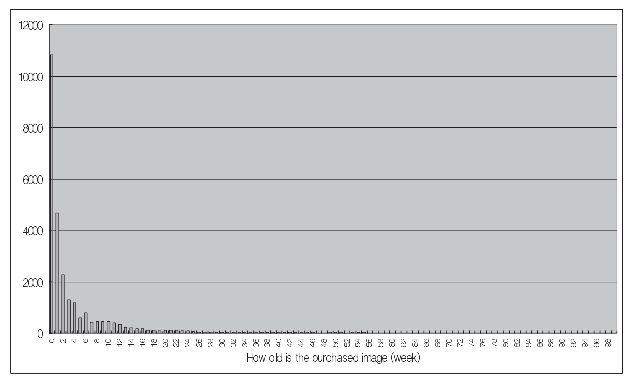


**Fig. 10** Statistics of purchased images.

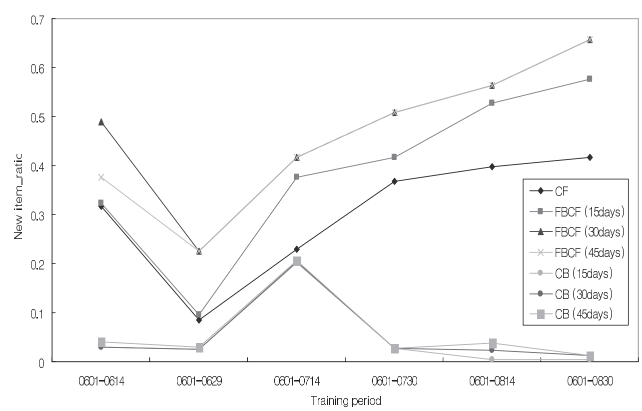


**Fig. 11** Effects of new items.

performance differences with respect to periods.

Figure 8 illustrates the variation as an increasing curve. As periods progressed, more rating information became available. Neighborhood formation becomes more accurate as the user profile contains more ratings and thereby the quality of CF recommendations is improved, whereas the quality of CB recommendations does not change as periods progress. This is because CB recommends items based only on analysis of its content rather than correlation between people with similar preferences. For the June 01-June 14 training period, FBCF's performance is lower than that of pure-CF because the rating information of FBCF is not sufficient to calculate the inter-cluster distance between user pairs. However, the average hit ratio of FBCF is about 24% higher than that of pure-CF and 290% higher than that of CB at a significance level of 1 percent in the June 01-Aug 30 training period. The result shows that FBCF can reflect a users' recent preference by computing the correlation between users with similar preference in the multi-dimensional visual feature space.

Figure 8 also shows that the rate of improvements of FBCF in average hit ratio over six periods is 168.9% whereas that of pure-CF is 27.4% and that of CB is -43.3%.

To evaluate the performance change over the number of clusters, the number of clusters used in this experiment is 1, 3, 5, 7, and 9. For example, FBCF_CL3 denotes feature based CF when the number of feature clusters for each user is less than or equal to three. Figure 9 illustrates that the performance of FBCF_CL3 is better than that of any other method.

For 25,680 transactions, we investigated the sales pattern of wallpaper images. Figure 10 shows that 67% of all target images are purchased within 2 weeks from the beginning and 22% of the remaining images are purchased within 4 weeks. In the next experiment, we consider items that are put on the market within 15 days or 30 days as new items, respectively.

Figure 11 shows the performance of the new item recommendation of both FBCF and pure-CF with respect to various training periods. In "FBCF (15days)", a new item is defined as an image put on the market within 15 days. In "FBCF (30days)" and "FBCF (45 days)", new items are defined as images put on the market within 30 days and 45 days, respectively. The new item ratio of "FBCF (15days)" is about 37.9% higher than that of pure-CF and those of "FBCF (30days)" and "FBCF (45days)" are about 57.5% higher than that of pure-CF when the training period is June 01- Aug 30, whereas the new item ratios of CB(15days), CB(30days), and CB(45days) are about 97% lower than that of pure-CF.

## 5. Conclusion

In this paper, we proposed a new FBCF recommendation algorithm using a real image database. FBCF reflects the user's preference using a feature-based user profile in a multi-dimensional vector space. The objective is not only to improve the performance of the recommendation for image contents, but also to enable the recommendation of new images because new image recommendation is an important issue for the multimedia contents market. We conducted a series of experiments using a system prototype and had very encouraging results.

For future research, we shall further improve image recommendation system performance using user interaction mechanism. We will also extend our method to different types of multimedia contents such as music and video. Successful application of the FBCF approach to these types of content will require research on the proper interfaces for contents of different types.

## Acknowledgements

This work was supported by Inha University Research Grant.

### References

[1] K. Barnard, P. Duygulu, R. Guru, P. Gabbur, and D. Forsyth, "The effects of segmentation and feature choice in a translation model of object recognition," Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), vol.2, pp.675–682, 2003.
[2] M. Balabanovic and Y. Shoham, "Fab: Content-based collaborative recommendation," Commun. ACM, vol.40, no.3, pp.66–72, 1997.
[3] C. Carson, M .Thomas, S. Belongie, J. Hellerstein, and J. Malik, "Blobworld: A system for region-based image indexing and retrieval," Proc. 3rd Intl. Conference on Visual Information Systems, 1999.
[4] H. Greenspan, G. Dvir, and Y. Rubner, "Region correspondence for image matching via EMD flow," IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVIL' 00), 2000.
[5] W. Hill, L. Stead, M. Rosenstein, and G. Furnas, "Recommending and evaluating choices in a communication of use," Proc. CHI 95, 1995.
[6] H. Hotelling, "Multivariate quality control," in Techniques of Statistical Analysis, ed. C. Eisenhart, M.W. Hastay, and W.A. Wallis, McGraw-Hill, N.Y., 1947.
[7] C.Y. Kim, J.K. Lee, Y.H. Cho, and D.-H. Kim, "VISCORS: A visual content recommender for the mobile Web," IEEE Intelligent Systems, pp.32–39, Nov./Dec. 2004.
[8] D.-H. Kim and S.-H. Yu, "A new region filtering and region weighting approach to relevance feedback in content-based image retrieval," J. System and Software, vol.81, no.9, pp.1525–1538, Sept. 2008.
[9] P. Melville, R.J. Mooney, and R. Nagarajan, "Content-boosted collaborative filtering for improved recommendations," Proc. 18th Nat'l Conf. Artificial intelligence(AAAI02), pp.187–192, AAAI Press, 2002.
[10] A. Natsev, R. Rastogu, and K. Shim, "WARLUS: A similarity retrieval algorithm for image databases," Proc. SIGMOD '99, pp.395–406, Philadelphia, PA, 1999.
[11] M.J. Passani, "A framework for collaborative, content-based and demographic filtering," Artifical Intelligence Review, 1999.
[12] P. Resnick, N. Iacovou, P. Bergstrom, and J. Riedl, "Grouplens: An open architecture for collaborative filtering of netnews," Proc. ACM 1994 Conference on Computer Supported Cooperative Work, pp.175–186, 1994.
[13] B. Sawar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for E-commerce," Proc. ACM Conf. Electronic commerce 2000, pp.158–167, ACM Press, 2000.
[14] B. Sawar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithm," Proc. Tenth International World Wide Web Conference, pp.285–295, 2001.
[15] U. Shardanand and P. Maes, "Social information filtering: Algorithms for automating 'world of mouth'," Proc. 1995 ACM/SIGCHI Conf. Human Factors in Computing Systems (CHI 95), pp.210–217, Addison-Wesley, 1995.
[16] J. Shi and J. Malik, "Normalized cuts and image segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.8, pp.888–905, Aug. 2000.
[17] J.Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive integrated matching for picture libraries," IEEE Trans. Pattern Anal. Mach. Intell., vol.23, no.9, pp.947–963, 2001.
[18] X.S. Zhou and T.S. Huang, "Relevance feedback for image retrieval: A comprehensive review," ACM Multimedia Systems Journal, vol.8, no.6, pp.536–544, 2003.

## Appendix: The Earth Mover's Distance Measure

$C = (c_1, w_{c_1}), \ldots, (c_m, w_{c_m})$ is the first signature with $m$ clusters of target user $c$, where $c_i$ is the cluster representative and $w_{c_i}$ is the weight of the cluster, and $A = (a_1, w_{a_1}), ..., (a_n, w_{a_n})$ is the second signature with $n$ clusters of other user $a$. When $T^2_{c_i a_j}$ is the ground distance between clusters $c_i$ and $a_j$, the earth mover's distance is defined as the work normalized by the total flow:

$$EMD(C, A) = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} T^2_{c_i a_j}}{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}}$$

subject to following constraints:

$$f_{ij} \geq 0, 1 \leq i \leq m, 1 \leq j \leq n$$

$$\sum_{j=1}^{n} f_{ij} \leq w_{c_i}, 1 \leq i \leq m$$

$$\sum_{i=1}^{m} f_{ij} \leq w_{a_j}, 1 \leq j \leq n$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = min(\sum_{i=1}^{m} w_{c_i}, \sum_{j=1}^{n} w_{a_j})$$

**Deok-Hwan Kim** was born in Repulic of Korea in 1964. He received M.S. and PhD from Korea Advanced Institute of Science and Technology. He is an associate professor of Inha University in Korea. His research interests includes embedded system software, multimedia information retrieval, and data mining.