A Fuzzy-Based Multimedia Content Retrieval Method Using Mood Tags and Their Synonyms in Social Networks

Chang Bae Moon[1], Jong Yeol Lee[2], Byeong Man Kim[2] *

[1] ICT-Convergence Research Center, Kumoh National Institute of Technology (Korea)
[2] Computer Software Engineering, Kumoh National Institute of Technology (Korea)

Received 4 March 2021 | Accepted 12 September 2021 | Published 3 October 2022

Abstract

The preferences of Web information purchasers are rapidly evolving. Cost-effectiveness is now becoming less regarded than cost-satisfaction, which emphasizes the purchaser's psychological satisfaction. One method to improve a user's cost-satisfaction in multimedia content retrieval is to utilize the mood inherent in multimedia items. An example of applications using this method is SNS (Social Network Services), which is based on folksonomy, but its applications encounter problems due to synonyms. In order to solve the problem of synonyms in our previous study, the mood of multimedia content is represented with arousal and valence (AV) in Thayer's two-dimensional model as its internal tag. Although some problems of synonyms could now be solved, the retrieval performance of the previous study was less than that of a keyword-based method. In this paper, a new method that can solve the synonym problem is proposed, while simultaneously maintaining the same performance as the keyword-based approach. In the proposed method, a mood of multimedia content is represented with a fuzzy set of 12 moods of the Thayer model. For the analysis, the proposed method is compared with two methods, one based on AV value and the other based on keyword. The analysis results demonstrate that the proposed method is superior to the two methods.

Keywords

Cost-Satisfaction, Mood Fuzzy Value, Multimedia Content Mood Tag, Multimedia Content Retrieval, Social Network.

DOI: 10.9781/ijimai.2022.10.005

I. Introduction

According to [1], consumption patterns are changing from cost-effectiveness, which refers to performance with respect to price, to cost-satisfaction, which emphasizes cost-effectiveness along with psychological satisfaction. The multimedia content search method needs to be changed according to this flow. It will not only increase psychological satisfaction through searching by mood or emotion rather than text content, but also provide various additional services. Several studies have been conducted in accordance with this trend. A set of works [2]-[4] define the mood model, another research [5] shows how to search for multimedia content using the mood, and other papers [6], [7] analyze the relationship between mood and color of multimedia content.

Psychological satisfaction is a new trend in modern society as the single-person society has developed recently, such as the Yolo [8] and the Solo [9], which means that consumer's consumption prioritizes consumer's satisfaction over content's price. That is, cost-effectiveness means price satisfaction on multimedia content and the psychological satisfaction means customer satisfaction on its service. In this context, we propose a multimedia content retrieval method using mood to increase user's psychological satisfaction.

There are four main types of multimedia content search methods applicable to SNS (Social Network Services). The first one is query-by-text that retrieves information matching the query based on textual information stored in the multimedia content database. The second one is query-by-part where a part of the content (for example, part of music) is served as a query and content similar to that part is found. The third one is the query-by-example, which accepts the content itself as a query to find similar content. Unlike query-by-part, the entire content is entered as a query rather than a part. The last fourth one is query-by-class, which retrieves the content of a given class as input using predefined genres or classes.

Among the four types, query-by-class is a common search method, but requires expert or curator participation. That is, when new multimedia content is provided, class information must be input. In a situation where new multimedia content is constantly and rapidly generated, this method is difficult to respond quickly. A possible solution is to categorize contents using tags according to the taxonomy, enter bibliographic information about the class, or assign a class code. However, this method still requires human curator intervention. In addition, if the knowledge of a specific topic is insufficient, the classification system cannot be expanded, and one or more classes can be assigned to some content, so separate processing is required.

The folksonomy is mentioned as an alternative to the taxonomy-based classification system. It has a flat structure and requires users to participate in lieu of professional management of a librarian or curator [10]. Famous SNS like Instagram, Facebook, Last.fm, IPTV and YouTube use this classification to provide services. The biggest advantage of the folksonomy is that it avoids expandability and monopoly, which is a problem with the taxonomy-based classification systems. However, the folksonomy has a synonymous problem between tags used to describe content. That is, it is necessary to identify similar tags as well

* Corresponding author.

E-mail addresses: cb.moon@kumoh.ac.kr (C. B. Moon), soyeum@kumoh.ac.kr (J. Y. Lee), bmkim@kumoh.ac.kr (B. M. Kim).

as content including query tags and recommend content with similar tags. Also, if you provide a new word as a query tag, the content is not searched. In this case, you need to identify new words and recommend related content using existing tags associated with the new words. To solve this, a method that considers analogues based on AV is proposed in the study of [10]. Unfortunately, this method achieves performance that is inferior to a method based on keywords at recall level 0.1. Users use search engines to get the information they want, and they usually tend to see only the first few pages of search results. Therefore, it is important to provide a high-accuracy search at low recall levels. In this context, this paper attempts to improve the precision at low recall levels in [10].

In this paper, a method is proposed that supplements the problem of SNS service and can approach the retrieval performance of the keyword-based approach at recall level 0.1. To achieve this, the mood of a multimedia content is represented with a fuzzy set of 12 moods of the Thayer model. Then, the association between the fuzzy vectors of contents and the fuzzy vectors of tags of contents is analyzed. In addition, for the performance comparison, the proposed method is compared with two methods, one based on AV value [10] and the other based on keyword (Last.fm approach). Additional experiments are also carried out for four different similarity measures between two fuzzy vectors [11], [12], [13], [14]. Although the proposed method in this paper is applicable to all types of multimedia content if they only include annotation tags, it is preferentially applied to music content like the previous study for performance analysis.

II. Relevant Works

Among the methods used to define the mood, there are Russell [2] and Hevner [3]. These methods may cause semantically overlapping or ambiguous problems due to the adjective expression. A model to solve this ambiguity problem is the model of Thayer [4]. In this model, 12 mood words are expressed as two-dimensional vectors of Arousal and Valence, and the mood of multimedia content is expressed using these mood vectors. Arousal represents the intensity of stimuli felt by multimedia consumers, and Valence represents the stability experienced by multimedia consumers (Fig. 1) [10], [15]. Among the multimedia content retrieval methods, some [5]-[7] used this model.

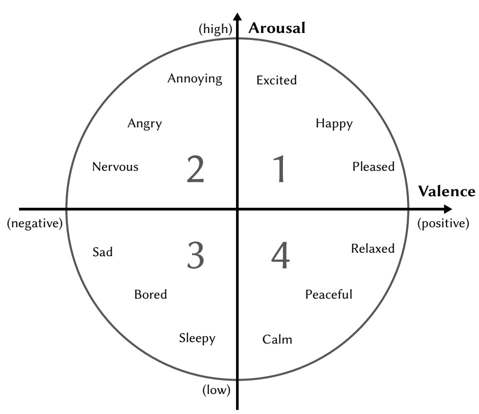


Fig. 1. Thayer's two-dimensional model.

By use of Thayer's mood vector (i.e., AV vector), Moon et al. [5] tried to solve the problems of music folksonomy-problems of tagging level (for example, happiness, very happy, too happy, etc.), synonym problems, and new words. In this study, in order to express music as AV vector values, subjects were asked to enter the Arousal and Valence values for the music they heard after listening to the music. So, both music and tags were expressed as AV values, so music with synonym tags could be searched. That is, the problem of synonyms in folksonomy could be solved naturally. However, this approach was difficult to apply to large amounts of data because the subjects had to assign a two-dimensional mood vector to each piece of music.

Other works [16]-[18] used SVM (Support Vector Machine), one of the machine learning methods, to grasp the mood of a new music piece. The learning data is composed of inputs and the corresponding output values. In these studies, the acoustic and musical characteristics of a music piece are used as input values, and the mood values are used as output values. For the mood values of music pieces, folksonomy mood tags provided by Last.fm were used. Some authors [19]-[21] studied the automatic tagging method for web content. In order to easily input the tags of web contents, tags related to web contents are presented and related tags can be selected among them. Others [22]–[27] conducted an automatic tagging study on images. After learning the association between the properties included in the content and the image tags using SVM, Bayes Point Machines [25], etc., the tags for the new image were automatically added using this.

Among existing AV-based retrieval methods to solve the problem of synonyms, there is one that [10] obtains mood vectors (strength of arousal and valence) of music from social and international mood data which were built for several years, and defines AV values of tags in order to make synonym-based search possible. The study, however, reveals that retrieval performance is worse than that of a keyword-based approach at recall level 0.1. In this paper, we propose a retrieval method that uses fuzzy-based multimedia mood which achieves the same retrieval performance as a keyword-based approach at recall level 0.1. In addition, the excellence of the proposed method is demonstrated as it was comparatively analyzed with the AV-based retrieval method of [10] and a keyword-based approach.

There are some works [28], [29], [30] on detecting mood based on fuzzy theory. In a research [28], the music mood detection using the AV model based on a fuzzy theory is proposed, where mood of music is predicted with the music features and the training data to construct the prediction model is given by a few volunteers. In another work [29], a method for detecting mood of a text message using fuzzy rules is proposed. However, in this paper, mood tags of multimedia contents given by participants from all over the world are used to constructing a mood fuzzy vector of the multimedia content.

III. A Fuzzy-Based Multimedia Content Retrieval Method Considering Synonyms

Fig. 2 outlines the proposed multimedia content-retrieval method considering synonyms based on fuzzy vectors. The method comprises the following four phases: (1) the construction of multimedia content

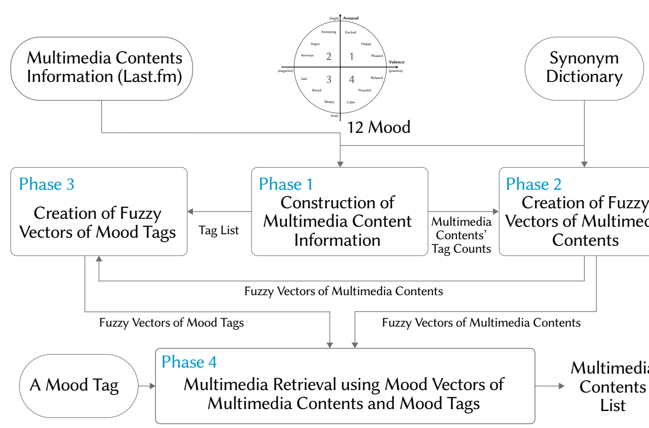


Fig. 2. Multimedia content retrieval structure.

information; (2) the creation of fuzzy vectors of multimedia content; (3) the creation of fuzzy vectors of mood tags; and (4) multimedia content retrieval based on these two fuzzy vectors.

A. Constructing Multimedia Content Information

Fig. 3 presents the process used to construct information for multimedia content divided into two stages: (1) collecting lists of multimedia content on the 12 moods; and (2) collecting tags of multimedia content and tag counts. Although the proposed method can be applied to all types of content, this study considers only music from Last.fm. Since the data collection method is the same as the method of Moon et al. [10], please refer to their work [10] for details.

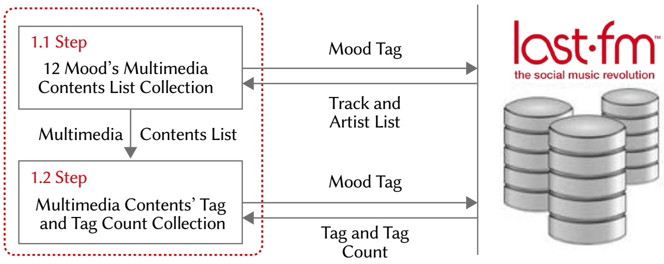


Fig. 3. Multimedia content information construction process.

B. Generating Fuzzy Vectors for Multimedia Content

The process employed to construct fuzzy vectors for multimedia content involves three stages: (1) collecting and analyzing synonyms; (2) calculating mood fuzzy vectors for the multimedia content; and (3) calculating fuzzy vectors of the multimedia content. To classify multimedia content tags into the 12 moods in Fig. 1 to construct a synonym mapping table, the outcome of a previous synonym analysis [1], [10] is used (see Table I).

For generating multimedia content fuzzy vectors, the count values of the 12 moods, called by mood vector, for multimedia content can be calculated using (1) by using the synonym mapping table and relevant tag information as follows:

$$V_k = (M^{1,k}, M^{2,k}, \ldots, M^{11,k}, M^{12,k})$$
$$M^{m,k} = \sum_{i=1}^{n} t_i^{m,k}$$
$$t_i^{m,k} = \begin{cases} c_i^k, & \text{if } Tag_i^k \in S^m \\ 0, & otherwise \end{cases} \quad (1)$$

where k is the index of the multimedia item, m the mood index ($1 \le m \le 12$), n the number of tags attached to multimedia item k, i the tag index ($1 \le i \le n$), Tag_i^k the i'th tag of multimedia item k, S^m the synonym list (in Table I) corresponding to mood m, and c_i^k the tag count of tag Tag_i^k; $t_i^{m,k}$ is the tag count of tag Tag_i^k for mood m, $M^{m,k}$ the count value of mood m for multimedia item k, and V_k the mood vector of multimedia item k.

For example, assume the tags and counts for the m' th multimedia item in Table II.

TABLE II. Tag and Count

Tag	Count	Tag	Count
annoy	30	**bothersome**	10
angered	15	**Tense**	10
anxious	7	**Excite**	15

TABLE I. Synonym Mapping Table

12 moods	Part	Synonym mood list
annoying	**verb**	annoy, rag, get to, bother, get at, irritate, rile, nark, nettle, gravel, vex, chafe, devil, displease
	adj	bothersome, galling, irritating, nettlesome, pesky, pestering, pestiferous, plaguy, plaguey, teasing, vexatious, vexing, disagreeable
angry	**adj**	aggravated, provoked, angered, enraged, furious, infuriated, maddened, black, choleric, irascible, hot under the collar(predicate), huffy, sore, indignant, incensed, outraged, umbrageous, irate, ireful, livid, smoldering, smouldering, wrathful, wroth, wrothful, raging, tempestuous, stormy, unhealthy, mad, wild
nervous	**adj**	tense, anxious, queasy, uneasy, unquiet, troubled, neural, skittish, flighty, spooky, excitable, aflutter, excited
sad	**adj**	bittersweet, doleful, mournful, heavyhearted, melancholy, melancholic, pensive, wistful, tragic, tragical, tragicomic, tragicomical sad, sorrowful, deplorable, distressing, lamentable, pitiful, sorry, bad
bored	**verb**	bore, tire, drill, cut
	adj	world-weary, tired, blase, uninterested
sleepy	**adj**	sleepy-eyed, sleepyheaded, asleep(predicate)
calm	**verb**	calm down, tranquilize, tranquillise, quieten, lull, comfort, soothe, console, solace, steady, becalm, stabilize, stabilise, cool off, chill out, simmer down, settle down, cool it, sedate, tranquillize, affect, change state, turn
	adj	unagitated, serene, tranquil, composed, placid, quiet, still, smooth, unruffled, settled, windless
peaceful	**adj**	peaceable, irenic, nonbelligerent, pacific, pacifist(prenominal), pacifistic, dovish, peace-loving, halcyon
relaxed	**verb**	relax, loosen, loosen up, unbend, unwind, decompress, slow down, loose, weaken, unstrain, unlax, make relaxed, act, behave, do, slack, slacken, slack up, decrease, lessen, minify, change, alter, modify, affect, change state, turn
	adj	degage, laid-back, mellow, unstrained
pleased	**verb**	please, delight, satisfy, gratify, wish, care, like
	adj	amused, diverted, entertained, bucked up(predicate), encouraged, chuffed, delighted, gratified, proud of(predicate), proud
happy	**adj**	blessed, blissful, bright, golden, prosperous, laughing(prenominal), riant, felicitous, fortunate, glad, willing, well-chosen, halcyon
excited	**verb**	excite, arouse, elicit, enkindle, kindle, evoke, fire, raise, provoke, stimulate, impact, bear upon, bear on, touch on, touch, stir, sensitize, sensitise, agitate, rouse, turn on, charge, commove, charge up, disturb, upset, trouble, sex, wind up, shake, shake up, energize, energise, change, alter, modify, affect
	adj	aroused, emotional, worked up, agitated, agog, crazy, fevered, intoxicated, drunk, overexcited, stimulated, stirred, stirred up, teased, titillated, thrilled, thrillful, delirious, frantic, unrestrained, activated, reactive, aflutter, nervous, mad, wild

Applying the synonyms in Table I to the above data collects "annoy" and "bothersome" (Table II) as synonyms of the "annoying" Thayer mood, giving a count of 40 for this mood. "Angered" is the only synonym of "angry", giving that mood a count value of 15. Likewise, "tense" and "anxious" are synonyms of "nervous", and thus the count of "nervous" is 17. "Excite" is the only term contributing to the count for "excited", which thus has a value of 15. Given that "excited", "annoying", "angry", and "nervous" are, respectively, the 3rd, 4th, 5th, and 6th moods in the Thayer model, the mood vector for multimedia item k is as follows:

$$V_k = (0, 0, 15, 40, 15, 17, 0, 0, 0, 0, 0, 0)$$

Finally, the mood fuzzy vector of a multimedia content is obtained, as indicated in (2). Each dimension value of the vector indicates membership of the mood that is pertinent to the dimension. For instance, the CF_k value of V_k explained above is (0.00, 0.00, 0.17, 0.46, 0.17, 0.20, 0.00, ..., 0.00).

$$CF_k = \left(\frac{V_k^1}{SV_k}, \frac{V_k^2}{SV_k}, \dots, \frac{V_k^{12}}{SV_k}\right),\ SV_k = \sum_{i=1}^{12} V_k^i \tag{2}$$

where k is the index of a multimedia content, V_k^i the count value of i' th mood of multimedia content k, SV_k the total mood count value of multimedia content k, and CF_k the mood fuzzy vector of multimedia content k.

In the work by Moon et al. [10], mood count values of 12 moods are converted to an AV value that is used as an internal tag of multimedia content. During the conversion process, the mood information is partially lost, resulting in degrading retrieval performance. For improving retrieval performance, in this paper, the possibility of each mood is calculated by equation (2). Also, considering this possibility, the similarity between the query and the multimedia contents is calculated.

C. Generating Mood Fuzzy Vector for Tags

The mood fuzzy vector of a tag is requisite to search multimedia contents by using mood fuzzy vectors of multimedia contents. The mood fuzzy vector of a tag is calculated by the process shown in Fig. 4. The mood fuzzy table of tags is generated as the outcome of this process. The mood fuzzy vector of a tag is calculated as shown in (3) as the mean of the mood fuzzy vectors of multimedia contents, which includes the mood tag:

$$TF_t = \frac{\sum_{i=1}^{l(t)} CF_i}{l(t)} \tag{3}$$

where TF_t is the fuzzy vector of tag t; $l(t)$ is the number of multimedia contents that includes tag t; and CF_i is the fuzzy vector of content i.

For example, in Fig. 4, the mood fuzzy vector of "Mood Tag (1)" can be calculated by the formula average($CF_1, CF_2, \dots, CF_i, CF_{l(1)}$), where n is the number of multimedia contents; and CF_i is the fuzzy vector of i'th multimedia contents that include "Mood Tag (1)".

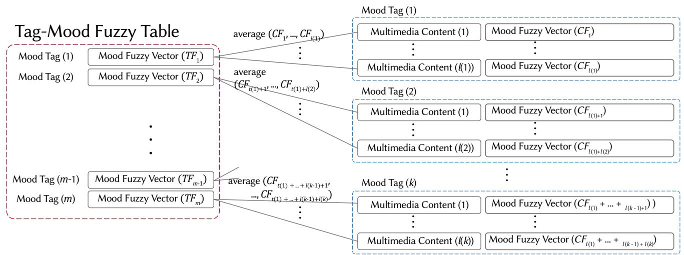


Fig. 4. Tag-mood fuzzy table creation process.

For example, assume the mood fuzzy vector of multimedia content in Table III. From the data, the mood fuzzy vector of the 'happy' tag is generated as follows: the 'happy' tag exists in the multimedia content 1 and 4, and thus the mood fuzzy vector is calculated as {(0.0, ..., 0.22, 0.45, 0.25, 0.07, 0.0) + (0.0, ..., 0.24, 0.63, 0.14)}/2 = (0.00, ..., 0.11, 0.22, 0.24, 0.35, 0.07). The 'sad' tag exists at 2, 5, 8, 9, and 10, so the mood fuzzy vector of the 'sad' tag is calculated as {(0.0, 0.0, 0.0, 1.0, ..., 0.0) + (0.0, 0.32, 0.32, 0.36, ..., 0.0) + (0.0, 0.42, 0.44, 0.14, ..., 0.0) + (0.0, 0.0, 0.17, 0.46, 0.17, 0.2, ..., 0.0) + (0.0, 0.0, 0.27, 0.41, 0.33, ..., 0.0)}/5 = (0.0, 0.15, 0.24, 0.47, 0.10, 0.04, ..., 0.0).

D. Fuzzy-Based Multimedia Content Retrieval

Fig. 5 shows the two stages of multimedia content retrieval based on mood fuzzy vectors of mood tags. The first stage involves searching the mood fuzzy vector of a tag, and the second comprises similarity calculation.

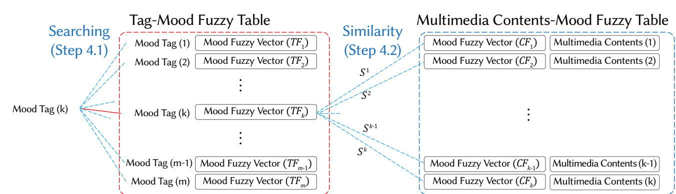


Fig. 5. Multimedia content retrieval method using multimedia content-mood fuzzy table and tag-mood fuzzy table.

After obtaining the mood fuzzy vector of the tag in the user query from the table of mood fuzzy vectors for tags, the similarity between it and the mood fuzzy vector of each multimedia item is calculated, as shown on the right-hand side of Fig. 5. Multimedia content with

TABLE III. Example of the Mood Fuzzy Vector of Multimedia Content

Multimedia Content No	Tag 1	Tag 2	Tag 3	Tag 4	...	Tag n	annoying	angry	nervous	sad	bored	sleepy	calm	peaceful	relaxed	pleased	happy	excited
1	peaceful	relaxed	pleased	tag	...	happy	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.45	0.25	0.07	0.00
2	sad				...		0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	rag	angry	nervous		...		0.38	0.31	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	amused	happy	excited		...		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.63	0.14
5	angry	nervous	sad		...		0.00	0.32	0.32	0.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	peaceful	relaxed	pleased		...		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33	0.33	0.33	0.00	0.00
7	rag	angry	nervous		...		0.21	0.57	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	angry	nervous	sad		...		0.00	0.42	0.44	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	sad	sleep	tense	tag	...	asleep	0.00	0.00	0.17	0.46	0.17	0.20	0.00	0.00	0.00	0.00	0.00	0.00
10	sad	tire	nervous		...		0.00	0.00	0.27	0.41	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00

high similarity is preferentially retrieved. Similarities are as shown in Table IV.

TABLE IV. Similarities and their Equations

Similarity	Equation
Similarity of [11]	$s_k(TF_k, CF_l) = \frac{1}{m}\sum^{m}\left(\frac{min(TF_k^i, CF_l^i)}{max(TF_k^i, CF_l^i)}\right)$ (4) where the denominator equalizes to zero, then $s_k(TF_k, CF_l) = 1$.
Similarity of [12]	$s_k(TF_k, CF_l) = \sum_{i=1}^{m}\frac{2 \times min(TF_k^i, CF_l^i)}{TF_k^i + CF_l^i}$ (5) where if $TF_k^i + CF_l^i = 0$, then $s_k(TF_k, CF_l) = 1$.
Similarity of [13]	$s_k(TF_k, CF_l) = \frac{\sum_{i=1}^{m} min(TF_k^i, CF_l^i)}{\sum_{i=1}^{m} max(TF_k^i, CF_l^i)}$ (6)
Similarity of [14]	$s_k(TF_k, CF_l) = \frac{1}{\sqrt{m}}\sqrt{\sum_{i=1}^{m}(TF_k^i - CF_k^i)^2}$ (7)

In the equation, $s_k(TF_k, CF_l)$ means similarity between mood fuzzy vector (TF_k) of tag k and mood fuzzy vector (CF_l) of l'th multimedia content; and TF_k^i and CF_l^i mean i'th value of each mood fuzzy vector.

IV. Experiments

To analyze the performance of the proposed method, approximately 50,000 multimedia items and their tags are collected from the API of the website Last.fm. In detail, the dataset contains the following numbers of each tag: 3,200 "angry"; 1,900 "annoying"; 40 "bored"; 10,000 "calm"; 70 "excited"; 10,300 "happy"; 170 "nervous"; 4,800 "peaceful"; 10 "pleased"; 1,900 "relaxed"; 10,200 "sad"; and 7,500 "sleepy". The dataset is used to construct the mood fuzzy tables of the multimedia content and of the tags. Then, the retrieval performance and distribution of the mood fuzzy vectors of the content and tags are analyzed using these tables. The recall and precision are calculated to evaluate the retrieval performance. After obtaining the interpolated precision at 10 recall levels [31], the average precision is calculated for the 12 Thayer mood words. The mood count vectors of multimedia contents and tags and the pseudocodes for calculating the mood fuzzy vectors from them are available at [32].

The recall and precision are calculated by (8), the interpolated precision is calculated by equation (9) and the average precision is calculated by (9).

$$P = \frac{|B|}{|B| + |C|},\ r = \frac{|B|}{|A| + |B|}, \tag{8}$$

where, r is tag recall, P precision, A false negative, C false positive and B true positive.

$$P(r_j) = max_{r_i \le r \le r_{i+1}} P(r) \tag{9}$$

where, $P(r)$ is a precision of recall level r and $r_i \le r \le r_{i+1}$ range of recall level. In this paper, r value range is 0.1.

$$\overline{P(r)} = \sum_{i=1}^{N_q}\frac{P_i(r)}{N_q} \tag{10}$$

where, N_q is number of queries, $P_i(r)$ precision of i'th query in recall r.

A. Analysis of the Mood Fuzzy Vectors of Multimedia Contents and Tags

Prior to analysis of retrieval performance, the fuzzy table of mood tags and the fuzzy table of multimedia contents are analyzed based on 12 mood tags of the Thayer two-dimensional model. According to the analysis, the membership value of the relevant mood of each mood tag (e.g., "pleased" mood for "pleased" tag, "happy" mood for "happy" tag, etc.) is high, as shown in Fig. 6. Among them, the tags "pleased", "excited" and "nervous" have a large membership value of relevant tag mood, and the membership values of the other moods are also large compared to other mood tags. This may be ascribed to the fact that the mood fuzzy vectors for the three tags are calculated by use of fewer than 200 multimedia content information, while more than approximately 1,000 for the residual nine tags.

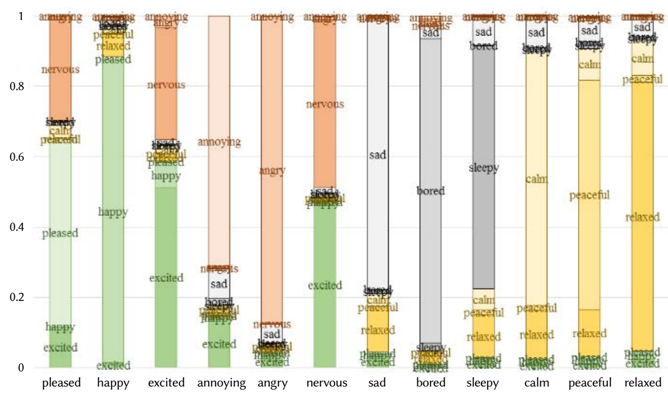
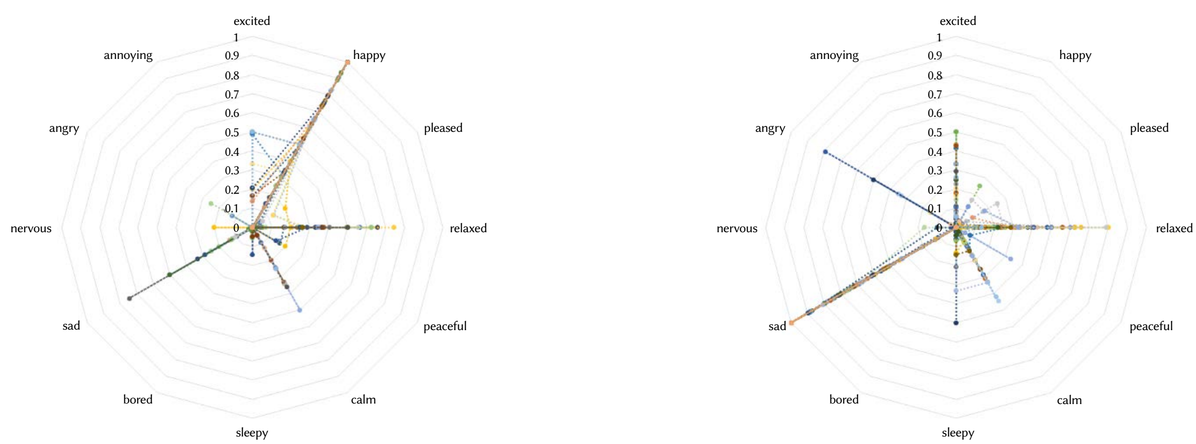


Fig. 6. Multimedia content retrieval method using multimedia content-mood fuzzy table and tag-mood fuzzy table.



(a) Multimedia content that includes the happy tag

(b) Multimedia content that includes the sad tag

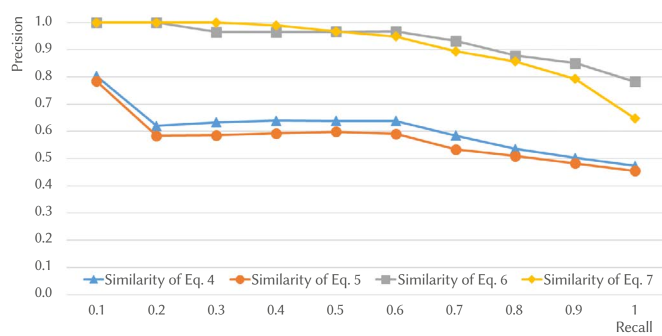
Fig. 7. Analysis of the mood fuzzy table of multimedia content.

TABLE V. ANOVA Test (MFV: Mood Fuzzy Vector)

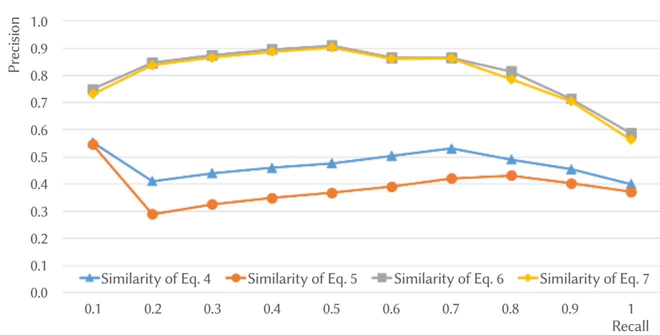
		Sum of Squares	DF	Mean Square	F	P-Value		Sum of Squares	DF	Mean Square	F	P-Value
Mood		700.730	11	63.703	11028.258	.000		3480.844	11	316.440	10382.104	.000
Error	**MFV1**	265.121	45898	.006			**MFV7**	1398.944	45898	.030		
Total		965.851	45909					4879.788	45909			
Mood		2037.450	11	185.223	36203.221	.000		1624.819	11	147.711	10769.730	.000
Error	**MFV2**	234.823	45898	.005			**MFV8**	629.508	45898	.014		
Total		2272.273	45909					2254.326	45909			
Mood		43.396	11	3.945	6167.032	.000		807.645	11	73.422	1984.547	.000
Error	**MFV3**	29.362	45898	.001			**MFV9**	1698.088	45898	.037		
Total		72.758	45909					2505.733	45909			
Mood		3800.000	11	345.455	10222.222	.000		3.038	11	.276	161.449	.000
Error	MFV4	1551.098	45898	.034			**MFV10**	78.505	45898	.002		
Total		5351.098	45909					81.543	45909			
Mood		27.451	11	2.496	4672.943	.000		5439.302	11	494.482	32791.637	.000
Error	MFV5	24.512	45898	.001			**MFV11**	692.120	45898	.015		
Total		51.963	45909					6131.422	45909			
Mood		2686.013	11	244.183	12243.833	.000		68.857	11	6.260	609.643	.000
Error	MFV6	915.360	45898	.020			**MFV12**	471.271	45898	.010		
Total		3601.373	45909					540.128	45909			

TABLE VI. Levene Test (MFV: Mood Fuzzy Vector)

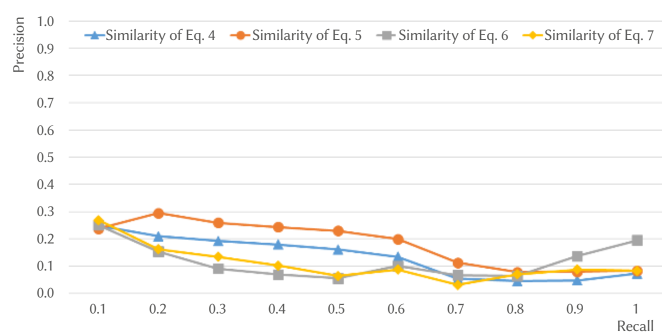
	MFV1	MFV2	MFV3	MFV4	MFV5	MFV6	MFV7	MFV8	MFV9	MFV10	MFV11	MFV12
Levene	10816.91	3610.548	454.9322	1061.525	746.1625	11109.76	4558.702	10067.41	908.7034	54.3630076	3547.89	778.6744
DF1	11	11	11	11	11	11	11	11	11	11	11	11
DF2	45898	45898	45898	45898	45898	45898	45898	45898	45898	45898	45898	45898
P-Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000



(a) with synonyms



(b) without synonyms



(c) calculated by (a)-(b)

Fig. 8. Retrieval performance comparison by 12 mood tags.

The fuzzy table of multimedia content is analyzed by randomly selecting 200 multimedia contents, including “happy” or “sad”, which have the highest number of multimedia contents. The analysis shows that “happy” has a large membership value as shown in Fig. 7 (a), and “sad” also has a large membership value as shown in Fig. 7 (b). For reference, the remaining 10 moods also exhibit a propensity that is similar to that of “happy” and “sad”.

This paper performs an ANOVA test and a test for equality of variances (the Levene test) to determine whether they have an independent distribution for mood fuzzy vectors of the fuzzy table of multimedia content. The 12 moods (happy, sad, annoying, pleased, excited, nervous, bored, sleepy, calm, peaceful, relaxed, and angry) of multimedia content are selected as independent variables, and the mood fuzzy vector of each multimedia content is selected as dependent variables. The outcome of the experiment is shown in Tables V and VI. Since null hypothesis H_0 can be rejected because all p-values are 0.000, alternative hypothesis H_1 can be adopted. Specifically, it can be concluded that the difference in distribution of mood fuzzy vectors and the difference in dispersion arise when the mood fuzzy vectors of multimedia content are classified into 12 mood groups.

B. Analysis of Fuzzy-Based Retrieval Performance Using Mood Tags

To analyze the retrieval performance of the proposed method, we perform three kinds of comparisons. The first is the comparison of retrieval performance for four similarities in Table IV, and the second is the comparison of the retrieval performance of the proposed method with the keyword-based method of Last.fm. The last is the comparison of the retrieval performance with the method based on the AV value [10]. The first experiment shows that when analogue is considered, methods using formulas (6) and (7), as shown in Fig. 8 (a), achieve performance of precision 1.0 at recall level 0.1 and 0.2; methods using formulas (4) and (5) achieve performance below precision 0.8 at recall level 0.1 and performance below precision 0.65 at recall level 0.2. Thus, the method that uses formulas (6) and (7) shows better performance than the method that uses formulas (4) and (5) overall. When analogue is not considered, however, the method that uses formulas (6) and (7) also achieves better performance than the method that uses formulas (6) and (7) (refer to (b) and (c)).

When comparing the keyword-based retrieval performance of Last.fm with the retrieval performance of the proposed method, the same performance is achieved with precision 1.0 at recall level 0.1 and 0.2, as shown in Fig. 9. However, the proposed method (the method that uses formulas (6) and (7)) shows better performance than the keyword-based retrieval method from recall level 0.3. Although the keyword-based retrieval performance approaches precision 0.0 when it approaches recall level 1.0, the proposed method remains above precision 0.6 even when it approaches recall level 1.0. Consequently, it can be concluded that the proposed method (the method that uses formulas (6) and (7)) shows better performance than the keyword-based retrieval method overall (refer to Fig. 9).

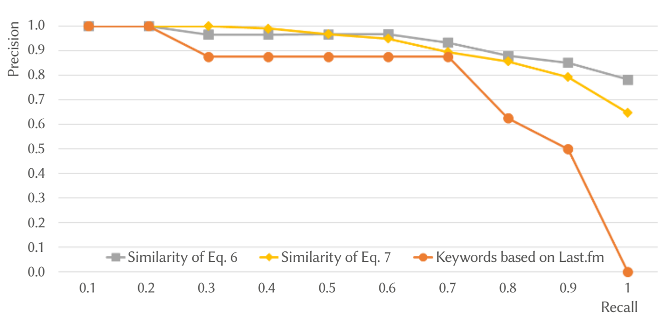


Fig. 9. Retrieval performance comparison between the proposed method and the Last.fm method.

Finally, comparing the retrieval performance of the AV-based method in [10] with that of the proposed method, the proposed method (the method that uses formulas (6) and (7)) achieves excellent performance, as shown in Fig. 10. The gap in retrieval performance is larger than precision of 0.1 at all recall levels. Importantly, the problem of the existing AV-based method at recall level 0.1 and 0.2, i.e., the reduction of retrieval performance, can be solved. In Figure 10, case #1 means retrieval performance in which tags having more than 1,000 multimedia contents are used, and case #2 means retrieval performance in which all 12 mood tags are used.

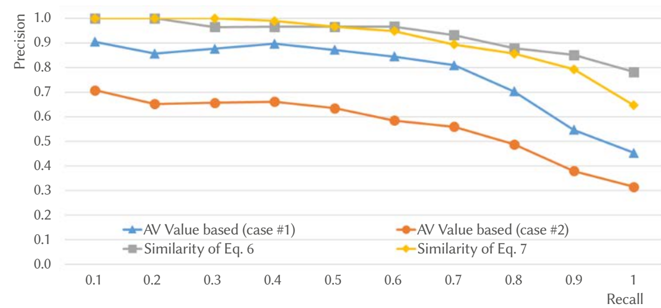


Fig. 10. Retrieval performance comparison between the proposed method and the AV-based method [10].

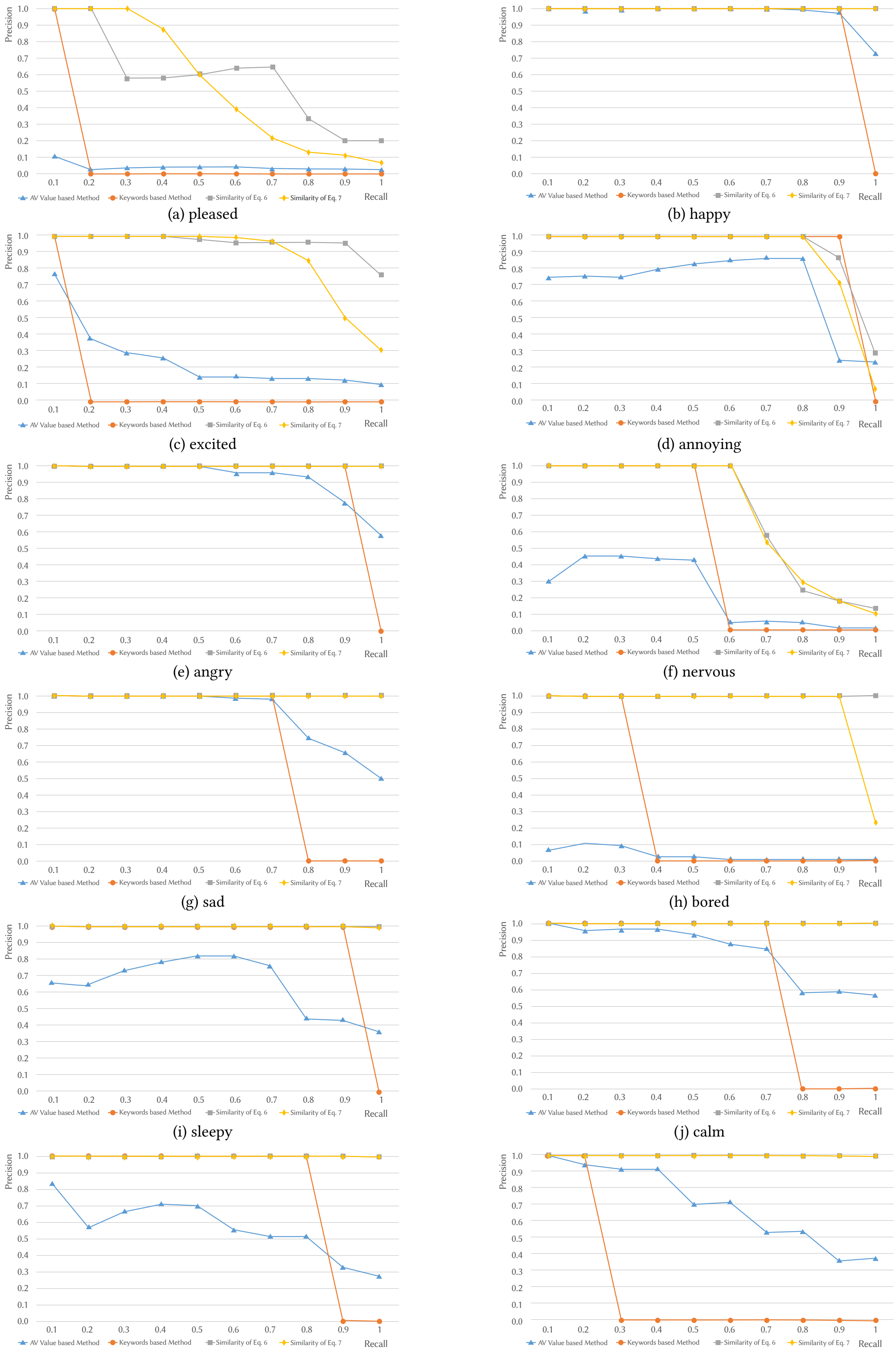
To analyze the retrieval performance in detail, the AV-based retrieval method [10], the keyword-based method of Last.fm, and the proposed method (the method that uses formulas (6) and (7)) are compared and analyzed on the basis of the 12 moods. According to the comparative analysis, the retrieval method of the proposed method shows excellent performance overall, as shown in Fig. 11. Moreover, the proposed method achieves excellent performance even with tags (pleased, excited, bored, and nervous) below 1,000 multimedia contents, for which the AV-based method [10] reveals performance degradation. Overall, the proposed method achieves superior performance compared to the extant methods.

V. Conclusion

The propensity of information purchasers on the Web is changing markedly from cost-effective to cost-satisfaction consumption. One of the methods to maximize cost-satisfaction constitutes utilizing multimedia contents’ mood. Some SNSs provide such a service through the use of folksonomy. However, folksonomy presents some problems, such as synonym and new coinage. To solve the synonym problem, a study [10] attempted to represent multimedia content with AV value (arousal and valence) of the Thayer model, but retrieval performance was inferior to the keyword-based method at recall level 0.1.

To overcome the challenge in [10], this paper proposed a method that uses fuzzy moods of multimedia content based on the Thayer mood model. This proposed method solved the synonym problem, as well as achieved retrieval performance that was the same as the keyword-based method at recall level 0.1. The keyword-based retrieval method does not take into account synonym and so suffer from performance degradation from recall level 0.3. In [10], some mood information is lost and thus performance degradation at recall level 0.1 and 0.2 is caused. However, the proposed method does not.

The method proposed in this paper was analyzed in two ways. The first is the analysis of the mood fuzzy table of tags and the mood fuzzy table of multimedia content that constituted intermediate outputs of the proposed method. The analysis demonstrated that the distribution of the mood fuzzy vectors of multimedia content is similar to the distribution of the mood fuzzy vector of tags. In particular, we conducted an ANOVA test and a test for equality of variances (the



(a) pleased
(b) happy
(c) excited
(d) annoying
(e) angry
(f) nervous
(g) sad
(h) bored
(i) sleepy
(j) calm
(k) peaceful
(l) relaxed

Fig. 11. Retrieval performance comparison per mood tag.

Levene test) using the mood fuzzy table of multimedia content. Subsequently, it was found that a difference of mood fuzzy vector distribution and dispersion occurred.

Second, the retrieval performances of the three methods – the keyword-based method of Last.fm, the AV-based retrieval method [10], and the proposed method - were analyzed. The analysis revealed that the problem of the AV-based retrieval method (the retrieval performance of multimedia content below 1,000 at recall level 0.1 and 0.2 is inferior to the retrieval performance of the keyword-based retrieval method) could be solved, and retrieved objects were increased by more than 50% when synonyms are considered. Overall, the method proposed in this paper is far superior to that of the keyword-based method of Last.fm and the AV-based retrieval method.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1F1A104833611, 2021R1I1A1A01042270).

References

[1] C. B. Moon, J. Y. Yi, D.-S. Kim, B. M. Kim, "Analysis of Overlapping Mood Tags Based on Synonyms," *Korea Computer Congress 2018 (KCC 2018), KIISE (2018)*, June 20-22; ICC JEJU, Korea, pp.667-669, 2018, doi: 10.9723/JKSIIS.2019.24.1.013.

[2] Russel, J. A., "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, Vol.39, No.6, pp.1161-1178, 1980, doi: 10.1037/h0077714.

[3] Hevner, K., "Experimental studies of the elements of expression in music," *The American Journal of Psychology*, Vol.48, No.2, pp.246-268, 1936, doi: 10.2307/1415746.

[4] Thayer, R. E., *The Biopsychology of Mood and Arousal*, Oxford University Press, 1990.

[5] C. B. Moon, H. S. Kim, B. M. Kim, "Music Retrieval Method using Mood Tag and Music AV Tag based on Folksonomy," *Journal of KIISE*, Vol.40, No.9, pp.526-543, 2013, doi: 10.9723/jksiis.2019.24.1.013.

[6] C. B. Moon, H.S. Kim, H. A. Lee, B. M. Kim, "Analysis of Relationships Between Mood and Color for Different Musical Preferences," *Color Research & Application*, Vol.39, No.4, pp.413-423, 2014, doi: 10.1002/col.21806.

[7] C. B. Moon, H.S. Kim, D. W. Lee, B. M. Kim, "Mood Lighting System Reflecting Music Mood," *COLOR research and application*, Vol.40, No.2, pp.201-212, 2015, doi: 10.1002/col.21864.

[8] Viskovich, Shelley Narelle, *Evaluation of a web-based Acceptance and Commitment Therapy (ACT) program to increase self-care and mental health skills in university students*, PhD Thesis, School of Psychology, The University of Queensland, 2019.

[9] Klinenberg, E., *Going solo: the extraordinary rise and surprising appeal of living alone*, London: Penguin Books, 2012.

[10] C. B. Moon, J. Y. Lee, D.-S. Kim, B. M. Kim, "Multimedia Content Recommendation in Social Networks using Mood Tags and Synonyms," *Multimedia Systems*, 26, pp.139–156, 2020, doi: 10.1007/s00530-019-00632-w.

[11] W.-J. Wang, "New similarity measures on fuzzy sets and on elements," *Fuzzy sets and systems*, vol. 85, pp. 305–309, 1997, doi: 10.1016/0165-0114(95)00365-7.

[12] D.-G. Wang, Y.-P. Meng, and H.-X. Li, "A fuzzy similarity inference method for fuzzy reasoning," *Computers and mathematics with applications*, vol. 56, pp. 2445–2454, 2008, doi: 10.1016/j.camwa.2008.03.054.

[13] C. P. Pappis and I. Karacapilidis, "A comparative assessment of measures of similarity of fuzzy values," *Fuzzy sets and systems*, vol. 56, pp. 171–174, 1993, doi: 10.1016/0165-0114(93)90141-4.

[14] L. Baccour, A. M. Alimi, R. I. John, "Some notes on fuzzy similarity measures and application to classification of shapes recognition of arabic sentences and mosaic," *IAENG International Journal of Computer Science*, vol. 41, no. 2, pp. 81-90, 2014.

[15] Huang, W., Wu, Q., Dey, N., Ashour, A., Fong, S. J., & González-Crespo, R., "Adjectives Grouping in a Dimensionality Affective Clustering Model for Fuzzy Perceptual Evaluation," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 10, pp. 28-37, 2020, doi: 10.9781/ijimai.2020.05.002

[16] Ness, S. R., Theocharis, A., Tzanetakis, G. and Martins, L. G., "Improving automatic music tag annotation using stacked generalization of probabilistic svm outputs," *Proc. of the 17th ACM international conference on Multimedia*, pp.705-708, 2009, doi: 10.1145/1631272.1631393.

[17] Laurier, C., Sordo, M., Serra, J. and Herrera, P., "Music mood representations from social tags," *Proc. of the 10th International Society for Music Information Conference*, Kobe, Japan, pp.381-386, 2009.

[18] Kim, J., Lee, S., Kim, S. and Yoo, W. Y., "Music mood classification model based on arousal-valence values," *Proc. of 13th International Conference on Advanced Communication Technology (ICACT*), 2011, pp.292-295, 2011.

[19] A. T. Ji, et al., "Collaborative Tagging in Recommender Systems," *AI 2007: Advances in Artificial Intelligence*, pp. 377–386, 2007, doi: 10.1007/978-3-540-76928-6_39.

[20] K. H. L. Tso-Sutter, et al., "Tag-aware Recommender Systems by Fusion of Collaborative Filtering Algorithms," *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*, pp. 1995–1999, 2008, doi: 10.1145/1363686.1364171.

[21] M. Vojnovic, et al., "Ranking and Suggesting Popular Items," *IEEE Transactions on Knowledge and Data Engineering*, vol.21, pp.1133-1146, 2009, doi: 10.1109/TKDE.2009.34.

[22] S. Yang, S. Kim and Y. M. Ro, "Semantic Home Photo Categorization," *IEEE Trans. Circuits and Systems for Video Technology*, Vol.17, No.3, March 2007, pp.324-335, 2007, doi: 10.1109/TCSVT.2007.890829.

[23] S. Yang, S. K. Kim, K. S. Seo, Y. M. Ro, J. Kim and Y. S. Seo, "Semantic categorization of digital home photo using photographic region templates," *International Journal of Information Processing and Management*, Vol.43, No.2, March 2007, pp.503-514, 2007, doi: 10.1016/j.ipm.2006.07.009.

[24] S. Yang, R. M. Ro, "Photo Indexing Using Person-Based Multi-feature Fusion with Temporal Context," *Proc. International Conference on Mobile Ubiquitous Computing, Systems*, Services and Technologies, Nov. 2007, pp.257-262, 2007, doi: 10.1109/UBICOMM.2007.29.

[25] E. Chang, G. Kingshy, G. Sychay, G. Wu, "CBSA: Content-Based Soft Annotation for Multimodal Image Retrieval using Bayes Point Machines," *IEEE Trans. Circuits and Systems for Video Technology*, Vol.13, No.1, January 2003, pp.26 – 38, 2003, doi: 10.1109/TCSVT.2002.808079.

[26] J. Li and J. Z. Wang, "Real-Time Computerized Annotation of Pictures," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.30, No.6, June 2008, pp.985-1002, 2008, doi: 10.1109/TPAMI.2007.70847.

[27] R. M ö rzinger, R. Sorschag, G. Thallinger and S. Lindstaedt, "Automatic Image Annotation using Visual Content and Folksonomies," *Multimedia Tools and Applications*, Vol.42, No.1, March 2009, pp.97-113, 2009, doi: 10.1007/s11042-008-0247-7.

[28] S.h. Rho, B.-J. Han, and E. Hwang, "A fuzzy inference-based music emotion recognition system," *In Proc. Visual Information Engineering*, pp. 673–677, 2008, doi: 10.1049/cp:20080398.

[29] S. Qamar, P. Ahmad, "Emotion Detection from Text using Fuzzy Logic," *International Journal of Computer Applications*, Vol. 121, No. 3, pp. 29-32, 2015, doi: 10.5120/21522-4501.

[30] Magdin, M., Držík, D., Reichel, J., & Koprda, S., "The Possibilities of Classification of Emotional States Based on User Behavioral Characteristics," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 4, pp. 97-104, 2020, doi: 10.9781/ijimai.2020.11.010

[31] D. Powers, "Evaluation: From precision recall and f-measure to roc informedness markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37-63, 2011, doi: 10.9735/2229-3981.

[32] https://drive.google.com/file/d/18KmHezs55_fMNj95XgceIDYOhhxu3HmZ/view

Chang Bae Moon

Chang Bae Moon received a BSc, an MSc, and a PhD from the Dept. of Software Eng. at Kumoh National Institute of Technology, Korea, in 2007, 2010, and 2013, respectively. He has been with the Kumoh National Institute of Technology since 2014 as a Research Professor in the ICT Convergence Research Center. From 2013 to 2014, he was a Senior Researcher in Young Poong Elec. Co. His current research areas include artificial intelligence, Web intelligence, information filtering, and image processing.

Jong Yeol Lee

Jong Yeol Lee received the BS and MS degree in Dept. of computer Eng. from Kumoh National Institute of Technology, Korea, in 1992 and 1994, respectively and the PhD candidate in software Eng. from Kumoh National Institute of Technology, Korea, in 2018. He has been with Kumoh National Institute of Technology since 2005 as a time lecturer of Computer Software Engineering Department. His current research areas include Artificial Intelligence, Machine Learning and Information Security.

Byeong Man Kim

Byeong Man Kim received the BS degree in Dept. of computer Eng. from Seoul National University (SNU), Korea, in 1987, and the MS and the PhD degree in computer science from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1989 and 1992, respectively. He has been with Kumoh National Institute of Technology since 1992 as a faculty member of Computer Software Engineering Department. From 1998 – 1999, he was a post-doctoral fellow in UC, Irvine. From 2005 - 2006, he was a visiting scholar at Dept. of Computer Science of Colorado State University, working on design of a collaborative Web agent based on friend network. His current research areas include artificial intelligence, Web intelligence, information filtering and brain computer interface.