Proceedings of the International MultiConference of Engineers and Computer Scientists 2016 Vol I,
IMECS 2016, March 16 - 18, 2016, Hong Kong

Automatic Recipe Metadata Generating Method by Considering Users' Various Moods

Natsuhiko Takata, Mayumi Ueda, Yukitoshi Morishita, and Shinsuke Nakajima

***Abstract*—Recently, users require the recipe recommendation system that provides the recipes reflecting their own purpose. In order to realize such a system, our system uses the metadata of the recipes. However, it is not realistic to add metadata for the huge number of recipes, manually. Therefore, in this paper, we propose an automatic recipe metadata generating method by considering user's various moods. Our system adds the metadata, which expresses the user's moods along to the five aspects, using a similarity of the recipes. We discuss the adequateness of our proposed method based on an evaluation experiment.**

***Index Terms*—cooking recipe recommendation, automatic metadata generation, similarity of recipes, five axis for the mood**

I. Introduction

RECENTLY, numerous cooking websites that recommend cooking recipes have been launched. For example, Cookpad[1] and Rakuten Recipe[2] are very popular in Japan. Cookpad contains 2.2 million recipes and 50 million monthly access users, and Rakuten Recipe contains 1 million recipes. This reflects the high demand for recipe providing services. In order to improve the accuracy of recipe recommendation, we consider that the approach, which add the various metadata to the recipes, is effective approach for the recipe recommendation system. For example, if the recipe have a metadata like a "good for the bedtime snack", the system can provide the effective recipes for user's purpose, as shown in Fig.1.

In our previous work, we developed the system that provides the recipes that suit user's mood[3]. The system contains the 493 recipes added the metadata relevant to the mood. Because the metadata was added manually by experts in nutrition, the metadata expresses precisely the moods. However, it is not realistic to add the metadata by manually for the huge number of recipes.

Therefore, in this paper, we propose an automatic recipe metadata generating method by considering user's various moods. Basic concept of our method is to add the metadata using the similarities between recipes. Our system extracts the feature vector from each recipe by analyzing the recipe name, ingredients and cooking steps. At first, our system adds the metadata using the similarities between the feature vector

Manuscript received January 31, 2016; revised February 11, 2016. This work was supported in part by the MEXT Grant-in-Aid for Scientific Research(C) (#26330351).

N. Takata and S. Nakajima are with Graduate School of Frontier Informatics, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto-city, Kyoto, 603-8555, Japan. e-mail: ({i1558110, nakajima}@cse.kyoto-su.ac.jp).

M. Ueda is with Faculty of Economics, University of Marketing and Distribution Sciences, 3-1 Gakuen-Nishimachi, Nishi-ku, Kobe, Hyogo 651-2188, Japan. e-mail: (Mayumi_Ueda@red.umds.ac.jp)

Y. Morishita is with Dai Nippon Printing Co., Ltd., Tokyo 141-8001, Japan. e-mail: (Morishita-Y4@mail.dnp.co.jp)

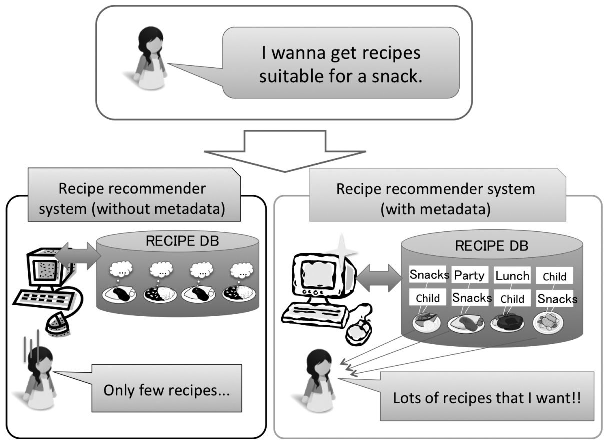

Fig. 1. The advantage of the recipe given metadata

of master recipes(recipes that have been given metadata manually) and the feature vector of target unlabeled recipes. And second, in order to improve the accuracy of adding the metadata, we extract the feature vector focusing on the five aspects related to the moods; for example, body, taste, time, money, and routine. Our system extracts the feature vectors for each aspect of the moods. Then calculate the similarities for each aspect between the feature vector of master recipes and the feature vector of unlabeled recipes.

This paper is structured as follows: the related work is give in Section II. Then in Section III, we describe our method for calculating the similarity of recipes, and the automatic metadata generation. In Section IV, we describe the method to add the metadata according to the user's various moods. And we offer our conclusions in Section V.

II. Related Work

There are many researches about recipe recommendation. In our previous work, we proposed a recipe recommendation method based on the user's culinary preferences and the quantity of each ingredient in a recipe[4]. The system estimates the user's preference using the user's recipe browsing and cooking history. Then, our method adds scores using the average and dispersion quantity of each ingredient in the recipe.

Karikome et al. proposed a system that helps users plan nutritionally balanced menu and visualize their dietary habits[5]. Their system calculates the nutritional value of each dish and records this information in a dietary log. The system then recommends recipes that foster nutrition. Shirai et al. developed the dictionary for the cooking actions[6]. This dictionary contains the information corresponding to the text in the recipe and cooking animation. Their final goal is to develop the system that generates cooking animation by analyzing recipes. Shidochi et al. proposed a method to find

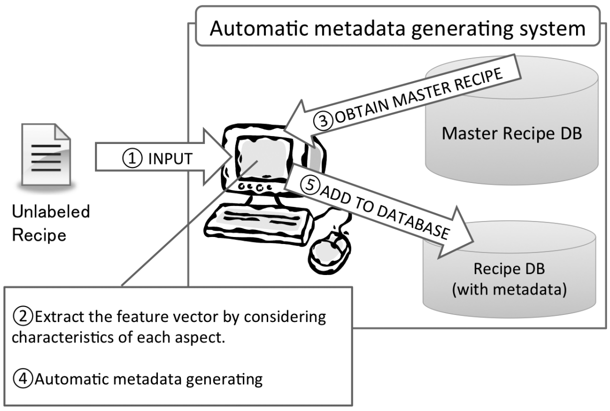

Fig. 2. Conceptual diagram of automatic recipe metadata generating system based on similarities between recipes

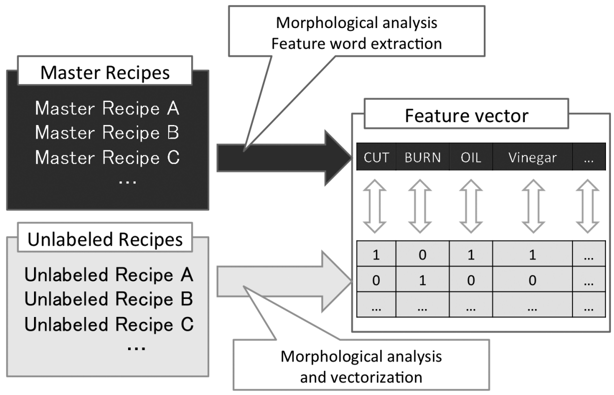

Fig. 3. Extracting the feature vector of recipes

replaceable materials in cooking recipe texts[7]. In order to find replaceable materials, they analyze the large amount of recipes. Then they extract materials and cooking actions in the same recipe group. In their method, the materials related to the same cooking action is the replaceable materials. Ueta et al. proposed a recipe recommendation system considering the nutritional information[8]. Their system accepts the natural language as a user's input, such as "want to cure my acne". In order to realize such system, they develop a co-occurence database that contains nutritional information and nouns in the recipes. Tachibana et al. proposed a method to extract the "Naming Concepts" for recipes that express the characteristics of the recipe[9]. Their method extracts the Naming Concepts by extracting the difference between the element of target recipe and the typical elements.

III. Evaluating recipe metadata genarated based on similarities between recipes

We explain the automatic metadata generating method based on similarities between recipes. In this research, we already have 493 recipes which have been already given metadata manually (i.e. master recipes). Our method can automatically generate recipe metadata for unlabeled recipes by using similar master recipes to the target unlabeled recipes.

A. *Recipe metadata generating method based on similarities between recipes*

Fig.2 shows a conceptual diagram of automatic recipe metadata generating system based on similarities between recipes. Our system uses the recipes that have been given metadata manually as master recipes. Metadata is expressed the user's mood along to the five aspects. The value of each metadata is from -5 to 5. The five aspects are given below.

- body （ Tired ↔ Cheerful ）
- taste （ Non-fatty ↔ Rich taste ）
- time （ Easy ↔ Genuine ）
- money （ Low priced ↔ Gorgeous ）
- modify （ Classic ↔ Modified ）

Our method generate recipe metadata using the similarities between the master recipe and the target unlabeled recipe. In order to calculate the similarities between recipes, our method uses the feature vector of the recipes.

Cooking recipe contains many elements, for example, recipe name, ingredients, cooking steps, nutrition, images. However, it is difficult to consider every elements for calculating similarities. Hence, our method extracts the feature vector using the recipe name, ingredients, and cooking steps. We define the dimension of the feature vector on the basis of the feature vector of master recipes.

Fig.3 shows the method for extracting the feature vector of recipes. Our method analyzes the master recipes using the Japanese language morphological analyser "MeCab[10]".

We extract morpheme from master recipes, and remove the "stopword" that have no relation to the feature of recipe. We define the dimension of the feature vector as the noun, verb, and adjunction extracted from the recipes. In addition, we adopt cosine similarity as the method for calculating similarity between a master recipe and a unlabeled recipes.

Next, we describe how to generate recipe metadata. The metadata for an unlabeled recipe is calculated using the following formula.

$$U(i) = \frac{S_1 \cdot M_1(i) + S_2 \cdot M_2(i) + S_3 \cdot M_3(i) + S_4 \cdot M_4(i) + S_5 \cdot M_5(i)}{S_1 + S_2 + S_3 + S_4 + S_5}$$

- $U(i)$: Score of an aspect of the moods suitable for the unlabeled recipe
- $M_n(i)$: Score of an aspect of the moods of the master recipe that has nth similarity ranking with the target unlabeled recipe.
- $i = \{\text{body, taste, time, money, modify}\}$
- S_m : Similarity between the target unlabeled recipe and a master recipe that has mth similarity ranking with the unlabeled recipe.

B. *Experimental evaluation*

This section describes an experimental evaluation to verify the appropriateness of metadata generated by method mentioned in section III-A.

1) Procedure of experimental evaluation: The procedure of the experimental evaluation is as follows:

1. The system shows a scenario, which has five kinds of metadata like {body, taste, time, money, modify} to a participant, and also recommends both 30 recipes out of 493 master recipes and 30 recipes out of 30,000 unlabeled recipes.

ISBN: 978-988-19253-8-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

2. The participant gets into the main character of the scenario and gives a score (i.e., gain score) between 0 to 30 to the each recommended recipe by considering if it is suitable for main character or not.

The scenarios and their metadata are as follows.

Scenario 1

Profile: 26 years old, a working person, male
The person would like to find a recipe suitable for a person, who is dead tired from working.
Values of metadata:
body : -5, taste : -3, time : -5, money : 0, modify : -5

Scenario 2

Profile: 33 years old, housekeeper, female
The person would like to find a recipe suitable for her child's birthday party.
Values of metadata:
body : 5, taste : 5, time : 3, money : 3, modify : 3

Scenario 3

Profile: 19 years old, student, male
The person would like to find a recipe, which can be made at a low price.
Values of metadata:
body : 0, taste : 0, time : 0, money : -5, modify : 0

Scenario 4

Profile: 22 years old, student, female
The person would like to find a recipe suitable for the supper after school.
Values of metadata:
body : 2, taste : -2, time : -1, money : 0, modify : 0

Scenario 5

Profile: 45 years old, teacher, male
The person would like to find a recipe suitable for snacks to go with the beer.
Values of metadata:
body : 1, taste : 5, time : -4, money : 0, modify : -3

Scenario 6

Profile: 20 years old, student, female
The person would like to find a recipe of a fancy breakfast, which can be made in a short time.
Values of metadata:
body : 5, taste : -4, time : -3, money : 0, modify : 2

Scenario 7

Profile: 28 years old, housekeeper, female
The person would like to find a recipe suitable for lunch box dishes, which can make her husband pleasant.
Values of metadata:
body : 4, taste : 3, time : 2, money : 4, modify : 5

Scenario 8

Profile: 22 years old, student, female
The person would like to find a recipe suitable for her supper that can be made easily.
Values of metadata:
body : -5, taste : 1, time : -5, money : -2.5, modify : 0

Scenario 9

Profile: 14 years old, student, female
The person would like to find a recipe, which can be easily made since she is poor at cooking.
Values of metadata:
body : 4, taste : 0, time : -5, money : -4, modify : 3.5

Scenario 10

Profile: 30 years old, homekeeper, female
The person would like to find a recipe for lunch, which can be made easily.
Values of metadata:
body : -1.5, taste : -4.5, time : -5, money : -4.5, modify : 2.5

2) Experimental result: Table I shows an example of the gain scores from 0 to 30 given to 30 recommended master recipes based on the each scenario by user.

TABLE I
AN EXAMPLE OF THE GAIN SCORES FROM 0 TO 30 GIVEN TO 30 RECOMMENDED MASTER RECIPES BASED ON THE EACH SCENARIO

Ranking by proposed method	master recipe	gain score
1	mozuku vinegar	30
2	shrimp grilled	15
3	Chirimenjako into grated radish	29
4	instant pickled	14
5	porridge	16
6	Hokke of dried fish	17
7	octopus and cucumber vinegared	13
8	chikuwa cucumber	28
9	octopus carpaccio	12
10	dried mushroom compote	2
11	smelt of dried fish	18
12	boiled tofu	11
13	horse mackerel and open	19
14	red and white trout	10
15	clam juice	9
16	Kasujiru	4
17	grated yam kelp juice	26
18	salt cucumber	27
19	fried squid and celery	3
20	tamago kake gohan	21
21	egg porridge	8
22	hot spring egg	7
23	natto rice	22
24	wheat Torogohan	20
25	cucumber with miso	25
26	seared cucumber	6
27	green beans sesame sauce	5
28	chiken hot pot	1
29	tofu	24
30	Grilled-Mushrooms	23

We adopt normalized Discounted Cumulated Gain (nDCG) in order to verify the metadata automatically generated based on our proposed mrthod. Here, nDCG is corresponding to normalized DCG, which is a evaluation criterion for ranking.

DCG is calculated based on the following formula.

$$DCG_n = rel_1 + \sum_{k=2}^{n} \frac{rel_k}{log_2(k)}$$

- DCG_n corresponds to the value of DCG of top n rankings.
- rel_1 corresponds to the gain score of the first ranking.
- rel_k corresponds to the gain score of the kth ranking.

$nDCG$ is calculated based on the following formula.

$$nDCG = \frac{DCG}{idealDCG}$$

- $idealDCG$ corresponds to the DCG of the correct rankings.

Next, we explain how to calculate DCG and $nDCG$. Table II shows an example of gain score for top 5 rankings. Table III shows an example of ranking result with gain score.

TABLE II
AN EXAMPLE OF GAIN SCORE FOR TOP 5 RANKINGS

Data Name	Gain Score
Data A	5
Data B	4
Data C	3
Data D	2
Data E	1

TABLE III
AN EXAMPLE OF RANKING RESULT WITH GAIN SCORE

Ranking	Data Name	Correct/Incorrect	Gain Score
1	Data B	Correct	4
2	Data A	Correct	5
3	Data X	Incorrect	0
4	Data E	Correct	2
5	Data Z	Incorrect	0

In the case of Table II and Table III, DCG is calculated as follows;
$DCG = 4 + \frac{5}{log_2(2)} + \frac{2}{log_2(4)} = 7.723$

The data X and the data Z are incorrect, so that their gain scores are 0. In this way, DCG can be evaluation criteria for not only correct answer ratio but also its correct ranking. The larger DCG value indicates higher accuracy of the recommendation. The scope of $nDCG$ value is from 0 to 1 because $nDCG$ value is calculated by dividing DCG value by $idealDCG$ value.

In this case, $idealDCG$ and $nDCG$ are calculated as follows;

$$idealDCG = 5 + \frac{4}{log_2(2)} + \frac{3}{log_2(3)} + \frac{2}{log_2(4)} + \frac{1}{log_2(5)} = 14.530$$

$$nDCG = \frac{7.723}{14.530} = 0.532$$

We perform an experimental evaluation. The procedure of the experiment is shown bellow;

1. We prepare 30,000 labeled recipes by generating recipe metadata for 30,000 unlabeled recipes.
2. Against 10 scenarios, system extracts top 30 recipe rankings from 493 master recipes and also top 30 recipe rankings from 30,000 labeled recipes.
3. We extract ideal rankings by ranking these 30 rankings manually.
4. We calculate and compare two kinds of $nDCG$s of both master recipes and labeled recipes.

Result of the experimental evaluation is shown in Fig. 4. Fig. 4 indicates $nDCG$s of labeled recipes are close to $nDCG$s of master recipes in most scenarios, though average $nDCG$ of master recipes is better than average $nDCG$ of labeled recipes. Consequently, we may say that the rankings of labeled recipes as proposed method has good potentials, however we feel that we have to improve the recommendation accuracy for labeled recipe.

IV. AUTOMATIC RECIPE METADATA GENARATING BY CONSIDERING CHARACTERISTICS OF EACH ASPECT

As concluded in section III-B2, we should improve the recommendation accuracy for labeled recipe. Therefore, we propose a method that can automatically generate recipe metadata by using not one feature vector but five different feature vectors to represent five kinds of users' various moods.

As mentioned in section III-A, metadata is expressed the user's mood along to the five aspects. The value of each metadata is from -5 to 5.

・body （Tired ↔ Cheerful）
・taste （Non-fatty ↔ Rich taste）
・time （Easy ↔ Genuine）
・money（Low priced ↔ Gorgeous）
・modify（Classic ↔ Modified）

We describes how to generate recipe metadata by using five different feature vectors to represent five kinds of users' various moods in following section.

A. Feature vector extraction for each aspect

1) Feature vector for "body aspect" : As shown in Fig. 5, the feature vector for "body aspect" is defined by using characteristic expressions appeared in both high-tiredness recipes and high-cheerfulness recipes. They can express the body aspect of recipes more accurately. Thus, we believe that the system can achieve recipe metadata generating with high-accuracy by using such feature vector.

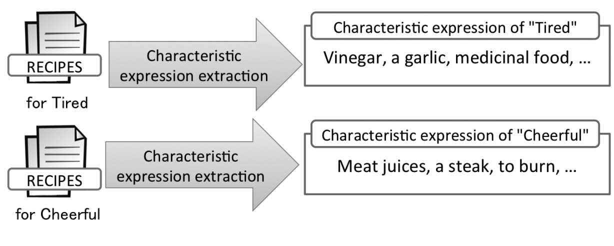

Fig. 5. Method of feature vector extraction for body aspect

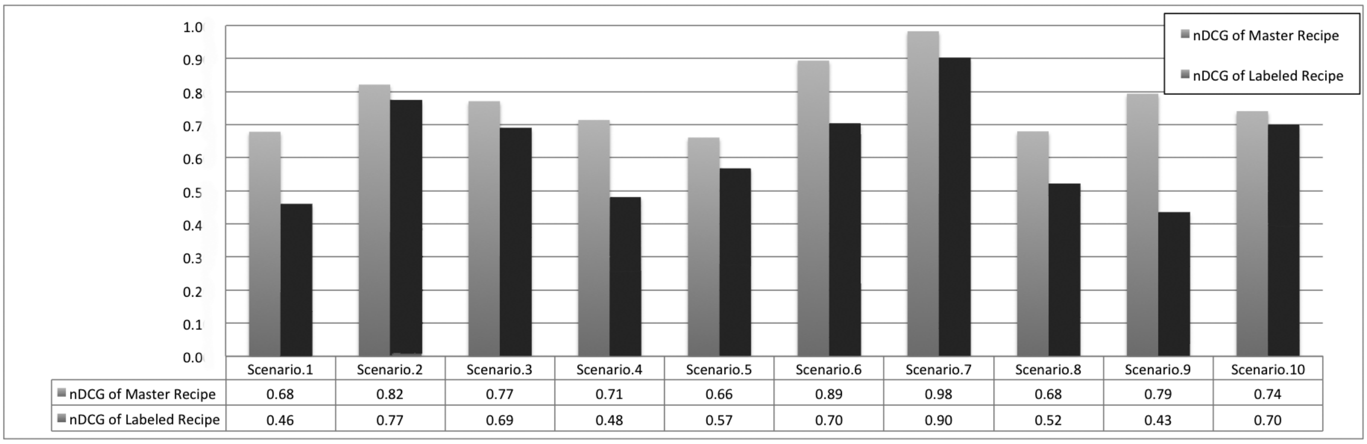

Fig. 4. Comparing $nDCG$s between Master recipe and Labeled recipe

2) Feature vector for "taste aspect": As shown in Fig.6, the feature vector for "taste aspect" is defined by using characteristic expressions appeared in both non-fatty recipes and rich-taste recipes. They can express the taste aspect of recipes more accurately. Thus, we believe that the system can achieve recipe metadata generating with high-accuracy by using such feature vector.

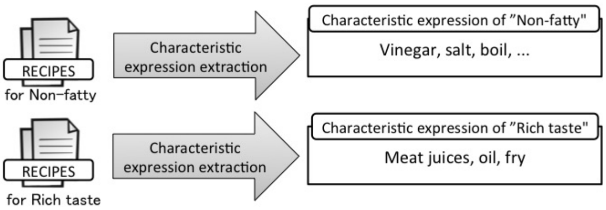

Fig. 6. Method of feature vector extraction for taste aspect

3) Feature vector for "time aspect" : As shown in Fig.7, we extract verbs expressing cooking actions and give a score to each verb based on the taking time for the action. For example, a cooking action "cut" does not take long time, and a cooking action "steam" long time. By using the table, the system can analyze a recipe text and estimate the expected cooking time for the recipe.

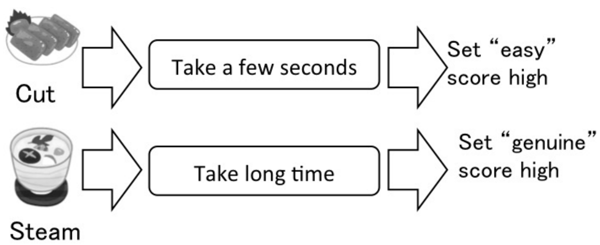

Fig. 7. Method of feature vector extraction for time aspect

4) Feature vector for "money aspect" : As shown in Fig.8, we survey a standard price of each ingredient and register the price into the standard price table at first. Therefore, system can calculate an estimated price of a dish based on the standard price table and the recipe. The value of feature vector for money aspect is normalized from -5 to 5 finally.

5) Feature vector for "arrangement aspect" : As shown in Fig.9, system searches by the name of unlabeled recipe

Fig. 8. Method of feature vector extraction for money aspect

and extracts the verbs and nouns from both unlabeled recipe and collected recipes with the same name. Then, value of arrangement aspect of unlabeled recipe is calculated based on the characteristic keywords and normalized from -5 to 5 finally.

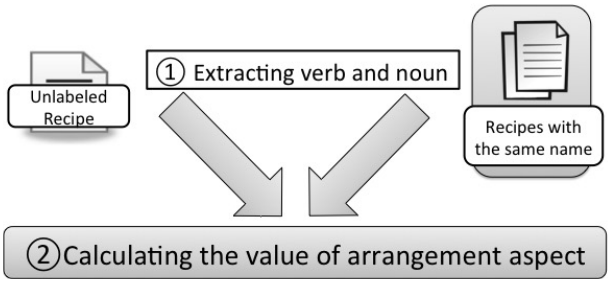

Fig. 9. Method of feature vector extraction for arrengement aspect

B. Preliminary experimental evaluation against feature vector of "body aspect"

The each value of body aspect is normalized from -5 to 5. Thus, we extract characteristics keywords for each body aspect score span that are $[-5,-4], [-4,-3], [-3,-2], [-2,-1], [-1,0], [0,1], [1,2], [2,3], [3,4], and\ [4,5]$ from master recipes, and calculate TF value of the keywords. Then, we adopt some characteristics keywords, whose TF value in high body aspect recipes is much more than the TF values in low body aspect recipes, as parameter of feature vector of body aspect.

We perform experimental evaluation for body aspect as a first step. At first, system generates feature vectors of body aspect against 25 unlabeled recipes based on 2 different methods. One is genarating recipe metadata by using one overall feature vector of a target recipe(i.s., the previous method). Another one is generating recipe metadata by using

ISBN: 978-988-19253-8-1
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

five feature vector customized to each aspect of a target recipe(i.s., the proposed method).

Then, we perform an experimental evaluation by conparing between those metadata generating methods. The participants compare two kinds of body aspect value and judge if which is appropriate to the target recipe. The number of participants is two. The number of trials are 50 (against 50 recipes).

TABLE IV
RESULT OF COMPARATIVE EXPERIMENT BETWEEN THE PREVIOUS METHOD AND THE PROPOSED METHOD

	The rate of average
The feature of a body axis is taken into consideration.	60%
The feature of a body axis is not taken into consideration.	40%

Table IV shows the result of comparative experiment between the previous method and the proposed method. According to the result, the proposed method got better result than the previous method. Namely, we may say that we should use the recipe metadata generaying method by using five feature vector customized to each aspect of a target recipe. As future work, we try to generate more accurate matadata based on improving the five feature vector customized to each aspect.

V. CONCLUSION

In this paper, we presented an automataic recipe metadata generating method for a cooking recipe recommendation system that considers users' various moods. Our method generates the metadata using the similarities between the feature vector of master recipes and the feature vector of the target unlabeled recipes.

At first, we developed a prototype system that recommends the recipes based on the metadata generating by one feature vector. In order to verify the effectivity of our proposed method, we conducted a preliminary experiment. We compared two recipe sets; (1) extract from master recipes (generate metadata manually), (2) extract from unlabeled recipes (generate metadata using our method). The nDCG values of our method is 0.78, and the nDCG values of master recipe is 0.84. In part of the result shows that our method can generate metadata equivaletly. In addition to the above, we considered a method to generate metadata by calculating the similarities using the feature vectors related to the five aspects of users' mood.

Our method generate the metadata for recommending the recipes according to users' mood. We will develop a prototype system to evaluate the accuracy of our recommenndation method that uses the metadata related to the five aspects of users' mood, in the future work.

REFERENCES

[1] Cookpad. https://cookpad.com/us/ (Accessed 21 January 2016)
[2] Rakuten Recipe. http://recipe.rakuten.co.jp/ (Accessed 21 January 2016)
[3] Yukitoshi Morishita and Tomiyo Nakamura, Evaluation of Search Axis of Recipe Recommender System based on Users' Moods and Market Needs of Food Sales Functions by Using Cooking Recipe, IEICE Technical Report, Vol.112,No.75, DE2012-14,pp.79-84, 2012.(in Japanese)
[4] Mayumi Ueda and Shinsuke Nakajima : "Cooking Recipe Recommendation Method Focusing on the Relationship Between User Preference and Ingredient Quantity", Transactions on Engineering Technologies, International MultiConference of Engineers and Computer Scientists 2014, pp.385-395 , Springer Netherlands, Jan. 2015.
[5] Shihono Karikome and Atsushi Fujii. A System for Supporting Dietary Habits: Planning Menus and Visualizing Nutritional Intake Balance. Proceedings of the 4th International Conference on Ubiquitous Information Management and Communication (ICUIMC 2010), pp.386-391, Jan. 2010.
[6] Kiyoaki SHIRAI and Hiroshi OOKAWA, Constructing a Lexicon of Actions for the Cooking Domain toward Animation Generation, IPSJ Natural Language Processing, Vol.2004 No.108, pp.123-128, 2004.(in Japanese)
[7] Yuka SHIDOCHI, Ichiro IDE, Tomokazu TAKAHASHI and Hiroshi MURASE, Finding Replaceable Materials by Cooking Recipe Mining, IEICE Transactions, Vol. J94-A No.7 pp.532-535, 2011.(in Japanese)
[8] Tsuguya Ueta, Masashi Iwakami, and Takayuki Ito, A recipe recommendation system based on automatic nutrition information extraction, Lecture Notes in Computer Science, 7091, pp.79-90, Springer, 2011.
[9] Akiho Tachibana, Shoko Wakamiya, Hidetsugu Nanba, and Kazutoshi Sumiya, Extraction of Naming Concepts Based on Modifiers in Recipe Titles, The 2014 IAENG International Conference on Internet Computing and Web Services, pp.507-512, 2014.
[10] MeCab:Yet Another Part-of-Speech and Morphological Analyzer, http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html